




## Pathway-specific enzymes from bamboo and crop leaves biosynthesize anti-nociceptive C-glycosylated flavones

Yuwei Sun<sup>1,4</sup>, Zhuo Chen<sup>1,2,4</sup>, Jingya Yang<sup>3,4</sup>, Ishmael Mutanda <sup>1</sup>, Shiyi Li<sup>3</sup>, Qian Zhang<sup>1,2</sup>, Ying Zhang<sup>1,2</sup>, Yulian Zhang<sup>1,2</sup> & Yong Wang<sup>1</sup> 

C-glycosylated flavones (CGFs) are promising candidates as anti-nociceptive compounds. The leaves of bamboo and related crops in the grass family are a largely unexploited bioresource with a wide array of CGFs. We report here pathway-specific enzymes including C-glycosyltransferases (CGTs) and P450 hydroxylases from cereal crops and bamboo species accumulating abundant CGFs. Mining of CGTs and engineering of P450s that decorate the flavonoid skeleton allowed the production of desired CGFs (with yield of 20–40 mg/L) in an *Escherichia coli* cell factory. We further explored the antinociceptive activity of major CGFs in mice models and identified isoorientin as the most potent, with both neuroanalgesic and anti-inflammatory effects superior to clinical drugs such as rotundine and aspirin. Our discovery of the pain-alleviating flavonoids elicited from bamboo and crop leaves establishes this previously underutilized source, and sheds light on the pathway and pharmacological mechanisms of the compounds.

<sup>1</sup>CAS-Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China. <sup>2</sup>University of Chinese Academy of Sciences, Beijing 100039, China. <sup>3</sup>College of Food Science and Technology, Shanghai Ocean University, Shanghai 201306, China. <sup>4</sup>These authors contributed equally: Yuwei Sun, Zhuo Chen, Jingya Yang. email: [yongwang@sibs.ac.cn](mailto:yongwang@sibs.ac.cn)

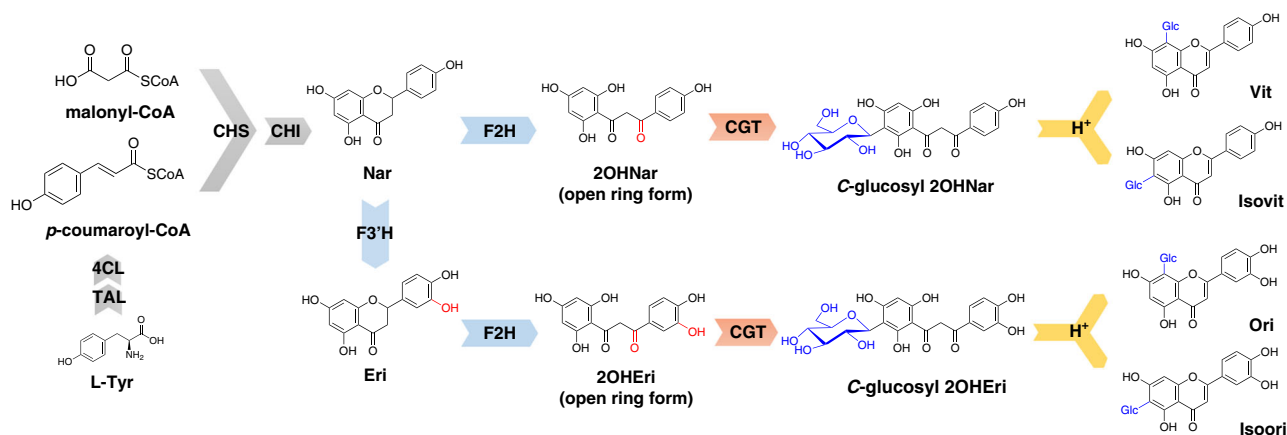
Plant parts have been used for treatment and management of pain and inflammation since time immemorial. Previous and ongoing work with plant extracts has identified some promising antinociceptive compounds, for example, C-glycosylated flavones (CGFs) that have been shown to exhibit analgesic activity in both partially pure extracts<sup>1,2</sup> and purified forms<sup>3,4</sup>. CGFs comprise one or more sugar moieties bound to an aglycone scaffold through a direct C–C glycosidic linkage (Supplementary Fig. 1). Compared to other flavone classes and closely related O-glycosylated flavonoids, less is known on C-glycosylated flavonoids in terms of their distribution pattern in the plant kingdom, the evolutionary history of pathway-specific genes and their pharmacological mechanism for anti-nociceptive activity and other biological functionalities.

As one of the richest sources of natural flavonoids, plants in the grass family provide a wide array of chemically diverse CGFs in their edible and inedible parts. For the past half a century, immense phytochemical studies on agriculturally important crops, i.e., rice<sup>5,6</sup>, wheat<sup>7</sup>, maize<sup>8</sup>, barley<sup>9</sup>, and even cash crops like sugarcane<sup>10,11</sup> have documented and profiled abundant C-glycosides in leaves, hulls, and grains, represented by vitexin (Vit), orientin (Ori), schaftoside, carlinoside, and maysin (Supplementary Table 1). Bamboo, although not cultivated for food production, is also one member of the grass family that is widely distributed in the world and has various products useful to humans. Edible parts of bamboo like shoots and seeds are commonly consumed in eastern and southern Asia as a rich dietary source of fiber and nutrients, as well as plant-specific phenolic metabolites<sup>12,13</sup>. However, as is the case with many food crops, the non-edible CGF-rich leaves of bamboo are typically viewed far less valuable and largely abandoned. Even though ancient Chinese did believe that some folk medicine or herbal tea made from bamboo leaves could help remove “toxic heat”, resulting in the alleviation of inflammation and pain<sup>14</sup>, a poor understanding of both the biosynthetic machinery of CGFs in bamboos and the inadequately defined pharmacological mechanisms severely hindered effective and extensive use of bamboo leaves as analgesics.

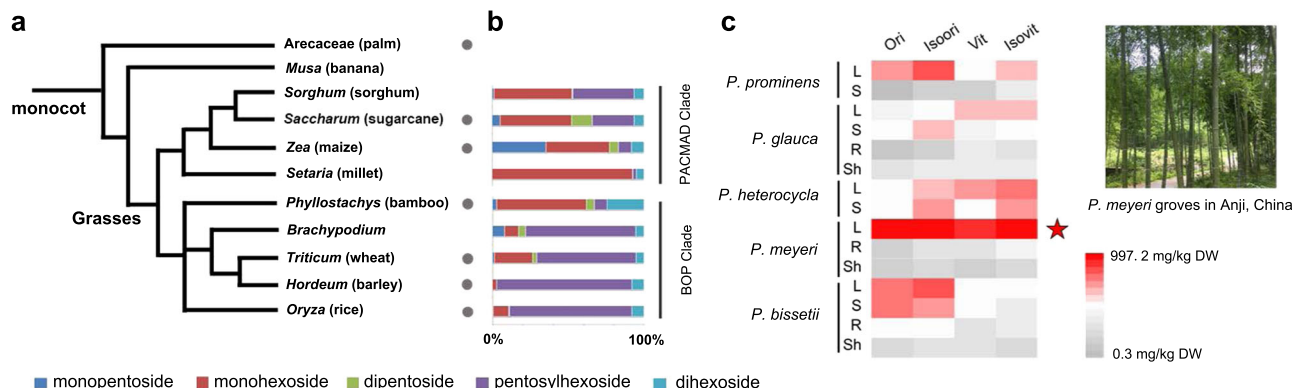
The in planta CGF biosynthesis (Fig. 1) has been so far partially established in a few monocotyledons and dicotyledonous plants<sup>15–23</sup>. The pathway is understood to involve upstream precursor biosynthesis and several downstream modification

enzymes including cytochrome P450s and glycosyltransferases. After formation by condensation of *p*-coumaroyl-CoA from the phenylalanine pathway and malonyl-CoA from fatty acid biosynthesis, naringenin (Nar), the common starter of all flavonoids, is diverted to hydroxylation on the flavanone scaffold by two P450 oxygenases, flavanone 2-hydroxylases (F2H) and flavanone 3'-hydroxylases (F3'H). While F3'H are generally present in most plants, F2H are highly specific to CGF pathways and the only few cases are reported in cereals<sup>19,20,24</sup>. Subsequently, a C-glycosyltransferase (CGT) that acts upon an open-ring form of 2-hydroxylated intermediate is key to C-glycoside formation (Fig. 1). CGTs belong to the family of the UDP-dependent glycosyltransferases (UGTs) that utilize activated UDP-sugars as donor. Despite that direct C-glycosylation on the final flavonoid skeleton have been documented in *Pueraria lobate*<sup>25</sup>, *Gentiana triflora*<sup>26</sup> and *Trollius chinensis*<sup>27</sup>, the other existing CGTs exclusively come from UGT708 subfamily and glycosylate 2-hydroxylated intermediates with UDP-glucose (UDP-Glc). Considering the chemical diversity of known CGFs bearing various sugars and different glycosylation patterns other than glucosyl substitution, there is still a huge gap between known CGTs and diverse structures of CGFs. The C-glycosylated intermediates are unstable and therefore undergo spontaneous dehydration in acidic solvent (Fig. 1), leading to a mixture of 6-C-β-D-glucoside)/isovitexin (Isovit, apigenin 6-C-β-D-glucoside) and Ori (luteolin 8-C-β-D-glucoside)/isoorientin (Isoori, luteolin 6-C-β-D-glucoside).

An improved understanding of the biosynthetic pathway of CGFs along with dissection of the pharmacological targets of biological activity can allow more efficient utilization of classically discarded parts of bamboo and crops, as well as hasten sustainable production of CGFs through synthetic biology. In this study, we took advantage of the recent availability of genomes in grass family to exhaustively uncover CGTs and P450 genes involved in CGF biosynthesis. Through comparative genomic and biochemical approaches, bifunctional C-arabinosyl/C-glucosyltransferring enzymes, a previously unknown branch in CGTs from cereals and bamboos were functionally profiled. We also demonstrated the success of heterologous expression of bamboo-originated proteins by metabolic engineering in an *Escherichia coli* chassis and further evaluated the efficacy of the major CGFs from bamboos in in vivo mouse model.



**Fig. 1 Overview of the proposed biosynthetic pathway of C-monoglucosylated flavone.** Gray arrows represent the major enzymes responsible for the biosynthesis of naringenin. Blue arrows are flavanone 2-hydroxylases (F2H) and flavanone 3'-hydroxylases (F3'H) decorating the skeleton of flavanone. C-glycosyltransferases (CGTs) are indicated in red arrows. After the formation of C-glycosylated intermediates, a dehydration reaction occurs spontaneously in acidic solvent (yellow arrows), producing a mixture of 6-C- or 8-C-glucosides. TAL tyrosine ammonia lyase, 4CL 4-coumarate CoA ligase, CHS chalcone synthase, CHI chalcone isomerase, Nar naringenin, Eri eriodictyol, 2OHNar 2-hydroxynaringenin, 2OHEri 2-hydroxyeriodictyol, Vit vitexin, Isovit isovitexin, Ori orientin, Isoori isoorientin.



**Fig. 2 C-glycosylated flavone profiling in grass species.** **a** Phylogenetic tree of monocot species (modified from CoGe). The grass family consists of two major branches: PACMAD clade and BOP clade<sup>29,30</sup>. Gray dots indicate the species that has been reported to accumulate C-glycosylated flavones. Common names are indicated in parentheses. **b** Distribution of C-glycosylated core structures (aglycones bearing only C-sugars) in the Gramineae plants. The proportion of different glycosylated forms are displayed in different colors. High percentage of pentosylhexoside is observed in BOP clade, while PACMAD clade contains more monohexosides. **c** Heat map visualization indicates metabolic spectrum of C-monoglucosylated flavone in different tissues of five *Phyllostachys* bamboos. The contents of major CGF compounds range from 0.3 to 997.2 mg/kg dry weight (DW). Monoglucosides are highly accumulated in the leaves of *P. meyeri* McClure. L leaf, S stem, R root, Sh Shoot.

## Results

### Leaf CGF spectrum of bamboo and related Gramineae crops.

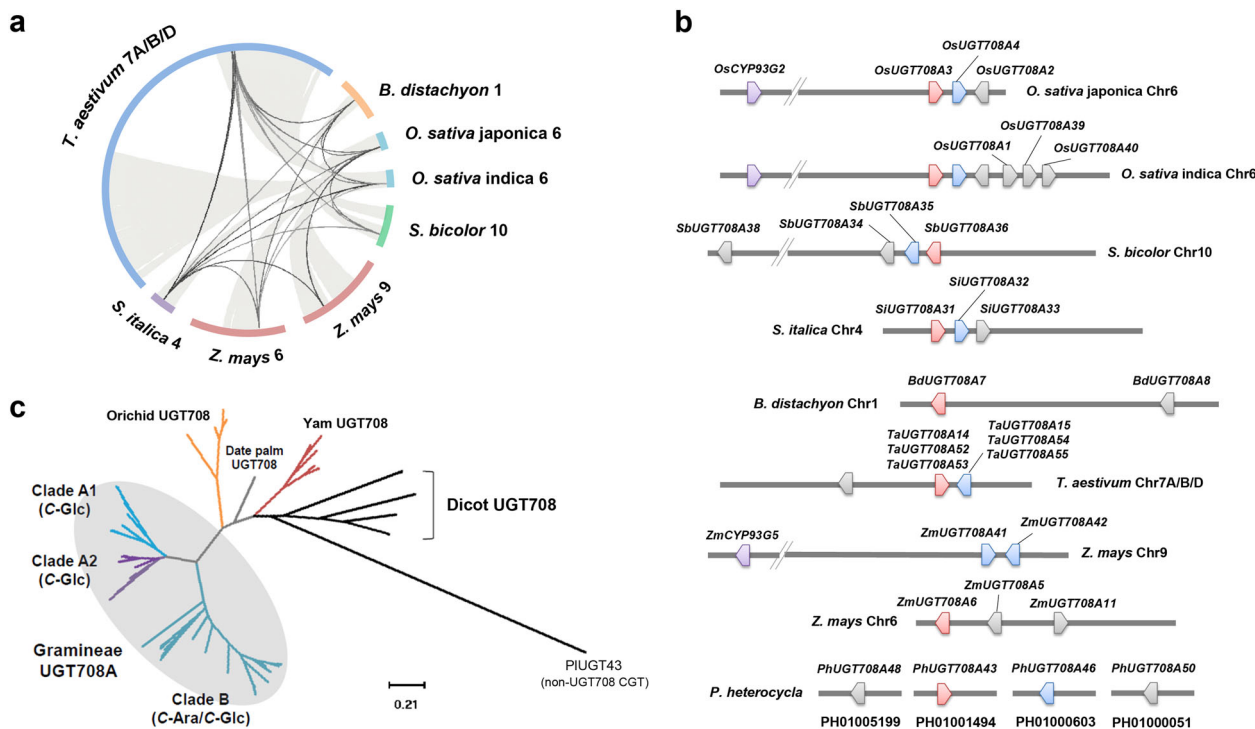
While CGFs are widely spread among monocot plants (Fig. 2a), especially in the leaves of the grass family<sup>28</sup>, a full picture of the CGF diverse spectrum still remains obscure. Here, we applied a liquid chromatography–high resolution–tandem–mass spectrometry (LC–HR–MS/MS) to reveal the profiling of C-glycosylated core structures in the selected Gramineae plants (Fig. 2b). The leaf extracts were treated with hydrochloric acid to remove all O-glycosidic linkages and retain aglycones bearing only C-sugars. Specific *m/z* values corresponding to different C-glycosylated forms of apigenin, luteolin, and chrysoeriol were extracted and calculated (Supplementary Fig. 2 and Supplementary Table 2), which clearly indicated that PACMAD clade crops (mainly in Panicoideae) accumulate a high proportion of monohexosyl flavones (Fig. 2b). Comparatively, the major glycosylated forms in typical BOP clade plants (rice, wheat, barley, and *Brachypodium*) are diglycosides with both C-pentose and C-hexose moieties, for example, schaftoside, carlinoside, and their isoform from *O. sativa*<sup>5</sup>.

Bamboos (*Bambusoideae*) are unique non-timber plants that belong to the BOP clade of the grass family<sup>29,30</sup>. Nevertheless, Moso bamboo (*Phyllostachys heterocyclus* cv. Pubescens, Synonym: *P. edulis*, Mao-zhu) possesses a C-glycoside profile that resembles those of PACMAD clade (Fig. 2b and Supplementary Fig. 2i). Characteristic fragments in MS/MS spectra also confirmed the presence of various CGF compounds in *P. heterocycle* extracts (Supplementary Fig. 3). Further investigation on the other *Phyllostachys* bamboos showed a diverse pattern of C-glycoside accumulation (Supplementary Fig. 4). Like *P. heterocycle*, *P. meyeri* McClure (a dominant species in Anji county, Zhejiang Province, China, Fig. 2c) also accumulates much more monohexosides than diglycosides, while *P. glauca* McClure, *P. prominens* W. Y. Xiong and *P. bissetii* McClure abundantly produce multiglycosylated forms. The content of major C-monoglucosides (Ori, Isoori, Vit, and Isovit) in different tissues of *Phyllostachys* species is visualized in Fig. 2c. We noticed a tendency that the leaf tissue stores most of the CGFs as expected. Among all the tested samples, *P. meyeri* accumulates 30–60 times higher (iso)vitexin and (iso)orientin than other bamboo species.

**Comparative genomics reveals a rich CGT reservoir.** The chemical diversity of CGFs in the grass family inspired us to explore

the genetic resource responsible for diverse CGF biosynthesis. Due to the close relationship and similar metabolite profiles between Gramineae crops and bamboos, we expected that these plants retained common C-glycoside biosynthesis-relevant genomic segments, which could be traced into their evolutionary history. We first targeted the short arm of rice (*Oryza sativa japonica*) chromosome 6, where the known CGF-biosynthesizing OsCGT<sup>15</sup> (OsUGT708A3) and OsF2H<sup>19</sup> (OsCYP93G2) are located. We carried out comparative genomic analyses and the *O. sativa* Chr6 was found collinear with those of sorghum (*Sorghum bicolor*, Chr10), foxtail millet (*Setaria italica*, Chr4), *Brachypodium distachyon* (Chr1), bread wheat (*Triticum aestivum*, Chr7A, 7B, and 7D) and maize (*Zea mays*, Chr6, and Chr9) (Fig. 3a). Multiple blocks linked to *T. aestivum* and *Z. mays* can be explained by their recent or ancient polyploidization events<sup>31,32</sup>. The scaffold PH01001494 of Moso bamboo genome<sup>33</sup> can also be partially aligned to these crops (Supplementary Fig. 5).

Discovered in the identified syntenic genome regions of grass family are tandem UGT-coding genes annotated to belong to UGT708 family (Fig. 3b). The rice cultivar Nipponbare (*O. sativa japonica*) possesses three tandem UGT708 genes, namely OsUGT708A2, OsUGT708A3, and OsUGT708A4. Their paralogues were observed widely existing in *O. sativa indica*, *S. bicolor*, *S. italic*, *B. distachyon*, *T. aestivum*, and *Z. mays* (Fig. 3b). The number of UGT708 copies is doubled in the long-grained rice (*O. sativa indica*) as compared to the japonica variety. By contrast, *B. distachyon* has only two UGT708 genes (*BdUGT708A7* and *BdUGT708A8*). We also observed an inversion of the tandem UGT708 segment on the *S. bicolor* Chr10 (Fig. 3b). In total, 40 monocot UGT708 genes (38 uncharacterized with 2 reported) were discovered by this approach (Supplementary Table 3). In order to gain better insights on the evolution of Gramineae UGT708 proteins, we constructed a maximum likelihood (ML) phylogenetic tree using amino acid sequences of the available UGT708 enzymes from both dicot and monocot species (Fig. 3c, for details, see Supplementary Fig. 6). The Gramineae-originated UGT708s form a monophyletic group (UGT708A), establishing a closer relationship with other monocotyledonous UGT708s from orchid, date palm, and yam (Supplementary Table 4). Notably, we found not only clear differentiation of UGT708 proteins between monocot and dicot species, but also the divergence of grass family-specific UGT708A family. UGT708A family is apparently



**Fig. 3 Identification of C-glycosyltransferases from bamboo-related Gramineae family.** **a** Syntenic blocks of japonica rice chromosome 6 are found across the genome of Gramineae crops, including *B. distachyon*, *O. sativa indica*, *Z. mays*, *S. bicolor*, *S. italica*, and *T. aestivum*. Locations of CGT-containing regions are linked by black lines. **b** Graphical map of tandem UGT708A genes within the syntenic blocks. Clade A1 CGT-encoding genes are indicated in red. Clade A2 CGTs are indicated in light blue. Bifunctional C-arabinosyl/glucosyltransferases in Clade B are indicated in gray color. Flavone-2-hydroxylases on the same chromosomes are indicated in purple. For details, see Supplementary Table 3. **c** **a** Maximum likelihood (ML) phylogenetic tree constructed using PIUGT43 (from *Pueraria lobate*, KU317801.2, as an out-group) and 68 available UGT708 protein sequences. Bar: 0.21 amino acid substitutions per site. Details of the tree are zoomed in Supplementary Fig. 6. C-Glc C-glucosylation, C-Ara C-arabinosylation.

separated into Clade A and Clade B, and Clade A can be subdivided into two branches (A1 and A2, Fig. 3c).

**Characterization of C-glucosyltransferases and C-arabinosyltransferases.** To investigate the putative function of these Gramineae UGT708 enzymes, the open-reading frames of the genes were amplified, cloned, and their proteins expressed. A total of 32 recombinant UGT708s were successfully expressed in *E. coli* (Supplementary Fig. 7) and tested through enzymatic assays. UDP-glucose (UDP-Glc) and UDP-arabinose (UDP-Ara) were used as sugar donors, and phloretin (Phr) and 2-hydroxynaringenin (2OHNar) were selected as potential acceptors (Fig. 4a).

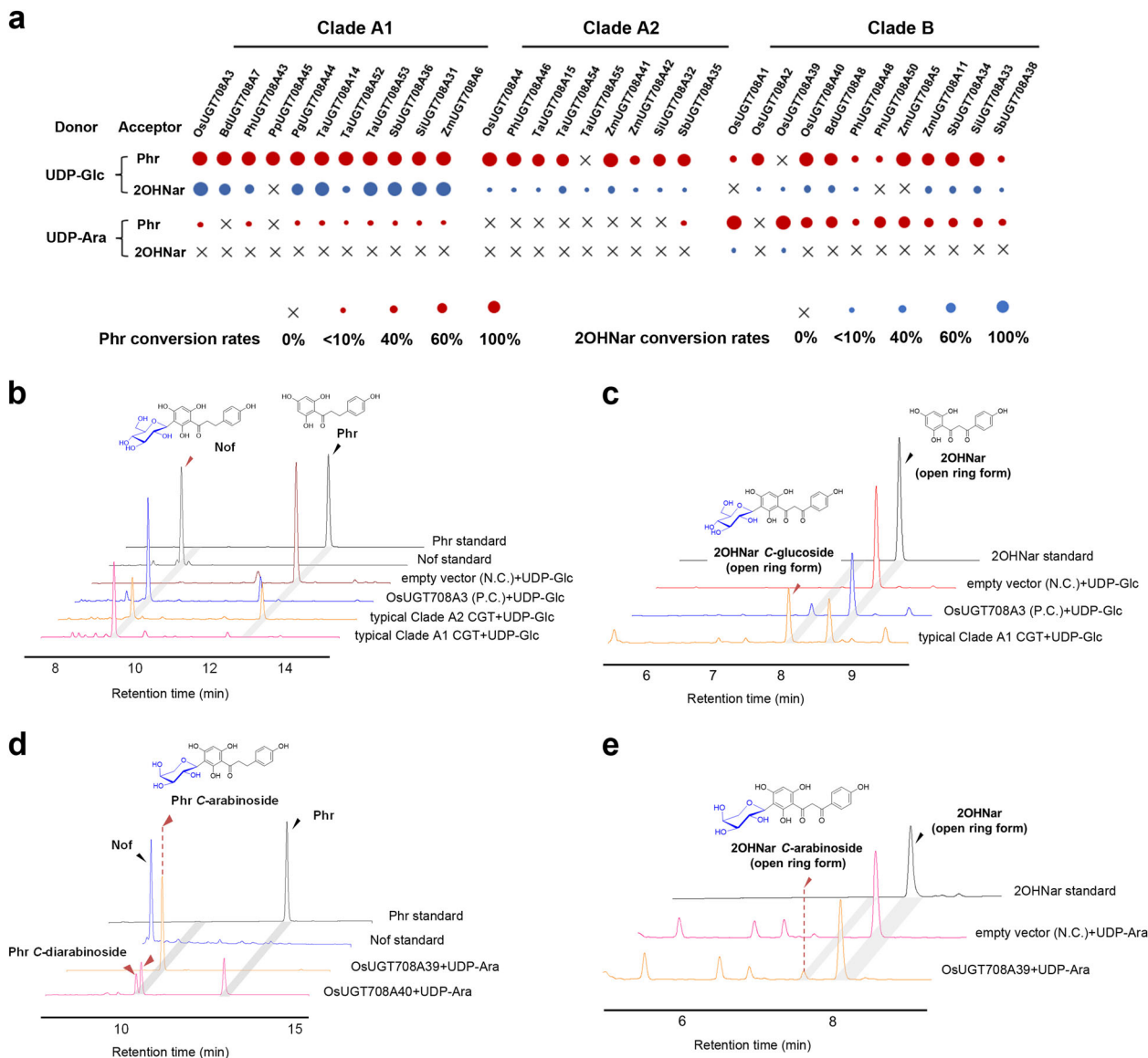
From the HPLC and LC-MS/MS analyses of the reaction product profiles (Fig. 4b and Supplementary Figs. 8, 9, 11, 12), all enzymes from clade A1 were verified as potent C-glucosyltransferases (CGTs) that utilized UDP-Glc to convert both Phr and 2OHNar to their corresponding glycosylated peaks that are identical to the products of positive control (*OsUGT708A3*-mediated reactions) (Fig. 4b, c). No trace of glycosylation was detected in the negative controls (empty vector-mediated reactions). NMR analyses of nothofagin (Nof, Phr 3'-C-glucoside) further verified the C-glucosylation catalyzed by Clade A1 enzymes (Supplementary Fig. 13). Bamboo CGTs orthologous to PhUGT708A43 were also successfully cloned from *P. meyeri*, *P. glauca*, *P. prominens*, and *P. bissetii*. Multiple variants were found in *P. glauca* and *P. prominens*, while other species seem to hold a conserved CGT completely identical to PhUGT708A43 (Supplementary Fig. 14). The  $K_m$  and  $k_{cat}/K_m$  values of PhUGT708A43 towards 2OHNar was measured to be  $1.75 \pm 0.66 \mu\text{M}$  (Supplementary Table 5) and  $4.28 \times 10^4 \text{ s}^{-1} \text{ M}^{-1}$ , which suggested a high affinity and high efficiency. With the exception

of bamboo CGTs and the reported *OsUGT708A3*, other clade A1 proteins generally favor Phr as indicated by their lower  $K_m$  (Supplementary Table 5). On the other hand, most enzymes from Clade A2 just glucosylate Phr to Nof with moderate conversion rate (Fig. 4b). When UDP-Ara was used as sugar donor, only trace of peaks corresponding to C-arabinosyl Phr (when 10 times intensified) were observed in the Clade A1/A2-catalyzed reactions (Supplementary Figs. 8b and 9b), which indicated that Clade A is a group of rigid CGTs that highly addict to UDP-Glc.

Subsequently, Clade B were proved to be promiscuous CGTs that consume both UDP-Ara and UDP-Glc, generating the glucosides and arabinosides of Phr and 2OHNar (Fig. 4d, e and Supplementary Figs. 10–13). Most of the Clade B enzymes tolerate both UDP-sugars (Fig. 4a). Interestingly, *OsUGT708A39* from *O. sativa indica* predominantly recognizes UDP-Ara, while *OsUGT708A2* from *O. sativa japonica* strictly selects UDP-Glc as its donor (Supplementary Fig. 10a, b). *OsUGT708A40*, *ZmUGT708A11*, and *BdUGT708A8* generated an additional product other than C-arabinosyl Phr, which was proposed to be C-di-arabinoside of Phr, as indicated in LC-MS/MS experiments (Supplementary Figs. 10e and 11c). Only *OsUGT708A1* and *OsUGT708A39* were found to slightly arabinosylate 2OHNar detectable in HPLC analyses (Supplementary Fig. 10d).

**Screening of bamboo CYP450 facilitate heterologous C-glycoside production in *E. coli*.** In order to achieve a high-yield heterologous C-glycoside production platform, we took advantage of the fast-growing and easy manipulation characteristics of *E. coli* system to screen Bamboo F2H and F3'H candidates responsible for C-glycoside biosynthesis. The constructed *E. coli* cell factory consisted of a heterologous flavone C-glycoside



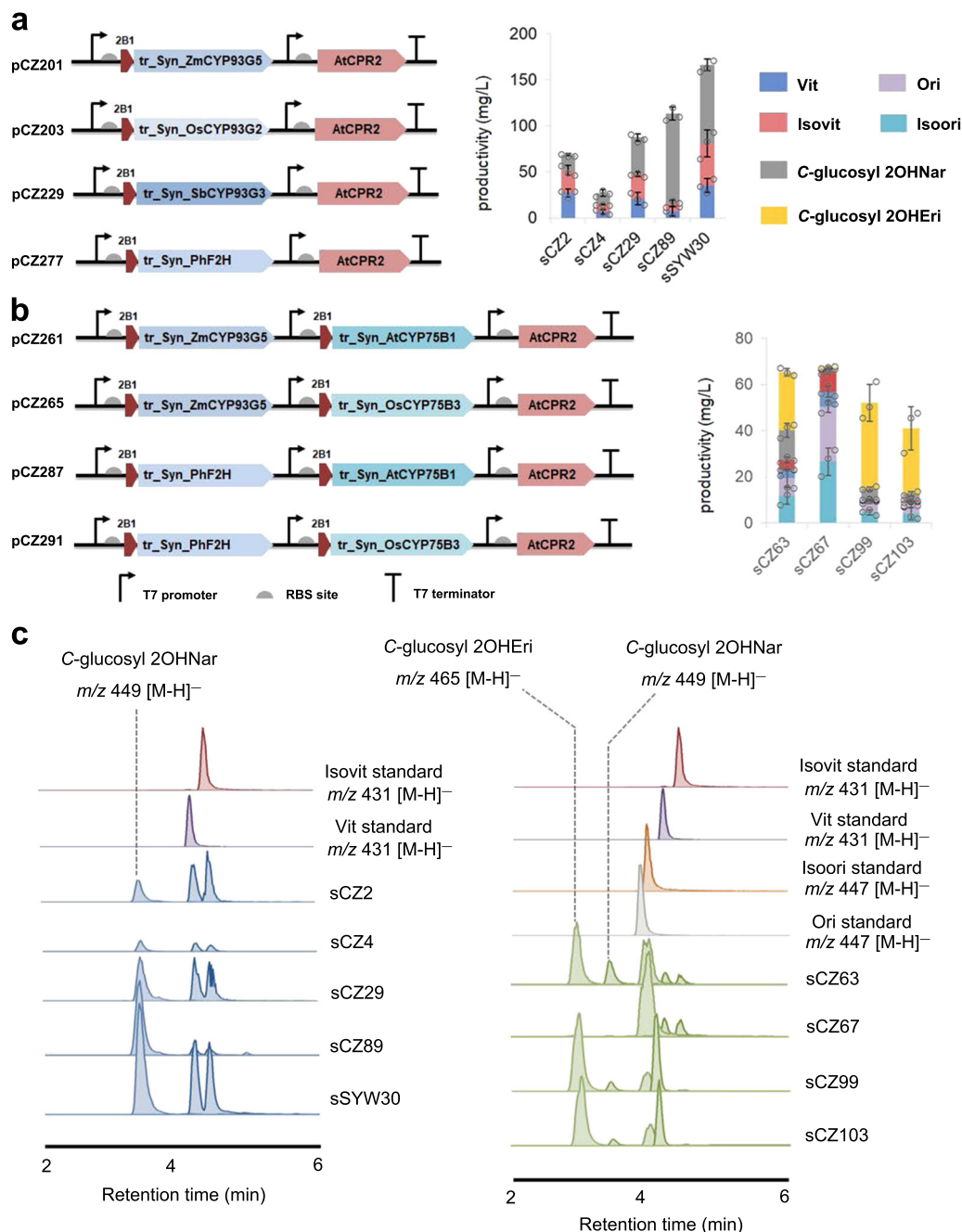


**Fig. 4 Functional characterization of C-glucosyl and C-arabinsyltransferases.** **a** Substrate preference of the Gramineae CGTs. The size of the dots indicate proportionally to the conversion rates of Phr and 2OHNar. Clade A enzymes are found to be rigid C-glucosyltransferases, while Clade B enzymes are promiscuous to different sugar donors (UDP-Glc/UDP-Ara). HPLC chromatograms of the reactions of Clade A CGTs with UDP-Glc and different receptors including phloretin (Phr, **b**) and 2-hydroxynaringenin (2OHNar, **c**). Clade B CGTs are proved to be able to recognize UDP-Ara, generating C-arabinside of Phr **d** and 2OHNar **e**. The absorption at 280 nm was monitored. For detailed HPLC chromatography of each CGTs (see Supplementary Figs. 8–10; For the MS/MS spectra, see Supplementary Figs. 11, 12).

pathway, which can be split into three modules: the Nar precursor module (Fig. 1, gray arrows), the flavonoid CYP450 decoration module (Fig. 1, navy arrows), and the flavonoid C-glycosylation module (Fig. 1, red arrows).

We first utilized a previously assembled pinocembrin-producing strain<sup>34</sup> (sCZ9, Supplementary Table 6) for Nar overproduction. The adopted phenylalanine ammonia lyase (PAL, from *Rhodotorula toruloides*) was proven to tolerate both tyrosine and phenylalanine precursors. To maximize Nar as the precursor of downstream decoration, gradient concentrations of tyrosine were supplemented to boost Nar production to 103 mg/mL (Supplementary Fig. 15). In the next step, three known F2H enzymes from Gramineous crops: the OsF2H (CYP93G2) from rice, the ZmF2H (CYP93G5) from maize, and the SbF2H (CYP93G3) from sorghum were codon optimized, N-terminal modified (transmembrane region truncation and replacement

with 2B1dH-tagged<sup>35</sup>) and tested for their ability to catalyze Nar to 2OHNar in vivo (Fig. 5a). To facilitate the evaluation of validity of the whole pathway, CGT module (PhUGT708A43-expressing plasmid) with confirmative effectiveness was also integrated. When the first three 2B1-infused F2H enzymes coupled with AtCPR2 partner were introduced into *E. coli* respectively, a mixture of Vit and Isovit was detected after a 72 h fermentation (Fig. 5c). An almost complete conversion of Nar was observed from the combination of truncated ZmF2H/AtCPR2 (sCZ2) and truncated SbF2H/AtCPR2 (sCZ29), yielding a highest productivity of 24 mg/L Vit and 27 mg/L Isovit, respectively. Moreover, the accumulation of C-glycosyl 2OHNar was also detected in these strains (Fig. 5c). We then attempted to clone F2H-encoding genes from *P. heterocyclus* and *P. meyeri* that have not been reported previously. The predicted CDS of *P. heterocyclus* CYP93G2-like gene (PH01000832G0680) was redefined through amino acid



**Fig. 5 De novo biosynthesis of flavone C-glycosides in *E. coli* cell factory.** The artificial pathway contains three modules including the naringenin precursor module, the flavonoid CYP450 decoration module, the C-glycosylation module. Details of multispecies genes involved are described in the Supplementary Table 6. **a** Engineered flavone-2-hydroxylases (F2H) from bamboo-related Gramineae family were used to evaluate their ability to produce Vit and Isovit. **b** The complete flavonoid P450 module were further assembled to produce Ori and Isoori. **c** De novo production of flavone C-glycosides. *E. coli* extracts were sampled 72 or 96 h after induction and analyzed by LC-MS. Target compounds were viewed by ion-extracted chromatogram. Standard samples were also analyzed for comparison.

sequence alignment, since it apparently lacks the N terminal when compared to its homologs in rice and maize. The amplified CYP93G genes in *P. heterocycla* and *P. meyeri* (PhF2H/PmF2H) are completely identical and shared 81% identity with the protein sequence of OsCYP93G2. The engineered bamboo CYP93G enzymes showed Vit/Isovit productivity comparable to that of OsCYP93G2, but also accumulated much more C-glycosyl 2OHNar than those of other strains, indicating a great potential of C-glycoside production. When feeding Nar directly into the strain harboring both PhF2H and PhUGT708A43 (sSYW30), Vit and Isovit increased to  $35.1 \pm 7.5$  and  $45.8 \pm 14.9$  mg/L (Fig. 5a).

For the optimization of F3'H, the codon-optimized CYP75B family genes from different species, namely CYP75B1 from *Arabidopsis thaliana*, CYP75B2 from *Petunia X hybrida*, CYP75B3 and CYP75B4 from rice (*O. sativa*) were N-terminal modified, assembled with CPR partners and introduced into *E. coli* to evaluate their production of eriodictyol. Fermentation of strains sCZ47, sCZ49, and sCZ51 (Supplementary Table 6) led to the accumulation of 30–69 mg/L eriodictyol, while CYP75B4-expressing sCZ53 failed to catalyze a 3'-hydroxylation reaction (Supplementary Fig. 16). Four F3'H-like genes were obtained from *P. meyeri*, which are identical to the predicted sequences of

PH0100052G1210, PH01000926G0660, PH01002485G0110, and PH01001730G0080 in *P. heterocyclus*. Among them, only the protein product of PH01001730G0080 (Ph/PmF3'H) exhibit comparable 3'-hydroxylase activity toward Nar with 40.7% conversion rate (Supplementary Fig. 16).

We finally constructed the complete flavonoid P450 module by inserting the most active F3'H-encoding genes (*AtCYP75B1* and *OsCYP75B3*) between F2H (*ZmCYP93G5* or PhF2H) and *AtCPR2* (Fig. 5b), and integrated the constructs into *E. coli* with a Nar module and CGT module. The resulting strains sCZ63, sCZ67, sCZ99, and sCZ103 (Supplementary Table 6) produced a mixture of four flavone C-glycosides. From simple precursors glucose and tyrosine, strain sCZ67 (*ZmCYP93G2-OsCYP75B3-CPR*) produced 26.6 mg/L Isoori and 23.9 mg/L Ori without accumulation of C-glycosylated flavone intermediates (Fig. 5c). The final products were proved by large-scale fermentation, isolation, and NMR-based identification (Supplementary Fig. 17).

**Isoorientin exhibits both neuro-analgesic and anti-inflammatory activities.** To preliminarily screen the antinociceptive effect of C-glycosides, the four major monoglucosides (Ori, Isoori, Vit, and Isovit) were assessed in acetic acid-writhing, hot plate and formalin-induced (first and second phase) nociception assays using ICR male mice. All the four compounds showed significant analgesic activity in acetic acid writhing and formalin-induced nociception tests, but no effects on the latency time on the hot plate (Supplementary Fig. 18). Generally, isoorientin (Isoori) exhibited the best effects among the tested compounds. Isoori (7.5–30 mg/kg) exerted analgesic effect in a dose-dependent manner, and 30 mg/kg Isoori inhibited acetic acid-induced writhing by 86.2%, licking time by 73.9% in the first phase and 48.3% in the second phase of formalin-induced nociception (Fig. 6a). The anti-nociception action of Isoori was equivalent to that of the positive reference drug rotundine (ROT) in the first phase and acetylsalicylic acid (ASA) in the second phase.

Since isoorientin was observed to be effective towards both the first and second phases of formalin-induced nociception, it was considered as a promising agent with both neuro-analgesic and anti-inflammatory activities. Next, we investigated its potential molecular targets through evaluating the effects on some of the neurotransmitter or physical/chemical stimuli-activating receptors. We selected the GluN2B subunit of N-methyl-D-aspartate receptors (NMDARs)<sup>36</sup> and transient receptor potential vanilloid 1 (TRPV1)<sup>37</sup> due to their important roles in pain transmission and modulation. Isoori (30 mg/kg) significantly suppressed TRPV1 protein expression in the spinal dorsal horn of mice and decreased the level of GluN2B in a dose-dependent manner (Fig. 6b). On the other hand, the anti-inflammatory effects of Isoori were indicated by its inhibition of TNF- $\alpha$ , IL-1 $\beta$ , and promotion of anti-hyperalgesic cytokine IL-10 production in formalin-induced mice paw inflammation (Fig. 6c).

Given that Isoori belongs to flavonoids that have been reported to have antioxidant activity<sup>38,39</sup>, its effects on the oxidative stress indicators like glutathione synthetase (GSH), malondialdehyde (MDA), superoxide dismutase (SOD), and ascorbic acid (VC) levels were also investigated. Compared to the vehicle solution, Isoori effectively increased the concentrations of GSH, SOD, and VC in paws injected with formalin, but suppressed formalin stimulated MDA production in a dose-dependent manner (Supplementary Fig. 19)

We also evaluated potential toxicity of Isoori in the same mice model through oral administration of a high dose in an acute toxicity test. After 14 days of a single administration (5000 mg/kg), 4 of the 14 mice were dead, and the living mice did not show any behavioral alterations. The survival curve (Supplementary

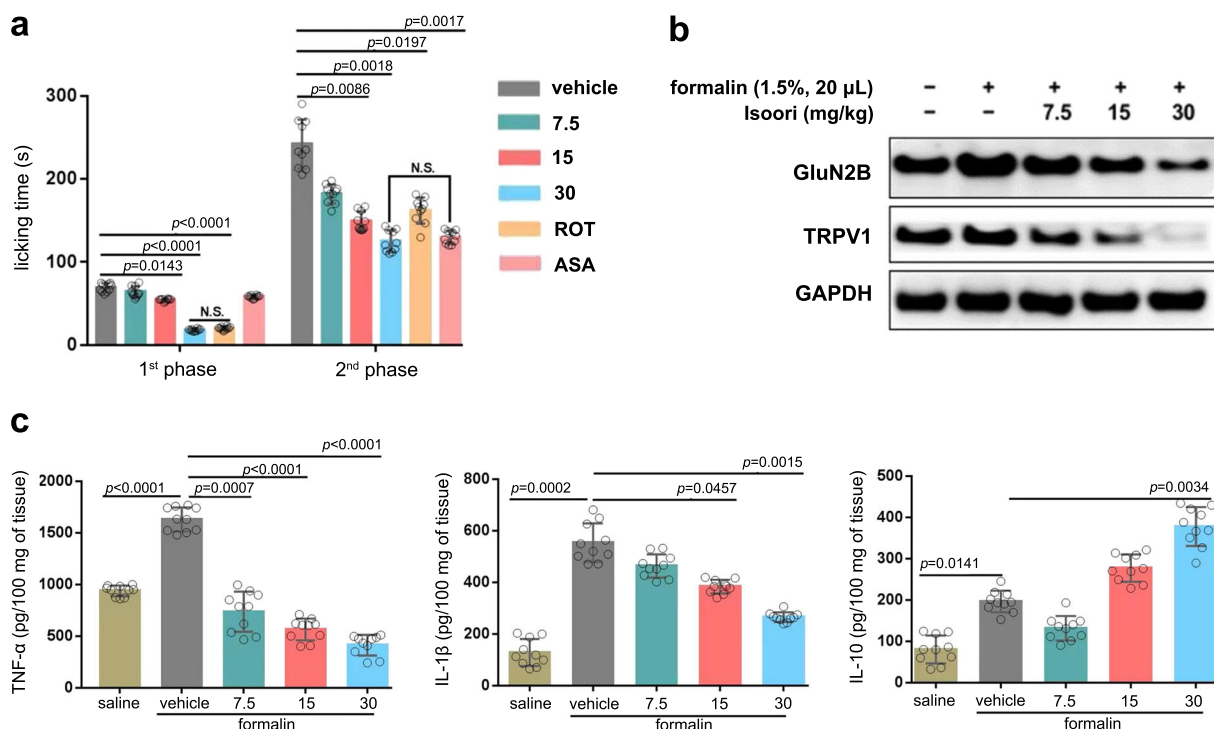
Fig. 20) shows that the LD<sub>50</sub> of Isoori on the ICR mice is more than 5000 mg/kg, which conforms to a low toxicity profile.

## Discussion

In light of the rising opioid crisis, development of natural analgesic drugs has been receiving increasing interest. Plants and microbes have provided humankind with some of the most clinically relevant drugs and secondary metabolites with beneficial effects on various ailments including pain relief. For example, some recent successful cases reported are indole alkaloids from *Nauclea latifolia*<sup>40,41</sup> and *Corydalis yanhusuo*<sup>42</sup>, and diterpene glucosides from *Rhododendron micranthum*<sup>43</sup>. Flavonoids are an important class of phytochemicals ubiquitously existing in the plant kingdom and have members that are considered promising candidates for natural analgesics<sup>44</sup>. To enable full utilization of flavonoids, there is a strong need to uncover more bioresources rich in flavones, and to improve our understanding of the biosynthetic machinery and mechanisms by which flavonoids confer their beneficial effects. Our present work revealed bamboo species as a rich source of CGFs, and elucidation of the biosynthetic pathway uncovered C-glucosyl/C-arabinosyl-transferring UGT708A enzymes. We further demonstrate the construction of the pathway in a heterologous host, and show that major CGFs in bamboo, especially isoorientin, has both neuroanalgesic and anti-inflammatory effects superior to clinical drugs used as positive controls.

Plants have evolved two pathways for glycosylating the flavone scaffold through O-glycosylation and C-glycosylation. Unlike O-glycosylation that terminates flavone biosynthesis, C-glycosylation occurs on unstable intermediates prior to the final scaffold formation. Except for PiUGT43<sup>25</sup>, GtUF6CGT1<sup>26</sup>, and TcCGT1<sup>27</sup>, all characterized CGTs to date belong to the UGT708 family<sup>15–18,22,23</sup>, indicating that the UGT708 family may be specialized for C-glycosylation of phenylpropanoids and flavonoids. Through this study, we have significantly expanded the knowledge on monocot UGT708A family by not only tripling the number of known CGTs from current 12 (2 monocot, 7 dicot UGT708, and 3 non-UGT708 CGTs) to 38, but also by revealing a group of hitherto unknown group that could recognize UDP-Ara (Figs. 3, 4 and Supplementary Table 3). C-arabinosides are unique to the leaves of grass family plants. Arabinosyltransferases of plant origin have been reported to be involved in the arabinoxylan biosynthesis<sup>45–47</sup> and in post-translational modification of glycoproteins<sup>48</sup>, but UDP-arabinosyltransferases decorating secondary metabolites are hardly reported with just two cases<sup>49,50</sup> (identified as O-arabinosyltransferases). Recent genetic-based studies on rice and maize metabolism have implied some gene candidates for C-arabinosylation<sup>51–54</sup>, but none have been functionally characterized to date. The Clade B CGTs found in this research are promiscuous C-arabinosyl/glucosyltransferases, which is considered as a group of specialized CGTs evolving for more diverse recognition of sugar donors. The grass family plant may take the advantage of these potent Clade B CGTs to diversify their C-glycoside spectrum, producing abundant C-arabinosylated flavone compounds.

Screening of P450 oxidases in this research revealed bamboo flavanone 2-hydroxylases (Ph/PmF2H) and flavanone 3'-hydroxylases (Ph/PmF3'H). The F2Hs from bamboo and crops are CGF pathway-specific enzymes that result in the formation of unstable chalconoid intermediates readily dehydrated to flavones<sup>19</sup>. In our in vivo *E. coli*-based cell factory, instead of 2-hydroxylated compounds, a great accumulation of C-glycosylated 2-hydroxynarigenin or 2-hydroxyeriodictyol was observed (up to 97 mg/L, Fig. 5). This phenomenon was not mentioned in the



**Fig. 6 Antinociceptive effects of Isoorientin.** **a** Effects of Isoori on the first and second phases of formalin-induced nociception in mice were tested in different concentrations (7.5, 15, 30 mg/kg, i.p.). Reference drug [acetylsalicylic acid (ASA, 200 mg/kg, p.o.), rotundine (ROT, 20 mg/kg, i.p.)] was administered to mice. Details are described in “Methods” section. Each column represents mean  $\pm$  SEM ( $n = 10$ ).  $P < 0.05$  was considered as statistically significant. **b** Isoori suppressed the synaptic accumulation of NMDA receptor GluN2B subunits and TRPV1 expression levels in the spinal dorsal horn of mice. The level of GluN2B and TRPV1 were checked using immunoblotting analysis with GAPDH as internal reference. Experiments were repeated three times with similar results. **c** Effects of Isoori (7.5, 15, 30 mg/kg) on TNF- $\alpha$  (left), IL-1 $\beta$  (mid), and IL-10 (right) production in formalin paw inflammation. Drug administration were described in “Methods” section. TNF- $\alpha$  tumor necrosis factor- $\alpha$ , IL-1 $\beta$  interleukin-1 $\beta$ , IL-10 interleukin-10.

prior reports of metabolic engineering on CGF biosynthesis in yeast systems<sup>55,56</sup>, which indicated that *E. coli* probably lacked endogenous dehydratases active on 2-hydroxylated flavanone. The successful bio-production of (iso)orientin and (iso)vitexin here was achieved through synthetic approach, which involves multiple bioparts from bamboo, maize, rice, and mouse-ear cress (*A. thaliana*). The ratio of 6-C- and 8-C- products is generally 6:4 that likely reflect the result of a non-enzymatic reaction.

The present study also demonstrates that the four CGFs dominant in leaves of bamboos and crops showed significant antinociceptive and anti-inflammatory effects (Fig. 6). Isoorientin exhibited the best analgesic activity in both the first and second phase of the formalin assay, displaying anti-nociceptive activity equivalent to the reference drug ROT, a dopamine receptor antagonist in the first phase, and ASA, an anti-inflammatory agent in the second phase (Fig. 6a). This was interesting, given that first phase pain responses are mainly neurogenic pain response<sup>57</sup>, while second phase pain responses are mediated through inflammatory mechanisms<sup>58</sup>. Opioid analgesics are known to exert their antinociception effects on phase I and phase II, while non-steroidal anti-inflammatory drugs (NSAIDs) only inhibit phase II pain alleviation. A previous study showed that Vit exhibits protection against isoflurane-induced increase in expression of GluN2B and TRPV1 in PC12 cells<sup>59</sup>. In the present work, in vivo evidence of suppression of TRPV1 and NR2B by Isoorientin in the spinal dorsal horn was provided. Compared to aforementioned examples and the mechanisms of most clinical compounds, Isoori is herein reported as a highly effective compound to exhibit both analgesic and anti-inflammatory activity which are equivalent to the clinical commonly used drugs, respectively.

Taken together, this study revealed the normally underutilized leaves of bamboos and other cereal crops as a rich source of CGFs. We leveraged the availability of sequenced plant genomes to exhaustively search CGTs and shed more light on the metabolic diversity of flavonoids in plants. In addition to the evolutionary relationships of CGTs unraveled here, this work tripled the number of known CGTs by successful characterization of specific CGTs with glucosyl and arabinosyl-transferring activities. Crucial insights in the biosynthesis of CGF in bamboos are presented, together with a demonstration of the potential to utilize these availed genomic resources in metabolic engineering in *E. coli*. Finally, we provide insights on the anti-nociceptive and anti-inflammatory activity of Isoorientin in in vivo models.

## Methods

**Extraction and profiling of flavone C-glycosides.** The frozen leaf samples (100 mg) were crushed and extracted overnight at room temperature with 5 mL 80% methanol. The crude extracts were filtered, concentrated in vacuo to 0.5 mL aqueous solution and treated with 1 M hydrochloric acid (HCl) in boiling water for 6 h. The acid-treated samples were extracted three times with *n*-butanol. The combined butanol extracts were further evaporated to dryness and re-dissolved in 1 mL 50% HPLC grade methanol before LC-MS/MS analysis.

Ultra-performance liquid chromatography (UPLC)-electrospray ionization (ESI)-high resolution (HR) tandem mass spectrometry (MS/MS) data were acquired using Q Exactive hybrid quadrupole-Orbitrap mass spectrometer (Thermo Scientific, USA) equipped with an Acquity UPLC BEH C18 column ( $\varnothing$  2.1  $\times$  50 mm, 1.7  $\mu$ m, Waters, USA). The mobile phase consisted of methanol (0.1% formic acid, solvent A) and H<sub>2</sub>O (solvent B). A linear gradient was set as follows: 0.0–10.0 min, 5–100% A in B; 10.0–12.5 min, 100% A; 12.5–15.0, re-equilibrate to the initial condition. The flow was 0.25 mL/min. The mass acquisition was performed in negative ionization mode with full scan (50–1000). The results are shown in Fig. 2 and Supplementary Fig. 2.



**Bioinformatics analysis.** Genomes of the grass family plants were downloaded from Ensembl Plants. Multiple collinearity of subgenomes was detected by MCScanx<sup>60</sup> with default parameters and plotted by circle plotter (Fig. 3a). The Gramineae *UGT708*-encoding genes were identified by Plant Compara within Ensembl Plants. Genome or transcriptome of banana (*Musa acuminata*), date palm (*Phoenix dactylifera*), yam (three *Dioscorea* species), and orchid (*Phalaenopsis equestris*) (Supplementary Table 4) were BLASTed for *UGT708*-encoding genes. Sequences longer than 300 amino acids were selected and subjected into phylogenetic tree construction, along with Gramineae *UGT708A* and dicot *UGT708* enzymes. The phylogenetic trees of *UGT708* proteins were generated by Mega X using ML method with bootstrap analyses of 1000 replications (Fig. 3c and Supplementary Fig. 6). The multiple alignments of amino acid sequences were generated by using Clustal Omega program provided by the European Molecular Biology Institute.

**In vitro enzymatic assay and kinetic studies of CGTs.** A total of 32 recombinant CGTs were successfully cloned using gene-specific primers (Supplementary Table 7) and expressed in *E. coli* (see Supplementary Method). A typical enzymatic assay was performed in a 100  $\mu$ L aliquot of reaction mixture containing buffer A (100 mM NaCl, 20 mM Tris-HCl, pH 8.0), 400  $\mu$ M UDP-glucose or 200  $\mu$ M UDP-Ara, 100  $\mu$ M chalconoid acceptors [Phr, 2-hydroxynaringenin (2OHNar)], 25  $\mu$ L crude (or 5  $\mu$ g purified) enzymes. The reaction mixtures were incubated at 37 °C for 2 h and finally quenched by 100  $\mu$ L methanol and vortexed vigorously. The kinetic parameters of CGTs were determined in a 50  $\mu$ L reaction solution consisting of buffer A, 1–10  $\mu$ g each purified enzyme and 1–100  $\mu$ M Phr and 2OHNar. The concentration of UDP-glucose was fixed at 400  $\mu$ M. The reactions were incubated for 20 min at 37 °C and stopped by the addition of 50  $\mu$ L methanol.

The quenched reaction mixtures were centrifuged (12,000 rpm) for 15 min and subjected to HPLC analyses (Dionex UltiMate 3000SD HPLC system, Thermo Scientific, USA). Separation was achieved on a C18 column [SilGreen ODS column ( $\varnothing$  4.6  $\times$  250 mm, S-5  $\mu$ M), Greenherbs Co., Ltd., Beijing, China] with a flowrate of 1 mL/min at 40 °C. Detection was performed with mobile phases containing acetonitrile (0.1% formic acid, solvent A) and H<sub>2</sub>O (0.1% formic acid, solvent B) under a linear gradient elution: 0–20 min, 5–100% A in B, 100% A maintained for 5 min. The injection volume was 20  $\mu$ L and absorption was monitored at  $\lambda$  = 280 nm (Fig. 4). The enzyme kinetic parameters were calculated by quantification of the formation of glucoside products. Michaelis–Menten curves were generated by GraphPad Prism 5.

**Cloning of flavonoid hydroxylases, construction of heterologous CGF synthetic pathway, and de novo production.** The candidate F2Hs from various plants (OsCYP93G2 XM\_015787468.2, SbCYP93G3 XM\_002461241.2, ZmCYP93G5 XM008661918.2, and Ph/PmF2H) and F3'Hs (AtCYP75B1 AF271651.1, PxCYP75B2 AF155332.1, OsCYP75B3 XM\_015757555.2, OsCYP75B4 XM\_015757228.2, and Ph/PmF3'H) were codon-optimized and synthesized by GenScript Biotech (Nanjing, China). Their N-terminal transmembrane helices were predicted through TMHMM Server v.2.0, truncated and replaced with 2B1dH tag. The engineered F2Hs, F3'Hs were assembled with AtCPR2 (NM\_179141.2) into pETDuet-1 by use of ClonExpress II One Step Cloning Kit (Vazyme, Nanjing, China) (Supplementary Table 6, for primers, see Supplementary Table 7). For the construction of dual flavonoid hydroxylases cassette, the engineered AtCYP75B1 and OsCYP75B3 were, respectively, inserted into *Bam*HI/*Not*I-double digested pCZ201 and pCZ277 through one-step cloning, resulting in pCZ261, pCZ265, pCZ287, and pCZ291 (Supplementary Table 6).

The plasmid pYH055<sup>34</sup> with spectinomycin resistance was transformed into the *E. coli* BL21(DE3) (designed as strain sCZ9) for the Nar production. The plasmid pCZ86 (pET28a-PhUGT708A43) with kanamycin resistance was chosen for the CGT module. For the de novo production of CGFs from tyrosine, the assembled flavonoid hydroxylases and CPR cassettes were cotransferred with pYH055 and pCZ86 into *E. coli* BL21(DE3). The seeds were precultured at 37 °C in LB medium overnight and then inoculated (1:100) into MOPS minimal medium supplemented with 5 g/L glucose and 2.5 g/L L-tyrosine. After the OD<sub>600</sub> reached to 1.0, IPTG (0.1 mM) was added to the cultures. Subsequently, the cultures were incubated at 22 °C and 220 rpm for 3–4 days. Aliquots of the cultures (500  $\mu$ L) were taken every day for monitoring the target products and subsequently supplemented with 0.5% glucose. The samples were extracted by 500  $\mu$ L *n*-butanol for three times. The combined supernatant was evaporated under vacuum and dissolved in 100  $\mu$ L methanol; 2  $\mu$ L was injected for UPLC–MS/MS analysis (Fig. 5c). Condition of LC–MS/MS was identical to that described above.

For the production of CGFs from Nar, only bamboo P450 (pCZ277) and CGT (pCZ86) module were incorporated to give strain sYW30. Nar (25 mg/L) was supplemented into the medium for every 24 h for four times. The extraction and chromatographic analyses were performed in the way as the samples of de novo production from tyrosine.

**Antinociceptive assay.** The acetic acid-induced test was carried out according to the method described previously<sup>61</sup>. Ori, Vit, Isovit (30 mg/kg, i.p.) and isoorientin (7.5, 15, 30 mg/kg, i.p.) prepared above or vehicle (5% DMSO in normal saline, i.p.) was administered to mice 1 h before intraperitoneal (i.p.) injection of acetic acid

(0.6% in distilled water, 10 mL/kg). ASA (200 mg/kg, p.o.) was employed as reference drug, which was dissolved in the same vehicle. The number of writhes was counted for 15 min. The hot-plate test measured response latencies according to the method described by Eddy and Leimbach<sup>62</sup>. Animals were placed on an PE34 (IITC, Life Science Instrument, USA) hot-plate maintained at 55  $\pm$  1 °C and the time between placement of the animal on the hot-plate and the occurrence of either the licking of the hind paws, shaking, or jump off from the surface was recorded as response latency. Mice with baseline latencies of <5 s or more than 30 s were eliminated from the study 24 h previously. Animals were treated with different compounds (Ori, Vit, Isovit, 30 mg/kg; isoorientin, 7.5, 15, 30 mg/kg, i.p.) or morphine (5 mg/kg, i.p.) 1 h before the experiments. Control animals received the same treatment as mentioned in the abdominal writhing test. The formalin test was similar to that described previously<sup>58</sup>. Twenty microlitres of 1.5% formalin were injected subcutaneously into the right hind paw of mice. The time (in seconds) spent licking and biting the injected paw was taken as an indicator of nociception response. Responses were measured for 0–10 min (first phase) and 10–60 min (second phase) after formalin injection, respectively. Ori, Vit, Isovit (30 mg/kg, i.p.), isoorientin (7.5, 15, and 30 mg/kg, i.p.), ROT (20 mg/kg, i.p.), or ASA (200 mg/kg, p.o.) was administered 1 h before formalin injection and control animals received the same volume of vehicle (0.5% DMSO in distilled water). All experiments were approved by the Institutional Animal Care and Use Committee (IACUC) of Shanghai Ocean University and were performed in accordance with national and institutional guidelines for the care and use of laboratory animals.

**Western blot.** In another cohort of animals with identical treatment, at 1 h after the injection of formalin when the pain sensitivity reached its peak, both ipsilateral and contralateral dorsal horns of lumbar spinal cord (from segments L3 to L5) were dissected after a laminectomy under deep anesthesia of formalin. The dorsal horn sample was homogenized in ice-cold Laemmli buffer, containing 50 mM Tris-HCl, pH 7.5, 0.5% sodium dodecyl sulfate (SDS), 5% 2-mercaptoethanol, and 1% protease inhibitor cocktail (Millipore Sigma, St. Louis, MO, USA). The proteins were separated on SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and then transferred to polyvinylidene difluoride membranes. Membranes were then incubated with different primary antibodies including GluN2B (Cell Signaling Technology, Danvers, USA) with a 1:1000 dilution, GAPDH (Proteintech, USA) with a 1:5000 dilution, Trpv1 (Abcam, England) using a 1:500 dilution overnight at 4 °C. Membranes were then incubated with goat anti-rabbit IgG (Beyotime Biotechnology, China) or goat anti-mouse IgG using a 1:1000 dilution for 1 h at room temperature. Proteins were detected by BCA protein quantification kit (Sangon biotech, Shanghai, China) and visualized on X-ray films. Densitometry was analyzed by ImageJ software (National Institutes of Health, Bethesda, MD, USA). The density of specific bands of Western blot was then normalized to corresponding loading control bands.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The gene sequences of *Phyllostachys* bamboos were deposited in GenBank under the following accession numbers: PhUGT708A43 (PhCGT1), MK616588; PhUGT708A46 (PhCGT2), MK616589; PhUGT708A48 (PhCGT3), MK616590; PhUGT708A50 (PhCGT4), MK616591; PhUGT708A44 (PpCGT1), MK616592; PpUGT708A45 (PpCGT1), MK616593; PhF2H, MK628906; PhF3'H, MK636710. Any other data not present in the manuscript or supplementary materials is available from the authors upon reasonable request.

Received: 13 July 2019; Accepted: 12 February 2020;

Published online: 06 March 2020

## References

1. Gorzalczany, S., Marrassini, C., Mino, J., Acevedo, C. & Ferraro, G. Antinociceptive activity of ethanolic extract and isolated compounds of *Urtica circularis*. *J. Ethnopharmacol.* **134**, 733–738 (2011).
2. Negri, G., Mattei, R. & Mendes, F. R. Antinociceptive activity of the HPLC- and MS-standardized hydroethanolic extract of *Pterodon emarginatus* Vogel leaves. *Phytotherapy* **21**, 1062–1069 (2014).
3. Borghi, S. M. et al. Vitexin inhibits inflammatory pain in mice by targeting TRPV1, oxidative stress, and cytokines. *J. Nat. Prod.* **76**, 1141–1149 (2013).
4. Zhu, Q. et al. Antinociceptive effects of vitexin in a mouse model of postoperative pain. *Sci. Rep.* **6**, 19266 (2016).
5. Besson, E. et al. C-glycosylflavones from *Oryza sativa*. *Phytochemistry* **24**, 1061–1064 (1985).
6. Yang, Z. G. et al. Toward better annotation in plant metabolomics: isolation and structure elucidation of 36 specialized metabolites from *Oryza sativa* (rice) by using MS/MS and NMR analyses. *Metabolomics* **10**, 543–555 (2014).

7. Wojakowska, A., Perkowski, J., Goral, T. & Stobiecki, M. Structural characterization of flavonoid glycosides from leaves of wheat (*Triticum aestivum* L.) using LC/MS/MS profiling of the target compounds. *J. Mass Spectrom.* **48**, 329–339 (2013).
8. Elliger, C. A., Chan, B. G., Waiss, A. C., Lundin, R. E. & Haddon, W. F. C-glycosylflavones from *Zea mays* that inhibit insect development. *Phytochemistry* **19**, 293–297 (1980).
9. Norbaek, R., Brandt, K. & Kondo, T. Identification of flavone C-glycosides including a new flavonoid chromophore from barley leaves (*Hordeum vulgare* L.) by improved NMR techniques. *J. Agric. Food Chem.* **48**, 1703–1707 (2000).
10. Colombo, R., Yariwake, J. H., Queiroz, E. F., Ndjoko, K. & Hostettmann, K. On-line identification of further flavone C- and O-glycosides from sugarcane (*Soccharum officinarum* L., Gramineae) by HPLC–UV–MS. *Phytochem. Anal.* **17**, 337–343 (2006).
11. Coutinho, I. D. et al. Metabolite profiling of sugarcane genotypes and identification of flavonoid glycosides and phenolic acids. *J. Agric. Food Chem.* **64**, 4198–4206 (2016).
12. Nirmala, C., Bisht, M. S., Bajwa, H. K. & Santosh, O. Bamboo: a rich source of natural antioxidants and its applications in the food and pharmaceutical industry. *Trends Food Sci. Technol.* **77**, 91–99 (2018).
13. Hoyweghen, L. V., Beer, T. D., Deforce, D. & Heyerick, A. Phenolic compounds and anti-oxidant capacity of twelve morphologically heterogeneous bamboo species. *Phytochem. Anal.* **23**, 433–443 (2012).
14. Panee, J. Potential medicinal application and toxicity evaluation of extracts from bamboo plants. *J. Med. Plant Res.* **9**, 681–692 (2015).
15. Brazier-Hicks, M. et al. The C-glycosylation of flavonoids in cereals. *J. Biol. Chem.* **284**, 17926–17934 (2009).
16. Ferreyra, M. L. F. et al. Identification of a bifunctional maize C- and O-glycosyltransferase. *J. Biol. Chem.* **288**, 31678–31688 (2013).
17. Nagatomo, Y. et al. Purification, molecular cloning and functional characterization of flavonoid C-glycosyltransferases from *Fagopyrum esculentum* M (*buckwheat*) *cotyledon*. *Plant J.* **80**, 437–448 (2014).
18. Ito, T., Fujimoto, S., Suito, F., Shimosaka, M. & Taguchi, G. C-glycosyltransferases catalyzing the formation of di-C-glycosyl flavonoids in citrus plants. *Plant J.* **91**, 187–198 (2017).
19. Du, Y., Chu, H., Chu, I. K. & Lo, C. CYP93G2 is a flavanone 2-hydroxylase required for C-glycosylflavone biosynthesis in rice. *Plant Physiol.* **154**, 324–333 (2010).
20. Morohashi, K. et al. A genome-wide regulatory framework identifies maize *pericarp color1* controlled genes. *Plant Cell* **24**, 2745–2764 (2012).
21. Casas, M. I. et al. Identification and characterization of maize salmon silks genes involved in insecticidal maysin biosynthesis. *Plant Cell* **28**, 1984–1987 (2016).
22. Chen, D. W. et al. Probing the catalytic promiscuity of a regio- and stereospecific C-glycosyltransferase from *Mangifera indica*. *Angew. Chem. Int. Ed.* **54**, 12678–12682 (2015).
23. Hirade, Y. et al. Identification and functional analysis of 2-hydroxyflavanone C-glycosyltransferase in soybean (*Glycine max*). *FEBS Lett.* **589**, 1778–1786 (2015).
24. Du, Y., Chu, H., Wang, M., Chu, I. K. & Lo, C. Identification of flavone phytoalexins and a pathogen-inducible flavone synthase II gene (*SbFNSII*) in sorghum. *J. Exp. Bot.* **61**, 983–994 (2010).
25. Wang, X., Li, C., Zhou, C., Li, J. & Zhang, Y. Molecular characterization of the C-glycosylation for puerarin biosynthesis in *Pueraria lobata*. *Plant J.* **90**, 535–546 (2017).
26. Sasaki, N. et al. Identification of the glycosyltransferase that mediates direct flavone C-glycosylation in *Gentiana triflora*. *FEBS Lett.* **589**, 182–187 (2015).
27. He, J.-B. et al. Molecular and structural characterization of a promiscuous C-glycosyltransferase from *Trollius chinensis*. *Angew. Chem. Int. Ed.* **131**, 11637–11644 (2019).
28. Harborne, J. B. & Williams, C. A. Flavonoid patterns in leaves of the Gramineae. *Biochem. Syst. Ecol.* **4**, 267–280 (1976).
29. Clark, L. G., Zhang, W. P. & Wendel, J. F. A phylogeny of the grass family (Poaceae) based on *ndhF*-sequence data. *Syst. Bot.* **20**, 436–460 (1995).
30. Zhang, W. P. Phylogeny of the grass family (Poaceae) from *rpl16* intron sequence data. *Mol. Phylogenet. Evol.* **15**, 135–146 (2000).
31. Swigonova, Z. et al. Close split of sorghum and maize genome progenitors. *Genome Res.* **14**, 1916–1923 (2004).
32. Marcussen, T. et al. Ancient hybridizations among the ancestral genomes of bread wheat. *Science* **345**, 1250092 (2014).
33. Peng, Z. et al. The draft genome of the fast-growing non-timber forest species moso bamboo (*Phyllostachys heterocycla*). *Nat. Genet.* **45**, 456–461 (2013).
34. Li, J. et al. Production of plant-specific flavones baicalein and scutellarein in an engineered *E. coli* from available phenylalanine and tyrosine. *Metab. Eng.* **52**, 124–133 (2019).
35. Scott, E. E., Spatzenegger, M. & Halpert, J. R. A truncation of 2B subfamily cytochromes P450 yields increased expression levels, increased solubility, and decreased aggregation while retaining function. *Arch. Biochem. Biophys.* **395**, 57–68 (2001).
36. McRoberts, J. A. et al. Selective knockdown of NMDA receptors in primary afferent neurons decreases pain during phase 2 of the formalin test. *Neurosci* **172**, 474–482 (2011).
37. Basbaum, A. I., Bautista, D. M., Scherrer, G. & Julius, D. Cellular and molecular mechanisms of pain. *Cell* **139**, 267–284 (2009).
38. Long, X. et al. Flavonoids composition and antioxidant potential assessment of extracts from Gannanzao Navel Orange (*Citrus sinensis* Osbeck Cv. Gannanzao) peel. *Nat. Prod. Res.* 1–5, <https://doi.org/10.1080/14786419.2019.1593162> (2019).
39. Chen, G. L., Fan, M. X., Wu, J. L., Li, N. & Guo, M. Q. Antioxidant and anti-inflammatory properties of flavonoids from *Lotus plumule*. *Food Chem.* **277**, 706–712 (2019).
40. Boumendjel, A. et al. Occurrence of the synthetic analgesic tramadol in an African medicinal plant. *Angew. Chem. Int. Ed.* **52**, 11780–11784 (2013).
41. Boucherle, B. et al. *Nauclea latifolia*: biological activity and alkaloid phytochemistry of a west African tree. *Nat. Prod. Rep.* **33**, 1034–1043 (2016).
42. Wu, L. et al. Identification of alkaloids from *Corydalis yanhusuo* W. T. Wang as dopamine D-1 receptor antagonists by using CRE-luciferase reporter gene assay. *Molecules* **23**, 2585 (2018).
43. Sun, N. et al. Rhodomicranosides A–I, analgesic diterpene glucosides with diverse carbon skeletons from *Rhododendron micranthum*. *Phytochemistry* **158**, 1–12 (2019).
44. Xiao, X., Wang, X., Gui, X., Chen, L. & Huang, B. Natural flavonoids as promising analgesic candidates: a systematic review. *Chem. Biodivers.* **13**, 1427–1440 (2016).
45. Porchia, A. C., Sorensen, S. O. & Scheller, H. V. Arabinoxylan biosynthesis in wheat. Characterization of arabinosyltransferase activity in Golgi membranes. *Plant Physiol.* **130**, 432–441 (2002).
46. Harholt, J. et al. ARABINAN DEFICIENT 1 is a putative arabinosyltransferase involved in biosynthesis of pectic arabinan in *Arabidopsis*. *Plant Physiol.* **140**, 49–58 (2006).
47. Schultink, A., Cheng, K., Park, Y. B., Cosgrove, D. J. & Pauly, M. The identification of two arabinosyltransferases from tomato reveals functional equivalency of xyloglucan side chain substituents. *Plant Physiol.* **163**, 86–94 (2013).
48. Ogawa-Ohnishi, M., Matsushita, W. & Matsubayashi, Y. Identification of three hydroxyproline O-arabinosyltransferases in *Arabidopsis thaliana*. *Nat. Chem. Biol.* **9**, 726–732 (2013).
49. Corcuera, L. J. & Bandurski, R. S. Biosynthesis of indol-3-yl-acetyl-myoinositol arabinoside in kernels of *Zea mays* L. *Plant Physiol.* **70**, 1664–1666 (1982).
50. Louveau, T. et al. Analysis of two new arabinosyltransferases belonging to the carbohydrate-active enzyme (CAZY) glycosyl transferase family1 provides insights into disease resistance and sugar donor specificity. *Plant Cell* **30**, 3038–3057 (2018).
51. Chen, W. et al. Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nat. Genet.* **46**, 714–721 (2014).
52. Wen, W. et al. Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. *Nat. Commun.* **5**, 4438 (2014).
53. Dong, X. K. et al. Comprehensive profiling and natural variation of flavonoids in rice. *J. Integr. Plant Biol.* **56**, 876–886 (2014).
54. Matsuda, F. et al. Metabolome-genome-wide association study dissects genetic architecture for generating natural variation in rice secondary metabolism. *Plant J.* **81**, 13–23 (2015).
55. García Vanegas, K. et al. Indirect and direct routes to C-glycosylated flavones in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **17**, 107 (2018).
56. Brazier-Hicks, M. & Edwards, R. Metabolic engineering of the flavone-C-glycoside pathway using polyprotein technology. *Metab. Eng.* **16**, 11–20 (2013).
57. Abdul-Wahab, I. R., Guilhon, C. C., Fernandes, P. D. & Boylan, F. Antinociceptive activity of *Pereskia bleo* Kunth. (Cactaceae) leaves extracts. *J. Ethnopharmacol.* **144**, 741–746 (2012).
58. Rabelo, A. S. et al. Antinociceptive, anti-inflammatory and antioxidant activities of aqueous extract from *Remirea maritima* (Cyperaceae). *J. Ethnopharmacol.* **145**, 11–17 (2013).
59. Chen, L., Zhang, B., Shan, S. & Zhao, X. Neuroprotective effects of vitexin against isoflurane-induced neurotoxicity by targeting the TRPV1 and NR2B signaling pathways. *Mol. Med. Rep.* **14**, 5607–5613 (2016).
60. Wang, Y. et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **40**, e49 (2012).
61. Koster, R., Anderson, M. & Debeer, E. J. Acetic acid for analgesic screening. *Fed. Proc.* **18**, 412–417 (1959).
62. Eddy, N. B. & Leimbach, D. Synthetic analgesics. 2. Dithienylbutylamines and dithienylbutylamines. *J. Pharmacol. Exp. Ther.* **107**, 385–393 (1953).

## Acknowledgements

This work was financially supported by the National Key R&D Program of China (2018YFA0900600), the Strategic Priority Research Program “Molecular mechanism of Plant Growth and Development” of CAS (XDB27020202, XDB27020103), the National Natural Science Foundation of China (Grant nos. 31670099, 31700261). This work was also financially supported by the Construction of the Registry and Database of Bioparts for Synthetic Biology of the Chinese Academy of Science (No. ZSYS-016), the International Partnership Program of Chinese Academy of Science (No. 153D31KY5B20170121) and the National Key Laboratory of Plant Molecular Genetics, SIPPE, CAS. We would like to thank Dr. Wenjuan Yuan, Wenzhi Zhou, Yuanhong Shan, and Shizhen Bu in the Core Facility Centre of SIPPE for technical assistance on high-resolution MS and NMR acquisition. We thank Dr. Lei Yang (Shanghai Chenshan Plant Science Research Center, CAS) for providing the *Phyllostachys* bamboo materials.

## Author contributions

Y.W., Y.S., and J.Y. conceived the study. Y.S. performed the bioinformatics analysis, biochemical experiments, and prepared the manuscript. Z.C. constructed the plasmids, strains, and carried out microbiological manipulations and extractions. J.Y. and S.L. accomplished the pharmacological study. Y.S., I.M., J.Y., and Y.W. wrote the paper. Q.Z., Ying Z. and Yulian Z. contributed to enzyme kinetic study and data analysis. All authors contributed to discussion of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s42003-020-0834-3>.

**Correspondence** and requests for materials should be addressed to Y.W.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020