Allylic hydroxylation of triterpenoids by a plant cytochrome P450 triggers key chemical transformations that produce a variety of bitter compounds

Shohei Takase¹, Kota Kera², Yoshiki Nagashima², Kazuto Mannen², Tsutomu Hosouchi², Sayaka Shinpo², Moeka Kawashima¹, Yuki Kotake¹, Hiroki Yamada¹, Yusuke Saga¹, Junnosuke Otaka¹, Hiroshi Araya¹, Masaaki Kotera³, Hideyuki Suzuki²*, and Tetsuo Kushiro¹*

From the ¹School of Agriculture, Meiji University, Kawasaki, Kanagawa 214-8571, Japan; ²Department of Research and Development, Kazusa DNA Research Institute, Kisarazu, Chiba 292-0818, Japan; ³Development Department of Chemical System Engineering, School of Engineering, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan

Running title: Cucurbitacin biosynthesis in M. charantia

*To whom correspondence should be addressed: Tetsuo Kushiro: School of Agriculture, Meiji University, Kawasaki, Kanagawa 214-8571, Japan; kushiro@meiji.ac.jp; Tel. 81 (44) 934-7105. Hideyuki Suzuki: Department of Research and Development, Kazusa DNA Research Institute, Kisarazu, Chiba 292-0818, Japan; hsuzuki@kazusa.or.jp; Tel. 81 (438) 52-3947; Fax. 81 (438) 52-3948.

Keywords: cytochrome P450, terpenoid, biosynthesis, plant, secondary metabolism, enzyme catalysis, multifunctional enzyme, cucurbitacin, allylic hydroxylation, Momordica charantia

ABSTRACT

Cucurbitacins are highly oxygenated triterpenoids characteristic of plants in the family Cucurbitaceae and responsible for the bitter taste of these plants. Fruits of bitter melon (Momordica charantia) contain various cucurbitacins possessing an unusual ether bridge between C5 and C19, not observed in other Cucurbitaceae members. Using a combination of next-generation sequencing and RNA-seq analysis and gene-to-gene co-expression analysis with the ConfeitoGUIplus software, we identified three P450 genes, CYP81AQ19, CYP88L7, and CYP88L8, expected to be involved in cucurbitacin biosynthesis.
CYP81AQ19 co-expression with cucurbitadienol synthase in yeast resulted in the production of cucurbita-5,24-diene-3β,23α-diol. A mild acid treatment of this compound resulted in an isomerization of the C23-OH group to C25-OH with the concomitant migration of a double bond, suggesting that a non-enzymatic transformation may account for the observed C25-OH in the majority of cucurbitacins found in plants. The functional expression of CYP88L7 resulted in the production of hydroxylated C19, as well as C5–C19 ether-bridged products. A plausible mechanism for the formation of the C5–C19 ether bridge involves C7 and C19 hydroxylations, indicating a multifunctional nature of this P450. On the other hand, functional CYP88L8 expression gave a single product, a triterpene diol, indicating a monofunctional P450 catalyzing the C7 hydroxylation. Our findings of the roles of several plant P450s in cucurbitacin biosynthesis reveal that an allylic hydroxylation is a key enzymatic transformation that triggers subsequent processes to produce structurally diverse products.

Terpenoids have been used as a source for medicines, cosmetics, materials, and food additives, making them one of the most important groups of compounds from natural sources. The diversity of the terpenoid structure originates from both the variation in the carbon skeleton and the variation in the modifications of the core carbon structure, mostly through oxidation and glycosylation. While studies on terpene synthases have been well documented (1), there are relatively limited studies on terpene modification enzymes. Oxidative modifications are often carried out by cytochrome P450 monooxygenases (P450s), which comprise one of the largest oxygenase families in biological systems. In plants, numerous P450 genes exist in the genome, being involved in various processes such as herbicides metabolism and the biosynthesis of plant hormones and secondary metabolites (2, 3). Owing to the large number of genes, the identification of a specific P450 gene involved in a specific process of interest has been hampered. Furthermore, the number of P450s involved in a particular transformation step is often unknown.

Cucurbitaceae is a plant family that is used worldwide and contains many economically important vegetables and fruits, such as the cucumber and melon. Cucurbitacins are characteristic terpenoids found among Cucurbitaceae. They are highly oxygenated triterpenoids that exhibit various biological activities, including anti-cancer, anti-inflammatory, anti-fertility, and
anti-diabetic activities (4-6). Moreover, cucurbitacins are famous for their bitterness, which is responsible for the bitter taste of most of the Cucurbitaceae products. Cucurbitacins derive from the tetracyclic triterpenoid cucurbitadienol, which is formed through the enzymatic cyclization of 2,3-oxidosqualene (Fig. 1). The enzyme cucurbitadienol synthase (CBS) is an oxidosqualene cyclase (OSC), which is responsible for the formation of all the cyclic triterpenes (7). Cucurbitadienol is further oxidized at numerous positions, such as C2, C7, C11, C16, C19, C20, C22, C23, C24, and C25, resulting in various cucurbitacins and indicating that numerous P450s are involved in these modifications (6). The recent genome sequencing of the cucumber (Cucumis sativus) has uncovered several P450s involved in cucurbitacins biosynthesis. These are CYP88L2 for cucurbitadienol C19 hydroxylase (8), CYP81Q58 for C25 hydroxylase (8), CYP87D20 for C11 carbonylase and C20 hydroxylase (9), and Cm180/C1180 (CYP number not assigned) for C2 hydroxylase (10). In addition to cucurbitacins, several P450s have been identified in different plants and are involved in triterpene biosynthesis. Most of those P450s target oleane type and belong to several CYP families such as CYP72A and CYP716A (11-21). Others involve ursane, lupane, and dammarane types as substrates (12, 18, 19, 22, 23).

We have initiated our studies on one of Cucurbitaceae plants, Momordica charantia, known as bitter melon. M. charantia is famous for its bitter taste and used as a vegetable worldwide, being especially popular in Asian countries. While M. charantia possesses numerous cucurbitacins, like other Cucurbitaceae plants, it also contains the characteristic cucurbitacins goyaglycosides and momordicosides, which contain an ether linkage between C5 and C19 that is not present in other cucurbitacins (Fig. 1) (24). Some of these goyaglycosides exhibit anti-diabetic activity and are promising natural agents for the treatment of diabetes (5). Here, we report the identification of three novel P450 genes from M. charantia and, through functional expression in yeast, reveal an unprecedented transformation in cucurbitacins biosynthesis.

**Results**

**Selection of candidate genes involved in cucurbitacins biosynthesis using RNA-seq datasets**

To identify the modification enzymes involved in cucurbitacins biosynthesis, we utilized the expression profiling datasets from ten different tissues of M. charantia obtained in our previous RNA-seq analysis (25). We are particularly
focused on P450s, which mainly contribute to the oxygenation reactions of triterpenoid backbones. Our previous studies revealed that OSC genes showed characteristic tissue-dependent expression patterns (25). In particular, *M. charantia* McCBS gene, encoding cucurbitadienol synthase, showed the highest expression in leaves rather than in fruits. Therefore, P450 genes involved in cucurbitacins biosynthesis were expected to show a similar expression pattern. First, 27,127 total contigs in the RNA-seq datasets were annotated with a BLASTX search. Next, contigs that showed a highly similar expression pattern with the McCBS gene were selected using the ConfeitoGUIplus software (ver1.2.3), which was designed to detect correlation networks (26). Eighteen contigs were obtained, that contained three P450 genes, namely M01465 (CYP81AQ19), M04110 (CYP88L7), and M00873 (CYP88L8) (Fig. 2 and Table S2). Therefore, we decided to determine the function of these candidate P450s using heterologous co-expression with McCBS in yeast cells. Among the other contigs obtained, one of them was squalene monoxygenase, which is the enzyme that converts squalene to 2,3-oxidolsqualene, a substrate for OSC. Moreover, two contigs were annotated as transporters, exhibiting high correlations with McCBS. These candidates were expected to be involved in triterpene biosynthesis and transport.

**Functional analysis of CYP81AQ19**

Heterologous expressions of candidate P450 genes were carried out in *Saccharomyces cerevisiae* strain GIL747. This budding yeast strain accumulates 2,3-oxidolsqualene, which is a ubiquitous substrate for OSC, due to a lack of an endogenous lanosterol synthase. *S. cerevisiae* strain GIL747 has been used for functional analysis of genes involved in triterpene biosynthesis (27). The strain GIL747 was transformed to express McCBS, each candidate P450, and a cytochrome P450 reductase from *Lotus japonicus* (*LjCPR*, GenBank accession no. AB433810). First, we examined the function of CYP81AQ19. The transformed yeast cells were cultured, expression induced by galactose, and the collected cells were extracted with hexane and ethyl acetate (3:1) and analyzed by high-performance liquid chromatography/tandem mass spectrometry (HPLC/MS-MS). As a result, a specific product 2 was detected in yeast extracts expressing CYP81AQ19, McCBS, and LjCPR (Figs. 3, and S1). The mass spectrum of this product did not show any ion peak that corresponded to the expected [M+H] at m/z 443 (M+H+16) that would arise from the introduction of an oxygen atom due to hydroxylation. Instead, it showed an ion peak
at m/z 425, which may derive from the dehydration of a newly introduced hydroxyl group (M+H+16-18) (Figs. S2, and S3). If so, this hydroxyl group seems to be rather labile. An additional dehydrated ion peak at m/z 407 was seen, which was most likely derived from the dehydration of 3β-hydroxyl group. Assuming that the hydroxylation of cucurbitadienol took place, a large-scale (3 L) culture of yeasts was carried out, and the product 2 was extracted and purified by silica gel column chromatography for nuclear magnetic resonance (NMR) analysis. The 1H-NMR spectrum of 2 exhibited eight methyl signals, including one secondary methyl and two vinylic methyls, a hydroxymethine signal at δ 3.47, and two olefinic signals at δ 5.19 and 5.58, all of which mostly coincided with that of authentic cucurbitadienol (Figs. S5, and S6) (7). Most notably, however, the appearance of a δ 4.46 (td) signal suggested the presence of an allylic alcohol moiety. Correspondingly, the olefinic signal for C24 (δ 5.09) and two vinylic methyl signals for C26 and C27 (δ 1.68 and 1.60) of cucurbitadienol were shifted to a lower magnetic field (δ 5.19, 1.70, and 1.68, respectively). These results suggested that a newly introduced hydroxyl group was located at C23. In support of this idea, the 13C-NMR spectrum showed a shift of the C23 signal of cucurbitadienol from δ 24.8 to δ 65.9 (Fig. S7). Therefore, the structure of 2 was determined to be cucurbita-5,24-diene-3β,23-diol, and was further verified by HMBC and HMQC (Figs. S8, and S9).

To determine the absolute configuration of the C23-hydroxyl group, we employed a modified Mosher’s method, using 2-methoxy-2-phenyl-2-(trifluoromethyl) acetic acid (MTPA) esters (28). Owing to the high reactivity of the allylic hydroxyl group in 2 under acidic conditions, we modified the method to react with MTPA esters under basic conditions using N,N-dimethyl-4-aminopyridine (DMAP). As a result, both (S)- and (R)-MTPA esters were obtained and measured for 1H-NMR in order to show the absolute configuration of C23 in 2 to be 23R (Figs. S10-S12). Therefore, the product of McCBS and CYP81AQ19 in yeast was determined as cucurbita-5,24-diene-3β,23α-diol and CYP81AQ19 acted as cucurbitadienol C23α hydroxylase (Fig. 4).

**Structural conversion of 2 under non-enzymatic conditions**

During our analysis of the absolute configuration at C23-hydroxyl group of 2, the reaction with MTPA chloride under acidic conditions showed that 2 was converted to an unexpected product but not the desired MTPA ester. To confirm the high
reactivity of the allylic hydroxyl group under mild acidic conditions, 2 was treated with 0.1 N HCl/MeOH or kept in an NMR sample tube in CDCl₃ for a few days. Previous studies on tirucallane triterpenes having a similar 23α allylic hydroxyl group were shown to be prone to dehydration, resulting in 23,25-diene under mild acidic conditions (29). As a result, we observed that 2 was converted to 3 and 4 in HCl/MeOH and to 4 in CDCl₃ without any enzymes. In the ¹H-NMR spectrum of 3, the allylic hydroxymethine (δ 4.46, C23) and olefinic signals (δ 5.18, C24) disappeared and a new exo-methylene signal (δ 4.85, C26) and another two olefinic signals (δ 5.62, C23 and δ 6.11, C24) were observed (Figs. S13-S15). Through the comparison with published data on related tirucallane triterpenes (29), the structure of 3 was determined to be cucurbita-5,23,25-trien-3β-ol. On the other hand, the ¹H- and ¹³C-NMR spectra of 4 indicated the loss of an allylic hydroxyl group and the presence of three olefinic protons overlapped at δ 5.58 (Figs. S16-S18). The presence of two olefins was supported by the ¹³C-NMR spectrum (δ 121.4/141.2 for C5-C6 and δ 125.5/139.3 for C23-C24). In addition, a hydroxyl bearing carbon signal was observed at δ 70.8, while two methyls attached to a carbon bearing a hydroxyl group were observed at δ 1.31 in the ¹H-NMR. All these data pointed to a structure with Δ²₃-C25-ol and, comparing with published data regarding related compounds with the same partial structure, the structure of 4 was determined to be cucurbita-5,23-dien-3β,25-diol. Taken together, these results demonstrated that C23-hydroxylated cucurbitadienol 2 produced by CYP81AQ19 was converted to C25-modified cucurbitadienol 3 and 4 under mild acidic conditions. Therefore, the C23 allylic alcohol was shown to be rather unstable and eager to either dehydrate or isomerize into C25 tertiary alcohol with a concomitant migration of a double bond from C24(25) to C23(24) (Fig. 4).

Functional analysis of CYP88L7

Next, we examined the function of CYP88L7 by co-expressing it with McCBS and LjCPR in yeast GIL747. The extracts from the culture showed mainly two peaks in the LC/MS-MS analysis (Fig. 3). These products exhibited ion peaks that corresponded to an expected [M+H] at m/z 443 (M+H+16) and a putative [M+H-2H] at m/z 441 (M+H+16-2), suggesting that a hydroxylation and a formal dehydrogenation had taken place (Figs. S3, and S4). A sample was prepared from a large-scale culture (12 L) and purified using a silica gel column chromatography to give products 5 and 6. The ¹H-NMR spectrum of 5 exhibited a similar spectrum to cucurbitadienol.
However, an AB system derived from the geminal protons of a hydroxymethyl group was seen at $\delta$ 3.36 and 3.53 ($J=10.5$ Hz), while the C19 methyl signal ($\delta$ 0.92 in cucurbitadienol) disappeared (Figs. S19, and S20). The $^{13}$C-NMR spectrum also showed the existence of a hydroxymethyl carbon at $\delta$ 68.9 (Fig. S21). These results led to the determination of the structure of 5 as cucurbita-5,24-diene-3β,19-diol, which was further confirmed by HMBC and HMQC (Figs. S22, and S23). On the other hand, the $^1$H-NMR spectrum of 6 exhibited similar but slightly shifted AB system geminal protons at $\delta$ 3.51 and 3.67, while 3α-H also showed a slight shift from a typical $\delta$ 3.47 to $\delta$ 3.40 compared to cucurbitadienol (Figs. S24, and S25). The C29 methyl signal also showed a downfield shift from a typical $\delta$ 1.12-1.13 to $\delta$ 1.20. Furthermore, a drastic change was observed in the olefinic region where C6 proton at $\delta$ 5.59 disappeared and signals at $\delta$ 5.64 and 6.04 appeared, while the C24 proton at $\delta$ 5.09 remained unchanged. In the $^{13}$C-NMR spectrum, two carbinol carbons at $\delta$ 79.8 and 87.5 had newly appeared (Fig. S26). These results suggest an ether bridge between C5 and C19, which is the most characteristic structural feature of the cucurbitacins of *M. charantia*. Through the comparison with data from goyaglycosides (24), and from HMBC and HMQC spectra (Figs. S27, and S28), the structure of 6 was confirmed to be 5β,19-epoxy-cucurbita-6,24-dien-3β-ol.

Therefore, CYP88L7 functions as cucurbitadienol C19 hydroxylase and, surprisingly, also catalyzed an ether bridge formation between C5 and C19 (Fig. 5).

We also tested the co-expression of CYP88L7 and CYP81AQ19. The yeast GIL747 expressing McCBS, CYP81AQ19, CYP88L7, and LjCPR was cultured and the extracts were analyzed by LC/MS-MS (Fig. 3). Several products were observed, one of which gave the expected ion peak [M+H] at $m/z$ 459 (M+H+16x2) that arise from the introduction of two oxygen atoms due to hydroxylations (Figs. S3, and S4). A large-scale culture (6 L) was carried out to isolate these products. After silica gel column chromatography, two fractions were obtained, which were each subjected to preparative HPLC separations. From a more polar fraction, three products, 7, 8, and 9, were obtained. Preliminary NMR spectra of the fraction containing the mixture of these compounds indicated the presence of both species having a C23-allylic hydroxyl group or Δ$^{23}$-C25-ol in the side chain (Figs. S29-S32). In addition, the presence of a C19 hydroxyl group (an AB system at $\delta$ 3.38 and 3.55) was observed. Moreover, an unassigned carbinol methine proton and an olefinic proton were seen. Most notably, the presence of a $\delta$ 9.73 signal pointed to an aldehyde
moiety. The $^1$H-NMR spectrum of isolated 7 exhibited a very similar spectrum to that of 5, indicating the presence of a C19 hydroxyl group (Figs. S33, and S34). The only difference was the presence of a δ 4.46 signal for C23, which indicated the presence of a C23 allylic hydroxyl group. The $^{13}$C-NMR, as well as the HMBC spectra, confirmed the structure (Figs. S35, and S36), and 7 was determined as cucurbita-5,24-diene-3β,19,23α-triol.

Compound 8 showed the presence of an aldehyde ($^1$H: δ 9.70, $^{13}$C: δ 187.8), C23 allylic hydroxyl group (δ 4.46), and $\Delta^{24}$ (δ 5.19) (Figs. S37, and S38). Furthermore, a C6 olefinic proton was shifted to downfield (δ 5.89) and a new oxymethine proton at δ 3.97 appeared. From the HMBC spectrum obtained from the mixture sample (Fig. S31), a correlation from the aldehyde proton to C9 (δ 50.1) was observed, indicating that the aldehyde was attached to C9. Also, correlations were observed between the new oxymethine proton (δ 3.97) and C9 and C5 (δ 145.7), suggesting that the hydroxyl group was attached to C7. Therefore, the structure of 8 was determined as 3β,7,23α-trihydroxy-cucurbita-5,24-dien-19-al. The stereochemistry of the C7 hydroxyl was undetermined. Compound 9 showed a similar spectrum to 8, but with a side chain having $\Delta^{23}$-C25-OH, as evidenced by olefinic protons at δ 5.58 and methyls attached to carbinol carbon at δ 1.31 (Figs. S39, and S40). Therefore, the compound 9 was determined as 3β,7,25-trihydroxy-cucurbita-5,23-dien-19-al. The $^{13}$C-NMR of compounds 8 and 9 were tentatively assigned based on the spectrum obtained for the mixture.

On the other hand, from a less polar fraction on the silica gel column, an inseparable mixture of 10 and 11 was obtained. The $^1$H-NMR spectrum of the mixture indicated the presence of a C5-C19 ether bridge with a double bond located between C6-C7 (Figs. S41-S43). The signals at δ 4.47 and 5.58 also indicated the presence of both $\Delta^{24}$-C23-OH and $\Delta^{23}$-C25-OH structures at the side chain. The $^{13}$C-NMR, as well as the HMBC and HMQC spectra of the mixture (Figs. S44-S46), confirmed the structure as 5β,19-epoxy-cucurbita-6,24-diene-3β,23α-diol for 10 and 5β,19-epoxy-cucurbita-6,23-diene-3β,25-diol for 11. Collectively, when expressed together with C23 hydroxylase CYP81AQ19, CYP88L7 was shown to be a multifunctional enzyme catalyzing the C19 oxidation, not only to a hydroxyl or an ether but also to an aldehyde, along with the C7 hydroxylation (Fig. 5).

**Functional analysis of CYP88L8**

Finally, we examined the function
of CYP88L8 by co-expressing it with McCBS and LjCPR in yeast GIL747. From the HPLC analysis of the extracts from the yeast culture, the specific product 12 was detected (Fig. 6A). A large-scale culture (8 L) was carried out in order to isolate the compound after silica gel column chromatography. The $^1$H-NMR spectrum showed a signal for a new oxymethine proton at $\delta$ 3.94 similar to 8 and 9, which suggested the presence of a hydroxyl group at C7 (Figs. S47, and S49). Other signals were very similar to cucurbitadienol. The $^{13}$C-NMR spectrum also confirmed a new hydroxymethine carbon at $\delta$ 68.1 (Fig. S50). The HMBC correlations were observed between this proton and C5, C6, and C9, supporting the structure having a hydroxyl group at C7 (Figs. S47, and S51). The stereochemistry of C7-OH was determined by NOE measurements, where NOE effects were observed between C7-H and C5-H, C7-H and C14-Me, and C14-Me and C6-H (Figs. S48, S53, and S54). Therefore, C7-H was determined to have $\alpha$ configuration, indicating the 7$\beta$-OH stereochemistry. Thus, the structure of 12 was confirmed as cucurbita-5,24-diene-3$\beta$,7$\beta$-diol (Fig. 6B).

**Expression analysis of P450s in M. charantia**

The expression profiles of McCBS, CYP81AQ19, and CYP88L7 were examined by quantitative RT-PCR in order to verify the data from the RNA-seq analysis. Consistent with previous data, all three genes showed the highest expression in leaves, while moderate expression was seen in the roots for McCBS and CYP81AQ19 (Fig. 7). All the genes showed very low expression in fruits.

**Discussion**

Our candidate selection of P450 genes using the newly developed ConfeitoGUIplus (26), which is based on the Total Gene value expression patterns among different organs from RNA-seq analysis synchronizing with a core skeleton forming enzyme CBS, was successful in identifying three genes responsible for cucurbitacin biosynthesis. The combination of RNA-seq and gene-to-gene correlation analyses using the ConfeitoGUIplus software has the potential to become a standardized analysis method for mining new enzymatic and regulatory genes involved in secondary metabolites biosynthesis in plants whose genome has not been sequenced, such as in many useful medicinal plants.

The three P450s that we identified belonged to the CYP81AQ and CYP88L families and were similar to P450s reported for cucumber, CYP81Q58 and CYP88L2 (8), validating our selection methods (Fig. S55). Both CYP81Q58 and CYP88L2 were
reported to catalyze C25 and C19 hydroxylation, respectively. Instead, our study demonstrated that *M. charantia* CYP81AQ19 catalyzed the C23 hydroxylation at the allylic position of cucurbitadienol. This was in sharp contrast to the function of CYP81Q58 in cucumber, which was reported to catalyze C25 hydroxylation (8). To our surprise, the C23 hydroxyl group was readily converted to a Δ^23^-C25-OH structure under mild acidic conditions, such as treatment with HCl/MeOH. The dehydration into 23,25-diene was also observed. In fact, a similar dehydroxylation triggered a double bond migration and a C25 methoxylation, as was previously reported (30). However, we did not observe any methoxylation at C25. Instead, only a hydroxylation was seen. In general, the allylic hydroxyl group is labile to dehydroxylation, which in this case resulted in a C23 allylic cation that undergone isomerization into a C25 tertiary cation accompanied by a double bond shift (Fig. 8A). Finally, a capture of the cation by a hydroxyl ion would produce C25-OH. This isomerization step should shift towards the C25 cation, as the tertiary cation is more stable than the C23 secondary cation. Based on these considerations, we concluded that 2 was a direct enzymatic product and not the result of the isomerization of C25-OH, and that the true function of CYP81AQ19 was C23 hydroxylation and not C25 hydroxylation. The isomerization of C23-OH was also observed for products that derived from the co-expression of CYP81AQ19 and CYP88L7, such as in 8 and 9, and in 10 and 11. This indicated that the isomerization could take place during the extraction and purification steps. These observations raised an important question regarding whether the structure of naturally occurring cucurbitacins truly has a C25 hydroxyl group. If such an isomerization readily took place, it is difficult to deduce when and where it occurred. Most of the cucurbitacins isolated from *M. charantia* possessed a C25-OH group, with the exception of goyaglycoside-f, which had a C23-OH group (24). Moreover, the majority of cucurbitacins found in other Cucurbitaceae plants or other natural sources possessed a C25-OH group (6). Only a few were found to have C23-OH. It is still early to speculate, but finding a C23-hydroxylase in *M. charantia* strongly suggests that the P450s responsible for hydroxylation at these positions initially introduce a hydroxyl group at C23, which then isomerizes to C25 non-enzymatically. Whether this isomerization took place inside the plant or during the extraction procedure is unknown. It also raises the possibility that the structures of the majority of cucurbitacins isolated so far might be an artifact produced during extraction procedures, and that the original
structure could have possessed a C23 allylic hydroxyl group. If a P450 introduces a hydroxyl group directly at C25, a rational mechanism of C25 hydroxylation would assume that a P450 should abstract a H radical from C23, generating an allylic radical, which then requires the migration of a double bond to give the C25 radical which the FeIV-oxo complex (compound I) attacks. Compared with this C25 scenario, hydroxylation at C23 seems more straightforward and do not require a double bond migration. In the case of C25, a question arises regarding what factor favors a double bond migration before the attack of the compound I to the radical center. It also requires a substrate radical intermediate to shift with regard to the Fe center of the distal side of the heme, right after the H abstraction, in order for compound I to attack a different carbon to where it originally abstracted the H radical. In this line, the function of CYP81Q58 should be carefully reexamined in order to see if it truly hydroxylates the C25 position with a double bond migration from Δ24 to Δ23, and whether or not the observed C25 product results due to an artifact from the isomerization of C23-OH. In either case, the reaction of CYP81AQ19 brings us a unique opportunity to study the allylic hydroxylation catalyzed by P450, which seems to be a rare case in P450 reaction.

The function of *M. charantia* CYP88L7 was shown to be cucurbitadienol C19-hydroxylation, similarly to that of cucumber CYP88L2 (8). However, to our surprise, *M. charantia* CYP88L7 also catalyzed the formation of the ether bridge between C5 and C19. A rational mechanism for the formation of the ether bridge can be envisioned as follows (Fig. 8B). Both C19 and C7 hydroxylations are followed by a dehydroxylation at C7, generating an allylic cation, which then triggers a double bond migration to C6(7). Finally, the capture of the resulting C5 cation with the C19 hydroxyl produces the ether bridge. Therefore, unlike CYP88L2, *M. charantia* CYP88L7 also catalyzed the C7 hydroxylation in addition to the C19 hydroxylation. This was also evident from the deduced structures of 8 and 9, which have a hydroxyl at C7. Moreover, CYP88L7 also catalyzed the oxidation of C19 hydroxyl to give C19 aldehyde. Such C19 aldehyde species were found in *M. charantia*, e.g., momordicoside K (24). In fact, the structure of 9 corresponded to an aglycone of this compound. Therefore, the presence of unique cucurbitacins having an ether bridge between C5 and C19 in *M. charantia* is due to the presence of a multifunctional P450 that hydroxylates both C7 and C19. The absence of cucurbitacins having an ether bridge in cucumber corroborated the presence of a monofunctional CYP88L2 that can only
hydroxylate C19. Whether this ether bridge formation in CYP88L7 was enzymatically catalyzed or not is unknown at this moment. Unlike in the CYP81AQ19 case, none of the intermediates possessing C7 and C19 diol have been isolated. A non-enzymatic ether bridge formation after a P450 hydroxylation is known in the CYP707A case, where a P450 catalyzes the 8'-hydroxylation of the plant hormone abscisic acid, and is followed by the concomitant Michael addition of the resulting hydroxyl in order to produce phaseic acid (31, 32). Further studies are needed to see if this ether bridge formation takes place in the active site cavity of CYP88L7 or not. Similarly, the key to the ether bridge formation is the allylic hydroxylation at C7, which triggers the dehydroxylation that leads to the subsequent transformations.

On the other hand, the function of CYP88L8 was shown to be cucurbitadienol C7β hydroxylase. Although CYP88L8 exhibited high amino acid sequence identity with CYP88L7, exhibiting 79% identity (Fig. 9), it was a monofunctional P450 catalyzing a hydroxylation only at C7 and not at C19. Despite having such a high sequence identity, the functional difference between CYP88L7 is especially intriguing. Structural studies would certainly illuminate these important points in the future. Functional differences between CYP88L8 and CYP88L2 are also noteworthy, despite having a sequence identity of 54%, and resulted in a hydroxylation at different carbons. Cucurbitacins having only the C7β hydroxyl group have been isolated from M. charantia (33), and CYP88L8 should be responsible for the production of such cucurbitacins.

The expression of two P450 genes was confirmed to be higher in leaves rather than in fruits (Fig. 7). A synchronized expression with McCBS indicated the common regulation of the core skeleton forming OSC and modifying enzymes. A similar transcriptional regulator previously identified in cucumber might be responsible for the regulation of the cucurbitacins pathway in M. charantia as well (8). This also suggested that cucurbitacins, at least the oxidized form rather than the precursor cucurbitadienol, are transported from leaves to fruits, where large amounts of cucurbitacins accumulate and produce strong bitterness. However, whether glycosides are transported or not still remains unknown, requiring further identification of glycosyltransferase genes.

Our current findings show that an allylic hydroxyl group in the cucurbitane skeleton is highly reactive, triggering non-enzymatic transformations. When complex structures are encountered in natural products, it is often not clear how many enzymes are needed to construct such
complexity. In this case, an ether bridge between C5 and C19 represents an ambiguous case. At most, three or four enzymes could be predicted to participate in the formation of this structure. However, only one enzyme was used to produce them. The number of genes required to construct a molecule might be smaller than the number of chemical steps predicted rationally. Organisms seem to take advantage of non-enzymatic transformations to construct complex structures that otherwise would require multiple enzymes. In these cases, allylic hydroxylation was a key step to trigger the subsequent non-enzymatic processes in order to construct complex structures. Furthermore, the P450 catalyzed allylic oxidation might play a pivotal role in constructing complex structures in the biosynthesis of other natural products as well. So far, the supply of cucurbitacins by chemical synthesis or extraction from plants is difficult from the viewpoint of costs, efforts, and quantities. Our findings are expected to lead to the production of pure bioactive cucurbitacins using yeasts expressing these biosynthetic genes.

**Experimental procedures**

**Plant materials**

All ten different types of tissues of *M. charantia*, namely old leaves, young leaves, stems, tendrils, male flowers, female flowers, fruits, seedling leaves, seedling stems, and seedling roots were harvested in a vegetable field (Tateyama, Chiba) of the Southern Prefectural Horticulture Institute in the Chiba Prefectural Agriculture and Forestry Research Center and in a greenhouse (Kazusa DNA Research Institute) from 2012 to 2015. All tissues were cut into small pieces, frozen by liquid nitrogen, and stored at -80 °C prior to RNA extraction.

**RNA isolation and library preparation for RNA-seq analysis**

Total RNA extraction, library construction, and Illumina sequencing were prepared according to a previous report (25).

**De novo transcriptome assembly**

The obtained Illumina reads were sequenced by Illumina’s next-generation sequencing instruments Genome Analyzer IIx (in 2012), HiSeq 1000 (in 2012), and HiSeq 1500 Rapid mode (in 2013) with 100 bp paired end (PE) reads. The reads were assembled using the commercially available CLC Genomics Workbench version 5.5.1 (CLC bio Japan, Tokyo, Japan), using a minimum contig length of 800, performing scaffolding. A total of 27,127 contigs were obtained. The assembled contigs were also used as queries against the non-redundant (nr) protein database using the BLASTX program (e-5). The Total Gene values were
recalculated using CLC Genomics Workbench version 11.0. The raw RNA-Seq reads obtained in this study have been submitted to the DDBJ Sequence Read Archive (DRA) under the accession number DRA007507.

**Gene-to-gene correlation analysis using the ConfeitoGUIplus software**

In order to identify the modification enzymes involved in cucurbitacins biosynthesis, we employed the commercial software ConfeitoGUIplus version 1.2.3 developed in the Kazusa DNA Research Institute (contact: Dr. Hideyuki Suzuki, biosup-met@kazusa.or.jp). The ConfeitoGUIplus version 1.2.3 is a standalone software to detect network modules from a correlation network composed of molecular biology multivariate data, based on the Confeito algorithm (26), which allows the adjustment of the network modules sizes by modifying a single parameter and can detect elements specifically related to the network modules even when they are weakly correlated. Before the gene-to-gene correlation analysis, we calculated mean averages against the Total Gene values of RNA-seq total genes from ten different tissues (old leaves, young leaves, stems, tendrils, male flowers, female flowers, fruits, seedling leaves, seedling stems, and seedling roots) obtained by the three sequencers (GAIIx, HiSeq, and Rapid) for using as input file for the ConfeitoGUIplus software. The correlation network analysis using ConfeitoGUIplus version 1.2.3 was performed with the following parameters: Cosine Similarity, 4; Cosine Min Correlation, 0.5; Min Elements, 2; Max Elements, 50; and Solid Bold, 0.9 on False-Positive-Out (FPO) Analysis, and Vertex Specificity (VS) Threshold, 0.5; Cosine Correlation Threshold, 0.5; Max Elements, 1000; and Dots Bold, 0.9 on False-Negative-In (FNI) Analysis, in order to obtain gene-to-gene correlation network modules for including the McCBS gene and related genes involved in cucurbitacins biosynthesis. First, we performed an FPO analysis of the Confeito algorithm, resulting in 15 selective contigs (Fig. 2 and Table S2) from a total of 7,127 contigs. The FPO module contained two P450 genes, two transporter genes, a squalene monooxygenase gene, and 1-deoxy-D-xylulose-5-phosphate synthase gene highly correlated with the McCBS gene. Next, we performed an FNI analysis of the Confeito algorithm, resulting in an additional five contigs (Fig. 2 and Table S2) from the total 7,127 contigs. The FPO and FNI modules contained three P450 genes in total.

**Heterologous expression using yeast strain GIL747**
PrimeSTAR HS DNA Polymerase (TaKaRa, Japan) was used for PCR according to the manufacturer’s instructions. The RNA was extracted from ten different tissues of *M. charantia*, and cDNA was prepared as in our previous study (25). The full length of each P450 gene was amplified using each designed primer (Table S1) with appropriate cDNA, where significant expression was found, as a template and subcloned into the yeast expression vectors pESC-URA or pESC-LEU (Agilent, Japan). *Saccharomyces cerevisiae* GIL747 was constructed by crossing GL7 (MATa gal2 hem3-6 erg7) (34) with W303 (MATα leu2-3, 112 trp1-1 can 1-100 ura3-1 ade2-1 his3-11, 15 rad5-535) (35), sporulating the resultant diploid, and selecting a segregant with the desired phenotype. Finally, GIL747 was transformed and cultured as previously described (27).

**Purification and analysis of triterpene products from yeast**

The products from transformant GIL747 were extracted with hexane and ethyl acetate (3:1) as previously described (25). The resulting extracts were injected into liquid chromatography (LC)-LTQ-FT-ICR-MS (LC, Agilent 1100 series (Agilent, Japan); LTQ-FT-ICR-MS (Thermo Fisher Scientific, Japan)). The products were separated using a C18 column (TSKgel ODS-100V; 4.6 × 250 mm, 3 µm; TOSOH Bioscience, Japan) and analyzed in a positive-ionization mode of APCI with a high resolution (Full mass scans, 100,000), as previously described (25). A HPLC analysis was carried out with a LC-2010A HT (Shimadzu, Japan) with a column TSKgel ODS-80TM (4.6 ×150 mm, 5 µm; TOSOH Bioscience, Japan). A low-pressure gradient with acetonitrile (50–100%)/35 min, and then, 100% acetonitrile was flowed for 20 minutes with a flow rate of 1.0 mL/min with column temperature at 40 ºC and UV detection at 210 nm. All samples were dissolved in acetone and 10 µl was injected. In order to isolate the product triterpenes obtained by the heterologous expression of candidate genes, a large-scale culture was carried out. The yeast cells were collected and disrupted by 20% KOH/50% EtOH under reflux conditions for 30 minutes, and the products were extracted with hexane and ethyl acetate (3:1) and evaporated to dryness. Compound 2 was extracted from 3 L culture yeast cells expressing McCBS, CYP81AQ19, and LjCPR. The crude extract was applied onto a silica gel column (vol. 50 mL) and eluted with hexane and ethyl acetate (7:1) as a solvent. The desired fractions containing compound 2 (5 mg) were collected and evaporated to dryness. A large-scale culture (12 L) expressing McCBS, CYP88L7, and LjCPR was similarly extracted and the crude
extract was applied onto a silica gel column (vol. 50 mL) eluted with hexane and ethyl acetate (3:1) as a solvent. The desired fractions containing compounds 5 (4 mg) and 6 (3 mg) were collected and evaporated to dryness. The crude extract from a large-scale culture (6 L) expressing McCBS, CYP81AQ19, CYP88L7, and LjCPR was subjected to a silica gel column (vol. 50 mL) and eluted with hexane and ethyl acetate (4:1) as a solvent. Two semi-purified fractions (Fr. 1 (5 mg) containing compounds 10 and 11, and Fr. 2 (6 mg) containing compounds 7, 8, and 9) were obtained. Further purification was performed with a preparative HPLC (SSC-3461 pump, Senshu Scientific, Japan) using a Cosmosil 5C18-PAQ column (10 × 250 mm, 5 µm, Nacalai Tesque, Japan) with 70% aq. acetonitrile (for Fr. 1) or 85% aq. acetonitrile (for Fr. 2) as a solvent, with a flow rate of 15 mL/min and UV detection at 210 nm, in order to obtain 4 mg of 10 and 11 from Fr. 1 and 3 mg of 7, 1 mg of 8, and 1 mg of 9 from Fr. 2. Compound 12 was purified from a crude extract of 8 L culture yeast cells expressing McCBS, CYP88L8, and LjCPR. The crude extract was applied onto a silica gel column (vol. 50 mL) eluted with hexane and ethyl acetate (6:1) as a solvent. The desired fractions containing compound 12 (4 mg) were collected and evaporated to dryness. Each purified compound was analyzed by NMR using JEOL ECP-500 (1H at 500 MHz, 13C at 125 MHz) with CDCl₃ (99.8% atom ²H, Kanto Chemical, Japan) as a solvent with a solvent signal of δ 7.26 ppm for 1H and δ 77.0 ppm for ¹³C as references for chemical shifts.

**Determination of absolute configuration of 2 using a modified Mosher’s method**

Two µl of DMAP and (S)-2-methoxy-2-phenyl-2-(trifluoromethyl) acetic acid (MTPA) or (R)-MTPA were added to 2 mg of compound 2 in dichloromethane. After 30 minutes at room temperature, 5 µl of diisopropylamine was added for quenching the reaction. The (S)- or (R)-esterified products were purified by preparative TLC and ¹H-NMR spectra were measured. The absolute configuration of 2 was determined by calculating the difference in chemical shifts on ¹H-NMR spectra between (S)- and (R)-esterified products.

**Isomerization of 2 under acidic conditions**

Two mg of 2 was dissolved in 0.5 mL of 0.12 N HCl/MeOH and stood for an overnight at 4 °C. After the reaction, products were evaporated to dryness and purified by silica gel column chromatography.

**Expression analysis in M. charantia**

cDNAs derived from the six organs of *M. charantia* described above were...
synthesized by SuperScript III First-Strand Synthesis System for RT-PCR. The qPCR was performed according to the Automatic Ct method of the 7900HT Fast Real-Time PCR System (Applied Biosystems, Japan) using DyNAamo HS SYBR Green qPCR Kits (Thermo Fisher Scientific, Japan). Gene-specific primers are shown in Supplemental Table S1. Transcription levels of target genes were normalized to a reference gene (McActin) and indicated as means +/- S.D. of three independent experiments.

Accession numbers
CYP81AQ19: LC456843
CYP88L7: LC456844
CYP88L8: LC456845
Acknowledgements: This work was supported by Grants-in-Aid for Scientific Research on Innovation Areas ‘Biosynthetic Machinery’ (KAKENHI) from the Japan Society for the Promotion of Science (Grant Numbers 23108527 to T.K., 23108528 to H.S., and 25108727 to H.S.), the Kazusa DNA Research Institute Foundation, and the Meiji University. We would like to thank YAE NOGEI CO., LTD, for giving us seeds of *Momordica charantia* (Erabu, cultivar) and Dr. Shinji Mizuno (Southern Prefectural Horticulture Institute in Chiba Prefectural Agriculture and Forestry Research Center) for his help with the cultivation of *M. charantia* (Erabu, cultivar). We also thank the Meiji University Advanced Plant Factory Research Center for the cultivation of *M. charantia* and Dr. Hideki Yashiroda (The University of Tokyo) and Dr. Yoko Yashiroda (RIKEN) for constructing yeast GIL747. We thank Dr. David Nelson (University of Tennessee) for the nomenclature of P450 genes. We also thank Editage (www.editage.com) for English language editing.

Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article.
References

1. Christianson, D. W. (2017) Structural and chemical biology of terpenoid cyclases. *Chem. Rev.* **117**, 11570-11648

2. Mizutani, M., and Sato, F. (2011) Unusual P450 reactions in plant secondary metabolism. *Arch. Biochem. Biophys.* **507**, 194-203

3. Werck-Reichhart, D., Bak, S., and Paquette, S. (2002) Cytochromes P450, in *The Arabidopsis Book*, Somerville, C. R., and Meyerowitz, E. M. (eds) Rockville, M. D., American Society of Plant Biologists 1, e0028

4. Rios, J. L., Andujar, I., Escandell, J. M., Giner, R. M., and Recio, M. C. (2012) Cucurbitacins as inducers of cell death and a rich source of potential anticancer compounds. *Curr. Pharm. Des.* **18**, 1663-1676

5. Tan, M. J., Ye, J. M., Turner, N., Hohnen-Behrens, C., Ke, C. Q., Tang, C. P., Chen, T., Weiss, H. C., Gesing, E. R., Rowland, A., James, D. E., and Ye, Y. (2008) Antidiabetic activities of triterpenoids isolated from bitter melon associated with activation of the AMPK pathway. *Chem. Biol.* **15**, 263-273

6. Chen, J. C., Chiu, M. H., Nie, R. L., Cordell, G. A., and Qiu, S. X. (2005) Cucurbitacins and cucurbitane glycosides: structures and biological activities. *Nat. Prod. Rep.* **22**, 386-399

7. Shibuya, M., Adachi, S., and Ebizuka, Y. (2004) Cucurbitadienol synthase, the first committed enzyme for cucurbitacin biosynthesis, is a distinct enzyme from cycloartenol synthase for phytosterol biosynthesis. *Tetrahedron* **60**, 6995-7003

8. Shang, Y., Ma, Y., Zhou, Y., Zhang, H., Duan, L., Chen, H., Zeng, J., Zhou, Q., Wang, S., Gu, W., Liu, M., Ren, J., Gu, X., Zhang, S., Wang, Y., Yasukawa, K., Bouwmeester, H. J., Qi, X., Zhang, Z., Lucas, W. J., and Huang, S. (2014) Biosynthesis, regulation, and domestication of bitterness in cucumber. *Science* **346**, 1084-1088

9. Zhou, Y., Ma, Y., Zeng, J., Duan, L., Xue, X., Wang, H., Lin, T., Liu, Z., Zeng, K., Zhong, Y., Zhang, S., Hu, Q., Liu, M., Zhang, H., Reed, J., Moses, T., Liu, X., Huang, P., Qing, Z., Liu, X., Tu, P., Kuang, H., Zhang, Z., Osbourn, A., Ro, D. K., Shang, Y., and Huang, S. (2016) Convergence and divergence of bitterness biosynthesis and regulation in Cucurbitaceae. *Nat. Plants* **2**, 16183

10. Zhang, J., Dai, L., Yang, J., Liu, C., Men, Y., Zeng, Y., Cai, Y., Zhu, Y., and Sun, Y. (2016) Oxidation of cucurbitadienol catalyzed by CYP87D18 in the biosynthesis of mogrosides from *Siraitia grosvenorii*. *Plant Cell Physiol.* **57**, 1000-1007
11. Moses, T., Pollier, J., Faizal, A., Apers, S., Pieters, L., Thevelein, J. M., Geelen, D., and Goossens, A. (2015) Unraveling the triterpenoid saponin biosynthesis of the African shrub *Maesa lanceolata*. *Mol. Plant* **8**, 122-135

12. Moses, T., Pollier, J., Almagro, L., Buyst, D., Van Montagu, M., Pedreno, M. A., Martins, J. C., Thevelein, J. M., and Goossens, A. (2014) Combinatorial biosynthesis of sapogenins and saponins in *Saccharomyces cerevisiae* using a C-16α hydroxylase from *Bupleurum falcatum*. *Proc. Natl. Acad. Sci. USA* **111**, 1634-1639

13. Yasumoto, S., Seki, H., Shimizu, Y., Fukushima, E. O., and Muranaka, T. (2017) Functional characterization of CYP716 family P450 enzymes in triterpenoid biosynthesis in tomato. *Front. Plant Sci.* **8:21**, doi: 10.3389/fpls.2017.00021.

14. Seki, H., Ohyama, K., Sawai, S., Mizutani, M., Ohnishi, T., Sudo, H., Akashi, T., Aoki, T., Saito, K., and Muranaka, T. (2008) Licorice β-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proc. Natl. Acad. Sci. USA* **105**, 14204-14209

15. Seki, H., Sawai, S., Ohyama, K., Mizutani, M., Ohnishi, T., Sudo, H., Fukushima, E. O., Akashi, T., Aoki, T., Saito, K., and Muranaka, T. (2011) Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. *Plant Cell* **23**, 4112-4123

16. Shibuya, M., Hoshino, M., Katsumbe, Y., Hayashi, H., Kushiro, T., and Ebizuka, Y. (2006) Identification of β-amyrin and sophoradiol 24-hydroxylase by expressed sequence tag mining and functional expression assay. *FEBS J.* **273**, 948-959

17. Fukushima, E. O., Seki, H., Sawai, S., Suzuki, M., Ohyama, K., Saito, K., and Muranaka, T. (2013) Combinatorial biosynthesis of legume natural and rare triterpenoids in engineered yeast. *Plant Cell Physiol.* **54**, 740-749

18. Fukushima, E. O., Seki, H., Ohyama, K., Ono, E., Umemoto, N., Mizutani, M., Saito, K., and Muranaka, T. (2011) CYP716A subfamily members are multifunctional oxidases in triterpenoid biosynthesis. *Plant Cell Physiol.* **52**, 2050-2061

19. Kim, O. T., Um, Y., Jin, M. L., Kim, J. U., Hegebarth, D., Busta, L., Racovita, R. C., and Jetter, R. (2018) A novel multifunctional C-23 oxidase, CYP714E19, is involved in asiaticoside biosynthesis. *Plant Cell Physiol.* **59**, 1200-1213

20. Han, J. Y., Chun, J. H., Oh, S. A., Park, S. B., Hwang, H. S., Lee, H., and Choi, Y. E. (2018) Transcriptomic analysis of Kalopanax septemlobus and characterization of KsBAS, CYP716A94 and CYP72A397 genes involved in hederagenin saponin biosynthesis. *Plant
21. Geisler, K., Hughes, R. K., Sainsbury, F., Lomonossoff, G. P., Rejzek, M., Fairhurst, S., Olsen, C. E., Motawia, M. S., Melton, R. E., Hemmings, A. M., Bak, S., and Osbourn, A. (2013) Biochemical analysis of a multifunctional cytochrome P450 (CYP51) enzyme required for synthesis of antimicrobial triterpenes in plants. *Proc. Natl. Acad. Sci. USA* **110**, E3360-E3367

22. Han, J. Y., Kim, H. J., Kwon, Y. S., and Choi, Y. E. (2011) The cyt P450 enzyme CYP716A47 catalyzes the formation of protopanaxadiol from dammarenediol-II during ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol.* **52**, 2062-2073

23. Han, J. Y., Hwang, H. S., Choi, S. W., Kim, H. J., and Choi, Y. E. (2012) Cytochrome P450 CYP716A53v2 catalyzes the formation of protopanaxatriol from protopanaxadiol during ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol.* **53**, 1535-1545

24. Murakami, T., Emoto, A., Matsuda, H., and Yoshikawa, M. (2001) Medicinal foodstuffs. XXI. Structures of new cucurbitane-type triterpene glycosides, goyaglycosides-a, -b, -c, -d, -e, -f, -g, and -h, and new oleanane-type triterpene saponins, goyasaponins I, II, and III, from the fresh fruit of Japanese *Momordica charantia* L. *Chem. Pharm. Bull.* **49**, 54-63

25. Takase, S., Kera, K., Hirao, Y., Hosouchi, T., Kotake, Y., Nagashima, Y., Mannen, K., Suzuki, H., and Kushiro, T. (2018) Identification of triterpene biosynthetic genes from *Momordica charantia* using RNA-seq analysis. *Biosci. Biotechnol. Biochem.* doi: 10.1080/09168451.2018.1530096

26. Ogata, Y., Mannen, K., Kotani, Y., Kimura, N., Sakurai, N., Shibata, D., and Suzuki, H. (2018) ConfeitoGUI: a toolkit for size-sensitive community detection from a correlation network. *PLoS One* **13**, e0206075

27. Kushiro, T., Shibuya, M., and Ebizuka, Y. (1998) β-amyrin synthase cloning of oxidosqualene cyclase that catalyzes the formation of the most popular triterpene among higher plants. *Eur. J. Biochem.* **256**, 238-244

28. Ohtani, I., Kusumi, T., Kashman, Y., and Kakisawa, H. (1991) High-field FT NMR application of Mosher’s method. The absolute configurations of marine terpenoids. *J. Am. Chem. Soc.* **113**, 4092-4096

29. Liu, H., Heilmann, J., Rali, T., and Sticher, O. (2001) New tirucallane-type triterpenes from *Dysoxylum variabile*. *J. Nat. Prod.* **64**, 159-163

30. Yasuda, M., Iwamoto, M., Okabe, H., and Yamauchi, T. (1984) Structures of momordicines I, II and III, the bitter principles in the leaves and vines of *Momordica*
charantia L. *Chem. Pharm. Bull.* **32**, 2044-2047

31. Kushiro, T., Okamoto, M., Nakabayashi, K., Yamagishi, K., Kitamura, S., Asami, T., Hirai, N., Koshiba, T., Kamiya, Y., and Nambara, E. (2004) The *Arabidopsis* cytochrome P450 CYP707A encodes ABA 8'-hydroxylases: key enzymes in ABA catabolism. *EMBO J.* **23**, 1647-1656

32. Saito, S., Hirai, N., Matsumoto, C., Ohigashi, H., Ohta, D., Sakata, K., and Mizutani, M. (2004) *Arabidopsis CYP707As* encode (+)-abscisic acid 8'-hydroxylase, a key enzyme in the oxidative catabolism of abscisic acid. *Plant Physiol.* **134**, 1439-1449

33. Fatope, M. O., Takeda, Y., Yamashita, H., Okabe, H., and Yamauchi, T. (1990) New cucurbitane triterpenoids from *Momordica charantia*. *J. Nat. Prod.* **53**, 1491-1497

34. Gollub, E. G., Liu, K. P., Dayan, J., Adlersberg, M., and Sprinson, D. B. (1977) Yeast mutants deficient in heme biosynthesis and a heme mutant additionally blocked in cyclization of 2,3-oxidosqualene. *J. Biol. Chem.* **252**, 2846-2854

35. Fan, H. Y., Cheng, K. K., and Klein, H. L. (1996) Mutations in the RNA polymerase II transcription machinery suppress the hyperrecombination mutant hpr1 delta of *Saccharomyces cerevisiae*. *Genetics* **142**, 749-759

36. Takahashi, S., Yeo, Y. S., Zhao, Y., O'Maille, P. E., Greenhagen, B. T., Noel, J. P., Coates, R. M., and Chappell, J. (2007) Functional characterization of *premnaspirodiene* oxygenase, a cytochrome P450 catalyzing regio- and stereo-specific hydroxylations of diverse sesquiterpene substrates. *J. Biol. Chem.* **282**, 31744-31754

**FOOTNOTES**

The abbreviations used in the text are: OSC, oxidosqualene cyclase; P450, cytochrome P450 monooxygenase, NMR, nuclear magnetic resonance; and HPLC, high performance liquid chromatography
Figure 1. Biosynthetic pathway of cucurbitacins in *M. charantia*. Carbon numbering is shown in the structure of cucurbitadienol (1).
Figure 2. (A) A network module representing correlations of eighteen contigs highly correlated with McCBS (blue circle, M01391) in the ConfeitoGUIplus analysis. Pink circles (M00873 (CYP88L8), M01465 (CYP81AQ19), and M04110 (CYP88L7)) indicate the candidate contigs annotated as P450 in the BLAST search. The annotation of eighteen selected contigs is described in Table S2. Solid lines and dotted lines indicate correlations obtained by the FPO and the FNI analysis, respectively (26). (B) RPKM (reads per kilobase per million mapped reads) values from the RNA-seq analysis of three candidate P450 genes in each tissue which were similar to those of McCBS. Red and Black bars indicate McCBS and candidate P450 genes, respectively. The abbreviations on the x-axes denote OL: Old leaves, YL: Young leaves, St:
Stems, Te: Tendrils, MF: Male flowers, FF: Female flowers, Fr: Fruits, SL: Seedling leaves, SS: Seedling stems, and SR: Seedling roots.
Figure 3. LC/MS-MS chromatogram of yeast extracts from the expression of empty vector, McCBS, McCBS/LjCPR/CYP81AQ19, McCBS/LjCPR/CYP88L7, and McCBS/LjCPR/CYP81AQ19/CYP88L7. For products 10 and 11, peaks could not be assigned due to an identical mass. Moreover, peaks for aldehyde 8 and 9 could not be found on the McCBS/LjCPR/CYP81AQ19/CYP88L7 chromatogram.
**Figure 4.** CYP81AQ19 catalyzed hydroxylation at the C23 of cucurbitadienol. The allylic hydroxyl group underwent a dehydration to produce 3 and an isomerization to give 4 under acidic conditions.
Figure 5. CYP88L7 catalyzed hydroxylation at C19, as well as the ether bridge formation between C5 and C19. Upon co-expression with CYP81AQ19, aldehyde products 8 and 9 were also identified.
Figure 6. (A) HPLC chromatogram of yeast extracts expressing empty vector (black) and McCBS/LjCPR/CYP88L8 (green). (B) CYP88L8 catalyzed C7β hydroxylation of cucurbitadienol.
Figure 7. Quantitative RT-PCR analysis of McCBS (blue; left), CYP81AQ19 (red; center), and CYP88L7 (green; right) mRNA levels in the leaves (L), stems (St), fruits (Fr), seedling leaves (SL), seedling stems (SS), and seedling roots (SR). McActin was utilized as a reference. Data represents mean ± SD from three independent experiments.
Figure 8. Allylic hydroxylation triggers further structural rearrangements. (A) Isomerization of C23-OH into C25-OH via allylic cation intermediate. (B) Ether bridge formation triggered by dehydroxylation of C7-OH.
Figure 9. Amino acid sequence alignment between CYP88L7 and CYP88L8 from *M. charantia*.

Putative substrate recognition sequences (SRS 1–6) are denoted with blue underlines (1–6) (36). The conserved heme binding sequence is indicated with a red underline.
Allylic hydroxylation of triterpenoids by a plant cytochrome P450 triggers key chemical transformations that produce a variety of bitter compounds

Shohei Takase, Kota Kera, Yoshiki Nagashima, Kazuto Mannen, Tsutomu Hosouchi, Sayaka Shinpo, Moeka Kawashima, Yuki Kotake, Hiroki Yamada, Yusuke Saga, Junnosuke Otaka, Hiroshi Araya, Masaaki Koteru, Hideyuki Suzuki and Tetsuo Kushiro

J. Biol. Chem. published online October 27, 2019

Access the most updated version of this article at doi: 10.1074/jbc.RA119.009944

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts