**Molecular cloning and Characterization of a Flavanone
3-Hydroxylase Gene from *Artemisia annua* L.**

Authors' full names:

Shuo Xiong^{1,a}, Na Tian^{1,2,a}, Jinhua Long², Yuhong Chen¹, Yu Qin¹, Jinyu Feng¹,
Wenjun Xiao^{1,2,*}, Shuoqian Liu^{*,1,2,3}

Institution addresses:

¹ Hunan Collaborative Innovation for Utilization of Botanical Functional Ingredients,
College of Horticulture and Hardening, Hunan Agricultural University, Changsha,
410128, China.

² National Research Center of Engineering Technology for Utilization of Functional
Ingredients from Botanicals

³ Key Lab of Tea Science, Ministry of Education, Changsha, 410128, China.

^a These authors contributed equally

Corresponding author:

W. Xiao:

E-mail addresses: xiaowenjun88@sina.com

S. Liu

Tel.: +86-731-4635304; Fax: +86-731-4635306;

E-mail addresses: shqianliu@sina.com;

Abstract

Flavonoids were found to synergize anti-malaria and anti-cancer compounds in *Artemisia annua*, a very important economic crop in China. In order to discover the regulation mechanism of flavonoids in *Artemisia annua*, the full length cDNA of flavanone 3-hydroxylase (F3H) were isolated from *Artemisia annua* for the first time by using RACE (rapid amplification of cDNA ends). The completed open read frame of *AaF3H* was 1095 bp and it encoded a 364-amino acid protein with a predicted molecular mass of 41.18 kDa and a pI of 5.67. The recombinant protein of AaF3H was expressed in *E. coli* BL21(DE3) as His-tagged protein, purified by Ni-NTA agrose affinity chromatography, and functionally characterized *in vitro*. The results showed that the His-tagged protein (AaF3H) catalyzed naringenin to dihydrokaempferol in the present of Fe^{2+} . The K_m for naringenin was 218.03 μM . The optimum pH for AaF3H reaction was determined to be pH 8.5, and the optimum temperature was determined to be 35 °C. The *AaF3H* transcripts were found to be accumulated in the cultivar with higher level of flavonoids than that with lower level of flavonoids, which implied that *AaF3H* was a potential target for regulation of flavonoids biosynthesis in *Artemisia annua* through metabolic engineering.

Keywords: flavanone 3-hydroxylase; flavonoids; naringenin; dihydrokaempferol; *Artemisia annua* L.

1

2 **1. Introduction**

3 *Artemisia annua* L. (*A. annua*) is Chinese traditional medicinal plant, which has
4 been used to treat fever, inflammation and malaria for a long time in China (Liu et al.
5 1992; Udaykumar 2014), and now it is growing in many European countries and
6 North America (Dhingra et al. 1999). *A. annua* contains hydroxylated flavonoids and
7 polymethoxylated flavonoids such as eupatin, cirsilineol, chrysopenetin,
8 chrysosplenol D, casticin, and artemetin (Bhakuni et al. 2001). Flavonoids from *A.*
9 *annua* were reported to have biological activities, such as anti-oxidation, anti-cancer
10 and anti-malaria (Iqbal et al. 2012; Kim et al. 2014; Kim et al. 2015). In addition, *A.*
11 *annua* flavonoids were found to synergize anti-malaria and anti-cancer compounds
12 (Weathers et al. 2014), especially artemisinin, an endoperoxide sesquiterpene lactone
13 isolated from *A. annua* leaves that is widely used to treat malaria, one of the world's
14 most important parasitic diseases, affecting approximately 300-500 million people
15 worldwide and causing more than one million deaths per year (Salako 1998).
16 Recently, more and more research indicate that *A. annua* flavonoids inhibit the
17 activity of CYP450 enzymes responsible for altering the absorption and metabolism
18 of artemisinin in the body, while it improve the immunomodulatory activity in
19 subjects afflicted with parasitic and chronic diseases (Cherng et al. 2008). Therefore,
20 it is very important to improve the content of flavonoids through engineering the
21 biosynthetic pathway of flavonoids in *A. annua*.

22 The flavonoids, a large group of plant secondary metabolites consist of a 15
23 carbon atom phenylpropanoid core. Two aromatic rings (2×6 carbons) are joined by
24 a heterocyclic ring (3 carbons). Therefore the flavonoid structure is also referred to as
25 C6-C3-C6 (Harborne and Williams 2000). The flavonoid biosynthetic pathway has
26 been extensively studied in plants by a combination of genetic and labeling studies
27 (Quattrocchio et al. 2006). Hydroxylation of (2S)-flavanone to
28 (2R,3R)-dihydroflavonol, which are catalyzed by a (2S)-flavanone 3-hydroxylase
29 (F3H), is a key step in the biosynthesis of flavonols, anthocyanin and catechins. The
30 activity of F3H enzyme was first reported in flowers of *Matthiola incana* (Forkmann
31 et al. 1980), while the first successful purification and characterization of this enzyme
32 was performed on a red-flowering mutant *Petunia hybrida* (Britsch and Grisebach

1986). After that numerous amounts of F3H genes have been cloned and characterized from many plants, including *Saussurea medusa* (Jin et al. 2005), *Ginkgo biloba* (Shen et al. 2006), *Scutellaria viscidula* (Lei et al. 2010). However, no gene involved in flavonoids biosynthetic pathway from in *A. annua* has been identified. Recently, we obtained a red-stem *A. annua* cultivar containing higher level of anthocyanin in comparison with the wild-type, which allowed us to isolate several flavonoids related gene fragments. In the present work, we described the full-length cDNA cloning and identification of a gene encoding a (2S)-flavanone 3-hydroxylase (AaF3H) from *A. annua*.

2. Results and discussion

2.1 Candidate gene selection, full-length cloning and sequence analysis

A differently expressed cDNA library was constructed between two *A. annua* cultivars, in one of which the total anthocyanin level (aerial part) was 12.55 ± 9.78 mg/g dried weight, about 2-fold higher than that in another one (6.36 ± 0.56 mg/g). From the library, 67 non-repeated sequences were isolated by using cDNA-AFLP. Among them, an EST of 126 bp was found to show a 77% identity to F3H from *Arabidopsis thaliana*. In order to further identify the putative F3H, a full-length cDNA of 1866 bp long was obtained by SMART-RACE-PCR. The AaF3H ORF was 1095 bp and encodes a 364-amino acid protein with a predicted molecular mass of 41.18 kDa and a pI of 5.67.

Subsequently, a BLASTP search of Genbank with the predicted amino acid protein sequence of AaF3H was performed and results showed that the deduced AaF3H shared 67%, 76, 77, and 78% of identity at the amino acid level with other plant F3Hs from *Ginkgo biloba*, *Arabidopsis thaliana*, *Camellia sinensis* and *Citrus maxima* (Fig. 1), respectively. As shown in figure 1, a highly conserved ferrous iron ligation motif HXDX_{55}H (HTDPGTITLLLQDQVGGLQATRDGGKTWITVEPIEGAFVVNLGD HGYYSLSNGRFKNADH) and a 2-oxoglutarate (2-ODD) binding motif RXS (RLS) were found in AaF3H, which implicated that this deduced protein represented a 2-oxoglutarate-dependent dioxygenase. In addition, five conserved motifs commonly contained in plant 2-oxoglutarate-dependent dioxygenase were present in AaF3H (Fig. 1). In motif 2 and motif 3, three strictly conserved prolines (Pro_{148, 204, 207}) were found

(as shown in Fig. 1), which were proposed to play important roles on the folding process of the protein.

2.2 Phylogenetic analysis

In order to investigate the evolutionary relationships among AaF3H and other 2-ODD proteins involved in the biosynthesis of flavonoids in plants, the phylogenetic tree was constructed by using neighbor-joining method. The results (Fig. 2) showed that AaF3H and F3Hs from dicotyledons, such as *Canarium album*, *Litchi chinensis*, *Gossypium barbadense*, *Solanum pinnatisectum*, *Arabidopsis thaliana* and *Ipomoea batatas*, formed a cluster. While F3Hs from monocotyledon, including *Narcissus tazetta*, *Dioscorea alata* and *Zea mays*, formed another cluster far from AaF3H cluster. At the same time, all of investigated FLSs and ANSs from different species shared a same cluster, correspondingly, which was coincident with the report by Prescott and John (Prescott and John 1996).

2.3 Characterization of recombinant AaF3H

To verify the hydroxylation catalytic activity of AaF3H, the CDS of AaF3H was cloned into the expression vector pET28a(+), and over-expressed in *E.coli* BL21(DE3) cells as N-terminal His-tag function protein. Different culture and induction conditions were tested resulted in a high level of soluble protein (Fig. 3). The recombinant AaF3H protein was extracted and purified by Ni-NTA affinity column chromatography (shown as Fig. 3). From 50 mL of induced culture, 9.31 mg/mL of purified recombinant protein was obtained (Fig. 3). Subsequently, the purified AaF3H protein was assayed for flavanone 3-hydroxylase activity with ($\pm$)-naringenin as substrate, and 2-oxoglutarate, FeSO_4 as cofactors. Analysis of the reaction products by HPLC (Fig. 4b) and LC-MS (Fig. 4c&d) indicated that dihydrokaempferol was produced from naringenin with the catalysis of AaF3H and both of cofactor 2-oxoglutarate and Fe^{2+} were required for the activity of AaF3H (data not shown). The negative control assay, in which the AaF3H was boiled for 10 min at 95 °C, did not yield significant amount of dihydrokaempferol or other products (Fig. 4a). The stereo selectivity of the enzyme AaF3H was not investigated due to the lack of stereo-specific substrates. However, we observed that more than 50% of conversion of racemic mixtures used in the assays could not be achieved under all of tested conditions, which might implicated that AaF3H showed a stereo selectivity to

1 substrate, agreeing with previous reported for F3H which catalyses the stereo-specific
2 (3R)-oxidation of (2S)- but not (2R)-naringenin (Britsch and Grisebach, 1986).

3 *2.4 Kinetic parameter*

4 To further characterize catalytic properties of AaF3H, the reaction products were
5 quantified at different time. And the results (data not shown) suggested that the
6 recombinant AaF3H showed linear catalytic activity from 5 to 25 min. The optimum
7 pH for the recombinant protein was determined to be 8.5 (Fig. 5a). The catalytic
8 activity of recombinant AaF3H was sharply decreased when the pH value was less
9 than 8.5. Only about 50% of activity of recombinant AaF3H was remained at pH 7,
10 and no activity was detected at pH 6. Relatively, the temperature did not significantly
11 affect catalytic activity of recombinant AaF3H at the temperature range from 28 to
12 37 °C although the maximum activity was observed at 35 °C (Fig. 5b), which
13 implicated that the enzyme of AaF3H was thermal stability to some degree. The
14 kinetic parameters (Fig. 5c) of the AaF3H catalyzed reactions were quantified and
15 yield a Km of 218.03 µM for naringenin.

16 *2.5 Semi-quantitative RT-PCR analysis of AaF3H gene expression level*

17 In order to investigate the relationship between anthocyanin content and *AaF3H*
18 gene expression level, a semi-quantitative RT-PCR was employed to compare *AaF3H*
19 gene expression in a high-anthocyanin level cultivar with that in a low-anthocyanin
20 level cultivar by using 18s rRNA as reference gene. PCRs were performed for 24, 26
21 and 28 cycles, respectively. No significant electrophoresis band was obtained when
22 the cycle number was 24 (Fig. 6). When the cycle number was more than 26,
23 high-anthocyanin level cultivar showed stronger electrophoresis band than low-
24 anthocyanin level cultivar did (as shown in Fig. 6). The results suggested that *AaF3H*
25 transcripts were much more abundant in the high-anthocyanin level cultivar of *A.*
26 *annua* compared to low-anthocyanin level cultivar, which was consistent with the
27 cDNA-AFLP assay.

28 **3. Conclusions**

29 To isolate a cDNA of the enzyme flavanone 3-hydroxylase, we previously used
30 the strategy of cDNA-AFLP to separate the differently expressed ESTs between a
31 high-anthocyanin level cultivar and a low-anthocyanin level cultivar. The approach

1 resulted in the isolation and characterization of a flavanone 3-hydroxylase gene from
2 *A. annua* denominated as AaF3H. In the present work we obtained the full-length
3 cDNA of *AaF3H* by using RACE. Bioinformatic analysis of AaF3H indicated that it
4 formed a cluster with F3Hs from dicotyledons and it contained all of the conserved
5 motifs for plant 2-ODD. The AaF3H was produced by heterogeneous expression of
6 *AaF3H* gene in *E. coli* cell and purified AaF3H protein was isolated by agarose
7 affinity chromatography. *In vitro* assays revealed that AaF3H performed the function of
8 flavanone 3-hydroxylase that catalyzed the hydroxylation reaction of flavanone to
9 dihydroflavonol with the Fe^{2+} as cofactor. The K_m for naringenin was 218.03 μM . The
10 semi-quantitative RT-PCR analysis showed *AaF3H* transcripts were much more
11 abundant in the high-anthocyanin level cultivar of *A. annua* compared to
12 low-anthocyanin level cultivar, which suggested that the *AaF3H* gene might be a key
13 target gene for regulation of flavonoids biosynthesis.

14 **4. Experimental**

15 *4.1 Plant Material*

16 A high-anthocyanin *A. annua* cultivar and a low-anthocyanin *A. annua* cultivar
17 were collected in the experiment field of HAU (Hunan Agricultural University,
18 Changsha, Hunan, China), and identified by Prof. Puhua Zhou of HAU. One voucher
19 specimen (No. HF2010001) for high-anthocyanin cultivar and one voucher specimen
20 (No. LF2009001) for low-anthocyanin cultivar have been deposited at the herbarium
21 of HAU. The plants were grown in a controlled-growth chamber of our laboratory
22 (Changsha, Hunan, China) in a light/dark cycle of 16/8 h by using fluorescent lamps
23 (with a light intensity of 2800 lx) at 25 °C with 70% relative humidity. The fully
24 expanded leaves of 2-month-old plant were collected and frozen in liquid nitrogen and stored at
25 -80 °C until use. The aerial part of 2-month-old plant was dried in room and was used
26 for quantification of total anthocyanin.

27 *4.2 Chemicals*

28 All organic solvents were purchased from Changsha Chemical Reagent Company
29 (Changsha, Hunan, China). All other authentic compounds were from Sigma-Aldrich
30 (Shanghai, China).

31 *4.3 RNA isolation*

Total RNA of *A. annua* was extracted by using Trizol reagent (Promega). Approximately 0.2 g of the frozen leaves was ground in a pre-cooled mortar with liquid nitrogen and 1 mL of Trizol reagent was added with an incubation of 5 min at room temperature. Then 200 μ L of chloroform was added and was shaken for 3 min. After incubation of 15 min at room temperature, it was centrifuged at 13000 rpm for 15 min at 4 °C. The supernatant was transferred to a fresh tube and 250 μ L of salt buffer (containing 0.5 M sodium citrate and 1.2 M Sodium chloride) was added. Afterward, 250 μ L of isopropanol was added and mixed, and then the mixture was incubated for 10 min at room temperature. Then it was centrifuged at 13000 rpm for 10 min at 4 °C. The pellet was washed with 75% ethanol for two times and then dried at room temperature, and then was dissolved in 40 μ L RNAase free water.

4.4 Isolation of differently expressed genes

cDNA-AFLP was performed according to the previously described method Vos (Vos et al., 1995), by using *EcoR* I and *Mse* I restriction enzymes (Fermentas, USA). The adapters for AFLP reactions were as follows: 5'-CTCGTAGACTGCGTACC-3' and 3'-CTGACGCATGGTTAA-5' for *EcoR* I; 5'-GACGATGAGTCCTGAG-3' and 3'-TACTCAGGACTCAT-5' for *Mse* I. Primers 5'-GACTGCGTACCAATTC-3' and 5'-GATGAGTCCTGAGTAA-3' were used as pre-amplification primers of *EcoR* I and *Mse* I, respectively. Sixteen additional primers were used for selective amplification (Table 1). Selective amplification products were mixed with an equal volume of loading buffer and heat-denatured for 5 min at 95 °C. Then 5 μ L of each denatured sample was loaded and separated on a 6% polyacrylamide gel at 80 W and 50 °C in Sequi-Gen GT Nucleic Acid Electrophoresis Cell (Bio-Rad, USA) and then was visualized by silver staining. Dried gels were scanned with a GS-800 Calibrated Densitometer (Bio-Rad, USA).

4.5 Sequencing of differently expressed genes

Bands corresponding to differentially expressed transcripts were cut from the gel with a sharp needle and then were crushed in 40 μ L sterile water, heated at 65 °C for 30 min and then centrifuged at 13,000 rpm for 20 min. Supernatant (1 μ L) was re-amplified under the conditions used for selective amplification. The re-amplified products were separated in a 1% agarose gel, and each single band was purified using

1 the Promega Wizard SV Gel and PCR Clean-up System (Promega, USA). The
2 purified products were cloned into the pGEM-T Easy vector (Promega, Madison, WI,
3 USA) and then were transformed into *E. coli* cells to obtain a differently expressed
4 genes library.

5 4.6 RACE PCR Cloning

6 The amplification of the 3' and the 5' ends of *AaF3H* was performed using
7 SMART™ RACE cDNA Amplification Kit (Clontech). 5' RACE PCR reactions were
8 performed with primer UPM (5'-CTAATACGACTCACTATAGGGCAAGCAG

9 TGGTATCAACGCAGAGT-3') and gene specific primer F3H SR (5'-
10 TTGCAGCCCAAGCCCATTAGATCCTC-3'). 3' RACE PCR reaction was
11 performed by using gene specific primer F3H SF (5'-
12 -CGAGCGGCCTAAGGTGCCATACAACA-3') and primer UPM
13 (5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3').

14 The touchdown PCR program was initially denatured at 94 °C for 30 s, annealing at
15 72 °C for 3 min, 5 cycles; denatured at 94 °C for 30 s, annealing at 70 °C for 3 min 5
16 cycles; 94 °C for 30 s, annealing at 68 °C for 30s, and extension at 72 °C for 3 min,
17 22 cycles. The PCR products were sub-cloned into the vector pGEM-T (Promega) and
18 sequenced by using an automatic ABI 377 sequencer (Perkin Elmer Applied
19 Biosystems, Foster City, CA, USA) in Shanghai Sangon Biological Engineering
20 Technology and Services Co., Ltd. (Shanghai, China).

21 4.7 Protein Expression and Purification

22 The ORF of *AaF3H* gene was amplified by PCR using forward (5'- CGCCATATG
23 GCACCTATATCCTTAAAATGGGAC-3') and reverse (5'-TGGGAGCTC
24 CTAAGCAAAGATACTTTCAATGGGC -3') primers containing restriction sites
25 *NdeI* and *SacI*, respectively. The thermal cycling conditions were as follows: 95 °C for
26 1min, followed by 40 cycles of 95 °C for 30 s, 55 °C for 30 s , 72 °C for 45 s , and
27 one cycle of 72 °C for 10 min. The resulting fragment was sub-cloned into pGEM-T
28 vector (Promega) by TA cloning, and sequenced for correction. The pGEM-T vector
29 containing *AaF3H* gene was then digested with *Nde I* and *Sac I*, and was subsequently
30 ligated to pET28a(+) vector. The resulted plasmid, pET28a(+)-*AaF3H*, was used to

transform competent *E. coli* BL21 (DE3) strain for recombinant protein expression. Singular clone of *E. coli* BL21 (DE3), containing pET28a(+)-AaF3H expression plasmid, was inoculated in 3 mL of selective Luria-Bertani (LB) medium supplied with 50 mg/L of kanamycin and incubated at 37 °C under vigorous shaking (200 rpm). Five hundred micro-liter of overnight culture of *E. coli* BL21 (DE3) containing pET28a(+)-AaF3H was transferred to 50 mL of fresh LB medium containing 50 mg/L of kanamycin. The culture was continued to be incubated at 37 °C with shaking until the OD600 reached the value of about 0.6. Zero point five micromoles of isopropyl β -D-1-thiogalactopyranoside (IPTG) were added to the culture and it was grown at 28 °C for an additional 8 h. The bacteria were harvested by centrifugation at 4000 rpm for 10 min and washed by using binding buffer (300 mM NaCl, 50 mM sodium phosphate buffer, 10 mM imidazole, pH 7.4). The bacteria pellet was suspended in 1 mL of binding buffer and the cells were sonicated for 25 times for 10 s with an interval of at 30 s. The supernatant was collected by centrifugation at 13500 rpm for 30 min at 4 °C. AaF3H was further purified by nickel-nitrilotriacetic acid (Ni^{2+} -NTA) agarose affinity chromatography. The purified protein was desalted through PD-10 column (GE Healthcare) and stored in 50 mM phosphate buffer (pH 7.5) containing 1 mM dithiothreitol (DTT) and 20% glycerol at -80 °C. The concentration of purified protein was determined by the Bradford assay, using a standard curve of bovine serum albumin. The purity of the recombinant enzyme was assessed by standard SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) analysis.

4.8 *In vitro* enzyme assays

Enzyme assays were performed in 100 μL of 80 mM Tris-HCl buffer, containing 200 μmol naringenin, 0.5 mmol DTT, 1 mmol sodium ascorbate, 0.2 mmol ferrous ion, 1 mmol 2-oxo-glutarate and 1 μg His-tagged protein. After the incubation at 37 °C for 40 min, the reaction product was extracted by ethyl acetate. The extraction solution was dried by nitrogen and the residue was re-suspended in 100 μL methanol and then was analyzed by HPLC (high performance liquid chromatography). Boiled His-tagged protein was used as control. The optimum temperature was determined by performing reactions at 23 °C, 25 °C, 28 °C, 30 °C, 35 °C and 37 °C. The optimum pH was investigated from a set of assays performed at different pH value from 6.0 to 10. To construct the Michaelis-Menten equation, the enzyme assays were performed at optimum temperature and pH for 1-5 min in 0.1 mL of Tris-HCl buffer with the

200-700 μM of naringenin. All assays were performed in triplicate. The kinetic parameters of the AaF3H catalyzed reactions were quantified a Michaelis-Menten equation constructed using Origin Pro 9.0.

4.9 Chromatography condition

Naringenin and its corresposive product analysis was performed with reversed-phase HPLC(Agilent 1260 series) coupled to an PDA and MS detector on a Ecosil C₁₈ column (4.6 μm ×150mm, 5 μm), mobile phases were acetonitrile (B) and H₂O (A), and the elution program was: 0-52 min, 8% B - 85% B; 52-52 min, 85% B - 8% B; 55-60 min, 8% B - 8% B. The flow rate of the mobile phase was 1 mL/min, column oven was set at 30 °C, the elution was monitored by the absorption profile at 280 nm.

4.10 Semi-quantitative RT-PCR analysis of AaF3H gene expression

First-strand cDNA was synthesized from 1 μg of total RNA in a volume of 20 μL containing 20 mM Tris-HCl (pH 8.3), 100 mM KCl, 2.5 mM dNTP, 20 units of RNase inhibitor, 5 mM MgCl₂, 5 units of AMV reverse transcriptase and 2.5 pmol of Oligo dT-Adaptor primer. One microliter of the first-strand solution was used for PCR reaction in a total volume of 50 μL with 20 mM Tris-HCl (pH 8.3), 100 mM KCl, 2.5 mM dNTP, 5 units of Taq DNA polymerase (Promega, Madison, WI, USA), 5 mM MgCl₂, 10 pmol of each gene-specific amplification primer (forward 5'-GGGCTGCAAGTTGCTAGAGGTGT -3'; reverse 5'-CTGGTGGTCCGCATTCTTAAATC -3'). Amplification was carried out for 28, 26 or 24 cycles of denaturation at 94 °C for 30 s, annealing at 56 °C for 30 s, and extension at 72 °C for 30 s, with a final extension at 72 °C for 10 min, in a T-Gradient thermocycler (Biometra, Gottingen, Germany). The PCR products were checked by gel electrophoresis. The 18S rRNA gene (forward primer: 5'-TTTCGTCTATCGTCGCATTT -3'; reverse primer: 5'-CGTGGAGCAAGTCGTGTAA -3') was used as an internal standard.

Acknowledgements

We sincerely thank Dr. Weiping Zhu (Hunan Agricultural University, Changsha, China) for providing plant material. This work was financially supported by the International Science & Technology Cooperation Program of China (2014DFA32160)

and the National Science Foundation of China (No. 31270335, No. 31501367).

References

Bhakuni R, Jain D, Sharma R, Kumar S (2001) Secondary metabolites of *Artemisia annua* and their biological activity. *Current science* 80: 35-48

BRITSCH L, GRISEBACH H (1986) Purification and characterization of (2S) - flavanone 3 - hydroxylase from *Petunia hybrida*. *European Journal of Biochemistry* 156: 569-577

Cherng J-M, Chiang W, Chiang L-C (2008) Immunomodulatory activities of common vegetables and spices of Umbelliferae and its related coumarins and flavonoids. *Food chemistry* 106: 944-950

Dhingra V, Rao KV, Narasu ML (1999) Current status of artemisinin and its derivatives as antimalarial drugs. *Life sciences* 66: 279-300

Forkmann G, Heller W, Grisebach H (1980) Anthocyanin biosynthesis in flowers of *Matthiola incana* flavanone 3-and flavonoid 3'-hydroxylases. *Zeitschrift fur Naturforschung. Section C: Biosciences*

Harborne JB, Williams CA (2000) Advances in flavonoid research since 1992. *Phytochemistry* 55: 481-504

Iqbal S, Younas U, Chan KW, Zia-Ul-Haq M, Ismail M (2012) Chemical composition of *Artemisia annua* L. leaves and antioxidant potential of extracts as a function of extraction solvents. *Molecules* 17: 6020-6032

Jin Z, Grotewold E, Qu W, Fu G, Zhao D (2005) Cloning and characterization of a flavanone 3-hydroxylase gene from *Saussurea medusa*. *Mitochondrial DNA* 16: 121-129

Kim MH, Seo JY, Liu KH, Kim J-S (2014) Protective effect of *Artemisia annua* L. extract against galactose-induced oxidative stress in mice. *PloS one* 9: e101486

Kim W-S, Choi WJ, Lee S, Kim WJ, Lee DC, Sohn UD, Shin H-S, Kim W (2015) Anti-inflammatory, Antioxidant and Antimicrobial Effects of Artemisinin Extracts from *Artemisia annua* L. *The Korean Journal of Physiology & Pharmacology* 19: 21-27

- 1 Lei W, Yao R, Kang X, Tang S, Qiao A, Sun M (2010) Isolation and characterization
2 of the anthocyanidin genes PAL, F3H and DFR of *Scutellaria viscidula* (Lamiaceae).
3 Genetics and molecular research: GMR 10: 3385-3402
- 4 Liu KC-SC, Yang S-L, Roberts M, Elford B, Phillipson J (1992) Antimalarial activity
5 of *Artemisia annua* flavonoids from whole plants and cell cultures. Plant Cell Reports
6 11: 637-640
- 7 Prescott AG, John P (1996) Dioxygenases: molecular structure and role in plant
8 metabolism. Annual review of plant biology 47: 245-271
- 9 Quattrocchio F, Baudry A, Lepiniec L, Grotewold E (2006) The regulation of
10 flavonoid biosynthesis. The science of flavonoids. Springer, pp 97-122
- 11 Salako L (1998) Artemisinin and its derivatives: the regulatory and policy
12 implications for African countries. Med Trop (Mars) 58: 82-84
- 13 Shen G, Pang Y, Wu W, Deng Z, Zhao L, Cao Y, Sun X, Tang K (2006) Cloning and
14 characterization of a flavanone 3-hydroxylase gene from *Ginkgo biloba*. Bioscience
15 reports 26: 19-29
- 16 Udaykumar P (2014) Discovery of artemisinin: The Chinese wonder drug. Muller
17 Journal of Medical Sciences and Research 5: 191
- 18 Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J,
19 Peleman J, Kuiper M (1995): AFLP: a new technique for DNA fingerprinting. Nucl
20 Acid Res. 23: 4407-4414
- 21 Weathers PJ, Elfawal MA, Towler MJ, Acquah-Mensah GK, Rich SM (2014)
22 Pharmacokinetics of artemisinin delivered by oral consumption of *Artemisia annua*
23 dried leaves in healthy vs. *Plasmodium chabaudi*-infected mice. Journal of
24 ethnopharmacology 153: 732-736

25

26 **Figure legends**

27 **Figure 1.** Sequence alignment of the deduced AaF3H protein with other plant
28 flavanone 3-hydroxylases. AaF3H, *A. annua* F3H; AtF3H (AEE78766.1), *A. thaliana*
29 F3H; CsF3H, *C. sinensis* (AAT68774.1), F3H; GbF3H (AAU93347.1), *G. biloba*
30 F3H; CmF3H (ACR77526.1), *C. maxima* F3H. The conserved motifs are labeled with

bold lines. Red frames indicate the conserved motifs required for ferrous iron (HXD) and 2-oxoglutarate binding (RXS). The conserved amino acids are indicated with white foreground and black background. Non-similar amino acids were indicated with black foreground and white background.

Figure 2. Phylogenetic tree of some known 2-ODDs, including F3Hs, ANSs and FLSs. The bars represent evolutionary distance. The reliability of the tree is measured by bootstrap analysis with 1000 replicates.

Figure 3. SDS-PAGE analysis of the protein content of the pET28a-AaF3H transformed *Escherichia coli* BL21(DE3). Lanes: 1, 6 μ L of soluble lysate fraction from *E.coli* BL21 harboring pET28a-AaF3H induced by IPTG; 2, 6 μ L of soluble lysate fraction from *E. coli* BL21 harboring pET28a-AaF3H non-induced by IPTG; 3, 1 μ L of purified recombinant enzyme. The arrow points to the over-expressed His-AaF3H.

Figure 4. HPLC analysis of AaF3H catalyzed reactions using substrates naringenin (①) to form dihydrokaempferol (②). (a) HPLC chromatogram after incubation of naringenin with boiled His-AaF3H in the presence of ferrous ion, 2-oxoglutarate and ascorbate. (b) HPLC chromatogram after incubation of naringenin with active His-AaF3H in the presence of ferrous ion, 2-oxoglutarate and ascorbate. (c) LC-MS chromatogram of naringenin from the reaction products. (d) LC-MS chromatogram of dihydrokaempferol from the reaction products.

Figure 5. Kinetic assay of AaF3H with naringenin as a substrate: (a) effect of pH on AaF3H activity, (b) effect of temperature on AaF3H activity and (c) velocity of AaF3H at increasing naringenin concentrations.

Figure 6. Semi-quantification of *AaF3H* expression level in the leaves of *Artemisia annua* with high (H) and low (L) level of anthocyanin.

Contribution

Shuo Xiong: writing the manuscripts, metabolite analysis

Na Tian: Protein extraction and analysis

Jinhua Long: construction of protein expression

Gehua Wu: gene cloning

Yuhong Chen: gene expression analysis

Yu Qin: proteins blast analysis

Jinyu Feng: RNA extraction

Wenjun Xiao: design assays, discussion

Shuoqian Liu: design assays, data analysis

Table 1. Primer sequences of EcoR I and Mse I adapters for selective amplification.

Primers	sequence
E-AAG	5'-GACTGCGTACCAATTCAAG-3'
E-ACG	5'-GACTGCGTACCAATTCACG-3'
E-AGG	5'-GACTGCGTACCAATTCAGG-3'
E-ATC	5'-GACTGCGTACCAATTCATC-3'
M-AAC	5'-GATGAGTCCTGAGTAAAAC-3'
M-AAG	5'-GATGAGTCCTGAGTAAAAG-3'
M-AAT	5'-GATGAGTCCTGAGTAAAAT-3'
M-AGA	5'-GATGAGTCCTGAGTAAAGA-3'
M-AGT	5'-GATGAGTCCTGAGTAAAGT-3'
M-CAA	5'-GATGAGTCCTGAGTAACAA-3'
M-CCC	5'-GATGAGTCCTGAGTAACCC-3'
M-CGA	5'-GATGAGTCCTGAGTAACGA-3'

Figure 1

CmF3H	1	MAP-STLTALAAEKTINP-----SFVRFQDERPKVAYNEFSNEIPVISLAGIDDVG
CsF3H	1	MAPTTTTLTALAEKSLQQ-----KFVRDEDERPKVAYNVFSNEIPVISLAGIDEIE
AtF3H	1	MAP-GTLTELAGE SKLNS-----KFVRDEDERPKVAYNVFSDEIPVISLAGIDDVD
AaF3H	1	MAP---ISLKWDDNSLHEN-----RFVRDEDERPKVPYNKFSNEIPVISLKGIDDVE
GbF3H	1	MAPVQSVSAAPVPKFLPTENEAKVLQSI FVRDEDERPKVAYNEFSKEIPIISLAGMEGEE
		Motif 1
CmF3H	50	-----GKRAEICKKIVEACEDWGIFQVVDHGVD AKLISDMTRLATEFFALPPEEKLKF
CsF3H	51	-----GRSEICRKIVEACEGWGVFQVVDHGVDANLIAEMTRLAREFFALPPEEKLRF
AtF3H	50	-----GKRGEICRQIVEACENWGIFQVVDHGVDTNLVADMTRLARDFFALPPEDKLRF
AaF3H	50	DSSGGIKSRRAEICEKIIKACEEWGIFQVVDHGVDTKLLSEMTRLAREFFEMPAAEKLRF
GbF3H	60	-----RGRVREEVRAACEEWGIFQVLHHGVPSDLVHRMSQLSRSFALPSHEKLKF
		Motif 2
CmF3H	104	DMSGGKKGGFIVSSHLQGEVVKDWREIVTYFSYPKQSRDYSRWPDKPEGWMEVTKKEYSDQ
CsF3H	105	DMSGGKKGGFIVSSHLQGEAVQDWREIVTYFSYPIRARDYSRWPDKPEGWRAVTETETSEK
AtF3H	104	DMSGGKKGGFIVSSHLQGEAVQDWREIVTYFSYPVRNRDYSRWPDKPEGWVKVTEEYSER
AaF3H	110	DMTGGKKGGFIVSSHLQGEAVQDWREIVTYFSYPIKARDYSRWPDKPKEWRAVTEKYSER
GbF3H	112	DMTGGKRGGFVSSHLQGESVLDWREIFTYFSYPLRSRDYSRWPNPDGWREVVEEYSQE
		Motif 3 Motif 4
CmF3H	164	LMGVACKLLEVLSEAMGLEKEALTKACVDMDDQEI VVNYYPKCPQPDLTGLGLKRHTDPGTI
CsF3H	165	LMDLACKLLEVLSEAMGLEKEALTKACVDMDDQKVVINFYPKCPQPDLTGLGLKRHTDPGTI
AtF3H	164	LMSLACKLLEVLSEAMGLEKESLTNACVDMDDQKIVVNYYPKCPQPDLTGLGLKRHTDPGTI
AaF3H	170	LMGLGCKLLEVLSEAMGLNKEALKDACVDMDDQKV VVNYYPKCPQPDLTGLGLKRHTDPGTI
GbF3H	172	LMKLACKLLEII SESLGLEPQAVTKACLDMDQKVVINFYPKCPQPDMTGLGLKRHTDPGTI
		Motif 5
CmF3H	224	TLLLQDQVGGLQATKDNGK TWITVQPIEGAFV VNLGDHGHYLSNGRFKNADHQAVVNSNS
CsF3H	225	TLLLQDQVGGLQATR DGGK TWITVQ PVEGAFV VNLGDHGHYLSNGRFKNADHQAVVNSNC
AtF3H	224	TLLLQDQVGGLQATR DNGK TWITVQ PVEGAFV VNLGDHGHYLSNGRFKNADHQAVVNSNS
AaF3H	230	TLLLQDQVGGLQATR DGGK TWITVEPIEGAFV VNLGDHGYYLSNGRFKNADHQAVVNSNT
GbF3H	232	TLLLQDQVGGLQATKDG-LNWITVEPVEGAFV VNLGDHMHYLSNGKYKNADHQAVVNADT
CmF3H	284	SRLSIATFQNPAPGATVYPLK--IREGEKPVLEKPIPFSEMYRRKMSKDLELARLKKLAN
CsF3H	285	SRLSIATFQNPAP EATVYPLK--IREGEKPILEEPITFADMYKRKMSKDIE LAKLKKLAK
AtF3H	284	SRLSIATFQNPAPDATVYPLK--VREGEKAILEEPITFAEMYKRKMGRDLELARLKKLAK
AaF3H	290	SRLSIATFQNPAPNAIVYPLK--VNEGEKSIMEEAITFTDMYKKKMSTDELAHLKKLAK
GbF3H	291	SRLSIATFQNP AQEAIVYPLEGVVKAEDKCVMEEPITFAQMYSHKMNRDLELARQKKLAK
CmF3H	342	EK--HQDS---EKAKLDTKPIEEILA
CsF3H	343	EKKLLQDQDIEKAKLEIKSTDEIFA
AtF3H	342	EE--RD-----HKEVDK PVDQIFA
AaF3H	348	AK--QQD-----LEKVKPIESIFA
GbF3H	351	LQ---QD-----ESK-----

Figure 2.

Figure 3.

Figure 4.

Figure 5

Fig 6.