Convergent evolution leading to the appearance of furanocoumarins in citrus plants
Mariana Limones-Mendez, Audray Dugrand-Judek, Cloé Villard, Victoire Coqueret, Yann Froelicher, Frédéric Bourgaud, Alexandre Olry, Alain Hehn

To cite this version:
Mariana Limones-Mendez, Audray Dugrand-Judek, Cloé Villard, Victoire Coqueret, Yann Froelicher, et al.. Convergent evolution leading to the appearance of furanocoumarins in citrus plants. Plant Science, Elsevier, 2020, 292, pp.110392. 10.1016/j.plantsci.2019.110392 . hal-02455296

HAL Id: hal-02455296
https://hal.univ-lorraine.fr/hal-02455296
Submitted on 21 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial| 4.0 International License
Convergent evolution leading to the appearance of furanocoumarins in citrus plants.

Mariana Limones-Mendez1,&, Audray Dugrand-Judek1,&, Cloé Villard1, Victoire Coqueret1, Yann Froelicher2, Frédéric Bourgaud3, Alexandre Olry1, Alain Hehn1,#

1- Université de Lorraine, INRA, LAE, F54000, Nancy, France.
2- CIRAD, UMR AGAP, F-34398 Montpellier, France.
3- Plant Advanced Technologies SA, F-54500 Vandœuvre-lès-Nancy, France.

& Equal contribution
# Corresponding author: Alain.Hehn@univ-lorraine.fr

Summary:

Furanocoumarins are defense molecules mainly described in four plant families that are phylogenetically distant. Molecular characterization of the biosynthetic pathway has been started for many years in Apiaceae and Rutaceae. The results obtained thus far in Apiaceae indicated a major role of cytochromes P450 (P450s) in the CYP71 family. In the present work, we describe the importance of another subfamily of P450s, CYP82D, identified by using a deep analysis of the citrus (Rutaceae) genome and microarray database. CYP82D64 is able to hydroxylate xanthotoxin to generate 5-OH-xanthotoxin. Minor and limited amino acid changes in the CYP82D64 coding sequence between Citrus paradisi and Citrus hystrix provide the enzyme in the latter with the ability to hydroxylate herniarin, but with low
efficiency. The kinetic constants of the enzyme are consistent with those of other enzymes of this type in plants and indicate that it may be the physiological substrate. The activity of the enzyme is identical to that of CYP71AZ6 identified in parsnip, showing possible evolutionary convergence between these two families of plants. It is highly possible that these molecules are derived from the synthesis of ubiquitous coumarins throughout the plant kingdom.

**Keywords:** Cytochrome P450, CYP82, Furanocoumarins, *Citrus*, Convergent evolution

**Author contributions:**
MLM and ADJ did the cloning and functional characterization of the enzymes. ADJ and VC did the qPCR experiments and metabolic analysis. AO did the enzymatic characterization. CV produced the phylogenetic tree. YF, FB and AH supervised the project. AH wrote the article.

**Introduction**
Furanocoumarins (FCs) are specialized metabolites that have been widely reported as being involved in the defense of plants against aggressors [1]. These molecules, although described in higher plants, are mainly distributed in 4 families: Apiaceae, Moraceae, Fabaceae and Rutaceae [2,3]. Two main kinds of FCs exist depending on the position of the furan ring on the coumarin core molecule. The tracks leading to either linear or angular molecules are catalyzed by enzymes belonging to the UbiA Prenyltransferases. A prenylation occurring on C6 leads to the synthesis of the linear demethylsuberosin whereas a prenylation on C8 gives ostheno1 (Figure 1) [4–6] The presence of both isomers are, thus far, only reported in Apiaceae and are reported to reflect a coevolutionary process between plants and caterpillars [7]. These four plant families are quite distant from a phylogenetic point of view, and the
independent appearance of these molecules remains unexplained at the molecular level. The
first enzymes involved in the synthesis of these molecules to be characterized were a psoralen
synthase (CYP71AJ1-3) and an angelicin synthase (CYP71AJ4) [8,9]. Both enzymes belong
to the cytochrome 450 enzyme superfamily. CYP71AJ4, identified in *Pastinaca sativa*, may
result from the duplication of CYP71AJ3 followed by neofunctionalization [8]. This
assumption is consistent with the high amino acid sequence identity shared by the two
proteins (69.8 %) and is reinforced by the results published by Roselli et al., who showed their
clustering in the parsnip genome (the genes are 7.6 kb apart) [10]. Relatively minor changes
in the peptidic sequence of psoralen synthase therefore allowed the appearance of the angular
FC angelicin and subsequent angular molecules, thus participating in the arms race between
plants and insects. An extensive study of this specific CYP71AJ subfamily has led to the
identification of numerous orthologous genes, but their function could not be demonstrated

Krieger et al have highlighted another cytochrome P450 subfamily, CYP71AZ, involved in
the synthesis of FC [12]. CYP71AZ4 is able to hydroxylate psoralen, leading to the synthesis
of xanthotoxol, whereas CYP71AZ1 and CYP71AZ6 catalyze the hydroxylation of
xanthotoxin, generating 5-OH-xanthotoxin (Figure 1). Interestingly, CYP71AZ6 seems quite
specific to a single molecule (xanthotoxin), while CYP71AZ4 can transform different types of
molecules, including a FC (psoralen) and coumarins (6-methoxycoumarin, scopoletin, 7-
methoxy-3-methylcoumarin, and 7-methoxycoumarin) [12]. Finally, Krieger et al also showed
that CYP71AZ3 seems to be specific and dedicated to the hydroxylation of esculetin, another
coumarin [12]. All higher plants are able to produce coumarins and the fact that some
members of the CYP71AJ and CYP71AZ families are involved in the synthesis of coumarins
and/or FCs [11,12] provides new insights for understanding the evolutionary processes
leading to the appearance of FCs that might have diverged from coumarins. To date, the
different enzymes described to be involved in the synthesis of FCs belong to the same large CYP71 cytochrome P450 family. This family is responsible for the synthesis of highly diversified and specialized metabolites. For example, members of the family are involved in the synthesis of artemisin [13], menthofuran [14], camalexin [15,16], cyanogenic glucosides [17] or DIMBOA [18]. The diversity of chemical structures, biosynthetic pathways and plant species make this enzyme family a major actor in the adaptation of plants to their environments.

The CYP71 family is not the only cytochrome P450 that seems to be involved in the synthesis of such molecules (Figure S5). The functional characterization of CYP82C2 and CYP82C4 isolated from *Arabidopsis thaliana* reinforces the assumption that FCs are derived from coumarins. Kruse et al showed that these enzymes convert *in vivo* xanthotoxin to 5-OH-xanthotoxin when overexpressing *Arabidopsis* mutants are incubated in a culture media containing the substrate (Figure 1) [19], while this plant is not otherwise able to synthesize this kind of molecule. In 2018, Rajniak et al. showed that CYP82C4 is also involved in the oxidation of fraxetin (a coumarin), leading to the synthesis of sideretin [20], which is described as being important in iron homeostasis [21]. Taken together, these results indicate that the functionalization of coumarins into FCs may be realized in different ways (CYP71 or CYP82) throughout higher plants, thus supporting the convergent evolution hypothesis.

Rutaceae are reported to produce coumarins and linear FCs [22–24], but less is known about the enzymes involved in this synthesis. However, a cytochrome P450 associated with an evolutionary process related to the synthesis of FC was described. Gravot et al. showed that the cinnamate 4-hydroxylase (C4H) of *Ruta graveolens* was adapted to the presence of FCs. Although inhibited by psoralen and xanthotoxin through mechanism-based inactivation, the C4H isolated from *R. graveolens* displayed a lower sensitivity to these molecules than the C4H of *Helianthus tuberosus* [25]. Citruses belong to the Rutaceae family. The consumption
of these fruits has a wide range of beneficial effects for human health, since they are rich in vitamin C, folates, alimentary fibers and minerals [26]. In addition to these positive effects, it has also been demonstrated that the presence of FCs generates deleterious effects such as phototoxicity and photosensitivity [27–30] and a drug interaction effect known as "the grapefruit juice effect" that can cause overdose in humans by inhibiting liver cytochromes P450 involved in the detoxification of xenobiotics [31–33].

In this work, we investigated the biosynthesis of FCs in citrus species, focusing on the CYP82 family. The expression pattern of four genes and their functional characterization showed that two of them perform the hydroxylation of xanthotoxin to produce 5-OH-xanthotoxin. The physiological relevance of this result is supported by a quantitative PCR approach that shows a clear correlation between gene expression and the amount of these molecules.

**Material and methods**

**Citrus data**

Citrus genomic and transcriptomic data were collected from the Phytozome project database (https://phytozome.jgi.doe.gov/) and the RNAseq database published by Voo et al [34].

**Phylogenetic analysis**

The tree was built with the maximum likelihood (PhyML) method. The alignment was performed with MUSCLE using the default settings [35]. The variable regions creating gaps, the C-terminal region and the N-terminal membrane anchor were removed. One hundred iterations were performed, and the bootstrap values were calculated. The tree was rooted with CYP71AJ3, CYP71AJ4, CYP71AZ3, CYP71AZ5 and CYP71AZ6 from *P. sativa,*
CYP71AU3, CYP71AU4 and CYP71AH4 from *Vitis vinifera*, and CYP83B1 and CYP83A1 from *A. thaliana*.

**RT-PCR and cloning strategy**

RNA was extracted from flavedo of *C. paradisi* Star Ruby and *C. hystrix* fruits using the E.Z.N.A.® Plant RNA kit following the manufacturer’s recommendations (Omega Bio-Tek). Residual genomic DNA was removed using amplification-grade DNase I (Sigma-Aldrich). Reverse transcription was performed with a High-Capacity RNA-to-cDNA Kit (Applied Biosystems). The coding sequences were amplified with high-fidelity PrimeSTAR® Max DNA polymerase (Takara) and ligated into the pCR8 plasmid using the pCR8 TA/GW/Topo kit (Invitrogen). The designed primers were as follows: *CYP82D64* Fw 5' ATG GAG TTC CTA TTC ACA GCT GTA ATA ATC ACT GCC C 3', *CYP82D64* Rev 5' TCA ATG ATG ATG ATG AAA GCC ATA GGG GTT GGC A 3'; *CYP82D83* Fw 5' ATG GAC ATT CTT GGT TCG TCT CAA AC 3', *CYP82D83* Rev 5' TCA ATG ATG ATG ATG ATG ATG GCC GTA CAA AGA AGA AAG AC 3'; *CYP82D92* Fw 5' ATG GAC TTC CTT GTT TTA ACA TCT CAA C 3', *CYP82D92* Rev 5' TCA ATG ATG ATG ATG ATG GCC GTA CAA AGA AGA AAG 3'; *CYP82D179* Fw 5' ATG GAC TTC GTT TCA CGT CAA AC 3', *CYP82D179* Rev 5' TCA ATG ATG ATG ATG ATG GCC GTA CAA AGA AGA AAG 3'. For each ORF, we added a 6xHIS tag before the stop codon. Finally, the ORF was transferred into the pYeDP60 GW [12] plasmid using Gateway LR Clonase II Enzyme mix (Invitrogen).

**Expression of proteins and enzymatic characterization**

The recombinant proteins were produced in a yeast expression system [36]. The culture conditions, the preparation of microsomes and the biochemical screening were as described by Krieger et al. (2018).
Reverse transcription was performed with a High-Capacity RNA-to-cDNA Kit (Applied Biosystems). The primers used were EF1α Fw 5’-TGG TGA TGC TGG GTT TGT TA-3’, EF1α Rev 5’-CAC GCT CTT GAT GAC TCC AA-3’ for the reference gene; CYP82D64 Fw (1) 5’-GGC GCC ATG CTT TCC A-3’, CYP82D64 Rev (1) 5’-ACG ACC GGG AAA TTC TTG TG-3’; and CYP82D64 Fw (2) 5’-CAC CAC GTG AAG CAA TGG AA-3’, CYP82D64 Rev (2) 5’-TGC CAG CCG GGA CAT G-3’. qPCR was performed with SYBR Premix Ex Taq (Takara) based on the manufacturer’s recommendations. The amplification program was 95°C for 30 sec, followed by 40 cycles of 95°C for 5 sec and 60°C for 34 sec. The PCR was followed by a single run at 95°C for 15 sec, 60°C for 60 sec, and 95°C for 15 sec in order to assess the melting curve and therefore the specificity of the PCRs.

Enzymatic characterization

All the enzymatic reactions were conducted in 0.1 M KPi buffer. pH and temperature range were set between 6.0 to 8.0 and 15 °C to 40 °C. For the kinetic determination of optimal temperature and pH, saturating concentrations of substrates were used (100 µM) in phosphate buffer. For km kinetic parameter determination, xanthotoxin concentration was set between 0.5 and 300 µM and the reaction was conducted at optimal pH and temperature. 20 µl of the yeast microsomal preparation was added to each enzymatic incubation and enzymatic reactions were running during 5 min and stopped by the addition of 10 µl HCl 1N. Substrates and the resulting products were analyzed by HPLC at a wavelength of 320 nm as described by Vialart et al. Kinetic parameters were calculated using the SigmaPlot 12.0 software. Identification of the chemical structure of the products was done by comparison with their elution time in comparison to 5-OH-xanthotoxin standard molecule and by LTQ orbitrap as described by Krieger et al. (2018)
Phytochemical analysis

Methanol extraction and phytochemical analyses were performed as previously described by Dugrand et al. [22].

Results

Identification of cytochrome P450 coding sequences belonging to the CYP82 subfamily in Citrus genomic and microarray databases.

The Phytozome database (https://phytozome.jgi.doe.gov/) contains the genomes of Citrus sinensis and Citrus clementina. These genomes were annotated [37,38], but regarding cytochromes P450, the annotation was limited to the superfamily level. According to Nelson, two P450s belong to the same family if they share more than 40% amino acid identity and are in the same subfamily if they share more than 55% (http://drnelson.utmem.edu/CytochromeP450.html). To seek candidate genes, we performed a sequence homology search using the BlastN and tBlastX algorithms and CYP82C2 and CYP82C4 from A. thaliana as bait sequences. This first investigation led us to identify 79 P450 candidate sequences that were further restricted to 42 CYP82 partial sequences in the two Citrus species (16 for C. sinensis and 26 for C. clementina). Most were redundant, which led us to finally consider only 8 different coding sequences (Table 1). All the sequences were affiliated with the CYP82C (CYP82C46) and CYP82D (CYP82D64, 81, 82, 83, 84, 92, and 179) subfamilies according to the cytochrome P450 annotation committee (http://drnelson.utmem.edu/CytochromeP450.html).

To complete this first analysis, we compared our results with the database provided by Voo et al [34]. This analysis showed that 2 sequences of the 8 selected previously, i.e., CYP82D64 (i.e., contigs Cit.29478 and Cit.2373 in the database) and CYP82D79 (i.e., Cit.5945 in the
database), had differential expression (Table 1) between flavedo oil cavities (tissue with a high concentration of FCs, [22,23]) and albedo tissue (lack of FCs). These first analyses supported these two sequences as strong candidates involved in FC biosynthesis in *Citrus spp.*

To obtain a more detailed overview of the distribution of this P450 subfamily, we performed a phylogenetic analysis including 91 members of the CYP82 subfamily available in databases in the literature and various other databases (Supplemental figure 1). The tree was rooted with 11 sequences belonging to the CYP71 or CYP83 subfamily. This analysis showed that the already-characterized CYP82C2 and CYP82C4 were located on different branches than the CYP82D identified in citrus plants (Figure 2). Within the more specific citrus clade, we could not distinguish different groups that could provide clues for functional characterization since the genes were distributed between *C. sinensis* (unable to produce FCs) and *C. clementina* (producing FCs) [23].

**Functional characterization of the CYP82D candidates.**

In our analysis, CYP82D64, CYP82D83, and CYP82D179 were detected in both plants, whereas CYP82D82 and CYP82D81 were specific to *C. sinensis* and CYP82D92 was found only in *C. clementina*. Since the aim of our study was to identify genes involved in the synthesis of FCs, we focused our attention on CYP82D92, CYP82D64, CYP82D83, and CYP82D179, detected either in *C. clementina* alone or in both plants. The study performed by Krieger et al showed that minor modification of the sequence of CYP71AZ can lead to significant changes in substrate specificity and/or enzyme activity. To limit the risk of cloning gene variants displaying low FC biosynthetic activities, we amplified and cloned the cDNA coding sequences from *C. paradisi*, a citrus species producing diverse and large amounts of FCs [22,23]. The amplified sequences of *C. paradisi* showed minor changes (98 to 100%
identity) in comparison to the genomic sequences of *C. sinensis* or *C. clementina*. These differences were, however, not located in the enzyme active sites [39].

To identify their function, the 4 His-tagged P450s were expressed in a yeast heterologous expression system, and their expression was assessed by Western blot (Supplemental figure 2). CYP82D64, CYP82D92 and CYP82D179 were efficiently produced and further incubated in the presence of NADPH and the different intermediates of the biosynthetic pathway (demethylsuberosin, marmesin, xanthotoxol, xanthotoxin, bergaptol, bergapten, 5-OH-xanthotoxin, 8-OH bergapten and isopimpinellin) (Figure 1). Only metabolization of xanthotoxin could be observed in the presence of CYP82D64 (Figure 4A). The elution time and the fragmentation spectrum (maximum fragmentation intensities for m/z of 172, 217 and 205, for a total m/z of 233) were identical to those of 5-OH-xanthotoxin (Figure 4 B-E). No enzymatic activity could be detected for any other protein tested.

**Plasticity of substrate specificity between two CYP82D64 variants isolated from *C. paradisi* and *C. hystrix*.**

Several studies have provided evidence that only a few modifications to an amino acid sequence can alter the specificity of an enzyme. Concerning the synthesis of coumarins and FCs, this has been shown for CYP71AZ3, 4 and 6, which are able to transform esculetin, psoralen and xanthotoxin, respectively [12]. We cloned the CYP82D64 orthologous gene from combava (*C. hystrix*) fruits, a citrus species known to be a very high producer of FCs [23]. The CDS that we obtained from this plant was named CYP82D64v2 and showed several differences compared with the *C. paradisi* coding sequence (CYP82D64v1), especially in the substrate recognition sites SRS3, 4 and 6, which were reported to be important for the enzymatic activity of cytochromes P450 [39] (Figure 4). Microsomes prepared from yeast
overexpressing CYP82D64v2 were able to convert xanthotoxin into 5-OH xanthone at the same rate as CYP82D64v1. Surprisingly, CYP82D64v2 was also able to transform herniarin (7-methoxycoumarin) at a rather low rate into a product that was identified as 5-hydroxy-7-methoxycoumarin (Figure 5).

**Enzymatic characterization of CYP82D64v1 and v2**

The optimal pH of both CYP82D64s was assessed in phosphate buffer with a pH ranging from 6.0 to 8.0. Under these conditions, we showed that the hydroxylation activity of xanthotoxin was effective over the entire pH range and was optimal at a pH of 7.3 (Supplemental figure 3). The activity in the presence of herniarin could not be detected below a pH of 6.4, but the optimal activity was obtained at a pH of 7.2. The optimal temperature of the hydroxylation of both molecules was close to 28°C (27.8°C for xanthotoxin and 28.1°C for herniarin) (Supplemental figure 4). Finally, the kinetic constants (km) of CYP82D64v1 and v2 were determined to be 4.6+/−1.7 µM and 7.4+/−4.0 µM for xanthotoxin, respectively. These km values were rather low and consistent with those reported in studies performed on other cytochromes P450. This specificity may reflect a physiological activity. Because of the too-low enzymatic activity, the km could not be determined for herniarin.

**In vivo expression of CYP82D64**

Although the km of the reaction was consistent with a physiological function, we wanted to assess if there was a relationship between the expression level of the mRNA coding for CYP82D64 and the amounts of xanthotoxin and 5-OH-xanthotoxin produced in *C. paradisi* fruit at different stages of maturity: young, intermediate and mature (Figure 6A). We observed that the expression level of *CYP82D64* mRNA significantly decreased during fruit
development (Figure 6B). It was reduced 2.3-fold between the young and intermediate fruits and 18-fold between the intermediate-stage and ripe fruits. Regarding the expression of CYP82D64 mRNA, the concentrations of xanthotoxin and 5-OH-xanthotoxin also decreased significantly. The average content of the 2 molecules decreased 2.5-fold between the maturation of young fruits to the intermediate stage and 7.4-fold during the transformation from intermediate-stage to ripe fruits.

Discussion

The Rutaceae and Apiaceae families are quite distant in terms of angiosperm evolution. Interestingly, however, they are both able to produce FCs [23,24], although the diversity of molecules differs between them. Apiaceae, including parsnip, are able to synthesize both linear (i.e., psoralen derivatives) and angular isomers (i.e., angelicin derivatives) [8]. Rutaceae and, in particular, the Citrus genus produce only a linear isomer [23]. The first step involved in the production of these molecules concerns the C2'-hydroxylation of coumaroyl-CoA, leading to the synthesis of umbelliferone [10,40,41]. Umbelliferone is commonly described in many higher plants. This step is catalyzed by 2-oxoglutarate-dependent dioxygenase in R. graveolens [40] and P. sativa [10], which belong to the same DIOX30 clade [41], and therefore seems to be well conserved across angiosperms. The prenylation of umbelliferone has not yet been described in citrus but was reported for Apiaceae plants and is catalyzed by a UbiA prenyltransferase [4]. A recent study by Munakata and colleagues showed that in Ficus carica, a Moraceae able to produce FCs, prenyltransferase had a different evolutionary origin than the enzyme involved in the synthesis of demethylsuberosin in Apiaceae. It might therefore be possible that convergent evolution occurred at this stage [42].
Studies based on biochemical approaches performed in the 1980s pointed out that in Apiaceae, many enzymatic steps happening after the synthesis of demethylsuberosin (Figure 1) were catalyzed by enzymes belonging to the cytochrome P450 superfamily [43,44]. These results were confirmed by the identification of the CYP71AJ [8,9] and CYP71AZ subfamilies [12], which are involved in the synthesis of psoralen, angelicin, xanthotoxol and 5-OH-xanthotoxin. To identify the genes involved in the same reaction in other plant species, we searched for CYP71AJ or CYP71AZ orthologous genes in citrus databases. Since this investigation was unsuccessful, we extended our investigation to other cytochrome P450 families. Based on the work performed by Kruse and colleagues [19], who described the hydroxylation of 8-methoxypsoralen (i.e., xanthotoxin) by a P450 isolated from Arabidopsis, we focused on the CYP82 family. Among several candidate genes, we characterized CYP82D64 as a xanthotoxin hydroxylase. This enzyme displayed the same enzymatic function as CYP71AZ6 isolated from parsnip [12]. The low km and the close relationship between the expression level of the gene and the accumulation of 5-OH-xanthotoxin suggest that this reaction also occurs in planta as a physiological reaction.

In this study, we provided evidence that CYP82D64 was able to hydroxylate herniarin. This result is in line with that from a study recently published by Rajniak supporting the role of CYP82C4 in the production of sideretin, another coumarin [20]. This report, in addition to our results, highlights that the CYP82 family participates in the synthesis of both FCs and coumarins, as does the CYP71AZ subfamily in Apiaceae. These elements provide new insights in favor of the evolution of the FC biosynthetic pathway from the coumarin biosynthetic pathway. These results are also consistent with the extreme diversity of the CYP71 and CYP82 subfamilies (53 CYP71s and 25 CYP82s reported to date), demonstrating the plasticity of these enzymes that seem to be deeply involved in the adaptation of plants to their environments. Indeed, similar to the CYP71 subfamily, CYP82 members are associated
with highly functionally diverse families of molecules with very different chemical structures such as alkaloids, terpenes or phenolic compounds. For example, CYP82N3 and CYP82N4 enzymes of Papaver somniferum hydroxylate alkaline benzylisoquinolines [45]. CYP82E4 from Nicotiana tabacum corresponds to a nicotine N-demethylase [46]. CYP82G1 of A. thaliana is involved in the synthesis of volatile homoterpenes [47]. CYP82D63 from Mentha piperita and CYP82D33 from Ocimum basilicum hydroxylate flavones [48]. CYP82D64 from C. paradisi and CYP82C2 from A. thaliana hydroxylate xanthotoxin, and CYP82C4 hydroxylates fraxetin. This functional diversity was distributed over the entire generated phylogenetic tree (Figure 2).

In this work, we provide some new lines of evidence in favor of a convergent evolution theory for explaining the presence of FCs in Rutaceae and Apiaceae. It is highly possible that the same process occurred in Moraceae and Fabaceae, two other plant families producing large amounts of these specialized molecules. It is therefore not excluded that other P450 families or subfamilies play an important role in the synthesis of FCs and the subsequent adaptation of these plants to their environments. Such convergent evolution is not unique in the plant kingdom. Genes encoding enzymes belonging to different families may have identical functions, such as the anthocyanin-5-O-glucosyltransferases of Dianthus caryophyllus and A. thaliana. In this case, the enzymes have different structures but share the same glycosylation of anthocyanin function [49]. Similarly, ferulate-5-hydroxylase (CYP788A1) from Selaginella moellendorffii shares only 37% identity with angiosperm P450s but plays the same role as these P450s in the synthesis of S units of lignin [50].

Acknowledgments
MLM was funded by CONACYT. The authors would like to thank the SATT SAYENS and the "Impact Biomolecules" project (ANR-15-IDEX-04-LUE Lorraine Université d'Excellence) for financial support. The authors also thank Clément Charles for technical assistance.

References


[45] G.A.W. Beaudoin, P.J. Facchini, Isolation and characterization of a cDNA encoding (S)-cis-N-methylstylopine 14-hydroxylase from opium poppy, a key enzyme in sanguinarine

https://doi.org/10.1073/pnas.0506581102.

https://doi.org/10.1073/pnas.1009975107.


Figures:

Figure 1: schematic representation of the furanocoumarin biosynthetic pathway and the enzymatic steps catalyzed by cytochrome P450s.

Red arrows: known steps; gray arrows: P450 putative steps; black arrows: catalyzed by other enzyme families.
Figure 2: phylogenetic tree of the CYP82 family.


Figure 3: enzymatic conversion of xanthotoxin by CYP82D64.

A) DAAD chromatogram, B) UV spectrum of the product, C) MS analysis of the product, D) MS analysis of commercial 5-OH-xanthotoxin, E) UV spectrum of commercial 5-OH-xanthotoxin.
Figure 4: sequence comparison of CYP82D64 variants isolated from C. clementina, C. paradisi and C. hystrix.

Red stars highlight the differences, and green stars show the substrate recognition sites [39]. The C. clementina sequence was predicted from the CDS available in Phytozome.

Figure 5: enzymatic conversion of herniarin by CYP82D64v2.

A) DAAD chromatogram, B) UV spectrum of the product, C) MS analysis of the product, D) MS analysis of commercial 5-hydroxy-7-methoxycoumarin, E) UV spectrum of commercial 5-hydroxy-7-methoxycoumarin.

Figure 6: comparison of the expression level of CYP82D64 mRNA to the xanthotoxin and 5-OH-xanthotoxin contents in C. paradisi fruits. A) Three stages of maturity were tested: immature, intermediate and mature fruits. B) Expression level of CYP82D64 and concentrations of xanthotoxin and 5-OH-xanthotoxin.

Supplemental Data

Supplemental figure 1: list of the P450 sequences used to create the phylogenetic tree in Figure 2

Supplemental figure 2: Western blot analysis of microsomes prepared from yeast expressing CYP82D64, CYP82D83, CYP82D92 and CYP82D179. Void: yeast transformed with empty pYeDP60 plasmid; Positive: yeast expressing CYP71AZ4 (Krieger et al, 2018).
Supplemental figure 3: optimal pH determination for CYP82D64v1 and v2

A) Optimal pH for CYP82D64v1 with xanthotoxin, B) optimal pH for CYP82D64v2 with xanthotoxin, C) optimal pH for CYP82D64v2 with herniarin.

Supplemental figure 4: optimal temperature determination for CYP82D64v1 and v2

A) Optimal temperature for CYP82D64v1 with xanthotoxin, B) optimal temperature for CYP82D64v2 with xanthotoxin, C) optimal temperature for CYP82D64v2 with herniarin.

Supplemental figure 5: Phylogenetic tree constructed with the predicted cytochrome P450s belonging to the Clan71 of *Citrus sinensis* and *Citrus clementina*. The tree has been rooted with CYP51G1. Sequence alignment was performed using MAFFT relying on Multiple sequence alignment default settings (https://mafft.cbrc.jp/alignment/software). The tree was built with IQ-TREE multicore version 1.6.11 for Linux 64-bit built Jun 6 2019 and FigTree.
Main figures
Figure 1:
Figure 3
Figure 4
Figure 5
Figure 6
Table 1

<table>
<thead>
<tr>
<th>Phytozone ID</th>
<th>C. clementina</th>
<th>C. sinensis</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP82C46</td>
<td>Ciclev10028201m</td>
<td>orange1.1g010622m</td>
</tr>
<tr>
<td>CYP82D64</td>
<td>Ciclev10025349m</td>
<td>orange1.1g009723m</td>
</tr>
<tr>
<td>CYP82D81</td>
<td>-</td>
<td>orange1.1g009844m</td>
</tr>
<tr>
<td>CYP82D82</td>
<td>-</td>
<td>orange1.1g009347m</td>
</tr>
<tr>
<td>CYP82D83</td>
<td>Ciclev10000845m</td>
<td>orange1.1g009786m</td>
</tr>
<tr>
<td>CYP82D84</td>
<td>Ciclev10000852m</td>
<td>orange1.1g009899m</td>
</tr>
<tr>
<td>CYP82D92</td>
<td>Ciclev10000846m</td>
<td>-</td>
</tr>
<tr>
<td>CYP82D179</td>
<td>Ciclev10000844m</td>
<td>orange1.1g026655m</td>
</tr>
</tbody>
</table>
>NP_001312791_(NT)
APEIPGAWPIQLGDSQKLDLGADKYGPFVTFLRGLMIGYPPYLVNWEAAKCLCTTHDSDKFARPSMEGSIGKYARFT

>NP_001312999_(NT)
PPKIPGGWPVIGHLFYFNNNGDFSQKLDGLGDADKYGPFVTFLRGLMIGYPPYLVNWEAAKCLCTTHDSDKFARPSMEGSIGKYARFT

>NP_001312768_(NT)
PPKIPGGWPVIGHLFYFNNNGDFSQKLDGLGDADKYGPFVTFLRGLMIGYPPYLVNWEAAKCLCTTHDSDKFARPSMEGSIGKYARFT

>NP_001312226_(NT)
PPKIPGGWPVIGHLFYFDDDSKLARKLGDLADKYGPFVTFLRGLMIGYPPYLVNWEAAKCLCTTHDSDKFARPSMEGSIGKYARFT

>NP_001312052_(NT)

Supplemental Figure 1

634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
ERYTTEECEVGRFNPAGTRLLNVWKIHDPVVENPSDFQPERFLCDKVGVDLYGQNYELIPFGAGRRVCPAIVSSLQTMHYAL

ARLIQGYEMKASDGTKVMENIAMSCHKMSPLEVIISPRE

>AFB74614_(PSo)

APNAAGAWPLIGHLRLLMNDKELYRALGSMADKYGPAFNIRLGNQEVFLVSNWEMVKQCFGQNDKLSNRQTTTALAKYMLNQT

TSGFAPYGYWRELKIMVQQLLSKSQLESWKHLKIKEMDASFSKLNELCNNNTLRMDWFAELTFNVIARNVFQYGSQGSKGER

YKKTLEALLHLMSIAFASDVFIPSLWDLGRLIRNMFGRGDLSNGLIEEIEHRRQKRLQDFVDVLSVAEKKDPDLVIKSMILEIV

SGSGETTSSSTWLACLLNNPHELKKAKEELDHVGRDKHVDEESSDPKLYVYINAIIKESMRYPNAGML-

DRLALEEECVEGFFHPAGGRLVNWKVQRIQRPSVWNPLEFKPERWLEKMDVDYKGNHHEIFIPGIRGRRMCAGMLWADEVHIL

VLPRLHGFDMKAASNGKVDMAEAGMVIFQIPLEMTLEVMPNVPRE

>AFB74617_(PSo)

LPEASGAPWVPHGLLFLMNENNDNHVLTHGMAKDYGPIFSLRFGRHRRLYLVSWSMEMVKECFTNDKLSNRPSLAVKLMFYDTES

YGFAPYGYKGYRELKIKHLSSNQQLKFKHLRILREDNVSNFKHLECSNLSLRMDWFAAYLTFENVIRIVGFGQSNTNSQEKYK

LAIDEVSNLMATFASDVPILGLWDLTLGTLGTMMNCGIKLDAVGDADEDHRQK-

LEDFFDVCALIMEQONHEISVESKIALDMSSGSDTTKLMWGSTSLNNHPDILDKAAKEEVETDVFYFGKKVDAADVLPNLVYIIAQIKES

SMRLYPASTLM-

ERMSTTDCDVGFFHPAPGTLLWNVWVKMRQDRPRVRKDPVLFLPERFLSDKGMVDVKQNYELIPFGGFALEVLHIL

VLTRLILEFEMKAPEG-KIDMRAPGFFHNKVKPLDQLTPT

>CP52824_(AGC92398)_(PSo)

APEASGRRPMGHLFLFNDGELTHRKLGMADATYPGVPNIRFGSHKTLLVSWSWEIKECFTNNDKLSNRPGTGLIKLMFYDASV

GYAPYAYWREDLRLKSTKLSSHRDITKHLRSSEVESCESELYQSWGRAPVRMDSWGLDLTFNVVARIAGKNNFDVGAQRYK

AAMDEAMRLMRFFASFDVIPSLSWDLRGLVREMKCASEIDSMATWVEJHVRKNN-

HDFIDVCLDIMEHDDPDVVLKSTCLDMGLGSSDDTTTVTLTWAMSLLNHQPVLQKAKELEETVQGKRNQVDDSDIPNLPIQAIKES

TMLYPAGPL-

ERRTMDCEVAYQVPGAPTRLLWNVWVMQRDGNVYKDPLEFPRDPRFLTSNADVDLKQHYELIPFGAGRRICPGVSFAVQMLHL

VLARLLHEFEITPETPK-VDMAESGGLLCCYKIMPLEVLIKPRL

>XP_013693485_(BN)

APMVPGAWPVVGLMPYLSKPPYVFPGAMADVYGPVFMTKMGYSYNTLISSEEVARECYTVHDKVL-

RRPGIIASKLLGYNLFLFSPQGAYWRAMRKIMISQLVSASVVDFTKGDHRAEADVAFRDLYVRWEHVLMKEGFIDITSNTT

LMMLAGKRYEARYGKLMREAVHYFGRYFVSDTIPGWLWDIR-

EKKKMKKMKAQELYVLESWEMEEHSKRKKDRKDYLDMLEIIEIQSDCRTTATKLFLMNFSQGDPVIAILWVAVSLLNNQHVNLTKA

QEELSDSTIQKRVVSESLKDVLQYAIKEFTRLYPPTAFNAYRETFESEDFVIAFGHVPAGTHLNVWKKVQRDPTFWPNEEFQPE

RFLTSHKEVDAGGKNLPPFLGRRACPASSLGTGKMQYILARFLHSDLDNPSSQNVDMSESNDLVNKLATELEVFITPR

>XP_013596097_(BO)

APMVPGAWPVVGLMPYLSKPPYVFPGAMADVYGPVFMTKMGYSYNTLISSEEVARECYTVHDKVL-

RRPGIIASKLLGYNLFLFSPQGAYWRAMRKIMISQLVSASVVDFTKGDHRAEADVAFRDLYVRWEHVLMKEGFIDITSNTT

LMMLAGKRYEARYGKLMREAVHYFGRYFVSDTIPGWLWDIR-

EKKKMKKMKAQELYVLESWEMEEHSKRKKDRKDYLDMLEIIEIQSDCRTTATKLFLMNFSQGDPVIAILWVAVSLLNNQHVNLTKA

QEELSDSTIQKRVVSESLKDVLQYAIKEFTRLYPPTAFNAYRETFESEDFVIAFGHVPAGTHLNVWKKVQRDPTFWPNEEFQPE

RFLTSHKEVDAGGKNLPPFLGRRACPASSLGTGKMQYILARFLHSDLDNPSSQNVDMSESNDLVNKLATELEVFITPR

>XP_013596097_(BO)

APMVPGAWPVVGLMPYLSKPPYVFPGAMADVYGPVFMTKMGYSYNTLISSEEVARECYTVHDKVL-

RRPGIIASKLLGYNLFLFSPQGAYWRAMRKIMISQLVSASVVDFTKGDHRAEADVAFRDLYVRWEHVLMKEGFIDITSNTT

LMMLAGKRYEARYGKLMREAVHYFGRYFVSDTIPGWLWDIR-

EKKKMKKMKAQELYVLESWEMEEHSKRKKDRKDYLDMLEIIEIQSDCRTTATKLFLMNFSQGDPVIAILWVAVSLLNNQHVNLTKA

QEELSDSTIQKRVVSESLKDVLQYAIKEFTRLYPPTAFNAYRETFESEDFVIAFGHVPAGTHLNVWKKVQRDPTFWPNEEFQPE

RFLTSHKEVDAGGKNLPPFLGRRACPASSLGTGKMQYILARFLHSDLDNPSSQNVDMSESNDLVNKLATELEVFITPR

>XP_013596097_(BO)
XP_017242880_(DC)

LPEAGAWPFQGHMHLLGPNQLTHYTLGTMADKYGPAFTIKLGLSHKVMILSNWEMAKECFLTHDKNFIGRPIVLASKLLGYNLAM

FGFAPYSSYWREIRKJVSLELLSVRRIETLKGLRVSEVTSQTSIREVKLWVEVLVDMDKQWSNITHNLIMGMVAGKRYYEAGAQFE

KMRVKFPGFLFGALFSDFPSLWLDKSAHKDRMDKTRAQELDFQIQWLEEHKQRRLLEDMDVMLTLKDFDADTINKATSLNLI

LAGSDSTVALLJAVSLLLLNRCMLERAAQALEDTFTVGKDHRVEDESINITLYLQAVKETLRLYSPSIPILRAAIEDCTISGYNIPA

GTRLMVNYWKHHRDVEWSEPKNFPERFLLTHKIDDLKGNHEFIEPFGSSRGACPGVHALAQVHLASLLHCFDITNISDEDVD

LTERSFGLTNFKATPLQVYLLPR

>CYP82C46_(CA)

PPEEPAWPQGHLHLLEKNNQLNHQTLGKMADKYGKAFMRIHLRGLQAVKETCFTTNDKFLSSPPKFLAVKLMGDHAM

LGAPYGYPYWRNRLATVELLSNHQTELKVRDTEFKFKECFTTNDKFLSRPKSLAVKLMGDHAM

ALGNFYLYGLLPNRLPSDPNFLGLDVNGYTQMKTAKLDFGLRLVHEHRKRLNKLIDVMVLDDEADTVKSTCLSLIL

GGGDTVTVTLTWALSLLLNNHVLKQAQDLSIHAGMNQRVEISDNKLYLQAVKETLRLYPLAPLSAPREAMEDCIGGFIHAP

GTRLNWLKMHRDPSIWNAPSEIFIPFRLNELNAKLĐNVQDFQELPFGSSRGRCPGSIFALQGQHLTLARLLHAFELGTVSDTLVD

MRESPGMTVPKATLEVTLPRL

>XP_006479372_(CS)

PPEAGAWPCGHLLLAGSNQIHQTGLGEMADKYGKAFLHLHILGHRLAVSSEVTECFTTNDKFLSFRPKSLAVKLMGDHAM

IGAPYGYSYWREVFRLSTVQSLNSHRLELHHRDTEFKFKECFTTNDKFLSRPKSLAVKLMGDHAM

QKALGDFLTYLLGFILSDNPPFLGLDVNGYTQMKRTAKLDDVFGRLVQEHHRKRLNEDFIDVMSLDDDEADTVKATCL

SLISGQKDVTVTITWALSLLLNNHVLKKAQDELHIHVKNQREVESDNSKLYLQAVKETLRLCPPPLASAPREAMEDCTIDGFHI

PGATRLNVLWKMHRDPSIWNAPSEIFIPFRLNELNKLĐNVQDFQELPFGSSRGRCPGSIFALQGQHLTLARLLHAFELGTVSDTLVD

VDMGSQPGTVKATPLEVTLTPRV

>XP_006426130_(CC)

PPEAGAWPCGHLLLAGSNQIHQTGLGEMADKYGKAFLHLHILGHRLAVSSEVTECFTTNDKFLSFRPKSLAVKLMGDHAM

IGAPYGYSYWREVFRLSTVQSLNSHRLELHHRDTEFKFKECFTTNDKFLSRPKSLAVKLMGDHAM

QKALGDFLTYLLGFILSDNPPFLGLDVNGYTQMKRTAKLDDVFGRLVQEHHRKRLNEDFIDVMSLDDDEADTVKATCL

SLISGQKDVTVTITWALSLLLNNHVLKKAQDELHIHVKNQREVESDNSKLYLQAVKETLRLCPPPLASAPREAMEDCTIDGFHI

PGATRLNVLWKMHRDPSIWNAPSEIFIPFRLNELNKLĐNVQDFQELPFGSSRGRCPGSIFALQGQHLTLARLLHAFELGTVSDTLVD

VDMGSQPGTVKATPLEVTLTPRV

>XP_017244377_(DC)

PPEPAPAGAWPIIGHLHLLGANKILHQLFGAMADKLGPVFCLNLGTKRRTLVVTSWEVAKECFTEHKVLSNRPITTLAVKLMGDHAM

GFLPNGPWTYLYKLVMELLSSRRLELKHPISEPSVFFVRLYLWELSPVVLKLTQFRGDLTNNVVRMVAAGKRYF-

EEARQFQKASEDFLRMVGLFVSDAVLPFGWIDTLKYGKGMKKTAAAMDNILEGWEMEEHKQKRLQDFMVHLMLNARDS

ISDITKGTCLSLISLGYDAMDVMITLWAVSSLLNNHRVLKKAQDLELHIHVKNQREVESDVKNTLYRALKIKTLRLYPAPLNGPH

EAIECDTSCGFIHISAGTFLVNLWKLQDPSIWSPDFELQPERFIEKHDVIDWGRNFLFPFGSRRACPGTTTLALQVHLTLACQLL

HVSGELTGVDLPMTESPGPLNTKACKPLTEVTFRPRL

>XP_017243930_(DC)

PPEPAPAGAWPIIGHLHLLGANKILHHTLGDMADKYGPAFTIKLGLSHKVMILSNWEMAKECFLTHDKNFIGRPIVLASKLLGYNLAM

FGFAPYSSYWREIRKJVSLELLSVRRIETLKGLRVSEVTSQTSIREVKLWVEVLVDMDKQWSNITHNLIMGMVAGKRYYEAGAQFE

KMRVKFPGFLFGALFSDFPSLWLDKSAHKDRMDKTRAQELDFQIQWLEEHKQRRLLEDMDVMLTLKDFDADTINKATSLNLI

LAGSDSTVALLJAVSLLLNNRCMLERAAQALEDTFTVGKDHRVEDESINITLYLQAVKETLRLYSPSIPILRAAIEDCTISGYNIPA

GTRLMVNYWKHHRDVEWSEPKNFPERFLLTHKIDDLKGNHEFIEPFGSSRGACPGVHALAQVHLASLLHCFDITNISDEDVD

LTERSFGLTNFKATPLQVYLLPR

>CYP82C46_(CA)
1010 >CYP82D64_(CH)
1011 APEAGGGWVYGHLHLLGGPEPLHKVGLKIADEKYGPIFTLDKVGVRRTLVSNWEIAKECFTLDKVATRQKTVADILCYDSSLG
1012 FVPYGHYWRQIRKIAITELLSYRLESKLKVRESEVRCTLCKELYNLDVAVLMDMKCWETYTVLNVRILMVGKRCNTSSDQGWK
1013 DELSRRFFFEMKGLVSDALPFLRWLDIGGERDLMKTARDELDMIQWRLEHRKRGDSDNFMGASILDDRDADTINAKTACLA
1014 LILGESDTSSTLWAISSLNLNNRHALLKQAQELDLQQQGVKERRAVEDSTRQNLVLQLAIKKETRLYPAQPILLAPREAMEDECTVSGY
1015 HVIPAGTRLMINAWKIQDPRVRWENPSAQPQPERFLPAHAVDVRQQQFELIPFGSRRSCPCASSALQLHLTLARRLHAFELATPL
1016 DQPVDMSESPGTVPKATPLKLVSPRL
1017 >CYP82D64_(CP)
1018 APEAGGGWVYGHLHLLGGPEPLHKVGLKIADEKYGPIFTLDKVGVRRTLVSNWEIAKECFTLDKVATRQKTVADILCYDSSLG
1019 FVPYGHYWRQIRKIAITELLSYRLESKLKVRESEVRCTLCKELYNLDVAVLMDMKCWETYTVLNVRILMVGKRCNTSSDQGWK
1020 DELSRRFFFEMKGLVSDALPFLRWLDIGGERDLMKTARDELDMIQWRLEHRKRGDSDNFMGASILDDRDADTINAKTACLA
1021 DNFMGAMLSILDDRDADTINAKTACLA
1022 >XP_006464300_(CS)
1023 APVAAGAWPLIGHLHLLGGQEPTHRILLGLVDVADKYGPIFTMKGMLNLRLVSNWEIAKECFTINDKPFANRPSMAMDLLGYNFL
1024 AIAASPQNYWRESKMATIELLSSNQLQKLLKVRESEVKASIKGLFMENCRKVSRCKHVEGILVDIMRHIAGKRTSQDQSSFQ
1025 EHATNFDFLICGFVVSDALPFLRWLDIGGKEKLKMTADEVFILVQGWLDEHKSRSEDFMYVMLSFDDDDTDVTIKAACL
1026 ILAASDTATTVALWTIAAILLNNRDLVKAQNELDVHVGTRQVNSEDMNLAYLQVLKETLRLYPAAPLILVPESIEDCTSTPGL
1027 PVAGTQLIVNAWGVQDGPVWEPCFKFLPERFLKDYDVLKQNFQFIFPGSRRMCPGSFAQVQMLILSLHGFDFATPSDA
1028 PVDMEAEKSLTYAKATPLEVLLTTPRL
1029 >XP_006464298_(CS)
1030 APVAAGAWPLIGHLHLLGGQEPTHRILLGLVDVADKYGPIFTMKGMLNLRLVSNWEIAKECFTINDKPFANRPSMAMDLLGYNFL
1031 AIAASPQNYWRESKMATIELLSSNQLQKLLKVRESEVKASIKGLFMENCRKVSRCKHVEGILVDIMRHIAGKRTSQDQSSFQ
1032 EHATNFDFLICGFVVSDALPFLRWLDIGGKEKLKMTADEVFILVQGWLDEHKSRSEDFMYVMLSFDDDDTDVTIKAACL
1033 ILAASDTATTVALWTIAAILLNNRDLVKAQNELDVHVGTRQVNSEDMNLAYLQVLKETLRLYPAAPLILVPESIEDCTSTPGL
1034 PVAGTQLIVNAWGVQDGPVWEPCFKFLPERFLKDYDVLKQNFQFIFPGSRRMCPGSFAQVQMLILSLHGFDFATPSDA
1035 DMEEAKSLTYAKATPLEVLLTTPRL
1036 >XP_006464299_(CS)
1037 APRAAGAWPYUGHLHLLGGQEPTHRILLGLVDVADKYGPIFTMKGMLNLRLVSNWEIAKECFTINDKPFANRPSMAMDLLGYNFL
1038 AIAASPQNYWRESKMATIELLSSNQLQKLLKVRESEVKASIKGLFMENCRKVSRCKHVEGILVDIMRHIAGKRTSQDQSSFQ
1039 EHATNFDFLICGFVVSDALPFLRWLDIGGKEKLKMTADEVFILVQGWLDEHKSRSEDFMYVMLSFDDDDTDVTIKAACL
1040 ILAASDTATTVALWTIAAILLNNRDLVKAQNELDVHVGTRQVNSEDMNLAYLQVLKETLRLYPAAPLILVPESIEDCTSTPGL
1041 PVAGTQLIVNAWGVQDGPVWEPCFKFLPERFLKDYDVLKQNFQFIFPGSRRMCPGSFAQVQMLILSLHGFDFATPSDA
1042 VDMEAEKSLTYAKATPLEVLLTTPRL
1043 >XP_024956574_(CS)
1044 APRAGGWVPYUGHLHLLGGQEPTHRILLGLVDVADKYGPIFTMKGMLNLRLVSNWEIAKECFTINDKPFANRPSMAMDLLGYNFL
1045 AIAASPQNYWRESKMATIELLSSNQLQKLLKVRESEVKASIKGLFMENCRKVSRCKHVEGILVDIMRHIAGKRTSQDQSSFQ
1046 EHATNFDFLICGFVVSDALPFLRWLDIGGKEKLKMTADEVFILVQGWLDEHKSRSEDFMYVMLSFDDDDTDVTIKAACL
1047 ILAASDTATTVALWTIAAILLNNRDLVKAQNELDVHVGTRQVNSEDMNLAYLQVLKETLRLYPAAPLILVPESIEDCTSTPGL
1048 PVAGTQLIVNAWGVQDGPVWEPCFKFLPERFLKDYDVLKQNFQFIFPGSRRMCPGSFAQVQMLILSLHGFDFATPSDA
1049 VDMEAEKSLTYAKATPLEVLLTTPRL
1050 >XP_024956574_(CS)
CYP82D179 (CP)

1124 >CYP82D179_(CP)
1125 APEARGSWPVIGHLLLGGPEPShVWLGKMADEGPFIKMKIMGNRALVSVSNWEAMAKECLTHDKVFASRPKTLAMEILGYNFSM
1126 FGSPGYSYWREKRKATTLLESLNHRGLKLRHVREYELKTLKLDYLWKLLEWEMRWRDLTTIRNVLRMIVGKCGSVDG-
1127 EEWVGLTRFSMEGKGVFSVDAPFLRLWLDIGGDERSMKIAKELDIVQGWLHEEHHKRRDSLDFMSVLMLSGDHDVTNKAT
1128 CMGLVLAAATENTMTLTVWSSLNHRDLNQANKELDILVGTKRQNVESDINKLVLYLQAILKEAMRLYPASPLSHIESPECTISGS
1129 YNVPTSTQLVINWVKLQRDPRVWEEPCKPQFPERFLTRHKIDVRQGQNFELLPGSAGRCPGVSFALQVMFTLALLHGFDFATP
1130 SNPEVDMGAEGLGSIQKARPVEVLAPRL
1131 >XP_006471473_(CS)
1132 APEAGGSWPVIGHLLLGGPEPPhVWGLKLADEYPFIFKMKIMGNRALVSVSNWEAMAKECLTHDKVFASRPKTLAMEILGYNFSM
1133 GFSPGYSYWRETRKIALESLNHRGLKLRHVREYELKTLKLDYLWKLLEWEMRWRDLTTIRNVLRMIVGKCGSVDG-
1134 EEWIELLTRFSMEGKGVFSVDAPFLRLWLDIGGDERSMKIAKELDIVQGWLHEEHHKRRDSLDFMSVLMLSGDHDVTNKATC
1135 LGVLVLAATDTTITTLTWVISLLNNYRDLTMAQNLEDILVGKRVVESDINKLVLYLQAILKEAMRLYPAAPILPHIESPECTISGYN
1136 YVPSTQILVINWVKLQRDPRVWEEPCKPQFPERFLTRHKIDVRQGQNFELLPGSAGRCPGVSFALQVMFTLALLHGFDFATPSE
1137 PDMGEGGLGSIQKARPVEVLAPRL
1138 >XP_006471495_(CS)
1139 APEAGGSWPVIGHLLLGGPEPPhVWGLKLADEYPFIFKMKIMGNRALVSVSNWEAMAKECLTHDKVFASRPKTLAMEILGYNFSM
1140 GFSPGYSYWRETRKIALESLNHRGLKLRHVREYELKTLKLDYLWKLLEWEMRWRDLTTIRNVLRMIVGKCGSVDG-
1141 EEWIELLTRFSMEGKGVFSVDAPFLRLWLDIGGDERSMKIAKELDIVQGWLHEEHHKRRDSLDFMSVLMLSGDHDVTNKATC
1142 LGVLVLAATDTTITTLTWVISLLNNYRDLTMAQNLEDILVGKRVVESDINKLVLYLQAILKEAMRLYPAAPILPHIESPECTISGYN
1143 YVPSTQILVINWVKLQRDPRVWEEPCKPQFPERFLTRHKIDVRQGQNFELLPGSAGRCPGVSFALQVMFTLALLHGFDFATPSE
1144 PDMGEGGLGSIQKARPVEVLAPRL
1145 >XP_006471496_(CS)
1146 APEAGGSWPVIGHLLLGGPEPPhVWGLKLADEYPFIFKMKIMGNRALVSVSNWEAMAKECLTHDKVFASRPKTLAMEILGYNFSM
1147 GFSPGYSYWRETRKIALESLNHRGLKLRHVREYELKTLKLDYLWKLLEWEMRWRDLTTIRNVLRMIVGKCGSVDG-
1148 EEWIELLTRFSMEGKGVFSVDAPFLRLWLDIGGDERSMKIAKELDIVQGWLHEEHHKRRDSLDFMSVLMLSGDHDVTNKATC
1149 LGVLVLAATDTTITTLTWVISLLNNYRDLTMAQNLEDILVGKRVVESDINKLVLYLQAILKEAMRLYPAAPILPHIESPECTISGYN
1150 YVPSTQILVINWVKLQRDPRVWEEPCKPQFPERFLTRHKIDVRQGQNFELLPGSAGRCPGVSFALQVMFTLALLHGFDFATPSNE
1151 PDMGEGGLGSIQKARPVEVLAPRL
1152 >CYP82D62_AGFP30366_(MP)
1153 PPEAGGAWFPLGLHNISSGTGLPHVSLGNLADKHKPIFIRIGVHRAVVSSEVIKELFTNTDAAVSSRSVKAKGHLAYDAMLG
1154 FSSYGYTRQMRKLVSLLELFARSERVLQNVRSETAHFENYLAWEVEVSEMKELGELMSVNLKVMAGKRFDAEEARCRR
1155 VMREFFHGLFLSVDAFPYLGLWDLGHEHHERRMTKTAEMDEMEVGLAEHRRKEAQRFPMDVMLESEVKEYDVITKATCGT
1156 LIAGGTDTTAVYMWALALLNNPHVQLQAKHELHDTHVGKQVRNVEDNLNLVNYLQAITRKEHLRYPGSLGTRLTQDCCHVGG
1157 YHIPKETLWNLWKLRHRPRVWSDPSEFRPERFLNGEKEDVKGQDFEFIPSGRRCPGQSFALQVMFTLALLHGFDFATPSNE
1158 NEEIDMSEASGLTNIKATPLDVLAPRL
1159 >CYP82D33_AGFP30364_(OB)
1160 PPEAGGAWFPLGLHNISSGTGLPHVSLGNLADKHKPIFIRIGVHRAVVSSEVIKELFTNTDAAVSSRSVKAKGHLAYDAMLG
1161 FSSYGYTRQMRKLVSLLELFARSERVLQNVRSETAHFENYLAWEVEVSEMKELGELMSVNLKVMAGKRFDAEEARCRR
1162 VMREFFHGLFLSVDAFPYLGLWDLGHEHHERRMTKTAEMDEMEVGLAEHRRKEAQRFPMDVMLESEVKEYDVITKATCGT
1163 LIAGGTDTTAVYMWALALLNNPHVQLQAKHELHDTHVGKQVRNVEDNLNLVNYLQAITRKEHLRYPGSLGTRLTQDCCHVGG
1164 YHIPKETLWNLWKLRHRPRVWSDPSEFRPERFLNGEKEDVKGQDFEFIPSGRRCPGQSFALQVMFTLALLHGFDFATPSNE
1165 NEEIDMSEASGLTNIKATPLDVLAPRL

48
LAGIVDSDFPMGWIDLSGMKARLRNFKEMDFYEELIDELKPRNRPELDVMLKNKGS---

MDSKAILLNVNGGTTGTSATLTVWAMTLRMRQVGQMMAQEDIRRRVGRGKVNSNEDDLIQNLPYLRAVKTMRMLYPTGPLLIP

RETVESSIIIGEYMIPKTKLTVYSWMAIGRDPEWKNPNMFVPERFLE--

RPDLNYKQGFYEYIPFAGRGRACIHLGVTTVELALANLLYTFDWEPPAFEDIDDETLNGLTLQKNALYRPKYYI

>CYP83A1_NP_193113_(AT)

PPGPGP-LPVIQLNNQPLK-FL-

PQRFAGWAKKYGIPSLYSRIGSTRMVTILSSSAELAKELKTTQDVFNAFPDRPHRHGHEFISYGRDMALNHYTPYREIRKMGMNHLFS

PTRVATFKHREEEEARMMDKINKAADKEVVDISELMTFTNSVVCRCQAFGKYNGEEMKRFIKILYGTQSVLGLIKFSDFFPPYGCG

FLDLGSGLTAYMKECFQERTYDQIEVNETLPKRSMIDMLMLGKIGKEETFDNVKAVILDVAGTDAAAVWWGMTYLMYK

PQVLLKAQAQVEYREMKEKGTVTEDDVKNPLYFRALVKETLRIPEVPIPLIPRACIQDTRKAGYDIPAGTTVTVNNAVAWSDEKEWG

NPDEFPERFLE--

KEVDKFVTDEYIPEFGGRMCPRMGLAAMLEVYALNLLLSFNFKLPNPDINMDVTGLAMHKSQHLKL--PEK

>CYP83B1_NP_194878(AT)

PPGPKG-LPHELNNHMEKFNQPHLFL-

RLSKLYGFIINTKMGIRGLRLAVISSAEELAKELKTTQDLNFARPLLKQGQTMYSQGERLFQGFQTYAYREMRMKCMVNFSPNRVA

SRFVPRVEEECQRMMDKINYKAADQGTVDLSELLLSTFTNCVWCRCQAFGKRNYGTEMKFRIDILEYETQALLGTFFSDLPFYGLLTDG

LSRLKLKAFKELDTYQLELDELDPNREFSPFLDLMQIWYKDFKKHTHKNAMKIDLVPQGPTDAAAVWVAMTYLJUKYMMKKA

QDEVRSVGDGYVSEEDIPNLPLYAKVSHSLRLEPVIIPLHIERTIDAKIGGYDPAKTIQHVNAWAVSRDAAWGNPNEFIPERF

MNEIHKGVDFQDFEPLLFGSGRRCPAMHGLIAMVIEIPANLLLKYKFDWSLKPEDIKMDDVTGLAMHKEHLVLAPTPI

>CYP82H1_AY532373 (AM)

PPEPGRWPLGHIHLGLGSKHLHILGDMADEYGPISLNLGINKTVSVTSWEVAKEFTTQDRVFATRPKSVQVQVGVYNSRVM

FFQYYGAYWREMRKLAELLSNRRLDLMLKHRESEVNLFLKYLFEYQWSDKVEMMKMFQFLDLITNVRTVAGKKYNEERSQFQ

KAMAFMHLGGLMMVSDALPLLGLWDVKGCCKGKMKTAAEIDHILGSLCEHKQQKRSEDDFIVMAMADGDNTDTAIAKGTCLS

LILGGYDTTSATLMWLASSLLNNRHLKLKAQDEMDOYQVRDRQVKVESDVKNLLTQLAIKETLRLYAPAALSQVHHKAMADTCV

AGFNIPAGTRLLVLNLWKMHRDPKWDSPLDQPERFPLQKHINVDWQGDNFELLPFGSGRRSCPGFITAMQVHLTLAQLLLHGFELG

TVLDSSIMTESSGTDPRATPLEVTLTPR

>CYP82S3_XP_002264790 (VV)

PPEPAPWFFUVHLHGLGKKQVTVRTFLMADGKDVPMGFPVPBFRLGMYRALVSNHEAAKECFNTNDKVFASRPSSSASKILGNYVA

FGLAPYGPLWERIKLMQLIMLRSQRLDLMLHVSHEFAUIGKEILYWLVNKVVISWEFEHNLNNVLRMVAGKRFYGEAGSA

TAVIKKLVLAPAIFVDSVIFLFLEWVLQDLHSLKMQVAKMDESVLESWEEHTGRLNSQFIDMLTMLKDYSIETIKAIATVMML

IVAGSDTTSITSTWLLSALLNRMHVKMHAQDELMLKVGDRWVWQSDIQNLVLKAIKVEVLRLYAPVLLVPEAMEDCHVVGY

HIPKGTRLLVNAWKLLHPAMPAWSNEEFQPRFLSTHTADVLDQIGNFELIPPFSGRSCPGINMALQMHLTIQALIQLQFDATMP

NSPVDMAEAISITMPLKTLEVPMTLPR

>CYP82S3_XP_002264790 (VV)

PPEPAGWFFUVHLHGLGKKQVTVRTFLMADGKDVPMGFPVPBFRLGMYRALVSNHEAAKECFNTNDKVFASRPSSSASKILGNYVA

FGLAPYGPLWERIKLMQLIMLRSQRLDLMLHVSHEFAUIGKEILYWLVNKVVISWEFEHNLNNVLRMVAGKRFYGEAGSA

TAVIKKLVLAPAIFVDSVIFLFLEWVLQDLHSLKMQVAKMDESVLESWEEHTGRLNSQFIDMLTMLKDYSIETIKAIATVMML

IVAGSDTTSITSTWLLSALLNRMHVKMHAQDELMLKVGDRWVWQSDIQNLVLKAIKVEVLRLYAPVLLVPEAMEDCHVVGY

HIPKGTRLLVNAWKLLHPAMPAWSNEEFQPRFLSTHTADVLDQIGNFELIPPFSGRSCPGINMALQMHLTIQALIQLQFDATMP

NSPVDMAEAISITMPLKTLEVPMTLPR

>CYP82S3_XP_002264790 (VV)

APEAAGAWPIGHLHLGLGKKQVTVRTFLMADGKDVPMGFPVPBFRLGMYRALVSNHEAAKECFNTNDKVFASRPSSSASKILGNYVA

MFAFPHGAYWREMRKLAELLSNRRLDLMLKHRESEVNLFLKYLFEYQWSDKVEMMKMFQFLDLITNVRTVAGKKYNEERSQFQ

RCQKSIREFFRLGLFVFDSDAPFLGLGVDVQHLHKLKATKTKDATEMIGDAQEWLEEHRKDSQDLMDVMVLSILAGYDADTINKATS

LILIAAGSDTTSITSTWLLSALLNRMHVKMHAQDELMLKVGDRWVWQSDIQNLVLKAIKVEVLRLYAPVLLVPEAMEDCHVVGY

HIPKGTRLLVNAWKLLHPAMPAWSNEEFQPRFLSTHTADVLDQIGNFELIPPFSGRSCPGINMALQMHLTIQALIQLQFDATMPS

NSPVDMAEAISITMPLKTLEVPMTLPR

>CYP82S3_XP_002264790 (VV)
PQEPSRWPPIGHILLLGADKPLHPLGADKLPHRTLGDMADKYGPIFCIHLGLRKALVVSSWEVAKECYTTNDKVFATRPRSSLAVKLGMGIDHAM

LGFAPYQPYWRDVRKLAMVELSSNRLQEMKLHVQDSEVEFLIKELYQYWQARALVEMKFERFNLMVNVMVMAIAIGKYFDDPEK

RGKKALDFMLLGVLFMVSAPFPLGWLDVKGYTTDDMKKIAKELDYLGLRWVEEHQPRSSLVDFLHVMLVSIDDRDPDTVIKAT

CLNLLAGYDTTSTLWALLSSNNHALKQAAAEILIEHVGKHQVDSGDKNLVYQLAIVIKTLRLYPGPGLVPEAMEDECTV

AGFHQQAGTRLMLNLWKLHREDPRVWLDPLEFQPERFLTNHAIGLDVRGKNYEALLPFGSRRVCPGIFSAELEHTHLARALLLHGFELG

VVADSPVDMTEGPGLSAPKATPLETVIPRL

>CYP82S6 XP_019075061 (VV)

APKPPGAWPIGHILLLSQVPFPITRGLGAMADKHGPVFMIQLGMPHAPVIVSSHEAVKECFTTNDKVFASPRRSSSVSKLGLGYNYAGF

GSAPYGLFWMERTKLSSVEILSARRNLNLKEDVRISELDACIQDLYSLSIEVVMSEFEHLTFNPVLRLMIAGKYFNEEARAGIAI

KKFLLSAGFVPSDVFPFLRLDLGKYLGSMKHLAEELDCLVGSWVEEHVMRLKSHDFDVLLASTVQDHTREVIKATGNLIVGG

SDSTSTSTWILSALLNNREAMKQAEELDLGKYLQVEEIDDQKLDYLRRAIKESLRLYAAPPLLVPEATQDCHVCYHIKPCT

RGLFVNAWLKHLRDPRVWSNPEEFEPERFGLSHANLDFVGHQFELIPGFSSRACPQINMALQLMLTFARLLQFDMATPSNAPVD

MTEGISFTMPKLTPLCVMLTTPRL

>CYP82S7 CBI37514(VV)

APKPPGAWPIGHILLLSQVPFPITRGLGAMADKHGPVFMIQLGMPHAPVIVSSHEAVKECFTTNDKVFASPRRSSSVSKLGLGYNYAMF

GSAPYGLFWMERTKLSSVEILSARRNLNLKEDVRISELDACIQDLYSLSIEVVMSEFEHLTFNPVLRLMIAGKYFNEEARAGIAI

KKFLLSAGFVPSDVFPFLRLDLGKYLGSMKHLAEELDCLVGSWVEEHVMRLKSHDFDVLLASTVQDHTREVIKATGNLIVGG

SDSTSTSTWILSALLNNREAMKQAEELDLGKYLQVEEIDDQKLDYLRRAIKESLRLYAAPPLLVPEATQDCHVCYHIKPCT

RGLFVNAWLKHLRDPRVWSNPEEFEPERFGLSHANLDFVGHQFELIPGFSSRACPQINMALQLMLTFARLLQFDMATPSNAPVD

MTEGISFTMPKLTPLCVMLTTPRL

>CYP71AU4 XP_010663080(VV)

PPSSPR-LPVGLNMHQLG--

IYPYRSLLCLAYCPILMLLLQLGVRVTLVSSPDAAEQIMKTMDLIFANRPKMLGKRLLLYDYKDVSVAPYGEYWRQMRSMICVLH

LLSNKRVQSNTVRREELISLKISSLQSLTSLDSMLGSMFMRLENDVIRVCAGFGRKSYSDEGKKFRRLGEVFELLLGFGNVYDYPW

LAWVYFNGWSSKVERVAKEDFDLFGLSVGVEHNDLDNEKDFDVDLLEIQRDMSDRDSIKALILDFAGTDDTTTVLEWAMTELLR

HPKAMKELQNEQVGRGKEHETDEKHMVYKAVIKETLRLHPPIPLLVPRESSQDVMGYHIPAGTMVINAWAICRDPMSSWD

EPEEFPRFNL--

TNIDFKGHDIPFAGRRGCPGISFAMATNELVLANLNVKFDWALPDAEDLDMETCGLTIHRKFPLAVSTPCF

>CYP71AU3 XP_010663079(VV)

PPSPPK-LPIJNMLNGQLG--

LLPHRSLWALAQRHPGMLLHFFGKVVPVVVIASAADAREIMKTNDVIFLNRPSFKAKLRLYDYKDVSVAPYGEYWRQMRSMICVLH

LSSNRVQSFGRVREETALLMEKSSSSSTPIDLSMFLSITNDLICRVALGRKSYSDEGKRYKLKEFVGLLGGFDVAYIPWLS

WVNINGDMETHODAKVEKEFDRPLDVEKKEHERRVKKDFDVNLLEIQREDSTTYYLTELWAMTMEYLH

PQVMRQLNEVRGIAQKGLLITEDDLDKMQLYKAVIKETLRLHPPVPLLPRESTRGAKIMGYDIEVTQGVTNINAWAICRDPMSSWD

EAEERPFRFNL--

SSIFGDKFELIPFAGRRGCPGISFAMATNELVLANLNVQDWEVGGREDLDMETCGLTIHRKFPLAVATPW

>CYP71AH4 XP_010663828 (VV)
PPGPKK-LPLIGNLHQLG--
SLPHVGLQRLSNEYGPLMYLKLGSVPTLVSSADMAREIFREHDLVFSSRPAPYAGKKLSYGCDNVVFAPYGEYWREVKRIVILEL
LSEKRVQSFQELREEEVTMLDVTH--
SSGPVYLSLETFFLSNNVICRVAFGKFKDGGGTPDPDILQETQNLGGFCIADFFPPWMGWFLNLGLDARLEKNFLELDKIYDKVIE
EHLDPERHEDLDVLIRVQKDDLSEIKIKGLTDMLIAGTDTSASLWVTMAELIRPSVMRKAQEEVRSAVRGKYQVEEDLSQLI
YKLNVKESLRHPPAPLLVPRKTNEDCTIRGYEVNPANTQVFVNGKSIATDPNYWENPNEFQPERFLD--
SAIDFRGQNFELLPGAGRRCAPAVNFLIELALANLHRFDWELADREDLDMESSAIGTVHKKNPLYLLATPAN
Supplemental Figure 2
Supplemental figure 4

A

B

C

1367

1368
Supplemental figure 5