Expression and Functional Analysis of a Transgenic Cytochrome P450 Monoxygenase in *Pueraria mirifica*  
(Analisis Ekspresi dan Fungsian Transgen Sitokrom P450 Monooksigenase dalam *Pueraria mirifica*)

**LI XI, RUETHAITHIP KAOPONG, SUREEPORN NUALKAEW, ARTHIYA CHULLASARA & AMORNAT PHONGDARA**

**ABSTRACT**

Isoflavonoids are the main compound in White Kwao Krua (*Pueraria mirifica*), which is an effective folk medicinal plant endemic to Thailand. It has been widely used for improving human physical and treating diseases. There are substances with estrogenic activities have been isolated from *P. mirifica*, such as puerarin, daidzein and genistein. Isoflavone synthase (*IFS*) is one of the key enzymes in Leguminous plants to convert liquiritigenin, liquiritigenin C-glucoside and naringenin chalcone to isoflavonoids. The aim of this research was to enhance the production of isoflavonoids by metabolic engineering. Transgenic plants were constructed by introducing P450 gene (*EgP450*) which is similar to *IFS* from oil palm (*Elaeis guineensis*), into *P. mirifica* by a biolistic method. After the transgenic plants had proved successfully, isoflavonoids of each group plants were determined by HPLC. The contents of daidzein and genistein in transgenic plants were higher than the control plants.

**Keywords:** *Elaeis guineensis*; isoflavonoids; *Pueraria mirifica*; P450 gene; transgenic plants

**ABSTRAK**


**Kata kunci:** *Elaeis guineensis*; gen P450; isoflavonoid; *Pueraria mirifica*; tumbuhan transgen

**INTRODUCTION**

*Pueraria mirifica* has a long using history in folk medicine in Thailand. This plant grows in the wild in northern, northeastern and western of Thailand (Cherdshewasart et al. 2007a). It belongs to *Leguminosae* family, locally called different names, but of which white Kwao Krua is more popular (Lakshnakara et al. 1952). *P. mirifica* has been claimed to improve the human physical appearances such as re-growing hair, promoting black hair, enhance the flexibility of the body and sexual performance, recovering smooth skin and prolonging life. In recent years, *P. mirifica* has been included into cosmetic, dietary supplement and pharmaceutical products (Malaviijitnond 2012).

More than 20 chemical substances with estrogenic activities, mainly isoflavones, have been isolated from *P. mirifica*, such as puerarin, daidzein, genistein, kwakhurin, tuberosin, and mirificin (Malaviijitnond 2012). Isoflavonoids, a class of flavonoid phenolic compounds, are particularly prevalent in *Leguminosae* (Dixon & Sumner 2003). By now, puerarin (Urasopon et al. 2008), miroestrol (Yusakul et al. 2011) and kwakhurin (Ingham et al. 2002; Ito et al. 2005) are found only in *P. mirifica*. The isoflavonoids of *P. mirifica* depended on genotype (Cherdshewasart & Sriwatcharakul 2007), environmental effect and plant age, but they are almost less than 0.2% of isoflavonoid contents (Cherdshewasart et al. 2007b; Malaviijitnond et al. 2004; Urasopon et al. 2008).

In *Pueraria*, the whole processes of isoflavonoids biosynthesis include two main stages (Dhaubhadel 2011; Franzmayr et al. 2012; He et al. 2011). The first stage is the biosynthesis of the common precursor of isoflavonoid, the 4-coumaroyl-CoA, which is synthesized orderly from phenylalanine by phenylalanine ammonia-lyase (PAL), cinnamic acid 4-hydroxylase (C4H) and 4-coumarate: CoA ligase (4CL). The biosynthesis genistein, daidzein, puerarin and other isoflavonoids from 4-coumaroyl-CoA and aroyl-CoA is the second stage. The naringenin chalcone synthesized from 4-coumaroyl-CoA and
The seeds of *P. mirifica* were used for *in vitro* plant cultures to obtain transgenic plants. The *P. mirifica* plants were discriminated and identified by its DNA barcodes.

**MATERIALS AND METHODS**

**PLANT MATERIALS**

The seeds of *P. mirifica* plants were brought from Lampang province of Thailand. They were used for *in vitro* plant cultures to obtain transgenic plants. The *P. mirifica* plants were discriminated and identified by its DNA barcodes.

**IDENTIFYING *P. MIRIFICA* SPECIES**

DNA EXTRACTION, AMPLIFICATION AND SEQUENCING

Leaf tissues of *P. mirifica* dried in silica gel. Total DNA was extracted using the plant genomic DNA Mini Kit (Gene aid, New Taipei, Taiwan). Four DNA barcodes (ITS region, *rbcL*, *matK*, and *psbA-trnH*) amplified by the polymerase chain reaction (PCR) technique. The total volume of reaction for 25 μL consisted of 2 μL (~60 ng) DNA, 1 μL of 50 mM MgCl₂, 2.5 μL of 10 × PCR buffer, 0.2 μL of 5 U/μL Pfu Taq DNA polymerase, 0.5 μL L of 10 mM dNTPs mix, 1.0 μL 2.5 μM of (primers), 17.8 μL deionized water. PCR products were first examined by 1.5% agarose gel electrophoresis and purified using the QIAquick® PCR Purification Kit (Qiagen, Hilden, Germany) and then sequenced in both directions with the primers used for PCR amplification on a sequencer ABI 310 (Applied Biosystems, USA). The sequences submitted to GenBank (supplement Table 1).

**SEQUENCING AND BIOINFORMATICS ANALYSIS**

DNA sequences alignments were performed using ClustalX2 and displayed using GeneDoc software (Larkin et al. 2007). The evolutionary relationships of taxa were inferred using the UPGMA method (Sneath & Sokal 1973), to construct phylogenetic trees by bootstrap analysis from 1000 replications of the data (Felsenstein 1985). The trees are drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the maximum composite likelihood method. All positions containing gaps and missing data were eliminated. Positions in the final dataset of evolutionary analyses were conducted in MEGA6 (Tamura et al. 2013). The divergence between sequences was analyzed and the number of base substitutions per site from between sequences was calculated. Analyses were conducted using the maximum composite likelihood model (Tamura et al. 2004).

**SUPPLEMENT TABLE 1. Sequences of the Primers for amplification**

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequences (5'→3')</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>rbcl</em>-F</td>
<td>ATGTCTCCCAAAAGAAGACAGACGTAAGGCA</td>
</tr>
<tr>
<td><em>rbcl</em>-R</td>
<td>ATGTCTCCCAAAAGAAGACAGACGTAAGGCA</td>
</tr>
<tr>
<td><em>matK</em>-F</td>
<td>GGTGACACTGCTGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
<tr>
<td><em>matK</em>-R</td>
<td>GGTGACACTGCTGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
<tr>
<td><em>psba</em>-trnH-F</td>
<td>GGTGACACTGCTGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
<tr>
<td><em>psba</em>-trnH-R</td>
<td>GGTGACACTGCTGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
<tr>
<td>ITS region-F</td>
<td>GGTGACACTGCTGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
<tr>
<td>ITS region-R</td>
<td>GGTGACACTGCTGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
<tr>
<td>18S-F</td>
<td>AGGACCTGCTGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
<tr>
<td>18S-R</td>
<td>AGGACCTGCTGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
<tr>
<td>GFP1303-F</td>
<td>GGTGACACTGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
<tr>
<td>GFP1303-R</td>
<td>GGTGACACTGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
<tr>
<td>EgP450-F</td>
<td>ATCCATGCTGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
<tr>
<td>EgP450-R</td>
<td>ATCCATGCTGCTGCTGCTGCTGCTGCTGCTG</td>
</tr>
</tbody>
</table>

**PRODUCTION OF TRANSGENIC PLANTS**

**PLANT MATERIALS**

The seeds of *P. mirifica* were cultivated in three flowerpots containing soil. Each flowerpot has nine plant seedlings which were treated in transformation while the spouts grow at 3-5 cm.
PREPARATION OF RECOMBINANT EG P450 AND TRANSGENIC PLANTS

The full-length cDNA of P450 enzyme from *E. guineensis* (EgP450, Genebank: ADM88550.1) was sub-cloned into pCAMBIA vector. Then, the particular primer set corresponding to the beginning of the ORF with the addition of an upstream in-frame are used to amplify EgP450 by PCR machine. The PCR products are ligated to a similarly digested pCAMBIA 1303 (CAMBIA, Australia), carrying the reporter gene (gus-gfp) controlled by a Cauliflower Mosaic Virus (CaMV35S) promoter. The inserted plasmid DNA was sequenced using the 3730 DNA sequencer to ensure the authenticity of the cloned nucleotide sequence (Supplement Figure 1).

The seedling, approximately 3-5 cm in size, were selected from the same seedling line and bombarded using the PDS-1000/He particle delivery system (BioRad, USA), the conditions for biolistic-mediated gene transfer for *P. mirifica* were optimized previously (Nakkaew et al. 2010) and the pressure of 1100 psi was chosen for bombardment transformation. Distance between stopping screen and target shelf was 90 mm and the vacuum was 27 mmHg. Gold particle size was 1.0 μm. Bombardments without DNA and with an empty vector were used as experimental control (Nakkaew et al. 2010).

IDENTIFICATION OF THE RESULT OF TRANSFORMATION

The expressions of the GFP gene in putative transgenic seedling were visualized using an Olympus BX51 compound microscope and a DP-71 digital camera fitted with a fluorescence unit cubes.

Genomic DNA from putative transgenic plants were isolated followed the user’s manual of DNA extraction kits. The amplification was performed in GeneAmp PCR System 2720 (Perkin-Elmer, USA). The CaMV35S promoter primer and EgP450 R SpeI were used in the amplification to confirm the insertion of recombinant EgP450 into the genome of the seedling. Thus GFP primers set will be used as a control to ensure a complete of transgenic *P. mirifica*.

ISOFLAVONOIDS PRODUCTS ANALYSIS VIA HPLC

Puerarin, genistein and daidzein of the 3 groups of isoflavonoids product were determined by high-performance liquid chromatography (HPLC) (Agilent 1200, America) from transgenic plants. It was performed using Hypersil BDS C18 column, mobile phase including methanol and water by gradient elution (Flow rate: 0.7 ml•min⁻¹; Injection volume: 10 μL; Column temperature: 25°C; detection wavelength: 250 nm).

MECHANISM ANALYSIS

If the expected result of this research has been obtained, the mechanism analysis of EgP450 and IFS (P450 from *P. mirifica*) will be performed using PROSITE (http://prosite.expasy.org/) and describing protein domains, functional sites as well as associated patterns and profiles to identify them.

STATISTICAL ANALYSIS

The contents of puerarin, genistein and daidzein among the 3 groups of *P. mirifica* were compared and the results were also expressed as the Mean±S.D. Data analysis was done base on analysis of variance (ANOVA) using R 2.50 version software (R Development Core Team. 2012) with Hmisc package (Harrell 2013). Significant level utilized in the data analysis was 0.05.

RESULTS

IDENTIFYING *P. MIRIFICA* SPECIES

A total of 15 DNA sequences of White Kwao Krua has been obtained (supplement Table 2). Choosing one of each barcode sequences was performed with BLAST analysis (http://blast.ncbi.nlm.nih.gov/Blast.cgi). *matK* and *ITS* region DNA sequences of *P. mirifica* could be found in BLAST results. The pairwise distances between the sequences of *P. mirifica* in NCBI and the sequences which obtained in the research plants are not more than

---

**SUPPLEMENT FIGURE 1. The diagram of pCAMBIA 1303-EgP450**
of variance (ANOVA). Genistein contents were statistically significant between T1 and T2 group plants \((p<0.01)\) and they were also statistically significant between T1 and T3 group plants \((p<0.05)\). Daidzein contents in the T1 group plants are statistical significant than other two groups \((p<0.05)\). Puerarin contents in three group trees show no significant differences \((F=3.445, \ p>0.05)\).

**MECHANISM ANALYSIS**

BLASTX alignments of EgP450 protein, and then chose 4 similar sequences were from other species to compare using clustalX2. The EgP450 proteins contain the similar characteristic functional in the six species, such as oxygen binding site region, aromatic region and active site (red box). From the structure characteristics, EgP450 and \textit{P. mirifica} P450 should present similar functions.

**DISCUSSION**

In our study, White Kwao Krua was bought from Lampang province of Thailand and identified by professional taxonomist; we further confirmed the species by DNA barcode technique. The \textit{rbcL}, \textit{psbA-trnH}, \textit{matK} and \textit{ITS} region sequence of White Kwao Krua were obtained in this research. The phylogenetic results of these sequences were constructed. In combination with the morphological characteristics of plants and other information, we can accurately identify that the White Kwao Krua which was bought from Lampang province were \textit{P. candollei} var \textit{mirifica} or \textit{P. mirifica}.

In previous work, a clone homologous to P450 (namely \textit{EgP450}) was selected from the list of highly expressed genes in our oil palm EST libraries (Phongdara et al. 2012). After an in-depth analysis, the P450 protein of oil

**SUPPLEMENT TABLE 2.** The length, percentage similarity of overlap section and GenBank accession of DNA fragments

<table>
<thead>
<tr>
<th>DNA fragments</th>
<th>Length (bp)</th>
<th>Pairwise distances scot</th>
<th>Voucher number</th>
<th>Gen Bank accession</th>
<th>GI</th>
</tr>
</thead>
<tbody>
<tr>
<td>matK</td>
<td>Sample 1</td>
<td>532</td>
<td>--</td>
<td>DLXY130517PM01</td>
<td>KM982345 874527190</td>
</tr>
<tr>
<td></td>
<td>Sample 2</td>
<td>532</td>
<td>--</td>
<td>DLXY130517PM02</td>
<td>KM982346 874527192</td>
</tr>
<tr>
<td></td>
<td>Sample 3</td>
<td>532</td>
<td>--</td>
<td>DLXY130517PM03</td>
<td>KM982347 874527194</td>
</tr>
<tr>
<td>18S</td>
<td>Sample 1</td>
<td>354</td>
<td>--</td>
<td>DLXY130517PM01</td>
<td>KM982348 874527196</td>
</tr>
<tr>
<td></td>
<td>Sample 2</td>
<td>354</td>
<td>--</td>
<td>DLXY130517PM02</td>
<td>KM982349 874527197</td>
</tr>
<tr>
<td></td>
<td>Sample 3</td>
<td>354</td>
<td>--</td>
<td>DLXY130517PM03</td>
<td>KM982350 874527198</td>
</tr>
<tr>
<td>ITS region</td>
<td>Sample 1</td>
<td>790</td>
<td>0.000</td>
<td>DLXY130517PM01</td>
<td>KM982351 874527199</td>
</tr>
<tr>
<td></td>
<td>Sample 2</td>
<td>790</td>
<td>0.000</td>
<td>DLXY130517PM02</td>
<td>KM982352 874527200</td>
</tr>
<tr>
<td></td>
<td>Sample 3</td>
<td>791</td>
<td>0.000</td>
<td>DLXY130517PM03</td>
<td>KM982353 874527201</td>
</tr>
<tr>
<td>psbA-trnH</td>
<td>Sample 1</td>
<td>723</td>
<td>--</td>
<td>DLXY130517PM01</td>
<td>KM982354 874527202</td>
</tr>
<tr>
<td></td>
<td>Sample 2</td>
<td>680</td>
<td>--</td>
<td>DLXY130517PM02</td>
<td>KM982355 874527204</td>
</tr>
<tr>
<td></td>
<td>Sample 3</td>
<td>724</td>
<td>--</td>
<td>DLXY130517PM03</td>
<td>KM982356 874527206</td>
</tr>
<tr>
<td>rbcL</td>
<td>Sample 1</td>
<td>729</td>
<td>$\leq 0.006$</td>
<td>DLXY130517PM01</td>
<td>KM982357 874527208</td>
</tr>
<tr>
<td></td>
<td>Sample 2</td>
<td>700</td>
<td>$\leq 0.006$</td>
<td>DLXY130517PM02</td>
<td>KM982358 874527210</td>
</tr>
<tr>
<td></td>
<td>Sample 3</td>
<td>700</td>
<td>$\leq 0.003$</td>
<td>DLXY130517PM03</td>
<td>KM982359 874527212</td>
</tr>
</tbody>
</table>

The results of pairwise distances scot was respectively compared with \textit{P. mirifica} in NCBI
FIGURE 1. The phylogenetic trees analysis of four barcodes (a; \textit{rbcL}, b; \textit{psbA-trnH}, c; \textit{matK} and d; ITS region) of White Kwao Krua and other species were analyzed by maximum likelihood method and evolutionary analyses were conducted in MEGA6.

FIGURE 2. The putative transgenic of \textit{P. mirifica} were cultivated in the small flowerpots containing soil until the sprouts height reach five centimeters and were used for transformation. The three sets of flowerpots of \textit{P. mirifica} were bombarded for each experiment with following DNA: pCAMBIA 1303-\textit{EgP450} (T1), empty pCAMBIA 1303 vector (T2) and gold particle control (T3).

FIGURE 3. Fluorescence microscope with cell sens standard software was chosen to identify the transgenic plants. Fluorescent light found in T1 plant (pCAMBIA 1303-\textit{EgP450}) and T2 plant (-empty pCAMBIA 1303 vector) but visualization of GFP fluorescence not found inside the T3 plants (1.0 μm gold particle control).
palm contains the same characteristic function domains as the CYP71 superfamily (Nelson et al. 1993). The sequence alignment demonstrates the closed relationships of EgP450 and P450 of protein from NCBI database. From the structure characteristics, EgP450, EgIFS and P. mirifica P450 should present similar domains of function.

IFS is one of the essential enzymes for engineering isoflavonoids production, because it is the entry point enzyme of biosynthesis. cDNAs encoding IFS have cloned from leguminous plants, and it has been transformed into Arabidopsis thaliana, tobacco and corn (Jung et al. 2000; Steele et al. 1999; Yu et al. 2000). Free genistein does not accumulate in Arabidopsis which express soybean IFS; then the isoflavone is converted to other compounds. The level of genistein conjugate production depends on the IFS activity level, substrate availability, substrate channeling and product turnover (Dixon & Ferreira 2002).

To avoid risks of homology-dependent gene silencing (HDGS) (Meyer & Saedler 1996), in this research EgP450 was chosen to transfer into P. mirifica, which directly cultivated in the flowerpots. The contents of isoflavonoids such as daidzein and genistein in transgenic P. mirifica, were higher than the control plants. There are some reasons about the mechanism. First, EgP450 and P. mirifica P450, which acting as Isoflavone synthase, should present similar functions in P. mirifica. The second, genetically modified donor plants from different classes may be able to avoid HDGS. The idea of choosing the synthetases from species of far genetic distance to create transgenic plants that were demonstrated here may provide a new way to obtain an excellent crop or herb in future works.

Daidzein of Pueraria mirifica has an anti-osteoporotic effect in postmenopausal women (Manonai et al. 2008). Genistein were used in both prophylaxis and treatment of diseases in recent years, acting such as antioxidant (Ganai & Farooqi 2015; Kanazawa et al. 1995), anticaner properties (Sarkar & Li 2002; Qi et al. 2011), cardioprotection (Deodato et al. 1999) and immunosupressor agent (Yellayi et al. 2002). Puerarin has been demonstrated to possess a variety of beneficial activities on alleviating angiocardiopathy (Pan et al. 2009), osteoporosis (Wong & Rabie 2007) and hangover (Penetar et al. 2012). In the process of daidzein biosynthesis, the important step is that...
trihydroxyisoflavanone is synthesized from liquiritigenin by IFS and then form daidzein and daidzein sequentially by HID and UGT. Genistein synthesized from naringenin by IFS and then forms genistin by UGT. The content of daidzein and genistein in transgenic plants of *P. mirifica* was higher than in the control plants. This phenomenon may be the cause from the *EgP450* having similar functional domains with IFS. Although the content of Puerarin in some transgenic plants of *P. mirifica* was higher than the control plants, the total puerarin contents were not statistically significantly (*p* > 0.05) high. Further researches are required to engineer these synthesis pathways. This work is the first attempt to implement the use of transgenic *P. mirifica* to enhance the production of some targeted compounds of isoflavonoids.

ACKNOWLEDGEMENTS

This work was supported by the Center for Genomics and Bioinformatics Research and the government budget of Prince of Songkla University (SCIS570197S). Scholarship to Mr. Li Xu was supported by the Graduated School, Prince of Songkla University, scholarship to Miss Ruethaithip Kaopong was supported by the National Research University Project of Thailand’s Office of the Higher Education Commission (SCIS540527M) and scholarship to Miss. Arthiya Chullasara was supported by the Research Assistant (RA) fund from the faculty of Science (1-2552-02-006).

REFERENCES


Ito, F., Iwasaki, M., Watanabe, T., Ishikawa, T. & Higuchi, Y. 2005. The first total synthesis of kwakhurin, a characteristic component of a rejuvenating plant, “kwao keur”: toward...


Center for Genomics and Bioinformatics Research Department of Molecular Biotechnology and Bioinformatics Faculty of Science, Prince of Songkla University Hatyai, Songkhla 90112 Thailand

*Corresponding author; email: pamorrna@yahoo.com

Received: 31 August 2016

Accepted: 17 April 2017