



Comparative Analysis of NADPH-Cytochrome P450 Reductases From Legumes for Heterologous Production of Triterpenoids in Transgenic *Saccharomyces cerevisiae*

Pramesti Istiandari¹, Shuhei Yasumoto^{1,2}, Pisanee Srisawat^{1,3}, Keita Tamura¹, Ayaka Chikugo¹, Hideyuki Suzuki⁴, Hikaru Seki^{1,2}, Ery Odette Fukushima^{1,5} and Toshiya Muranaka^{1,2*}

OPEN ACCESS

Edited by:

Mehran Dastmalchi,
McGill University, Canada

Reviewed by:

Tessa Moses,
The University of Edinburgh,
United Kingdom

Lucas Busta,
University of Minnesota Duluth,
United States

*Correspondence:

Toshiya Muranaka
muranaka@bio.eng.osaka-u.ac.jp

Specialty section:

This article was submitted to
Plant Metabolism
and Chemodiversity,
a section of the journal
Frontiers in Plant Science

Received: 22 August 2021

Accepted: 25 November 2021

Published: 16 December 2021

Citation:

Istiandari P, Yasumoto S, Srisawat P, Tamura K, Chikugo A, Suzuki H, Seki H, Fukushima EO and Muranaka T (2021) Comparative Analysis of NADPH-Cytochrome P450 Reductases From Legumes for Heterologous Production of Triterpenoids in Transgenic *Saccharomyces cerevisiae*. *Front. Plant Sci.* 12:762546. doi: 10.3389/fpls.2021.762546

¹ Department of Biotechnology, Graduate School of Engineering, Osaka University, Osaka, Japan, ² Industrial Biotechnology Initiative Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Osaka, Japan, ³ RIKEN Center for Sustainable Resource Science, Wako, Japan, ⁴ Department of Research and Development, Kazusa DNA Research Institute, Kisarazu, Japan, ⁵ Plant Translational Research Group, Universidad Regional Amazónica IKIAM, Tena, Ecuador

Triterpenoids are plant specialized metabolites with various pharmacological activities. They are widely distributed in higher plants, such as legumes. Because of their low accumulation in plants, there is a need for improving triterpenoid production. Cytochrome P450 monooxygenases (CYPs) play critical roles in the structural diversification of triterpenoids. To perform site-specific oxidations, CYPs require the electrons that are transferred by NADPH-cytochrome P450 reductase (CPR). Plants possess two main CPR classes, class I and class II. CPR classes I and II have been reported to be responsible for primary and specialized (secondary) metabolism, respectively. In this study, we first analyzed the CPR expression level of three legumes species, *Medicago truncatula*, *Lotus japonicus*, and *Glycyrrhiza uralensis*, showing that the expression level of CPR class I was lower and more stable, while that of CPR class II was higher in almost all the samples. We then co-expressed different combinations of CYP716As and CYP72As with different CPR classes from these three legumes in transgenic yeast. We found that CYP716As worked better with CPR-I from the same species, while CYP72As worked better with any CPR-IIs. Using engineered yeast strains, CYP88D6 paired with class II GuCPR produced the highest level of 11-oxo- β -amyrin, the important precursor of high-value metabolites glycyrrhizin. This study provides insight into co-expressing genes from legumes for heterologous production of triterpenoids in yeast.

Keywords: legumes, cytochrome P450 monooxygenases (CYP), NADPH-cytochrome P450 reductases (CPR), triterpenoids, heterologous production, *Saccharomyces cerevisiae*

INTRODUCTION

Triterpenoids are one of the most diverse groups of plant-specialized metabolites, with more than 20,000 compounds known to exist in nature (Oldfield and Lin, 2012). One of the triterpenoid groups, soyasaponin, is a structurally complex oleanane triterpenoid commonly found in legume plants. Soyasaponins have diverse biological functions such as anti-inflammatory, anti-carcinogenic, and cardiovascular-protective activities (Guang et al., 2014). Some triterpenoid saponins are unique in certain species within the legume family, such as glycyrrhizin, which is found only in the *Glycyrrhiza* species. Glycyrrhizin shows high pharmacological activities and is used as a natural sweetener because its sweetness is 150 times higher than that of sucrose (Kitagawa, 2002).

Cytochrome P450 monooxygenases (CYPs) are the most important enzymes responsible for the structural diversity of triterpenoids (Xu et al., 2004). CYPs are classified based on their sequence similarities (Moktali et al., 2012). The CYP716A and CYP72A subfamily are widespread in plants and show diverse functions in triterpenoid biosynthesis, while other CYP subfamilies such as CYP93E and CYP88D are known to be unique to legumes (Seki et al., 2015; **Figure 1**). To perform catalysis, CYPs require the electrons that are transferred by nicotinamide adenine dinucleotide phosphate (NADPH)-cytochrome P450 reductases (CPRs). Electrons from NADPH flow through flavin adenine dinucleotide (FAD) to the flavin mononucleotide (FMN) domains of CPRs, which are finally transferred to the heme center of CYPs (Paine et al., 2005). CPR and CYP are both membrane-bound proteins that have been reported to be present in microsomes in a ratio of 1:15 (Shephard et al., 1983). This high ratio of CYP to CPR implies that CYPs compete for CPRs and that they must be particularly organized in order to ensure electron supply from CPRs to huge numbers of CYPs at once (Shephard et al., 1983). A previous study on human CPR (hCPR) showed that specific mutations in the FMN domain improved interactions with a specific CYP (Esteves et al., 2020). Several point mutations were found to be co-localized with mutations found in naturally occurring hCPR variants that cause CYP isoform-dependency (Pandey and Flück, 2013; Burkhard et al., 2017) and negatively charged residues that have been previously suggested to be important in CPR:CYP interactions (Hamdane et al., 2009; Hong et al., 2010; Jang et al., 2010).

Unlike mammals and fungi with one CPR gene, plants have multiple CPR genes, depending on the species (Mizutani and Ohta, 1998; Jensen and Møller, 2010). Plant CPRs are branched into two classes, CPR class I and class II (Qu et al., 2015), which in this paper will be referred to as CPR-I and CPR-II, respectively. CPR-I generally has a shorter N-terminal membrane sequence than CPR-II (Parage et al., 2016). CPR-I is reported to be constitutively expressed and plays a role in primary or basal constitutive metabolism, while CPR-II is inducible by environmental stimuli and involves more defense mechanisms through plant secondary metabolism (Ro et al., 2002; Yang et al., 2010; Rana et al., 2013; Qu et al., 2015; Parage et al., 2016). As reported in previous study that knocking down the *CPR2* gene in *Catharanthus roseus* resulted in a significant decrease in

the monoterpene indole alkaloid content. In contrast, knocking down the *CPR1* gene did not result in any modification of the monoterpene indole alkaloid content (Parage et al., 2016).

Many CYPs from legumes have already been used for heterologous production of triterpenoids. Recently, glycyrrhizin has been successfully produced *de novo* in yeast by the combinatorial expression of different enzymes (Chung et al., 2020). Pairing different CPRs with different CYPs has also been reported to result in different productivity of plant specialized metabolites in heterologous yeast (Zhu et al., 2018). In previous studies, CPRs were randomly selected without consideration of the classes (Zhu et al., 2018; Jin et al., 2019). Therefore, the CPR screening was still a subject of trial and error, which might lead to an ineffective strategy for improving heterologous production.

In this study, we performed phylogenetic and gene co-expression analysis of plant CPR classes from two model legumes, *Medicago truncatula* and *Lotus japonicus*, and a medicinal legume, *Glycyrrhiza uralensis*, to clarify the differences in sequence and gene expression patterns. Through co-expression of different CPR classes and CYP families in heterologous yeast, we observed different preferences of plant CPR classes toward specific CYPs. This is the first comparative analysis of different CPR classes of legumes, which gave insight into the differential functions of CPR classes toward the heterologous production of triterpenoids.

MATERIALS AND METHODS

Phylogenetic and Structural Analysis of Plant CPR Classes

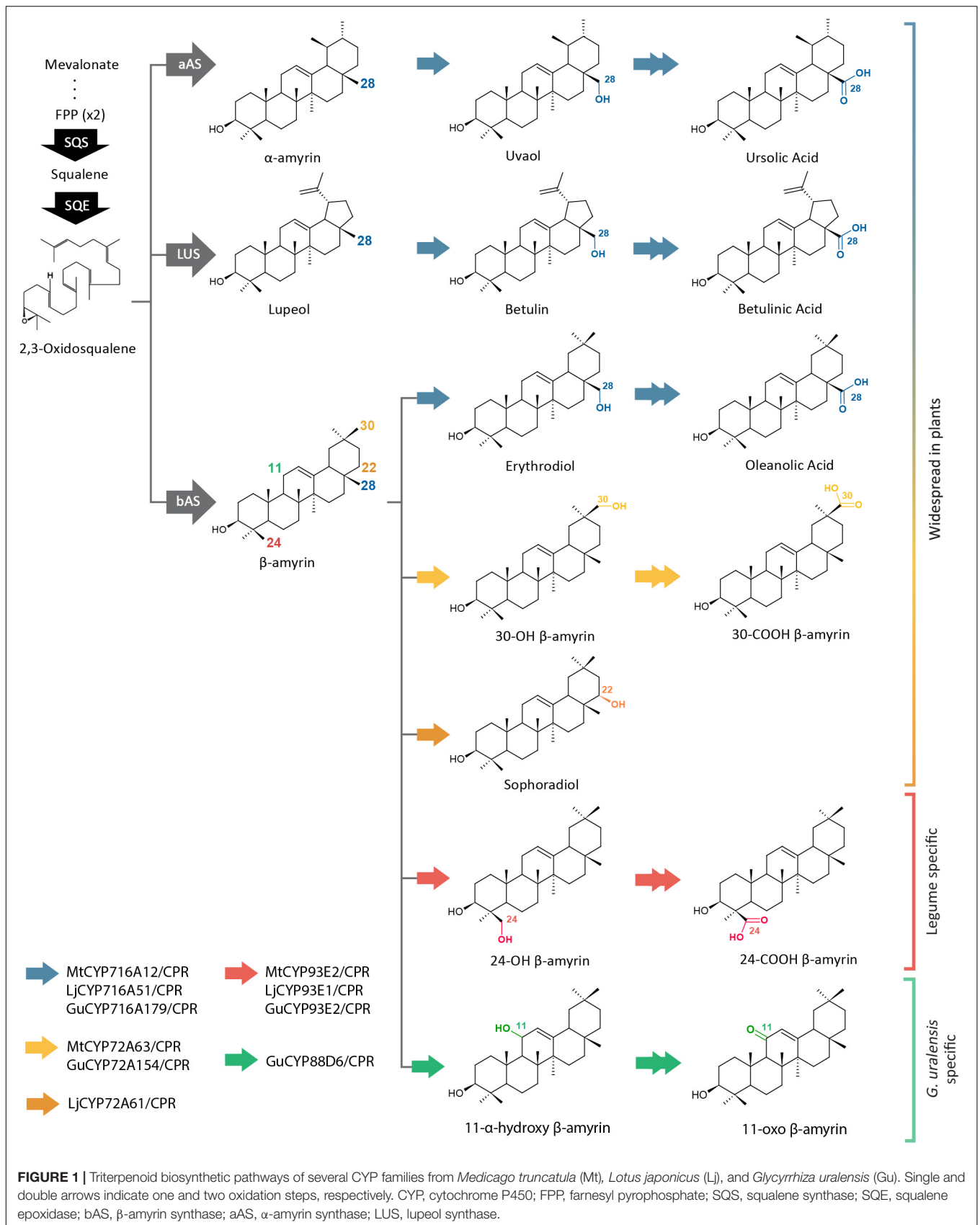
Plant CPR sequences from 20 plant species were obtained from the web-based resource for Arabidopsis P450, cytochrome b5, NADPH-cytochrome P450 reductases, and family 1 glycosyltransferases¹, NCBI², and each plant genome database. Accession numbers of CPR genes and genome databases used in this study are listed in **Supplementary Table 1**. Multiple sequences were aligned using ClustalW and were used for tree construction using MEGA7 with the neighbor-joining method. The position of protein helices was predicted using an *in silico* transmembrane helix prediction tool³ (**Supplementary Figure 1**). The conserved membrane anchor, FMN-, FAD-, CYP-, and NADPH-binding domains were predicted as described by Qu et al. (2015). Motif analysis of the FMN domain of legume CPR classes I and II was performed using the MEME Suite online software⁴ using a minimum of 37 sequences for each CPR class from 24 different legume species (**Supplementary Figure 2**). The structure model of MtCPRs was constructed using 5gxu.1.A (crystal structure of Arabidopsis ATR2) as a template with >71% identity in the SWISS-MODEL (Arnold et al., 2006). PyMol (DeLano, 2020) was used to visualize the amino acids at the active site of the enzyme.

¹www.P450.kvl.dk

²https://www.ncbi.nlm.nih.gov/

³http://www.cbs.dtu.dk/services/TMHMM/

⁴http://meme-suite.org



Chemicals

β -amyrin, erythrodiol, oleanolic acid, α -amyrin, uvaol, ursolic acid, and lupeol were purchased from Extrasynthese (Genay, France). Betulin, methyl jasmonate, and salicylic acid were purchased from Sigma-Aldrich (St. Louis, MO, United States). Betulinic acid and gibberellin were purchased from Tokyo Chemical Industry (Tokyo, Japan). Sophoradiol, 24-hydroxy- β -amyrin, 24-carboxy- β -amyrin, 11-oxo- β -amyrin, 30-hydroxy- β -amyrin, and 11-deoxoglycyrrhetic acid were kindly gifted by Dr. Kiyoshi Ohyama (Tokyo Institute of Technology).

RNA Sequencing

Methods of *G. uralensis* sample preparations, RNA extraction, and library construction for *de novo* assembly have been described in **Supplementary Methods**.

Co-expression Analysis

Co-expression analysis of *M. truncatula* and *L. japonicus* was performed using an online transcriptomic database from the gene expression atlas web servers <https://lipm-browsers.toulouse.inra.fr/pub/expressionAtlas/app/mtgeav3> and [https://lotus.au.dk/expat/\(MiyakojimaMG20v3.0\)](https://lotus.au.dk/expat/(MiyakojimaMG20v3.0)), respectively. Transcriptomic data for *M. truncatula*, *L. japonicus*, and *G. uralensis* were obtained from 274, 74, and 37 different samples, respectively. Pearson's correlation coefficients (PCCs) of different CPRs and CYPs were calculated using Excel (Microsoft Corp., Redmond, WA, United States).

N-Terminal Domain Swapping of CPR-I and II

The position of protein helix inside, transmembrane, and outside of the endoplasmic reticulum (ER) were predicted using *in silico* transmembrane helix prediction (see foot note 3) (**Supplementary Figure 1**). In this study, the N-terminal domain are represented by the protein located inside ER together with the transmembrane helix. To swap the N-terminal domain of MtCPR1 and MtCPR2, the first 46 amino acids of MtCPR1 are swapped with the first 60 amino acids of MtCPR2 using NEBuilder[®] HiFi DNA Assembly cloning to produce N-terminal sequence of MtCPR1 fused with truncated-N-terminal MtCPR2 (M1N-M2C) and N-terminal sequence of MtCPR2 fused with truncated-N-terminal MtCPR1 (M2N-M1C). Two different fragments of each designed chimeric CPRs (M2N-M1C and M1N-M2C) were amplified using pENTR-MtCPR1 and pENTR-MtCPR2 as template and primers listed in **Supplementary Table 2**. A minimum of 20-bp overlapping nucleotides of each CPR mutant and pENTR[™] as vector backbone were designed for fusion cloning.

Cloning and Vector Construction

Cytochrome P450 reductases class I and II genes from *M. truncatula*, *L. japonicus*, and *G. uralensis* were amplified from the respective cDNA using PCR with PrimeSTAR[®] Max DNA Polymerase (Takara Bio, Kyoto, Japan) and named MtCPR1, MtCPR2, LjCPR1, LjCPR2, GuCPR1, and GuCPR2.

The primers used in this study are listed in **Supplementary Table 2**. Wild-type and chimeric CPR genes were cloned into the pENTR[™] vector using the D-TOPO[®] (Thermo Fisher Scientific, Waltham, MA, United States) or NEBuilder[®] HiFi DNA Assembly (Ipswich, MA, United States) cloning method, to produce entry clones of CPR genes. Yeast expression clones of CPR, using pAG415-GAL-ccdB (plasmid number 14145, Addgene, Watertown, MA, United States) as the destination vector, were constructed using LR reaction with LR Clonase[™] II Enzyme mix (Thermo Fisher Scientific), to produce pAG415-CPR. pAG415GAL-ccdB was a gift from Susan Lindquist. Yeast expression clones of *MtCYP716A12*, *MtCYP72A63*, *LjCYP716A51*, *LjCYP93E1*, *LjCYP72A61*, and *GuCYP88D6* were constructed using LR reaction with LR Clonase[™] II Enzyme mix (Thermo Fisher Scientific) into pYES-DEST52 (Invitrogen, Carlsbad, CA, United States) and a Gateway[™]-compatible version of pESC-HIS generated previously in our laboratory (Yasumoto et al., 2016) as the destination vectors, to produce pYES-DEST52-CYP and pESC-HIS-CYP (Seki et al., 2008, 2011; Fukushima et al., 2011; Suzuki et al., 2019).

Yeast Strain Construction

Two different yeast strains were used in this research, INVSc1 (MATa his3D1 leu2 trp1-289 ura3-52; Thermo Fisher Scientific) and PSIII strain. PSIII strain was generated by transforming pESC-TRP[P_{GAL10}/tHMG1-T2A-upc2-1][P_{GAL1}/LjBAS]) as described in **Supplementary Methods 2** into PSI strain (BY4742/P_{erg7}:P_{MET3}-ERG7/trp1:P_{ACT1}-Gal4dbd-ER-VP16) constructed in a previous study (Srisawat et al., 2020). INVSc1 strain carrying pYES3-ADH-aAS, pYES3-ADH-OSC1, or pYES3-ADH-LUS (Fukushima et al., 2011) and PSIII strain were then co-transformed with pAG415-CPR, pYES-DEST52-CYP, and pESC-HIS-CYP consecutively using Frozen-EX Yeast Transformation II[™] (Zymo Research, Irvine, CA, United States).

Yeast Cultivation

In case of INVSc1, yeast sample with cell density at an OD₆₀₀ value in the range of 2.2–2.5 (late logarithmic phase) were cultured in 5 mL appropriate synthetic defined medium (Clontech, Palo Alto, CA) containing 2% glucose at 30°C and 220 rpm for 24 h. Yeast cells were collected by centrifugation, resuspended in 5 ml of appropriate synthetic defined medium (Clontech) containing 2% galactose and then incubated at 30°C and 220 rpm for 48 h. In case of the PSIII strain, yeast strains with cell density at an OD₆₀₀ value in the range of 1.4–1.6 (mid-logarithmic phase) were cultured in 5 mL appropriate synthetic defined medium (Clontech) containing 2% glucose, 100 nM β -estradiol, and 1 mM L-methionine and then incubated at 30°C and 220 rpm for 48 h. For the feeding assay, INVSc1 yeast was cultured as described previously and supplemented with 10 μ M of erythrodiol at the same time as galactose addition, and yeast were then cultured at 30°C and 220 rpm for 60 h.

Metabolite Extraction

Before extraction, 20 ppm of ursolic acid or uvaol were added to 5 mL yeast culture as internal standards. Triterpenoid

metabolite extraction was performed as previously reported (Fanani et al., 2019). The dried extract was dissolved in 500 μ L of chloroform:methanol 1:1 (v/v). Upon Gas chromatography–mass spectrometry (GC-MS) analysis, 100 μ L of the sample was evaporated to dryness and trimethylated with 50 μ L of *N*-methyl-*N*-(trimethylsilyl) trifluoro acetamide (Sigma-Aldrich) for 30 min at 80°C.

GC-MS Analysis

GC-MS was performed as previously reported (Fanani et al., 2019). For peak identification, authentic standards were used to confirm the peak in the sample and compare the retention time and mass fragmentation patterns. The relative amounts of triterpenoids in each strain were compared by calculating the area of each peak from the extracted chromatogram based on *m/z* values of 203 (Kim et al., 2018). Each sample of the different CYP-CPR combinations was analyzed using three biological replicates.

Statistical Analysis

The significance of difference between two yeast strains harboring different CPR classes from same species was determined by single-factor analysis of variance (ANOVA). One-way ANOVA was used to test the difference of triterpenoid levels among different yeast strains and the significance for the means were separated by Tukey's test. *p*-values less than 0.05 ($P < 0.05$) were considered to be significant for this study. All the statistical analyses were performed with SPSS 20.0 for Windows (IBM, Chicago, IL, United States).

RESULTS

Phylogenetic and Structural Analysis of Plant CPRs

The phylogenetic analysis results from 20 different plant species showed clear branching of CPR class I and II sequences (Figure 2). Each dicot plant species was found to have a minimum of one CPR sequence in each class. Monocots, gymnosperms, and mosses possess several copies of CPR-II genes, but no CPRs belonging to CPR class I have been found to date (Jensen and Møller, 2010). Amino acid sequences of plant CPRs from different plant families were compared to create a table of similarity matrices (Supplementary Table 3). Same CPR classes of the same plant family were found to have 75–95% similarity, while same CPR classes of different plant families were found to have 64–75% similarity. On the other hand, different CPR classes of the same or different plant families were found to have 55–65% similarity. Motif analysis was performed on different plant CPR classes (Supplementary Figure 2).

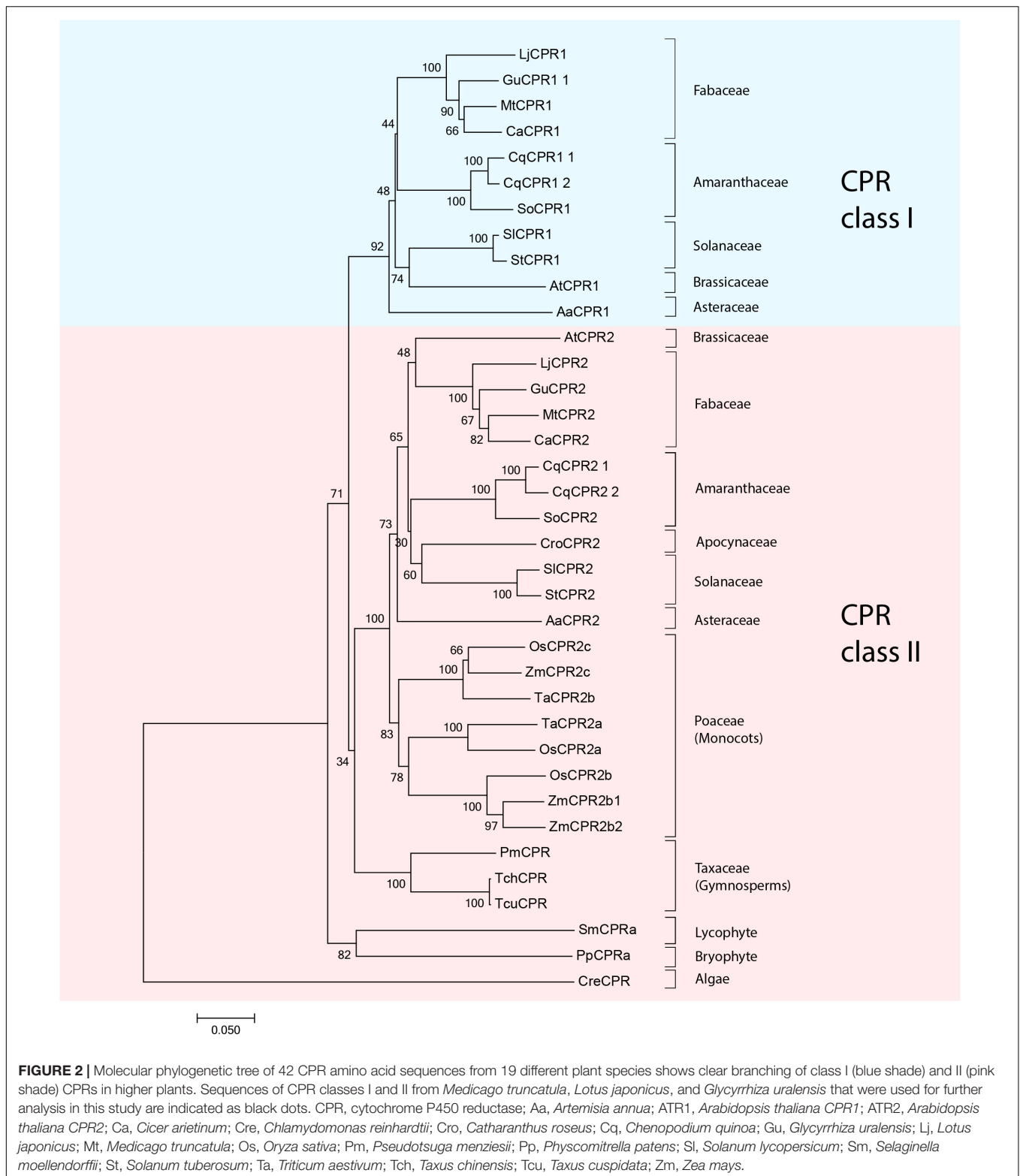
Multiple sequence alignments from several plant CPRs showed that CPR-I has approximately 10–20 amino acids shorter N-terminal membrane sequences than CPR-II (Figure 3). In this study, we focused on the structural analysis of the FMN domain of legume CPRs using CPR from a model legume, MtCPR. Based on motif analysis of amino acids from CPR classes I and II from 24 legume species, there were some differences in the

amino acid residues between CPR classes I and II in the FMN domain, which are conserved in each CPR class (Supplementary Figure 2). By multiple sequence alignment with hCPR, we highlighted eight amino acid residues of MtCPR1 and MtCPR2, which are co-localized with the position of hCPR mutations and negatively charged amino acid residues in the FMN domain. These amino acids have been reported to play a role in CPR:CYP-specific interactions (Campelo et al., 2018; Esteves et al., 2020; Figure 3). The location of each highlighted amino acid in the protein structure was observed through structural analysis of MtCPRs (Figure 4).

Gene Co-expression Analysis of Legume CPR Classes I and II

We selected two model legumes, *Medicago truncatula* and *Lotus japonicus*, and one of the high-value medicinal legumes, *Glycyrrhiza uralensis*, to analyze the gene expression levels of different CPR classes in legumes. Available transcriptomic data of *M. truncatula* and *L. japonicus* were retrieved from the respective transcriptome database, while transcriptome data of *G. uralensis* were obtained in this study (Supplementary Tables 4.1–4.3). RNA-seq analysis from 37 data-sets of *G. uralensis* resulted in 351,138,706 reads and were assembled into 226,599 unigenes (Supplementary Table 5). Co-expression analysis of CPR-I and CPR-II genes of *M. truncatula*, *L. japonicus*, and *G. uralensis* showed distinct expression patterns (Figure 5). CPR-I genes were found to be constitutively expressed with lower and more stable expression levels (Figure 5). In contrast, CPR-II genes were generally found to have higher expression levels than CPR-I genes, which varied depending on tissues and treatments (Figure 5). Overall, CPR-I showed higher expression levels than that of CPR-II in roots or nodules, while CPR-II showed higher expression levels than of CPR-I in chemically or biologically treated samples (Supplementary Table 6). Global analysis of the genes from each legume species was performed by calculating the PCC of each CPR class with each of the transcripts found in the assembly (Supplementary Table 7). CPR-I showed strong correlation ($PCC > 0.70$) with genes involved in basal metabolism or molecular functions, such as ATP-binding/-carrier proteins (Supplementary Table 7). In contrast, CPR-II correlated strongly with receptor genes such as the ubiquitin, glutamate, and jasmonate receptors, which are known to play a role in plant defense responses (Guo et al., 2018; Qiu et al., 2020; Doroodian and Hua, 2021) (Supplementary Table 7). Interestingly, despite its critical role in providing electrons for CYPs, the global transcriptomic analysis of genes correlated with CPRs showed no correlation with CYPs, at a $PCC > 0.4$ (Supplementary Table 7).

In order to analyze the correlation strength between the different CPR classes and CYPs in each species, we calculated the PCC values of different CPR classes in combination with CYP families 716A, 72A, and 93E from the three legume species, in addition to CYP88D6 from *G. uralensis*. MtCYP716A12, LjCYP716A51, and GuCYP716A179 catalyze the same reaction of three steps of oxidation at the C-28 position of α -amyrin, β -amyrin, and lupeol to produce the final product of ursolic acid, oleanolic acid, and betulinic acid, respectively,



(Fukushima et al., 2011; Tamura et al., 2017; Suzuki et al., 2019; **Figure 1**). Both MtCYP72A63 and GuCYP72A154 oxidize the C-30 position of β -amyirin to produce 30-OH β -amyirin, but only MtCYP72A63 catalyzes the further oxidation to produce

30-COOH β -amyirin (Seki et al., 2011). LjCYP72A61 catalyzes the one-step oxidation at the C-22 position of β -amyirin to produce sophoradiol (Ebizuka et al., 2011). MtCYP93E2, LjCYP93E1, and GuCYP93E3 catalyze the same reaction of three steps of

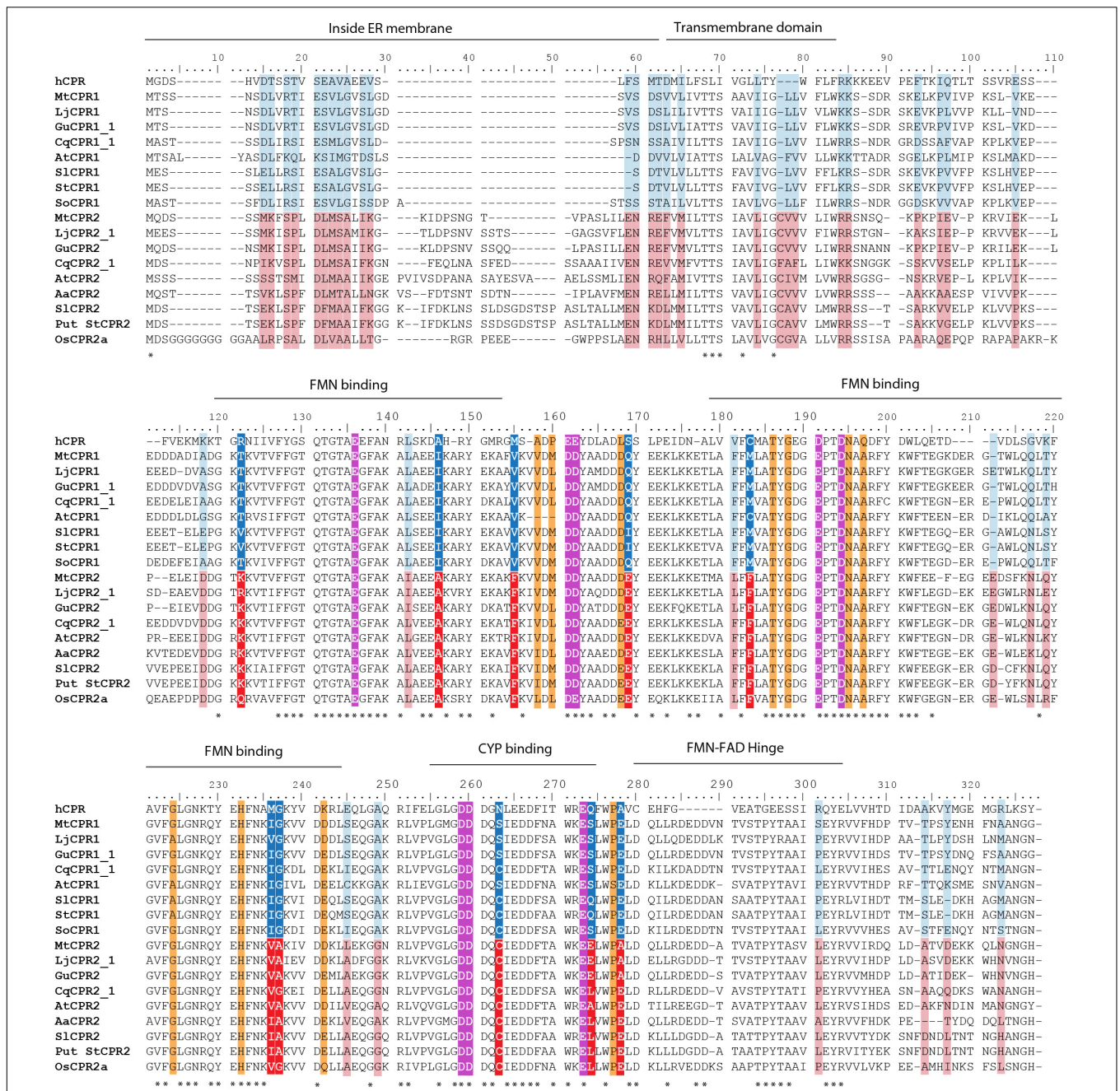
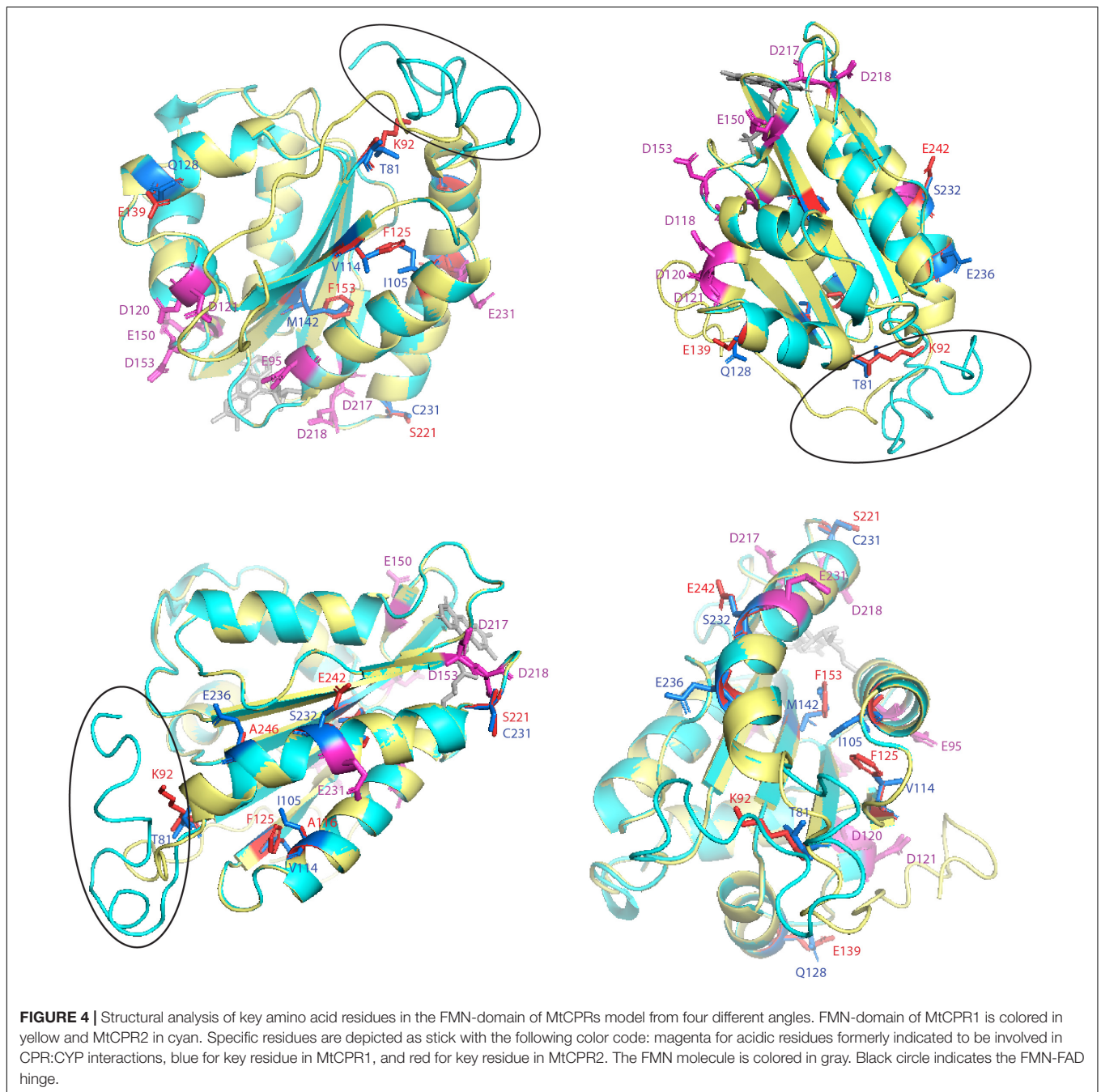


FIGURE 3 | Multiple sequence alignment of first 327 amino acids sequence of CPR class I and II containing N-terminal and FMN domains using the ClustalX color scheme for amino acid alignments. Dark blue, red, light blue, and pink colors indicate the different conserved amino acids in CPR class I and II in legume family based on motif analysis (Supplementary Figure 2). Dark blue and red colors are the eight highlighted key residue in this study. Magenta color indicates acidic residue formerly reported to be important in CYP:CPR interaction in human CPR (hCPR). Orange color indicates point mutations in hCPR that are reported to improve interaction with a specific CYP. Asterisks indicate the positions which have a single, fully conserved residue. CPR, cytochrome P450 reductase; Aa, *Artemisia annua*; ATR1, *Arabidopsis thaliana* CPR1; ATR2, *Arabidopsis thaliana* CPR2; Ca, *Cicer arietinum*; Cq, *Chenopodium quinoa*; Gu, *Glycyrrhiza uralensis*; Lj, *Lotus japonicus*; Mt, *Medicago truncatula*; Os, *Oryza sativa*; Sl, *Solanum lycopersicum*; St, *Solanum tuberosum*.

oxidation on the C-24 position of β -amyrin to produce the final product of 24-COOH β -amyrin (Seki et al., 2008; Fukushima et al., 2011; Suzuki et al., 2019). GuCYP88D6 catalyzes the two-step oxidation at the C-11 position of β -amyrin to produce 11-oxo β -amyrin (Seki et al., 2008).

The PCC table shows that different CPR classes had different correlation values with different CYPs (Table 1). The probeset ID used for PCC calculation is presented in Supplementary Table 8. CYP716A12 had a higher positive correlation value with MtCPR1, while CYP72A63 had a higher positive value with



MtCPR2. However, from the data-set, no significant difference was observed in the correlation of *MtCPRs* with *CYP93E2* (Table 1A). From the *L. japonicus* database MG20 v3.0, *LjCPR1* showed higher correlation values with *CYP716A51*, *CYP72A61*, and *CYP93E1*, as compared to *LjCPR2* (Table 1B). Based on the transcriptome data of *L. japonicus* in the Gifu version 2.0, *LjCPR1* also showed higher positive correlation value with *CYP716A51* and *CYP93E1*, as compared to *LjCPR2*. However, *CYP72A61* did not have strong correlation with any of the *LjCPRs*, but showed very high correlation value (>0.5) with *CYP93E1* (Supplementary Table 9). *GuCPR2* showed higher correlation

with *CYP716A179*, *CYP72A154*, and *CYP93E3*, as compared to *GuCPR1*. In contrast, *CYP88D6* correlated negatively with *GuCPR1* (Table 1C).

CYP Shows Preference Toward Certain CPR Class

To analyze the effect of different CPR classes on heterologous triterpenoid production, different CYPs and CPR classes from *M. truncatula* and *L. japonicus* were co-expressed in transgenic *S. cerevisiae* INVSc1, a commonly used strain for recombinant protein expression. By comparing the relative

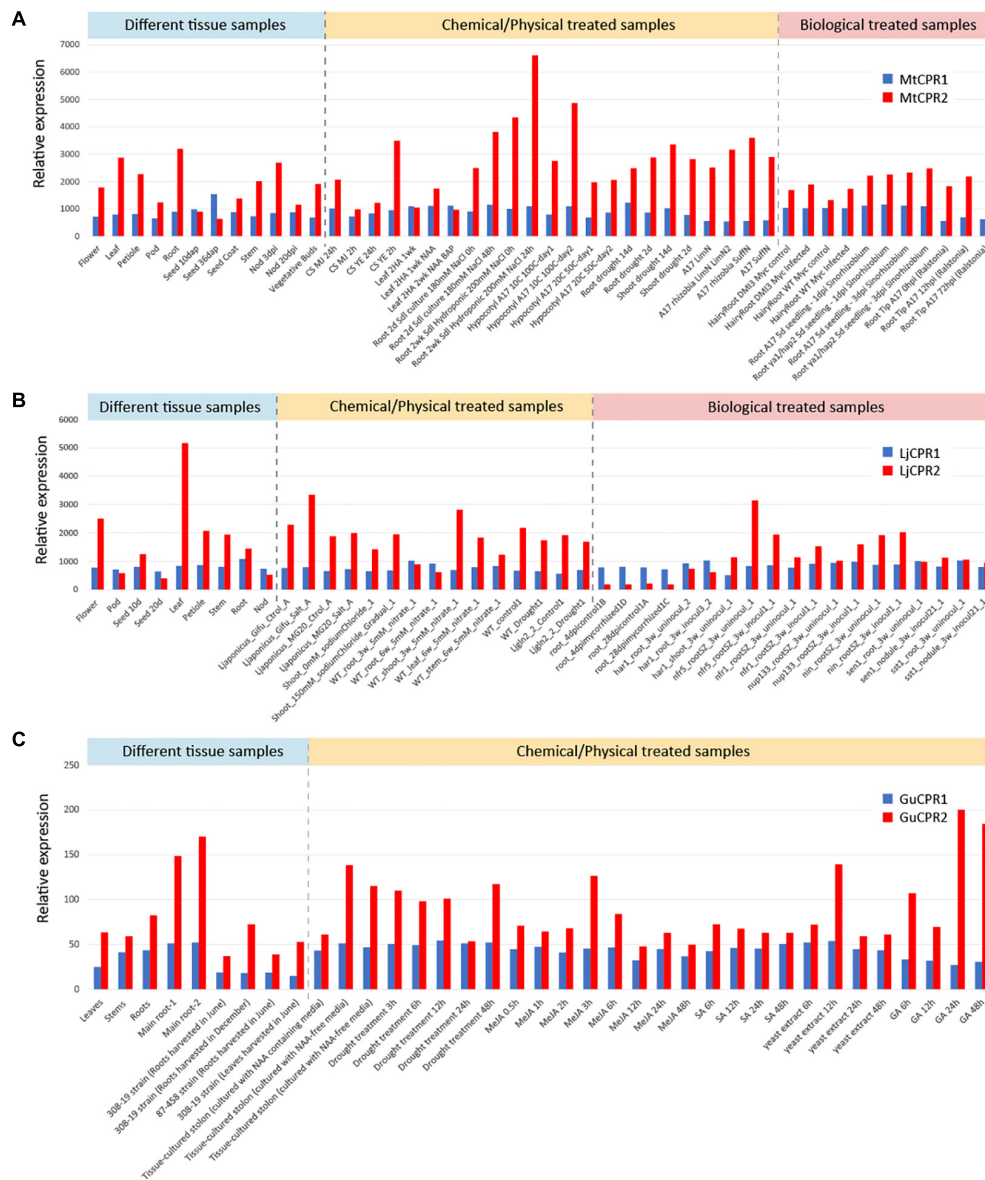


FIGURE 5 | Co-expression analysis of CPR class I and II from **(A)** *M. truncatula*, **(B)** *L. japonicus*, and **(C)** *G. uralensis* using selected transcriptomic data-sets representing samples from different tissue, chemical/physical treatments, and biological treatments. Transcriptomic data-sets of *M. truncatula* and *L. japonicus* were obtained from <https://mtgea.noble.org/v3/index.php> and <https://lotus.au.dk/ofMG-20version3.0>. The *G. uralensis* transcriptomic data-set was obtained using RNA-seq analysis in this study.

value of each metabolite using a semi-quantitative method with an internal standard, we compared the accumulation of oxidized triterpenoids resulting from the heterologous expression of different CPR-CYP pairs in transgenic yeast (**Figure 6**). The amount of triterpenoid backbone in the control samples was considered to be 100%. In yeast strains expressing MtCYP716A12 and LjCYP716A51, the intermediate compounds oleanolic and ursolic aldehyde could be detected based on the mass spectrum profile (**Supplementary Figure 3**) compared to previously reported mass spectra of oleanolic and ursolic aldehyde authentic standards (Misra et al., 2017).

However, betulinic aldehyde could not be detected in the yeast extract and was co-eluted with betulinic acid in the yeast extract; (Suzuki et al., 2018) therefore, the amount of betulinic acid shown in **Figure 7** corresponds to the sum of both compounds. C-30 and C-24 aldehydes were excluded because of the instability of the compound, lack of authentic standards, and no previous reports on their mass spectrum profile.

The results show that upon using INVSc1 β -amyrin-producing yeast, MtCYP716A12 and LjCYP716A51 showed higher relative amount of oleanolic acid when paired with

TABLE 1 | Correlation strength^a between different CPR classes and CYP families in (A) *M. truncatula*, (B) *L. japonicus*, and (C) *G. uralensis*.

(A)				
Gene	MtCYP716A12	MtCYP72A63	MtCYP93E2	
MtCPR1	0.12	0.22	0.25	
MtCPR2	-0.03	0.41	0.17	
(B)				
Gene	LjCYP716A51	LjCYP72A61	LjCYP93E1	
LjCPR1	0.54	0.59	0.70	
LjCPR2	-0.35	-0.35	-0.29	
(C)				
Gene	GuCYP716A179	GuCYP72A154	GuCYP93E3	GuCYP88D6
GuCPR1	-0.23	-0.33	0.34	-0.49
GuCPR2	0.33	0.29	0.70	-0.02

^aCorrelation strength of different CPR classes with different CYP families based on PCC. PCC values range from -1 to 1. Positive values were positively correlated, whereas negative values were negatively correlated. PCC, Pearson's correlation coefficient; Mt, *Medicago truncatula*; Lj, *Lotus japonicus*; Gu, *Glycyrrhiza uralensis*.

MtCPR1, as compared to when they were paired with MtCPR2, with a comparable amount of erythrodiol and oleanolic aldehyde (Figures 6A,B). Pairing MtCYP716A12 with LjCPR1 showed higher conversion of β -amyrin into erythrodiol compared to pairing it with LjCPR2. However, pairing MtCYP716A12 with LjCPR2 showed a higher conversion rate of erythrodiol into oleanolic aldehyde and oleanolic acid than with LjCPR1 (Figure 6A). Conversely, pairing LjCYP716A51 with LjCPR2 resulted in a higher conversion rate of β -amyrin into erythrodiol than with LjCPR1, while the conversion rate of erythrodiol into oleanolic acid was higher when pairing LjCYP716A51 with LjCPR1 than LjCPR2 (Figure 6B). These trends were also similar in the yeast feeding assay results of INVSc1 strain supplemented with erythrodiol as a substrate, although the difference was not as significant as in the *in vivo* assay (Supplementary Figure 4). In contrast, pairing MtCYP72A63 with CPR-II resulted in significantly higher conversion of β -amyrin into 30-OH β -amyrin and 30-COOH β -amyrin levels (Figure 6C) than pairing it with CPR-I. Similar to MtCYP72A63, pairing LjCYP72A61 with CPR-IIs resulted in significantly higher conversion of β -amyrin into sophoradiol (Figure 6D). In case of the legume-specific CYP subfamily, CYP93E1 paired with CPR-II showed higher conversion of β -amyrin into 24-COOH β -amyrin, as compared to CYP93E1 paired with CPR-I (Figure 6E).

CYP716As can also catalyze oxidations on different triterpenoid backbone. Therefore, to confirm that the effect of the CPR on the conversion rate can also be applied to different triterpenoid backbone, we then co-expressed MtCYP716A12 in α -amyrin and lupeol-producing strains (Figure 7). The result showed that MtCYP716A12 paired with MtCPR1 and LjCPR1 showed higher conversion of α -amyrin into ursolic acid and uvaol, respectively (Figure 7A). A similar trend was

observed in the lupeol-producing strain in which MtCYP716A12 paired with MtCPR1 showed a higher conversion rate of betulin into betulinic acid compared to that with MtCPR2, and MtCYP716A12 paired with LjCPR1 showed a higher conversion rate of lupeol to betulin and betulinic acid compared to LjCPR2 (Figure 7B). All the chromatograms of yeast *in vivo* are shown in Supplementary Figures 5–14.

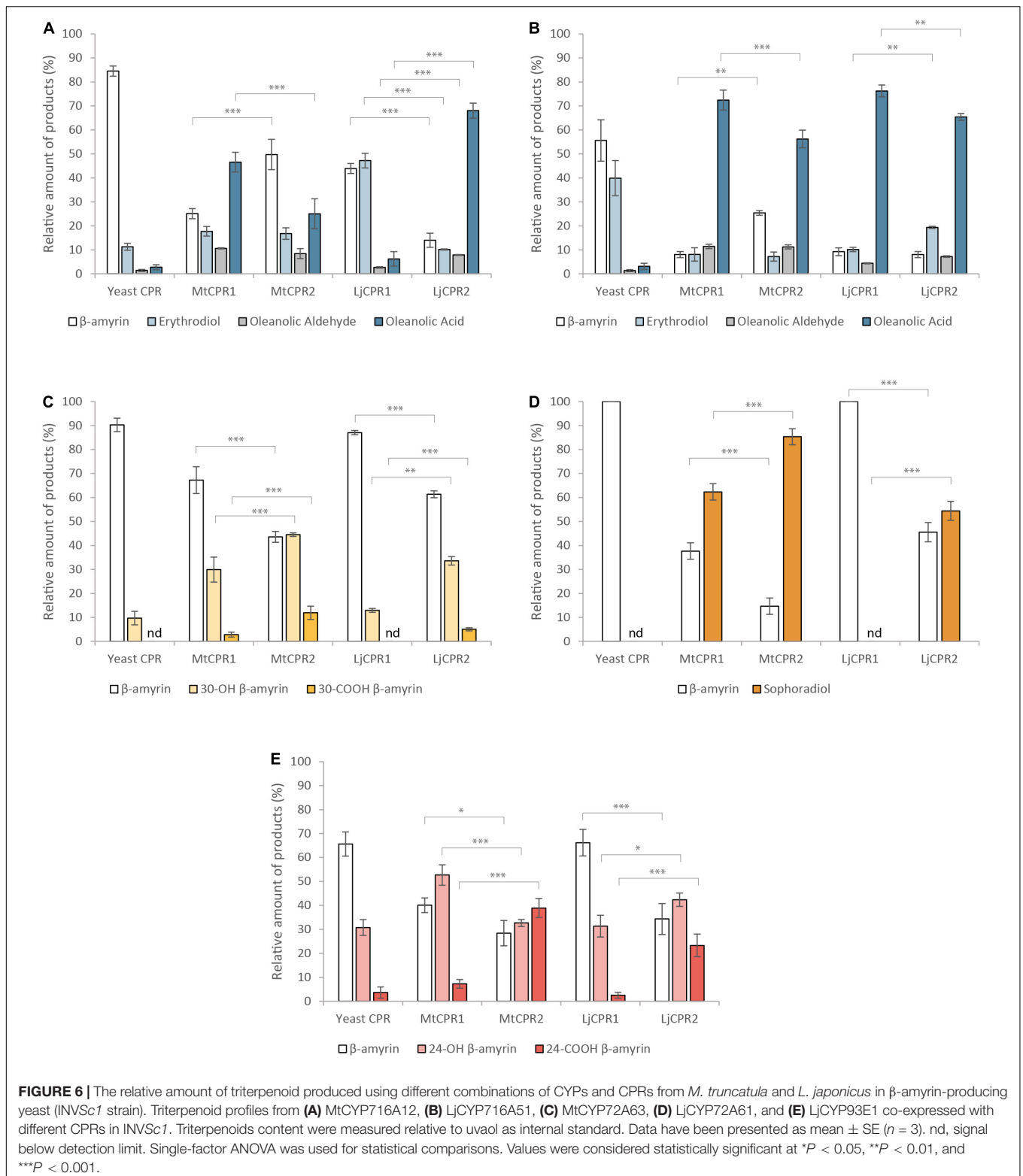
This yeast *in vivo* assay might provide insight into CPR-CYP preferences, which are better justified by their protein expression levels. Quantifying the CPR protein itself would not be sufficient because triterpenoids are the products of CYP. However, protein quantification of CYP remains challenging. We propose this study to analyze the effect of CPR class on heterologous production in yeast, but not on the catalytic activity of CPR or CYP. We admit that this is a limitation of yeast *in vivo* studies. Therefore, further in-depth *in vitro* analysis of the effect of CPR class on CYP activity might be needed in future experiments to provide more justifiable results.

N-Terminal Domain Swapping of MtCPR1 and MtCPR2

Based on the multiple sequence alignment result, the main difference of CPR class I and II is located in the N-terminal domain. CPR class I has shorter amino acid sequence than CPR class II (Figure 3). To analyze whether this N-terminal domain is responsible for the differences in the conversion rate of β -amyrin when CPR was co-expressed with different CYPs in transgenic yeast (Figure 6), we investigated the effect of this domain by swapping the N-terminal membrane sequence of MtCPR1 and MtCPR2. MtCPRs were used as the representative since *M. truncatula* is a model legume. We speculated if the N-terminal domain of CPR is indeed responsible for the CPR activity, swapping the N-terminal domain of CPR class I and II will affect the triterpenoids profile.

The location of transmembrane helix of MtCPR1 and MtCPR2 was predicted *in silico* (Supplementary Figure 1). Amino acids of MtCPR1 number 1–26 are located inside the ER, number 27–46 are the transmembrane helix, and number 47–692 are located outside the ER (Supplementary Figure 1A). Amino acids of MtCPR2 number 1–40 are located inside the ER, number 41–60 are the transmembrane helix, and number 61–701 are located outside the ER (Supplementary Figure 1B). In this study, the first 46 amino acids of MtCPR1 located inside the ER together with the transmembrane helix were swapped with the first 60 amino acids of MtCPR2. The resulting chimeric MtCPRs were N-terminal membrane sequence of MtCPR1 fused with truncated-N-terminal MtCPR2 (M1N-M2C) and N-terminal membrane sequence of MtCPR2 fused with truncated-N-terminal MtCPR1 (M2N-M1C) (Figure 8A).

Yeast *in vivo* assay was then performed by co-expressing the chimeric MtCPRs with CYP716As and CYP72As to observe its effect on triterpenoids conversion rate (Figures 8B–E). The results showed that by swapping the N-terminal domain of CPR class I with that of class II, the product from CYP716A12 and CYP716A51 were significantly reduced (Figures 8B,C). The production of erythrodiol by CYP716A51



co-expressed with mutant CPRs was even lower than the background level observed in native yeast CPR (Figure 8C). A similar effect was also seen in CYP72A61, where the chimeric CPR caused a significant decrease in conversion

of β -amyirin into sophoradiol (Figure 8D). Interestingly, only CYP72A63 was not affected by the chimeric MtCPRs. Moreover, the results showed that CYP72A63 paired with the N-terminal membrane sequence of MtCPR2 fused with

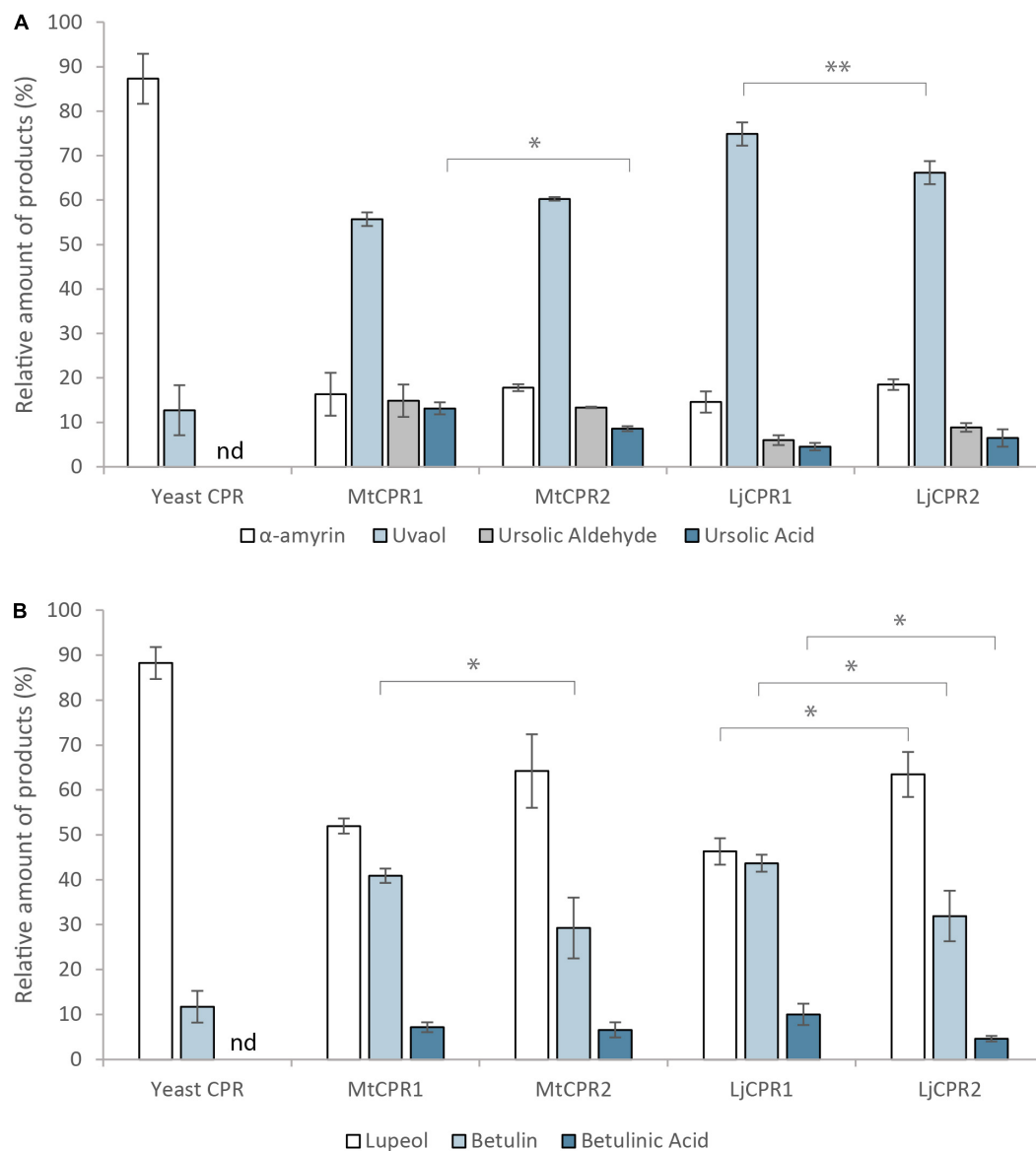


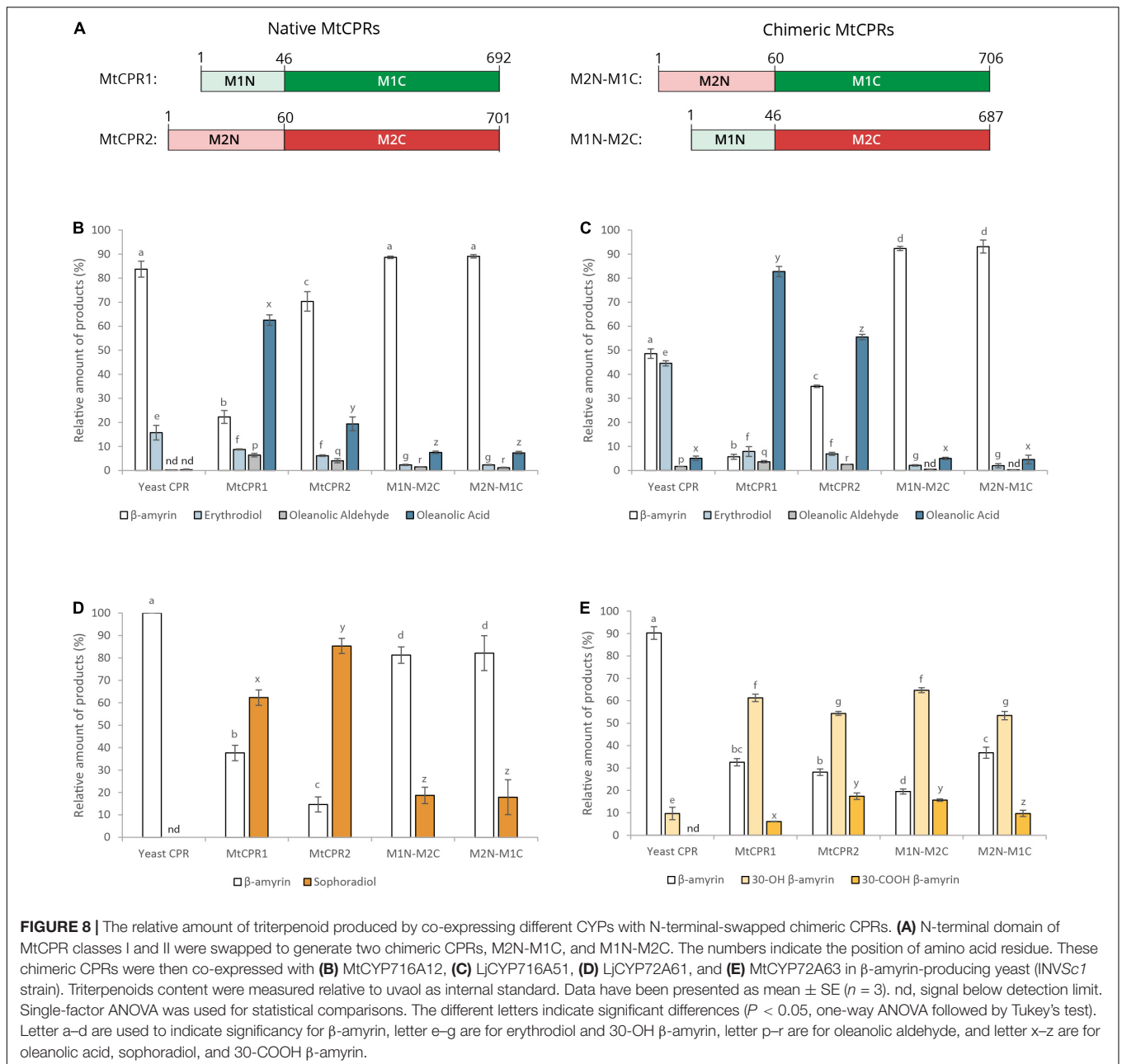
FIGURE 7 | The relative amount of triterpenoid produced by co-expressing MtCYP716A12 with different CPRs from *M. truncatula*, *L. japonicus*, and *G. uralensis* in **(A)** α -amyrin-producing and **(B)** lupeol-producing yeast (INVSc1 strain). Triterpenoids content were measured relative to erythrodiol as internal standard. Data have been presented as mean \pm SE ($n = 3$). nd, signal below detection limit. Single-factor ANOVA was used for statistical comparisons. Values were considered statistically significant at * $P < 0.05$ and ** $P < 0.01$.

truncated N-terminal MtCPR1 (M2N-M1C) showed higher conversion of β -amyrin into 30-OH β -amyrin and 30-COOH β -amyrin (Figure 8E).

Improving Triterpenoid Production Using an Engineered Yeast

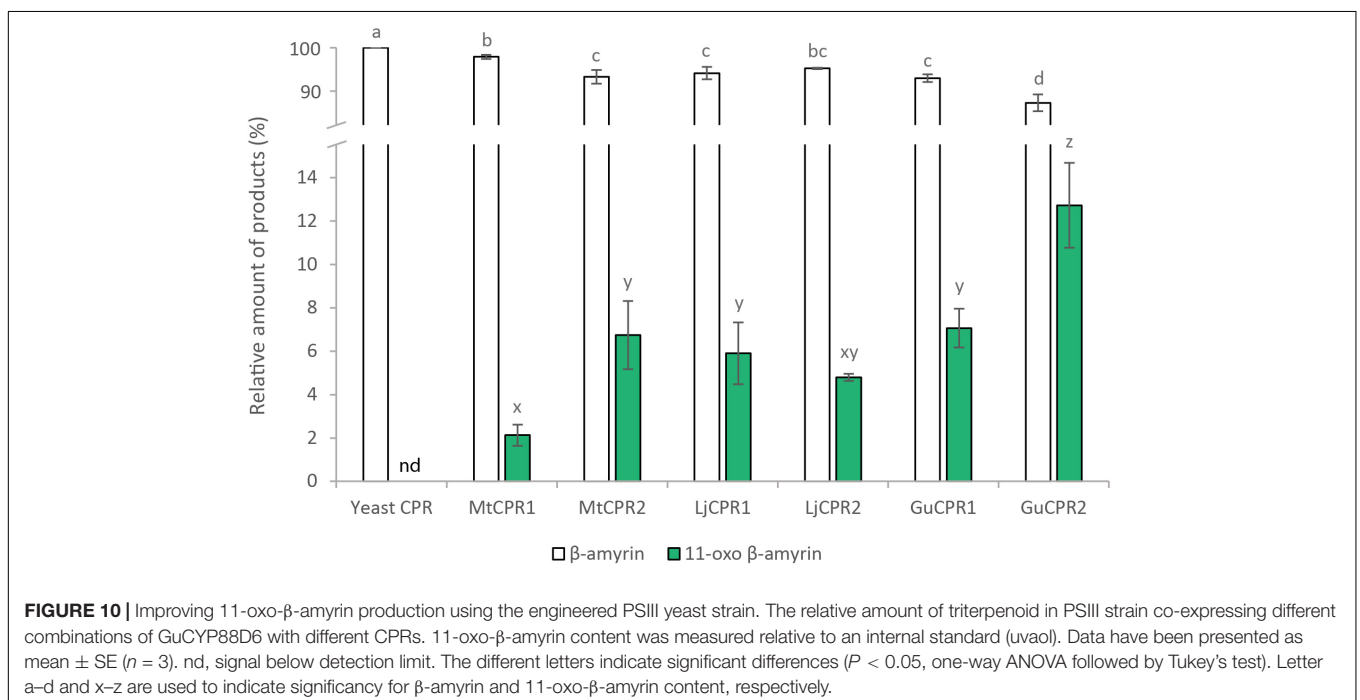
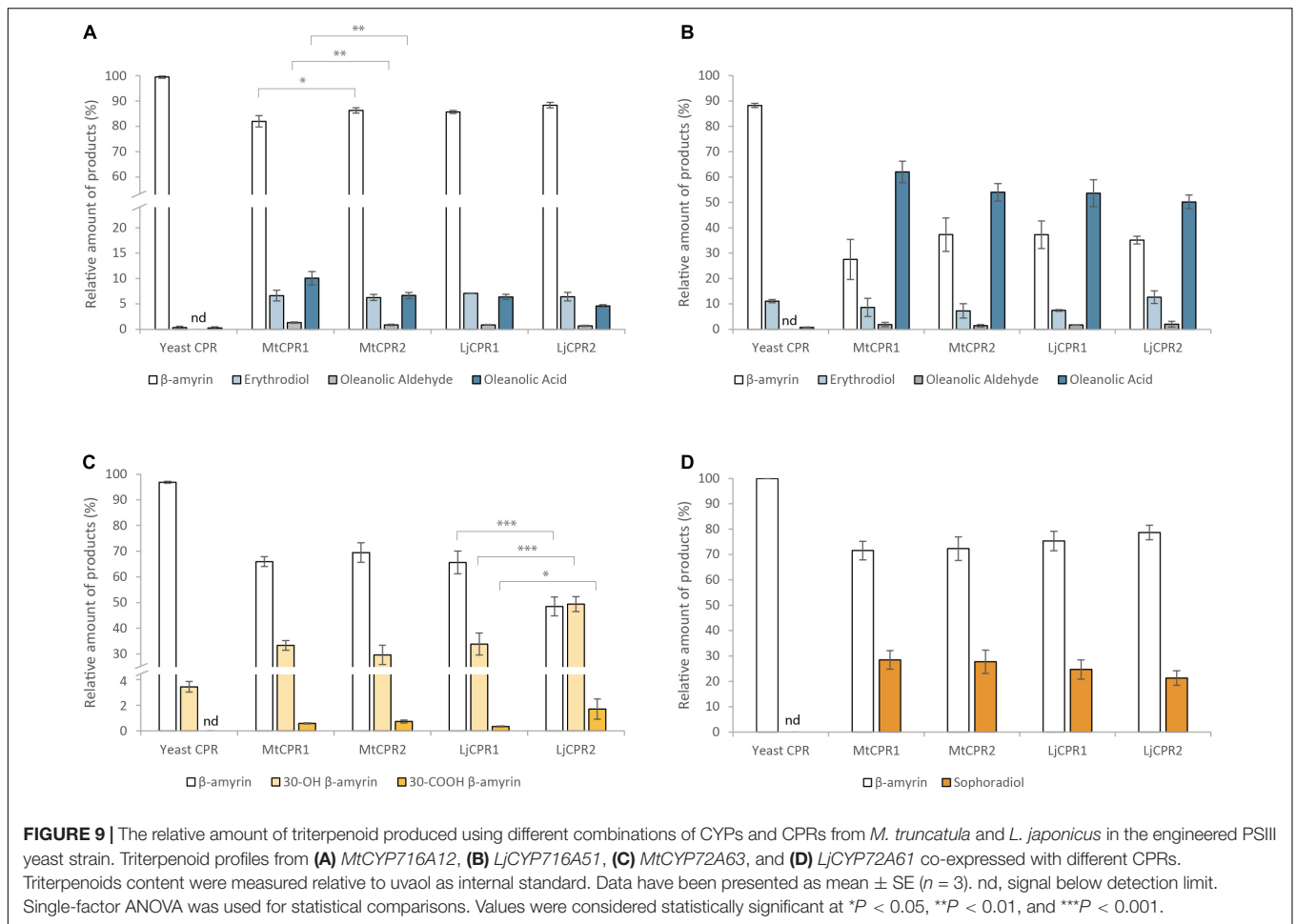
To further analyze the effect of different CPR classes on improving heterologous production, different CPR classes were co-expressed with different CYP families in the PSIII strain. PSIII is an engineered yeast strain that has been optimized for triterpenoid biosynthesis by means of some modifications in the

upstream mevalonate pathway. This strain is modified from the PSII strain constructed in previous research, which can produce a 100 \times times higher triterpene backbone as a substrate than its parental strain (Srisawat et al., 2020). In contrast to the INVSc1 results, in the PSIII strain, almost all the triterpenoid sapogenin production profiles had a lower ratio of β -amyrin, and the trends of CYP-CPR preferences observed in the INVSc1 strain seem to be less obvious in this strain (Figure 9). Nevertheless, some notable changes were observed in the PSIII strain. MtCYP716A12 paired with LjCPR1 that previously showed different conversion ratios of triterpenoids with LjCPR2 in the INVSc1 strain, showed similar conversion ratio profiles in the PSIII strain



(Figure 9A). The ratio of oleanolic acid to erythrodiol was lower in PSIII strain co-expressing MtCYP716A12 (Figure 9A). The PSIII strain co-expressing MtCYP72A63 resulted in much lower 30-COOH β -amyryn to 30-OH β -amyryn ratio, as compared to the INVSc1 strain (Figure 9C). Co-expressing all CPRs with LjCYP72A61 and MtCYP716A51 resulted in similar conversion ratio in PSIII strain (Figure 9D). However, similar to the INVSc1 strain, co-expressing MtCYP716A12 with MtCPR1 in the PSIII strain showed higher oleanolic aldehyde and oleanolic acid conversion rates than MtCPR2. In addition, co-expressing MtCYP72A63 with LjCPR2 in the PSIII strain showed higher 30-OH β -amyryn and 30-COOH β -amyryn conversion rates compared to that with LjCPR1.

Co-expression of different CYP and CPR classes from model legumes, *Medicago truncatula* and *Lotus japonicus*, have suggested that CYP shows a preference for certain CPR classes. Therefore, to further study the effect of CYP-CPR pairs, we added CPRs from a non-model legume species, *Glycyrrhiza uralensis*, which produces glycyrrhizin, a highly valuable metabolite found only in *Glycyrrhiza* sp. Furthermore, unlike CYP716As and CYP72As, which are commonly found in plant species, GuCYP88D6 uniquely presents in *Glycyrrhiza* sp. and is very specific in function. Therefore, it is interesting to facilitate contrasting results. GuCYP88D6 catalyzes the two-step oxidation at the C-11 position of β -amyryn to produce 11-oxo β -amyryn, the precursor for glycyrrhizin (Seki et al., 2008). The PCC value



of *GuCYP88D6* with *GuCPRs* also indicates that *GuCYP88D6* might have different preferences for *GuCPR* classes I and II, which shows a negative correlation with *GuCPR1* (Table 1C). Corresponding to this hypothesis, the results showed that co-expressing *GuCYP88D6* with class II *GuCPR* in the PSIII strain showed the highest conversion of β -amyrin into 11-oxo β -amyrin, as compared to the other CPRs tested in this study (Figure 10).

DISCUSSION

Based on the phylogenetic results of plant CPRs, it was identified that dicotyledons possess both CPR-I and CPR-II; thus far, only CPR-II genes have been identified in monocotyledons, gymnosperms, and mosses. Among angiosperms, as compared to dicotyledons, monocotyledons are well known to be more tolerant to abiotic stresses (Agrawal and Agrawal, 1999; Rao et al., 2006; Wright et al., 2010). CPR-II is believed to play a more important role in defense and adaptation mechanisms compared to CPR-I (Rana et al., 2013; Parage et al., 2016). The fact that multiple copies of CPR-II were found in all monocotyledons, while some dicotyledons might only possess a single CPR-II (Figure 2; Jensen and Møller, 2010), suggest that the role of CPR-II is very prominent in supporting abiotic resistance in monocots. In addition, the more sophisticated translocation and root system of dicotyledons (Scarpella and Meijer, 2004) might be supported by the presence of CPR-I found in dicotyledons. Nevertheless, it remains to be determined whether monocots possess only CPR-II. However, this might also explain why the expression of CPR-I is more abundant in roots and nodules with higher meristematic activity, while the expression of CPR-II is higher in leaf, shoot, and treated plants (Supplementary Table 4.4). Global co-expression analysis results also support this deduction, that CPR-I supports primary or basal metabolism, while CPR-II plays a more important role in defense mechanisms (Supplementary Table 5).

The FMN domain of CPRs has several conserved patches of acidic residues located on its solvent-exposed surface, suggesting that they are involved in the electrostatic interaction with its redox partners, including CYPs (Waskell and Kim, 2015). It was shown that in humans, every CYP interacts with CPR in a specific manner, based on the difference in the binding motifs located in the FMN domain (Esteves et al., 2020). The motif analysis of 24 different legume species showed that there are some differences in the conserved amino acid residues between legume CPR classes I and II in the FMN domain. Among them, eight residues either located on the conserved binding domain or near the previously suggested residue to be important for CYP:CPR interactions in humans were highlighted. CPR from a model legume, MtCPRs, was used for 3D structural analysis. The first highlighted amino acids, T81 in MtCPR1 and K92 in MtCPR2, are located on the first residue in the FMN binding domain (Figure 4). This difference in the residues is conserved in CPR I and II of legumes and other species. Based on the structural analysis, these residues were located near the FMN-FAD hinge (Figure 4). Interflavin electron transfer (ET) has been

previously reported to be the rate-limiting step in plant CPR (Simtchouk et al., 2013; Whitelaw et al., 2015). This highly flexible hinge also has been suggested to play a role in the formation of multiple open conformations to form complexes with CYP in an isoform-dependent manner (Campelo et al., 2018). Therefore, the difference in size and charge properties of T81 in MtCPR1 and K92 in MtCPR2 might greatly affect the conformational flexibility of the interflavin hinge during ET. The next highlighted residues are I105, V114, and M142 in MtCPR1 and A116, F125, and F153 in MtCPR2 (Figure 4). Even though these amino acid residues are all hydrophobic and only exhibit small differences in the chain length, these different residues of MtCPR1 and MtCPR2 are conserved in CPR I and II of legumes and other plant families (Figure 3). These residues are located in the α -helix of the FMN-binding domain, are close to the acidic residues important for ET, and face each other (Figure 4). The difference in these residues might cause minor changes in the geometrical shape of the α -helix of the FMN-binding domain, leading to improved interactions with specific CYPs, in which small conformational perturbations of the α -carbon backbone might accommodate changes in substrate binding (Bart and Scott, 2018). One of the most important residues is Q128 in MtCPR1 and E139 in MtCPR2 (Figure 4), which are also conserved in legumes and in other plant families (Figure 3). These residues are located in the acid-rich residues on the solvent-exposed surface, which are highly likely to serve as binding motifs for CPR:CYP interactions (Figure 4). The glutamate (E139) residue in CPR II provides a stronger negative charge to form ionic bonds with the respective CYP. Changing this to an uncharged glutamine residue, as in CPR I (Q128), will greatly affect the CYP interaction. The last three highlighted residues are S221, S232, and E236 in MtCPR1 and C231, E242, and A246 in MtCPR2 (Figure 4). Interestingly, these residues seem to be conserved only in CPR I and II of the legume family (Figure 3 and Supplementary Figure 2). These residues are located on the solvent-exposed surface of the FMN domain and near the acidic residue important for CPR:CYP interaction (Figure 4). C231 in MtCPR2 can form disulfide bridges with the interacting CYP, which cannot be achieved by S221 in MtCPR1. Serine-to-cysteine mutations have also been reported to have stabilizing effects because of optimization of van der Waals interactions and increased packing (Santos et al., 2007). Hence, C231 in CPRII might account for more stable interactions with specific CYPs, compared to that of CPRI. The last two residues are acidic residues of E242 and A246 in MtCPR2 and S232 and E236 in MtCPR1. The difference between the negatively charged glutamate (E) residue and short aliphatic alanine (A) or uncharged serine residue (S) greatly affects the ionic interaction between CPR and its ET partner. The positions of the acidic residues in CPR I and II are different but parallel in the α -helix (Figure 4), which is highly likely to lead to CYP isoform-specific interactions. The fact that these positions are conserved only in legume CPRs suggests that CPR from each plant family might exhibit unique CYP binding motifs that interact with specific CYPs.

This study emphasizes that different CYPs show a preference for certain CPR classes. Based on the yeast *in vivo* assay in the INVScl β -amyrin-producing strain, CYP716As paired

with CPR-I generally showed a higher conversion ratio of triterpenoid backbone when they were paired with CPR-II, except that MtCYP716A12 paired with LjCPR1 showed a lower conversion ratio of triterpenoids compared to LjCPR2 (Figure 6). This might be due to the incompatibility of CYP716A12 from *M. truncatula* with LjCPR1 owing to species differences. However, the preference of CPR-I over CPR-II of MtCYP716A12 was also observed in the INVSc1 α -amyrin and lupeol-producing strains (Figure 7). On the other hand, CYP72A63, CYP72A61, and CYP93E1 showed higher conversion ratio when co-expressed with CPR-II (Figure 6). It has previously been reported that the N-terminal of ATR2 (*Arabidopsis thaliana* CPR2), which contains high number of Ser/Thr residues, significantly enhances the ATR2 activity (Urban et al., 1997), which might lead to higher conversion rate. However, N-terminal domain swapping of MtCPR1 and MtCPR2 showed that the longer N-terminal membrane sequence of MtCPR2 did not result in increase of conversion rate of β -amyrin. In contrast, these chimeric CPRs caused a reduction in the β -amyrin conversion rate of the CYP716As and CYP72A61, while only showed little or no effect when co-expressed with CYP72A63. Interestingly, chimeric CPR M2N-M1C showed a higher conversion ratio of 30-COOH β -amyrin than native MtCPR1, implying that the N-terminal domain of MtCPR2 might correspond to an increase in the conversion ratio. This result suggests that there might be a specific CYP-CPR protein-protein interaction that was affected by the sequence-structural relationship in the N-terminal domain of CPRs. A previous study on human CPR-FMN-domain mutants has also shown that different CYP isoforms interact with CPR in a specific manner due to docking elements in the binding motifs of the FMN-domain, which affect the CYP-CPR affinity (Ritacco et al., 2019). The N-terminal domain swapping between MtCPR1 and MtCPR2 might cause a structural change in the CYP:CPR-specific binding motif, as shown in Figure 3. Further studies, such as mutagenesis experiments on CPR, are needed to understand the mechanisms underlying this phenomenon.

There are several aspects that might affect heterologous gene expression between INVSc1 and PSIII strain. The INVSc1 strain is a diploid, whereas BY4742, the parental strain of PSIII, is a haploid strain. Diploid yeast co-expressing similar plasmids resulted in more stable and approximately two times higher copy number than the haploid strain (Karim et al., 2013). The engineered pathway of PSIII strain also significantly affects triterpenoid production. Benefiting from the GEV chimeric transcriptional activator system, this strain can utilize glucose as a carbon source, while the triterpenoid biosynthetic genes are being activated by β -estradiol (McIsaac et al., 2011). Meanwhile, INVSc1 utilizes galactose as its growth medium to active GAL promoters, which presents the drawback of catabolic repression in the presence of glucose (Escalante-Chong et al., 2015). Upon growing in glucose, the PSIII strain exhibited a much higher growth rate than INVSc1, which might affect its ability to produce triterpenoids. One of the factors affecting enzyme activity is substrate concentration (Zhu et al., 2018). In the PSIII strain, the abundant β -amyrin concentration might drive the CYP-CPR pairs, which otherwise show lower conversion ratio

of β -amyrin in the INVSc1 strain. The fact that CYP716As and CYP72As are both widespread in plants and possess diverse functions in triterpenoid site-specific oxidations might be the reason they work well with all CPR classes in the engineered yeast strain.

Despite many variables affecting gene expression and metabolic regulation in yeast, this study achieved the highest conversion of β -amyrin into 11-oxo- β -amyrin by pairing *GuCYP88D6* with *GuCPR2*. Previous research has attempted to screen different CPRs for glycyrrhetic acid production and reported that *GuCPR1* is the best fit for *CYP88D6* (Zhu et al., 2018). However, this study revealed that pairing *CYP88D6* with *GuCPR2* showed 1.8 times higher conversion ratio of β -amyrin into 11-oxo- β -amyrin than pairing with *GuCPR1*. This result also reveals that the conversion of β -amyrin in PSIII is still inefficient, as compared to the INVSc1 strain, as shown in the lower ratio of the oxidized products of β -amyrin. This implies that there is still an opportunity to improve triterpenoid production in yeast by increasing the conversion efficiency of β -amyrin.

In conclusion, this study revealed that pairing different plant CYP families with different plant CPR classes results in different triterpenoid conversion ratio in heterologous yeast. This is the first study to compare different pairs of CPR classes and CYPs from legumes in transgenic yeast. The comparison data of this study did not cover enough CPR-CYP pairs from different species to generalize which CPR class is better for which CYP. Nevertheless, this study highlighted that CYP does have a preference for certain CPR classes, rejecting the previous notion that CPR-II is responsible for specialized metabolism. A strategy that could be proposed through this study is that performing screening for both CPR classes from the same species as their target CYPs would increase the chance of finding the best CPR pair for their heterologous production system. This study cannot be generalized to other species, but at least for legume species, the CYP72A family has a higher chance of working better with CPR class II. Furthermore, CYP716A has a higher probability of working better with CPR class I of the same species. Initiating research based on this fact might help other researchers find the best CPR pairs.

However, the underlying mechanisms of how a CPR prefers a specific CYP have not yet been fully elucidated. Therefore, more CPRs and CYPs from different plant families need to be studied in order to gain a deeper understanding of CPR class preferences toward CYPs. *In planta* and *in vitro* enzyme studies are necessary to further analyze the mechanisms of CYP-CPR interactions. Nevertheless, the observation that legume CPR classes support different CYPs in the heterologous production of triterpenoids in transgenic yeast may facilitate research on improving heterologous production of other beneficial compounds by co-expressing the proper CPR with the target CYPs.

DATA AVAILABILITY STATEMENT

Raw RNA-Seq reads of *G. uralensis* obtained in this study have been submitted to the DNA Data Bank of Japan

(DDBJ) Sequence Read Archive (DRA) under the accession numbers of DRA012266 and run accession numbers as listed in **Supplementary Table 10**.

AUTHOR CONTRIBUTIONS

PI, EF, and TM designed the experiments. PS constructed the PSIII yeast strain. KT, AC, HSu, and HSe conducted experiments on RNA-sequencing. PI conducted the rest of the experiments, analyzed the results, and wrote the whole manuscript. EF, SY, HSe, and TM conceived and supervised the study. PS, SY, HSe, and TM made the manuscript revisions.

FUNDING

This study was financially supported in part by the Grant-in-Aid for Scientific Research no. JP19H02921 awarded to TM and 17K07754, 20H02913 awarded to HSe, and the Monbukagakusho Scholarship from MEXT awarded to PI.

REFERENCES

- Agrawal, S. B., and Agrawal, M. (1999). *Environmental Pollution And Plant Responses*. Boca Raton, Fla: Lewis Publishers.
- Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22, 195–201. doi: 10.1093/bioinformatics/bti770
- Bart, A. G., and Scott, E. E. (2018). Structures of human cytochrome P450 1A1 with bergamottin and erlotinib reveal active-site modifications for binding of diverse ligands. *J. Biol. Chem.* 293, 19201–19210. doi: 10.1074/jbc.RA118.005588
- Burkhard, F. Z., Parween, S., Udhane, S. S., Flück, C. E., and Pandey, A. V. (2017). P450 oxidoreductase deficiency: Analysis of mutations and polymorphisms. *J. Steroid Biochem. Mol. Biol.* 165, 38–50. doi: 10.1016/j.jsbmb.2016.04.003
- Campelo, D., Esteves, F., Palma, B. B., Gomes, B. C., Rueff, J., Lautier, T., et al. (2018). Probing the role of the hinge segment of cytochrome P450 oxidoreductase in the interaction with cytochrome P450. *Int. J. Mol. Sci.* 19:3914. doi: 10.3390/ijms19123914
- Chung, S. Y., Seki, H., Fujisawa, Y., Shimoda, Y., Hiraga, S., Nomura, Y., et al. (2020). A cellulose synthase-derived enzyme catalyses 3-O-glucuronosylation in saponin biosynthesis. *Nat. Commun.* 11:5664. doi: 10.1038/s41467-020-19399-0
- DeLano, W. L. (2020). *The PyMOL Molecular Graphics System, Version 2.3*. New York: Schrödinger LLC.
- Doroodian, P., and Hua, Z. (2021). The ubiquitin switch in plant stress response. *Plants* 10:246. doi: 10.3390/plants10020246
- Ebizuka, Y., Shibuya, M., and Wakita, E. (2011). *C-22 hydroxylase*. U.S. Patent No 2011/0171698 A1. Washington, DC: U.S. Patent and Trademark Office.
- Escalante-Chong, R., Savir, Y., Carroll, S. M., Ingraham, J. B., Wang, J., Marx, C. J., et al. (2015). Galactose metabolic genes in yeast respond to a ratio of galactose and glucose. *Proc. Natl. Acad. Sci. U. S. A.* 112, 1636–1641. doi: 10.1073/pnas.1418058112
- Esteves, F., Campelo, D., Gomes, B. C., Urban, P., Bozonnet, S., Lautier, T., et al. (2020). The role of the FMN-domain of human cytochrome P450 oxidoreductase in its promiscuous interactions with structurally diverse redox partners. *Front. Pharmacol.* 11:299. doi: 10.3389/fphar.2020.00299
- Fanani, M. Z., Fukushima, E. O., Sawai, S., Tang, J., Ishimori, M., Sudo, H., et al. (2019). Molecular basis of C-30 product regioselectivity of legume oxidases involved in high-value triterpenoid biosynthesis. *Front. Plant Sci.* 10, 1–16. doi: 10.3389/fpls.2019.01520

ACKNOWLEDGMENTS

We thank Ms. Sayaka Shinpo (Kazusa DNA Research Institute) for technical support with the library construction and Illumina sequencing. We thank Dr. Kiyoshi Ohyama (Tokyo Institute of Technology) for providing sophoradiol and 24-hydroxy- β -amyrin standards. We also thank Aarhus University for providing seeds of *L. japonicus* Gifu B-129 accession (LORE1 insertion lines). We thank Susan Lindquist for the kind gift of pAG415GAL-ccdB yeast expression vector. A *de novo* assembly of *G. uralensis* RNA-seq analysis was performed on the NIG supercomputer at the ROIS National Institute of Genetics. This study will be included in a dissertation submitted by PI to Osaka University in partial fulfillment of the requirement for her doctoral degree.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.762546/full#supplementary-material>

- Fukushima, E. O., Seki, H., Ohyama, K., Ono, E., Umamoto, N., Mizutani, M., et al. (2011). CYP716A subfamily members are multifunctional oxidases in triterpenoid biosynthesis. *Plant Cell Physiol.* 52, 2050–2061. doi: 10.1093/pcp/pcr146
- Guang, C., Chen, J., Sang, S., and Cheng, S. (2014). Biological functionality of soyasaponins and soyasapogenols. *J. Agric. Food Chem.* 62, 8247–8255. doi: 10.1021/jf503047a
- Guo, Q., Yoshida, Y., Major, I. T., Wang, K., Sugimoto, K., Kapali, G., et al. (2018). JAZ repressors of metabolic defense promote growth and reproductive fitness in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* 115, E10768–E10777. doi: 10.1073/pnas.1811828115
- Hamdane, D., Xia, C., Im, S. C., Zhang, H., Kim, J. J. P., and Waskell, L. (2009). Structure and function of an NADPH-cytochrome P450 oxidoreductase in an open conformation capable of reducing cytochrome P450. *J. Biol. Chem.* 284, 11374–11384. doi: 10.1074/jbc.M807868200
- Hong, Y., Li, H., Yuan, Y. C., and Chen, S. (2010). Sequence-function correlation of aromatase and its interaction with reductase. *J. Steroid Biochem. Mol. Biol.* 118, 203–206. doi: 10.1016/j.jsbmb.2009.11.010
- Jang, H. H., Jamakhandi, A. P., Sullivan, S. Z., Yun, C. H., Hollenberg, P. F., and Miller, G. P. (2010). Beta sheet 2-alpha helix C loop of cytochrome P450 reductase serves as a docking site for redox partners. *Biochim. Biophys. Acta Proteins Proteom.* 1804, 1285–1293. doi: 10.1016/j.bbapap.2010.02.003
- Jensen, K., and Möller, B. L. (2010). Plant NADPH-cytochrome P450 oxidoreductases. *Phytochemistry* 71, 132–141. doi: 10.1016/j.phytochem.2009.10.017
- Jin, C. C., Zhang, J. L., Song, H., and Cao, Y. X. (2019). Boosting the biosynthesis of betulinic acid and related triterpenoids in *Yarrowia lipolytica* via multimodular metabolic engineering. *Microb. Cell Fact.* 18:77. doi: 10.1186/s12934-019-1127-8
- Karim, A. S., Curran, K. A., and Alper, H. S. (2013). Characterization of plasmid burden and copy number in *Saccharomyces cerevisiae* for optimization of metabolic engineering applications. *FEMS Yeast Res.* 13, 107–116. doi: 10.1111/1567-1364.12016
- Kim, O. T., Um, Y., Jin, M. L., Kim, J. U., Hegebarth, D., Busta, L., et al. (2018). A novel multifunctional C-23 Oxidase, CYP714E19, is involved in asiaticoside biosynthesis. *Plant Cell Physiol.* 59, 1200–1213. doi: 10.1093/pcp/pcy055
- Kitagawa, I. (2002). Licorice root. A natural sweetener and an important ingredient in Chinese medicine. *Pure Appl. Chem.* 74, 1189–1198. doi: 10.1351/pac200274071189

- McIsaac, R. S., Silverman, S. J., McClean, M. N., Gibney, P. A., Macinskas, J., Hickman, M. J., et al. (2011). Fast-acting and nearly gratuitous induction of gene expression and protein depletion in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 22, 4447–4459. doi: 10.1091/mbc.E11-05-0466
- Misra, R. C., Sharma, S., Sandeep, S., Garg, A., Chanotiya, C. S., and Ghosh, S. (2017). Two CYP716A subfamily cytochrome P450 monooxygenases of sweet basil play similar but nonredundant roles in ursane- and oleanane-type pentacyclic triterpene biosynthesis. *New Phytol.* 214, 706–720. doi: 10.1111/nph.14412
- Mizutani, M., and Ohta, D. (1998). Two isoforms of NADPH:cytochrome p450 reductase in *Arabidopsis thaliana* gene structure, heterologous expression in insect cells, and differential regulation. *Plant Physiol.* 116, 357–367. doi: 10.1104/pp.116.1.357
- Moktali, V., Park, J., Fedorova-Abrams, N. D., Park, B., Choi, J., Lee, Y. H., et al. (2012). Systematic and searchable classification of cytochrome P450 proteins encoded by fungal and oomycete genomes. *BMC Genomics* 13:525. doi: 10.1186/1471-2164-13-525
- Oldfield, E., and Lin, F. Y. (2012). Terpene biosynthesis: modularity rules. *Angew. Chem. Int. Ed.* 51, 1124–1137. doi: 10.1002/anie.201103110
- Paine, M. J. I., Scrutton, N. S., Munro, A. W., Gutierrez, A., Roberts, G. C. K., and Roland Wolf, C. (2005). “Electron transfer partners of cytochrome P450,” in *Cytochrome P450: structure, Mechanism, and Biochemistry: third edition*, ed. P. R. Ortiz de Montellano (Boston, MA: Springer US), 115–148. doi: 10.1007/0-387-27447-2_4
- Pandey, A. V., and Flück, C. E. (2013). NADPH P450 oxidoreductase: structure, function, and pathology of diseases. *Pharmacol. Ther.* 138, 229–254. doi: 10.1016/j.pharmthera.2013.01.010
- Parage, C., Foureau, E., Kellner, F., Burlat, V., Mahroug, S., Lanoue, A., et al. (2016). Class II cytochrome P450 reductase governs the biosynthesis of alkaloids. *Plant Physiol.* 172, 1563–1577. doi: 10.1104/pp.16.00801
- Qiu, X. M., Sun, Y. Y., Ye, X. Y., and Li, Z. G. (2020). Signaling role of glutamate in plants. *Front. Plant Sci.* 10:1743. doi: 10.3389/fpls.2019.01743
- Qu, X., Pu, X., Chen, F., Yang, Y., Yang, L., Zhang, G., et al. (2015). Molecular cloning, heterologous expression, and functional characterization of an NADPH-cytochrome P450 reductase gene from *Camptotheca acuminata*, a camptothecin-producing plant. *PLoS One* 10:e0135397. doi: 10.1371/journal.pone.0135397
- Rana, S., Lattoo, S. K., Dhar, N., Razdan, S., Bhat, W. W., Dhar, R. S., et al. (2013). NADPH-cytochrome P450 reductase: Molecular cloning and functional characterization of two paralogs from *Withania somnifera* (L.) Dunal. *PLoS One* 8:e57068. doi: 10.1371/journal.pone.0057068
- Rao, K. M., Raghavendra, A., and Reddy, K. J. (2006). *Physiology And Molecular Biology Of Stress Tolerance In Plants*. Dordrecht: Springer, doi: 10.1007/1-4020-4225-6
- Ritacco, I., Spinello, A., Ippoliti, E., and Magistrato, A. (2019). Post-translational regulation of CYP450s metabolism as revealed by all-atoms simulations of the aromatase enzyme. *J. Chem. Inf. Model.* 59, 2930–2940. doi: 10.1021/acs.jcim.9b00157
- Ro, D. K., Ehrling, J., and Douglas, C. J. (2002). Cloning, functional expression, and subcellular localization of multiple NADPH-cytochrome P450 reductases from hybrid poplar. *Plant Physiol.* 130, 1837–1851. doi: 10.1104/pp.008011
- Santos, J., Rizzo, V. A., Sica, M. P., and Ermácora, M. R. (2007). Effects of serine-to-cysteine mutations on β -lactamase folding. *Biophys. J.* 93, 1707–1718. doi: 10.1529/biophysj.106.103804
- Scarpella, E., and Meijer, A. H. (2004). Pattern formation in the vascular system of monocot and dicot plant species. *New Phytol.* 164, 209–242. doi: 10.1111/j.1469-8137.2004.01191.x
- Seki, H., Ohyama, K., Sawai, S., Mizutani, M., Ohnishi, T., Sudo, H., et al. (2008). Licorice β -amyirin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proc. Natl. Acad. Sci. U. S. A.* 105, 14204–14209. doi: 10.1073/pnas.0803876105
- Seki, H., Sawai, S., Ohyama, K., Mizutani, M., Ohnishi, T., Sudo, H., et al. (2011). Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. *Plant Cell* 23, 4112–4123. doi: 10.1105/tpc.110.082685
- Seki, H., Tamura, K., and Muranaka, T. (2015). P450s and UGTs: key players in the structural diversity of triterpenoid saponins. *Plant Cell Physiol.* 56, 1463–1471. doi: 10.1093/pcp/pcv062
- Shephard, E. A., Phillips, I. R., Bayney, R. M., Pike, S. F., and Rabin, B. R. (1983). Quantification of NADPH:cytochrome P-450 reductase in liver microsomes by a specific radioimmunoassay technique. *Biochem. J.* 211, 333–340. doi: 10.1042/bj2110333
- Simtchouk, S., Eng, J. L., Meints, C. E., Makins, C., and Wolthers, K. R. (2013). Kinetic analysis of cytochrome P450 reductase from *Artemisia annua* reveals accelerated rates of NADH-dependent flavin reduction. *FEBS J.* 280, 6627–6642. doi: 10.1111/febs.12567
- Srisawat, P., Yasumoto, S., Fukushima, E. O., Robertlee, J., Seki, H., and Muranaka, T. (2020). Production of the bioactive plant-derived triterpenoid morolic acid in engineered *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* 117, 2198–2208. doi: 10.1002/bit.27357
- Suzuki, H., Fukushima, E. O., Shimizu, Y., Seki, H., Fujisawa, Y., Ishimoto, M., et al. (2019). Lotus japonicus triterpenoid profile and characterization of the CYP716A51 and LjCYP93E1 genes Involved in their biosynthesis in planta. *Plant Cell Physiol.* 60, 2496–2509. doi: 10.1093/pcp/pcz145
- Suzuki, H., Fukushima, E. O., Umemoto, N., Ohyama, K., Seki, H., and Muranaka, T. (2018). Comparative analysis of CYP716A subfamily enzymes for the heterologous production of C-28 oxidized triterpenoids in transgenic yeast. *Plant Biotechnol.* 35, 131–139. doi: 10.5511/plantbiotechnology.18.0416a
- Tamura, K., Seki, H., Suzuki, H., Kojima, M., Saito, K., and Muranaka, T. (2017). CYP716A179 functions as a triterpene C-28 oxidase in tissue-cultured stolons of *Glycyrrhiza uralensis*. *Plant Cell Rep.* 36, 437–445. doi: 10.1007/s00299-016-2092-x
- Urban, P., Mignotte, C., Kazmaier, M., Delorme, F., and Pompon, D. (1997). Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. *J. Biol. Chem.* 272, 19176–19186. doi: 10.1074/jbc.272.31.19176
- Waskell, L., and Kim, J.-J. (2015). “Electron transfer partners of cytochrome P450,” in *Cytochrome P450*, ed. P. R. Ortiz de Montellano (Basel: Springer), 33–68.
- Whitelaw, D. A., Tonkin, R., Meints, C. E., and Wolthers, K. R. (2015). Kinetic analysis of electron flux in cytochrome P450 reductases reveals differences in rate-determining steps in plant and mammalian enzymes. *Arch. Biochem. Biophys.* 584, 107–115. doi: 10.1016/j.abb.2015.09.002
- Wright, T. R., Shan, G., Walsha, T. A., Lira, J. M., Cui, C., Song, P., et al. (2010). Robust crop resistance to broadleaf and grass herbicides provided by aryloxyalkanoate dioxygenase transgenes. *Proc. Natl. Acad. Sci. U. S. A.* 107, 20240–20245. doi: 10.1073/pnas.1013154107
- Xu, R., Fazio, G. C., and Matsuda, S. P. T. (2004). On the origins of triterpenoid skeletal diversity. *Phytochemistry* 65, 261–291. doi: 10.1016/j.phytochem.2003.11.014
- Yang, C. Q., Lu, S., Mao, Y. B., Wang, L. J., and Chen, X. Y. (2010). Characterization of two NADPH: cytochrome P450 reductases from cotton (*Gossypium hirsutum*). *Phytochemistry* 71, 27–35. doi: 10.1016/j.phytochem.2009.09.026
- Yasumoto, S., Fukushima, E. O., Seki, H., and Muranaka, T. (2016). Novel triterpene oxidizing activity of *Arabidopsis thaliana* CYP716A subfamily enzymes. *FEBS Lett.* 590, 533–540. doi: 10.1002/1873-3468.12074
- Zhu, M., Wang, C., Sun, W., Zhou, A., Wang, Y., Zhang, G., et al. (2018). Boosting 11-oxo- β -amyirin and glycyrrhetic acid synthesis in *Saccharomyces cerevisiae* via pairing novel oxidation and reduction system from legume plants. *Metab. Eng.* 45, 43–50. doi: 10.1016/j.ymben.2017.11.009

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Istiadari, Yasumoto, Srisawat, Tamura, Chikugo, Suzuki, Seki, Fukushima and Muranaka. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.