

Discover the Maze-like Network for Glabridin Biosynthesis

Corresponding Author: Professor Chun Li

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

This study elucidated the four types of reactions in glabridin biosynthetic pathway, featured 6 functional routes from 16 theoretical possibilities, and reconstructed the biosynthetic pathway from glucose to glabridin in yeast. Noteworthy, the authors discovered a new mechanism mediated by a dynamic methylation-demethylation cycle, which challenges the traditional linear perception of secondary metabolite biosynthesis and presents a novel paradigm for natural product biosynthesis. The manuscript is novel and creative in couple of ways, such as the first demethylated enzyme reported in plants, and the surprising result of high yield of multi-route network compared with single-route pathway. The multi-route network production of glabridin in yeast mimics the possibly actual situation in plant. However, we have some comments that might help to enhance the research rigor and to increase the impact of the article in nature product biosynthesis and biomanufacturing.

Major comments:

1. It will help a broad range of reader to understand the screening process of the 16 theoretical pathways if the authors could provide more details.
2. Please supplement organelle-specific markers and co-localization experiments for key enzymes. This information will give an in-depth knowledge of these genes, and facilitate other related research of these enzymes.
3. We suggest the authors to conduct in vitro enzyme activity assays for newly identified enzymes.

Minor comments:

4. Page 3, Line 14: How were the P1-P16 pathways in Extended Data Fig 1d streamlined from the metabolic pathway network in Extended Data Fig 1b? Apparently, in Fig 1b, there are far more than 16 possible pathways for synthesizing the end-product 15.
5. Page 3, Line 24; Page 4, Line 11; Page 6, Line 16; Page 7, Line 2; Page 23, Line 3; Page 34, Line 4-6, etc.: The genus name "Glycyrrhiza" can be abbreviated.
6. Page 5, Lines 7-11: Please standardize the writing of 2'OMT, 4'OMT, I2'H, etc. with respect to the apostrophe.
7. Page 5, Line 23: In the phylogenetic tree in Extended Data Fig 2b, some taxa are only labeled at the genus level. Please uniformly label them to the species level, especially for *G. glabra*, *G. inflata*, and *G. uralensis*, which are of interest in this article.
8. Page 6, Line 1: "the overall transcriptomic differences among species were relatively minor", please note Extended Data Fig 3c.
9. Page 6, Line 3: Please supplement in the main text briefly how the candidate genes of 7 PTRs, 18 PTs, 39 OCs, and 6 DMTs were screened, and list the methods in detail in the Methods section.
10. Page 6, Line 5: "Notably, these genes are primarily distributed in open and active chromatin regions." What is the basis for this conclusion? Please provide auxiliary figures or tables to prove it.
11. Page 6, Line 12: "This analysis prioritized candidate enzymes, including 1 PTR, 7 PTs, 2 OCs, and 2 DMTs." Please note "Fig. 2d" at the end.
12. Page 7, Line 17: Based on Fig 1a, I am curious whether reactions from 26 to 11, 18 to 14, and 19 to 15 were also attempted when detecting the catalytic activity of PTR. Whether these steps are valid affects the authenticity of the glabridin maze-like network. Please supplement this in the main text.
13. Page 7, Line 22: "we further evaluated the remaining 6 PTR genes". Why are not all 6 PTR genes fully shown in Fig 3a? Do Chr06 2577684 and Chr06 2580801 represent GgPTR6 and GgPTR7?
14. Page 7, Line 26: Change "Fig. 3b" to "Fig. 3a"; Page 7, Line 29: Change "Fig. 3a" to "Fig. 3b".
15. Page 8, Line 5: Since the catalytic activity of GgPTR2 from 25 to 10 was detected, why wasn't the reaction from 8 to 25

verified based on this when designing the activity test for candidate PT enzymes?

16. Page 8, Line 9: "a reported PcM4DT can catalyze the production of 10 from 25". Considering the context, this should be corrected to "8 from 25".

17. Page 8, Line 15: "the expression of OC is highly dependent on the transcriptional activity of the upstream genes(i.e. GgI2'H1)". Is there any additional evidence to support this? This conclusion cannot be drawn solely based on the high correlation from co-expression analysis.

18. Page 8, Line 19: "GgOC1 catalyzes the oxidative cyclization of 4'-O-methylpreglabridin (10) and preglabridin (14)". Why is it presented as 10 and 14 here, and in Fig 3e as well? Instead, why weren't the activities of OC catalyzing 10 or 14 detected separately?

19. Page 9, Line 8: Change "Fig3h" to "Fig. 3h".

20. Page 9, Line 10: The localization of GgPT1 and GgOC1 is only presented through mCherry fluorescence, which remains relatively vague. If classic organelle-specific marker genes (e.g., for the nucleus, mitochondria, or endoplasmic reticulum) are supplemented and co-localization experiments are performed with the mCherry-tagged GgPT1 and GgOC1, it will enable a more accurate determination of whether they reside within the nucleus or associate with other organelles. This would provide direct evidence for their "membrane-associated" status proposed in Fig. 3g and clarify the functional implications of their spatial distribution.

21. Page 9, Line 20: Change "Extended Data Fig. 6c-e" to "Extended Data Fig. 6c-d".

22. Page 10, Line 2: The newly identified PTRs, PT, OC, and DMT presented in the main text are all results of in vivo feeding experiments in yeast. To ensure the accuracy and reliability of the enzyme activity results, it is recommended to supplement in vitro enzyme activity experiments.

23. Page 11, Line 8: Change "Box plots in (i)" to "Box plots in (h)".

24. Page 12, Lines 1-11: It was mentioned above that NhpDA1 can convert 8 to 12 with a not-low conversion rate. Why does the pathway of 8 converting to 12 not appear in Fig 4d? How to explain this phenomenon? Similarly, GgDMT1 showed broad substrate recognition above and demonstrated good catalytic activity when using substances 9, 10, and 11 as substrates. Why does it only catalyze the reaction from 9 to 13 here?

25. Page 12, Line 25: Change "Fig. 4h, S9b-c" to "Fig. 4h, Extended Data Fig. 9b-c"

26. Page 13, Line 2: Change "Fig. 4i, Extended Data Fig. d-e" to "Fig. 4i, Extended Data Fig. 10d-e".

27. Page 22, Line 25: Italicize "G. glabra".

28. Page 23, Line 12: Change "Table 5" to "Table S5"; Page 27, Line 21: Change "Table 8" to "Table S8"; Page 31, Line 3: Change "Table 9" to "Table S9".

29. Page 27, Line 6: Change "100-mL" to "100 mL".

30. Page 27, Line 28: Change "24 hr" to "24 h".

31. Page 33, Line 10: "GgVR1 can convert 2'-hydroxydihydrodaidzein (17) to Demethylmedicarpin (12)". Considering the context and Extended Data Fig. 6b, this should be corrected to "GgVR3".

Reviewer #2

(Remarks to the Author)

This work examined enzymes capable of modifying isoflavan scaffolds. Authors reconstructed a network in yeast solidified that the metabolic redundancy and interconnectivity confers unexpected robustness and enables higher production efficiencies compared to a single-route design. My major concerns comprise:

1. Can the biosynthetic pathway of glabridin inferred from the metabolic pathway search algorithm based on known enzymatic reactions and the retrosynthetic algorithm based on reaction rules represent the actual biosynthetic pathway present in licorice?
2. Since its isolation from licorice in 1976, has no researcher explored the biosynthetic pathway of glabridin using isotope tracing methods?
3. Fig. 1a shows that most intermediates were not detected. Can the synthetic pathway be simplified for clearer presentation?
4. Extended Data Fig. 2b does not explicitly state the Latin names of the three species, only indicating their genus names, making it impossible to discern their phylogenetic relationships.
5. Fig. 2d-e do not correspond to the results and figures. It is recommended to improve them in the results section. Additionally, Fig. 2e is redundant and should be deleted or placed in the appendix.
6. The authors compared the enzymatic activities of GgPTR1 and GgPTR5, observing significantly lower activity in GgPTR5 without further mechanistic explanation. It is suggested to perform functional rescue mutations on GgPTR5 to determine whether its activity can be restored to that of GgPTR1.
7. It is recommended to further analyze the specific reasons why GgPT1 is active while the other six GgPTs are not. This can be done through amino acid sequence alignment and site-directed mutagenesis.
8. The test showed that GgOC3 had no corresponding activity, but this was not described in the results section. Further analysis was then conducted on the differences in sequence and active sites between GgOC1, GgOC2, and GgOC3 to further explain the reason for the lack of activity.
9. The article analyzed the expression of two GgDMTs, but only tested the activity of GgDMT1. Why were the results of DMT2's activity not presented? Does it lack corresponding activity? If there is no activity, it is recommended that the authors include the results and explain the reasons for the difference in activity between GgDMT1 and GgDMT2.
10. It is recommended to select PTR, PT, OC, and DMT for in vivo gene silencing experiments to further determine their involvement in the biosynthesis of glabridin.
11. Inconsistent reference formats, uniform verification required.

Reviewer #3

(Remarks to the Author)

The authors have successfully elucidated the biosynthetic process of glabridin, and this work does provide a blueprint for glabridin production. However, the narrative of a "maze-like" metabolic network appears unfounded. Theoretically, there may be multiple potential pathways for producing the target compound, but in nature—particularly in plants where such compounds are enriched—evolution typically favors a specific, dominant route. Thus, the core premise of this study is questionable.

Moreover, the research only identifies several enzymes involved in the conversion from compound 8 to glabridin, and the yield of glabridin remains excessively low. As such, the study made a unreal story and lacks genuine novelty. Given these limitations, this work does not meet the criteria for publication in the journal.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have satisfactorily addressed all my comments. I recommend acceptance for publication.

Reviewer #2

(Remarks to the Author)

I have no further concern.

Open Access This Peer Review File 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.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File 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 <https://creativecommons.org/licenses/by/4.0/>

Thank you for your letter and the reviewers' valuable comments regarding our manuscript entitled "Discover the Maze-like Network for Glabridin Biosynthesis". We have made corrections to the original manuscript according to these comments and suggestions. All changes made to the text in the manuscript are marked in red. The detailed corrections are listed below, point by point. In particular, we clarified the rationale for pathway selection, expanded enzyme validation experiments, refined data presentation, and adjusted the wording to ensure accurate and balanced conclusions. We highly appreciate your time and patience in editing our manuscript. Hopefully, you will find the revised manuscript suitable for publication in *Nature Communications*.

Referee #1 (Remarks to the Author):

This study elucidated the four types of reactions in glabridin biosynthetic pathway, featured 6 functional routes from 16 theoretical possibilities, and reconstructed the biosynthetic pathway from glucose to glabridin in yeast. Noteworthy, the authors discovered a new mechanism mediated by a dynamic methylation-demethylation cycle, which challenges the traditional linear perception of secondary metabolite biosynthesis and presents a novel paradigm for natural product biosynthesis. The manuscript is novel and creative in couple of ways, such as the first demethylated enzyme reported in plants, and the surprising result of high yield of multi-route network compared with single-route pathway. The multi-route network production of glabridin in yeast mimics the possibly actual situation in plant. However, we have some comments that might help to enhance the research rigor and to increase the impact of the article in nature product biosynthesis and biomanufacturing.

We thank Reviewer #1 for the positive evaluation of our work and their constructive suggestions, which significantly improved the manuscript. Below, we address each comment in detail.

Major Comments:

1.1. It will help a broad range of reader to understand the screening process of the 16 theoretical pathways if the authors could provide more details.

Thank you for your kind suggestion. 16 theoretical routes (P1-P16) were identified by systematically traversing the compound-conversion network presented in Fig. 1a. In the revised manuscript, we provided a concise theoretical overview in the main text. Briefly, the network in Fig. 1a was derived from a comprehensive retrosynthesis-guided reaction network that was first generated through systematic reaction-rule expansion and MCTS-assisted manual extension. From this exhaustive compound-reaction space, we manually curated a subset of intermediates and transformations based on known or predictable enzymatic logic in plant specialized metabolism. Only reactions consistent with established biosynthetic mechanisms (e.g., O-methylation, hydroxylation, reduction, prenylation, cyclization, and demethylation) were retained, yielding the curated compound-conversion network shown in Fig. 1a. We additionally streamlined the network by retaining only intermediates that are endogenous to licorice. Under this criterion, the number of compounds in Fig. 1a was reduced from 33 to 23,

and the compound numbering throughout the manuscript was updated accordingly. To enumerate all theoretically feasible routes within this refined network, we constructed a directed graph using the curated intermediates and transformations. To avoid redundant branching near the early precursor **4**, the conversion from **4** to **5** was excluded during graph construction. A depth-first traversal of all possible paths from compound **1** to the final product **15** yielded 13 candidate pathways. Each pathway corresponds to a distinct sequence of reactions that could plausibly lead to the final product, glabridin, from early pathway intermediates. Fig. 1 and Extended Data Fig. 1 have been improved to better illustrate the screening process. Corresponding revisions have also been made in the Results section (Page 3, Lines 14–24), the Methods section (Page 22, line 9–22) and the caption of Extended Data Fig. 1 (Page 33, line 7–Page 34, line 5) to clarify these procedures for readers.

*Page 3, line 14–24: The reaction-rule-based pathway search is designed to enumerate all biochemically plausible routes supported by known enzyme reaction types, including both pathways that could represent the biosynthetic route in licorice and others that licorice may not actually use. To focus on routes relevant to licorice, the network is further refined by retaining only intermediates detected in licorice (Fig. 1a). Using this curated compound–conversion network, we constructed a directed graph and performed a systematic depth-first traversal, which identified 13 potential biosynthetic pathways leading to the production of glabridin, involving L-phenylalanine (**1**) as precursor and 21 intermediate metabolites (Fig. 1a, Extended Data Fig. 1d). Among these, 9 compounds are located at pathway branch nodes, and over 69% of the predicted pathways include more than four branch nodes.*

Page 22, line 9–22: Pathway screening and curation

*A comprehensive retrosynthesis-guided reaction network was generated through systematic reaction-rule expansion and MCTS-assisted manual extension, as described above. From this exhaustive compound-reaction space, we manually curated a subset of intermediates and transformations based on known or predictable enzymatic logic in plant specialized metabolism. Only reactions consistent with established biosynthetic mechanisms (e.g., O-methylation, hydroxylation, reductation, prenylation, cyclization, demethylation) were retained. To improve biochemical relevance, we further filtered intermediates by requiring that they have been detected or reported in licorice. Application of this criterion yielded the refined network. To enumerate all theoretically feasible routes in network, we constructed a directed graph using its curated intermediates and transformations. To avoid redundant branching near the early-stage precursor **4**, the conversion from **4** to **5** was excluded during graph construction. We then performed a depth-first search to traverse all possible paths from compound **1** to the final product **15**, resulting in 13 candidate pathways.*

*Page 33, line 7–Page 34, line 5: d) Summary of Retrobiosynthesis results for the glabridin. Blue nodes represent intermediate compounds, and the red node indicates the final product, glabridin. The 13 candidate pathways (P1–P13) were obtained by traversing the refined reaction subgraph in Fig. 1a, which was manually curated from the broader retrosynthesis-derived network in Extended Data Fig. 1b and further streamlined by retaining only intermediates endogenous to licorice. To avoid redundant branching of early pathways, the conversion from **4** to **5** was omitted.*

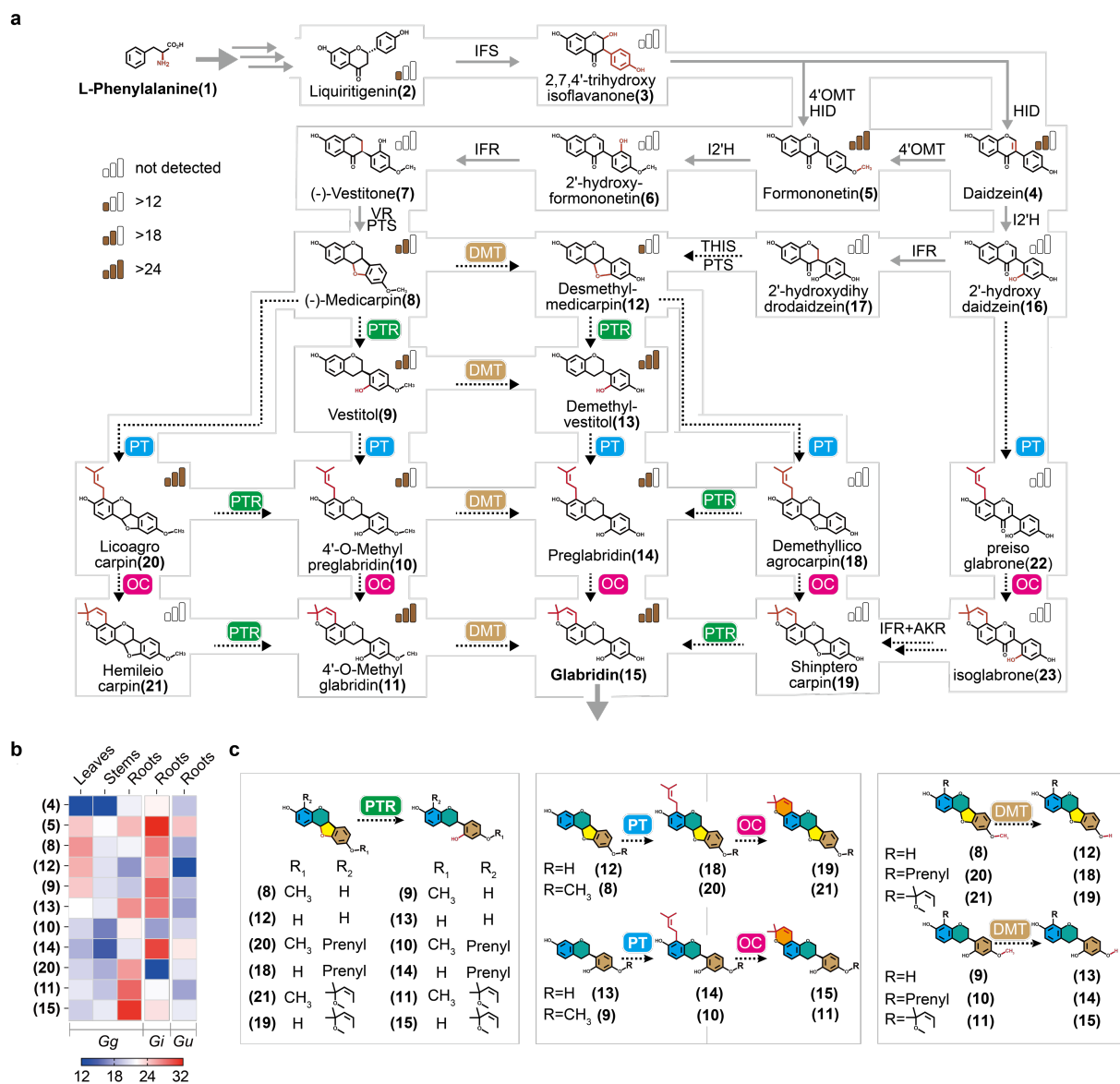
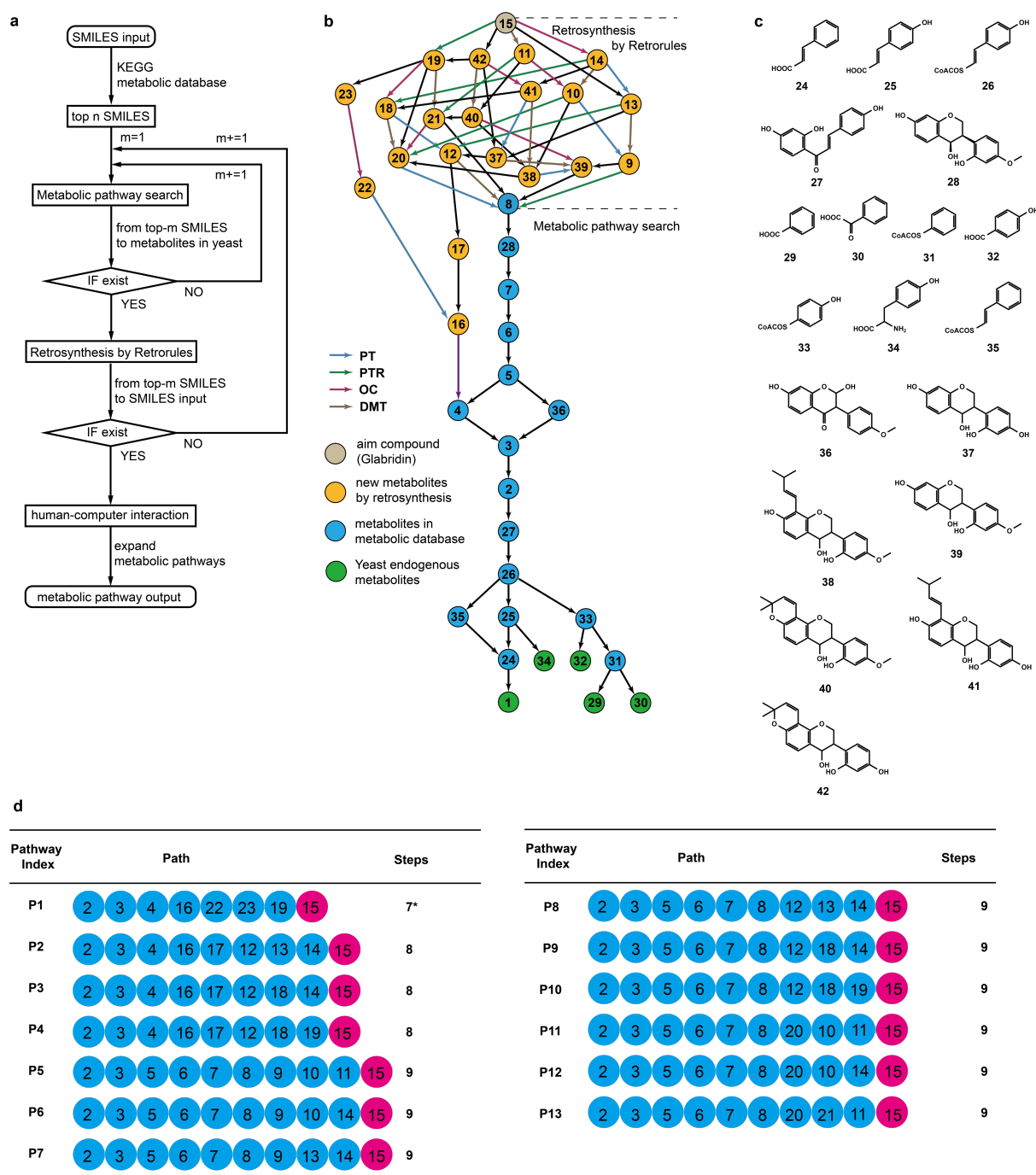


Fig. 1



Extended Data Fig. 1

1.2. Please supplement organelle-specific markers and co-localization experiments for key enzymes. This information will give an in-depth knowledge of these genes, and facilitate other related research of these enzymes.

We appreciate this valuable suggestion. In response, we performed co-localization experiments using Sec61-mNeonGreen (P.N.A.S. 2020, 117(31)), a well-established marker of the endoplasmic reticulum (ER), to determine the subcellular localization of GgPT1 and GgOC1 (Fig. S8). Confocal imaging revealed clear fluorescence overlap, indicating that both enzymes are partially associated with the ER membrane. These results substantiate their proposed

membrane localization and provide additional spatial insight into their catalytic roles. The corresponding descriptions (Page 9, line 34–Page 10, line 2), methodological details (Page 28, line 8–13) and the caption of Fig. S8 (Supplementary information Page 7, line 16–20) have been incorporated into the revised manuscript.

Page 9, line 34–Page 10, line 2: Furthermore, subcellular localization studies showed that GgDMT1 is cytoplasmic, whereas GgOC1 and GgPT1 co-localized with Sec61–mNeonGreen36, indicating that both enzymes are localized to the endoplasmic reticulum (Fig. 3g, Fig. S8).

Page 28, line 8–13: Subcellular localization and co-localization analysis

*For subcellular localization, the coding sequences of GgPT1 and GgOC1 were cloned into 2 μ -based multicopy plasmids under the control of constitutive promoters P_{THD3} and fused to fluorescent tags mScarlet. The constructs were transformed into a *Saccharomyces cerevisiae* strain integrated with Sec61–mNeonGreen, a canonical endoplasmic reticulum (ER) marker. Co-localization was analyzed using a confocal fluorescence microscope.*

Supplementary information Page 7, line 16–20: Supplementary Figure 8. Yeast confocal microscopy. Confocal images of yeast cells expressing mScarlet-GgPT1 (a) and GgOC1-mScarlet (b), co-localized with the ER marker Sec61-mNeonGreen. Merged images show co-localization of the red and green signals, indicating ER localization. Scale bar, 5 μ m.

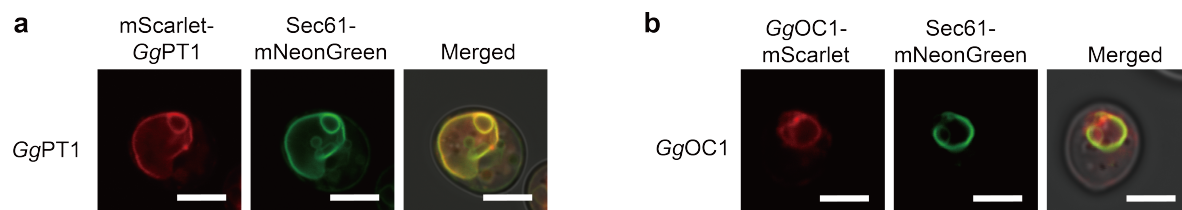


Fig. S8

1.3. We suggest the authors to conduct *in vitro* enzyme activity assays for newly identified enzymes.

We thank the reviewer for this valuable suggestion. In response, we performed *in vitro* enzyme activity assays for GgPTR1, GgPTR2 and GgDMT1 using purified recombinant proteins, while yeast microsomal fractions were used for the membrane-associated enzymes GgPT1 and GgOC1. Clear catalytic activities were detected for GgPTR1, GgPTR2, GgOC1, and GgDMT1 toward their corresponding substrates (Fig. S2). These results are consistent with the *in vivo* yeast assays and further strengthen the functional validation of the proposed biosynthetic pathway. For GgPT1, however, no detectable activity was observed in the microsomal system. We note that GgPT1 contains nine predicted transmembrane domains, whereas GgOC1 possesses only an N-terminal signal peptide (Fig. S9) and remains active in the same system. Multi-pass plant prenyltransferases are known to lose activity outside their native membrane environment due to improper folding, lack of essential lipids, or absence of required partner proteins (Current Opinion in Structural Biology 2001, 11(4); Biochemical Journal 2009, 421(2); Applied Microbiology and Biotechnology 2024, 108(1)). Thus, the lack of *in vitro* activity likely reflects technical limitations of reconstitution rather than the absence of catalytic function. Importantly, GgPT1 exhibits robust and reproducible activity *in vivo*, which we

consider strong functional evidence for its role in the pathway. This limitation and its biochemical rationale have been explicitly discussed in the revised manuscript (Page 16, line 12–16). The new *in vitro* assay data (Page 7, line 34–Page 8, line 3; Page 9, line 4–6; Page 9, line 17–18), methodological details (Page 28, line 27–Page 29, line 5; Page 28, line 7–20; Page 28, line 22–Page 29, line 7) and the caption of Fig. S2 (Supplementary information Page 3, line 2–7) have been incorporated into the revised manuscript.

Page 7, line 34–Page 8, line 3: To further validate their activities, we performed in vitro assays for both GgPTR1 and GgPTR2 (Fig. S2a-e). Both enzymes showed clear catalytic activity toward their respective substrates. These results are consistent with the in vivo yeast assays and further confirm the functions of the two enzymes.

Page 9, line 4–6: The catalytic activity of GgOCL was further validated in vitro using yeast microsomes (Fig. S2f-g), and the results were consistent with the in vivo yeast assays, confirming the function of GgOCL.

Page 9, line 17–18: The activity of GgDMT1 was further confirmed in vitro using purified protein (Fig. S2h-l).

Page 16, line 12–16: In addition, although we attempted to reconstitute GgPT1 activity in vitro using yeast microsomes, no measurable activity was detected. As multi-pass aromatic prenyltransferases are prone to loss of activity outside their native membrane environment⁴⁵, this likely reflects limitations of the current reconstitution system rather than a lack of catalytic function.

Page 28, line 27–Page 29, line 5: Protein expression and purification

Recombinant GgPTR1, GgPTR2 and GgDMT1 were cloned into the pET28a vector using Gibson assembly and transformed into E. coli BL21. Overnight seed cultures grown in TB medium supplemented with kanamycin were used to inoculate 1-L TB cultures, which were shaken at 37 °C to an OD600 of ~0.6. Protein expression was induced with 0.5 mM IPTG, followed by incubation at 16 °C for 20 h. Cells were harvested by centrifugation and resuspended in lysis buffer (50 mM Tris-HCl, 300 mM NaCl, 5% glycerol, 10 mM imidazole, pH 7.5). After sonication, the lysate was clarified (10,000 g, 20 min, 4 °C) and the supernatant was incubated with Ni-NTA resin for 20 min at 4 °C. Bound proteins were washed with lysis buffer and eluted with elution buffer containing 500 mM imidazole. Imidazole was removed by ultrafiltration (10 kDa cutoff) with two rounds of dialysis against storage buffer (20 mM HEPES, 150 mM NaCl, 20% glycerol, 1 mM DTT, pH 7.4). The purified proteins were used directly for enzymatic assays.

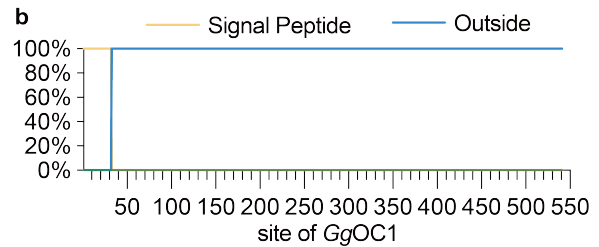
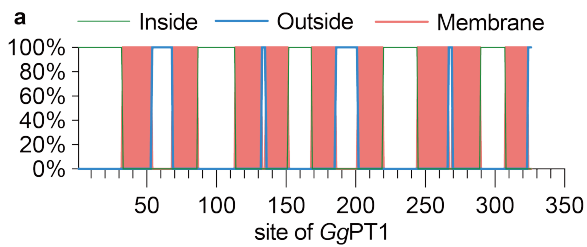
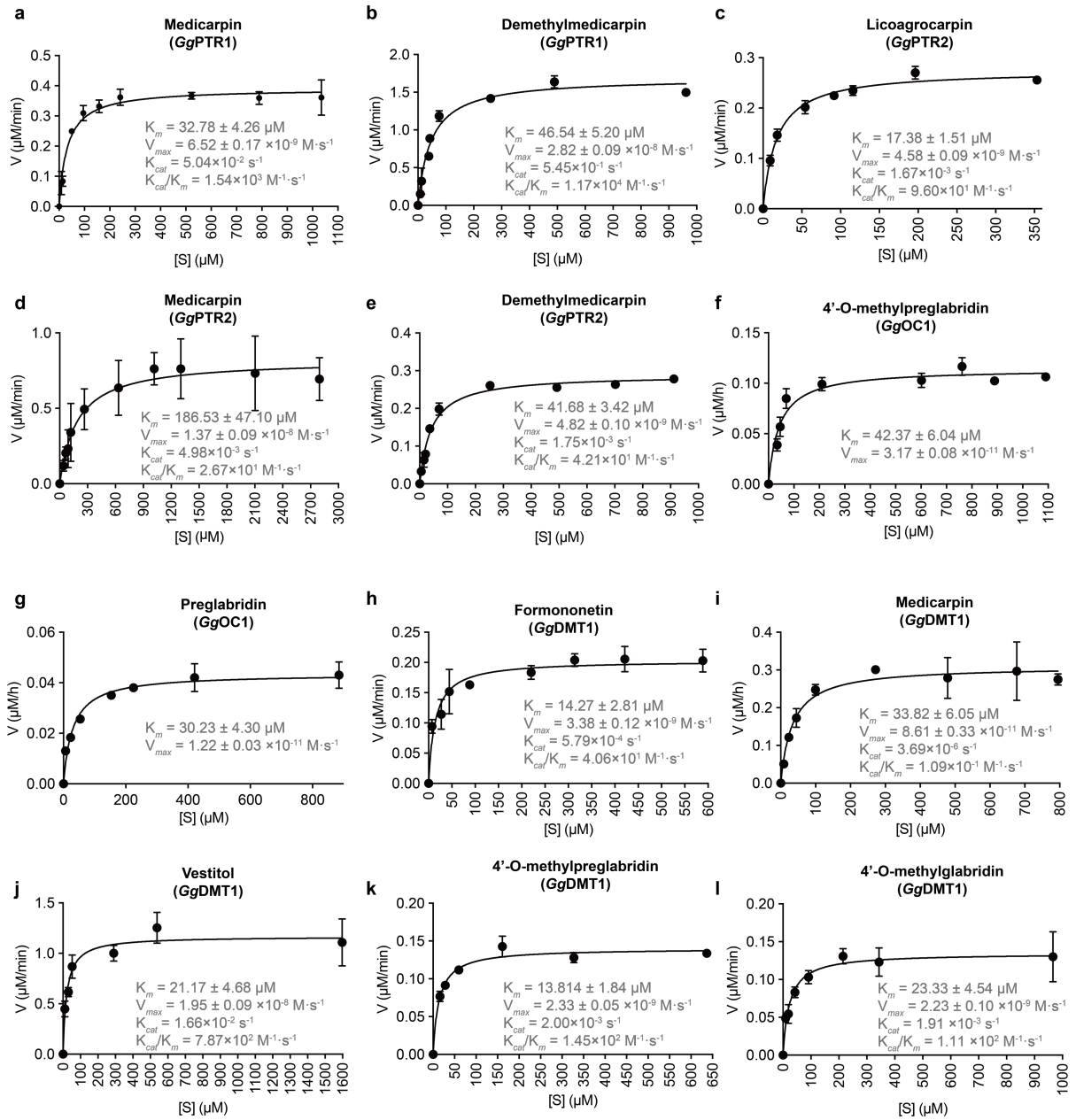
Page 28, line 7–20: Microsomes preparation

Microsomal fractions containing GgOCL were prepared from S. cerevisiae expressing the corresponding constructs. Seed cultures were grown in YPD medium at 30 °C overnight and used to inoculate 20 mL and subsequently 1 L YPD cultures, which were incubated for 2 days at 30 °C with shaking. Cells were harvested by centrifugation (4,000 × g, 10 min, 4 °C), gently resuspended in 100 mL ice-cold extraction buffer (20 mM Tris-HCl pH 7.5, 0.5 M sucrose, 5 mM MgCl₂, 1 mM DTT, 1 mM PMSF), incubated for 10 min at room temperature, and pelleted again under the same conditions. The cell pellet was washed once more with extraction buffer and subjected to high-pressure homogenization (three cycles at 1,200 bar) on ice. The homogenate was clarified by centrifugation (10,000 × g, 10 min, 4 °C), and the resulting supernatant was ultra-centrifuged at 100,000 × g for 1 h at 4 °C to collect the microsomal membrane fraction. The microsomal pellet was gently resuspended in 5 mL storage buffer (100 mM Tris-HCl pH 7.5, 0.25 M sucrose, 10 mM MgCl₂, 20% glycerol, 1 mM DTT), aliquoted, flash-frozen in liquid nitrogen, and stored at -80 °C until use.

Page 28, line 22–Page 29, line 7: Enzymatic assays

GgPTR1 and GgPTR2 were assayed in 50 mM Tris-HCl (pH 7.5), supplemented with 200 μ M NADPH under initial-rate conditions. For GgPTR1, medicarpin (8, 10–1000 μ M) and demethylmedicarpin (12, 10–1000 μ M) were assayed using 0.5 μ g purified enzyme. For GgPTR2, medicarpin (8, 40–3000 μ M), demethylmedicarpin (12, 10–1000 μ M), and licoagrocarpin (10–400 μ M) were assayed under initial-rate conditions. GgDMT1 were performed in 50 mM Tris-HCl (pH 7.5), supplemented with 100 mM NaCl, 1 mM 2-oxoglutarate, 1 mM ascorbate, 1 mM TCEP, and 100 μ M FeSO₄. Five methylated substrates were evaluated under initial-rate conditions with substrate ranges and enzyme amounts as follows: 4'-O-methylglabridin (11, 10–1000 μ M, 0.5 μ g enzyme); 4'-O-methylpreglabridin (10, 10–700 μ M, 10 μ g enzyme); vestitol (9, 10–1600 μ M, 5 μ g enzyme); medicarpin (8, 10–800 μ M, 50 μ g enzyme); and formononetin (5, 10–600 μ M, 25 μ g enzyme). Microsomal GgOC1 (5 μ L) was assayed using 4'-O-methylpreglabridin (10, 30–1000 μ M) or preglabridin (14, 10–1000 μ M) as substrates in 50 mM Tris-HCl (pH 7.5). All kinetic assays were performed in 100 μ L reaction mixtures at 30 °C, and reactions were quenched with an equal volume of ice-cold methanol, clarified by centrifugation (10,000 $\times$ g, 10 min), and analyzed by UPLC using a modified gradient: 10% B (0 min), 10–20% B (0–2.5 min), 20–60% B (2.5–10 min), 60–100% B (10–12.5 min), 100% B (12.5–13.5 min), 100–10% B (13.5–15 min), and 10% B (15–16 min). Kinetic parameters (K_m , V_{max} , k_{cat} , and k_{cat}/K_m) were obtained by nonlinear regression based on three independent replicates.

Supplementary information Page 3, line 2–7: Supplementary Figure 2. In vitro enzyme activity. Michaelis–Menten kinetic curves for purified recombinant enzymes GgPTR1 (a, b), GgPTR2 (c–e), GgDMT1 (h–l), and for yeast microsomal preparations expressing GgOC1 (f, g). Kinetic parameters (K_m , V_{max} , k_{cat} , and k_{cat}/K_m) were obtained by nonlinear regression fitting. Data represent mean $\pm$ s.d. from three independent biological replicates.



Minor comments:

1.4. Page 3, Line 14: How were the P1-P16 pathways in Extended Data Fig 1d streamlined from the metabolic pathway network in Extended Data Fig 1b? Apparently, in Fig 1b, there are far more than 16 possible pathways for synthesizing the end-product 15.

We thank the reviewer for this insightful comment. As noted in our response to point 1.1, the P1-P16 pathways were obtained by systematically traversing the reaction network shown in Fig. 1a, which itself represents a curated subgraph derived from the full compound-reaction space in Extended Data Fig. 1b. These 16 pathways were chosen to represent the most biochemically plausible and non-redundant reaction routes from early precursors to the final product (compound 15), based on compound connectivity, structural feasibility, and enzymatic logic.

Importantly, Extended Data Fig. 1b includes a much broader and more exhaustive network, generated through a retrosynthesis-guided reaction expansion. From this comprehensive network, we manually selected and refined the intermediate compounds and reactions shown in Fig. 1a, focusing on those with known or predictable enzymatic transformations. The traversal of this refined network yielded the 16 theoretical pathways (P1-P16) presented in Extended Data Fig. 1d.

To further improve clarity, we additionally retained only intermediates known to occur in licorice, which narrowed the candidate set to thirteen representative pathways. These clarifications have been added in the Results section (Page 3, Lines 14–24), the Methods section (Page 22, line 9–22) and the caption of Extended Data Fig. 1 (Page 33, line 7–Page 34, line 5) to clarify these procedures for readers.

1.5. Page 3, Line 24; Page 4, Line 11; Page 6, Line 16; Page 7, Line 2; Page 23, Line 3; Page 34, Line 4–6, etc.: The genus name "*Glycyrrhiza*" can be abbreviated.

We appreciate the reviewer's careful suggestion. The genus name *Glycyrrhiza* has been abbreviated to *G.* where appropriate throughout the manuscript, as following:

Page 2, line 23-25: Currently, the global supply chain for glabridin depends on extraction and purification from roots of the rare plant G. glabra,

Page 3, line 29-30: we conducted metabolomic profiling of three Glycyrrhiza species: G. glabra, G. uralensis, and G. inflata (Fig. 1a-b).

Page 3, line 3: including G. glabra (Gg), G. uralensis (Gu), and G. inflata (Gi).

Page 7, line 2–3: Genomic distribution of biosynthetic enzyme candidates for glabridin in G. glabra.

Page 7, line 4–5: The transcriptional profiles of these genes across five organs of G. glabra.

Page 17, line 8: which is essential for the wildlife conservation of rare G. glabra.

Page 22, line 24: Samples of G. glabra were collected from the Alar region of Xinjiang, China.

Page 23, line 21–22: A heatmap was generated to visualize the transcriptional profiles of candidate genes across five organs of G. glabra.

Page 26, line 1–2: Dried roots, stems, and leaves of G. glabra were ground into coarse powder under liquid nitrogen

Page 38, line 3–4: Locations of target genes in the G. glabra, G. inflata, and G. uralensis genomes and their corresponding evolutionary relationships.

1.6. Page 5, Lines 7–11: Please standardize the writing of 2'OMT, 4'OMT, I2'H, etc. with respect to the apostrophe.

The notation of all enzyme abbreviations containing apostrophes, such as 2'OMT, 4'OMT, and I2'H, has been standardized throughout the manuscript using the proper prime symbol (') for consistency and clarity, as following:

Page 5, line 11–14: *isoflavone synthase (IFS)*, *2,7,4'-trihydroxyisoflavanone 4'-O-methyltransferase (4'OMT)*, *2-hydroxyisoflavanone dehydratase (HID)*, *isoflavone 2'-hydroxylase (I2'H)*, *isoflavone reductase (IFR)*, *vestitone reductase (VR)*, *pterocarpan synthase (PTS)*, and *2'-hydroxydihydrodaidzein synthase (THIS)*.

Page 7, line 8–12: *Co-expression analysis of the G. glabra transcriptome using GgI2'H1 as a bait gene. The co-expressed genes were ranked based on their Pearson's correlation coefficient (PCC) with GgI2'H1. Genes shown in black are newly identified glabridin biosynthetic genes (this study) (Table S1). Homologs of the blue-labeled genes (I2'H, IFR, VR, and PTS) were previously identified²⁰.*

Page 7, line 32: *GgPTR2 can convert licoagrocarpin (25) to 4'-O-methylpreglabridin (10).*

Page 8, line 12–14: *catalyzing the conversion of compounds 9 and 13 into 4'-O-methylpreglabridin (10) and preglabridin (14) (Fig. 3d).*

Page 8, line 28–29: *the expression of OC shows a correlation with the transcriptional activity of the upstream genes(i.e. GgI2'H1),*

Page 8, line 31–33: *Indeed, functional assays of the two OCs demonstrated that only GgOC1 catalyzes the oxidative cyclization of 4'-O-methylpreglabridin (10) and preglabridin (14), yielding 4'-O-methylglabridin (11) and glabridin (15) (Fig. 3e)..*

Page 9, line 15–16: *we found GgDMT1 shows high promiscuity and efficiently demethylates both vestitol (9) and 4'-O-methylpreglabridin (10)..*

Page 9, line 22–23: *To further expand the enzyme toolbox for specific 4'-O-demethylation,.*

Page 9, line 26–27: *displaying promiscuity as well across toward multiple 4'-O-methylated substrates (5, 8, 9, 10, 11) (Fig. 3f)..*

Page 10, line 8–9: *we found that the I2'H1–IFR1–VR1 module is capable of catalyzing the conversion from intermediate 4 to 12.*

Page 15, line 16–17: *four more enzymes-GgI2'H1, GgIFR1, GgVR1, and GgPTS1-were expressed in yeast.*

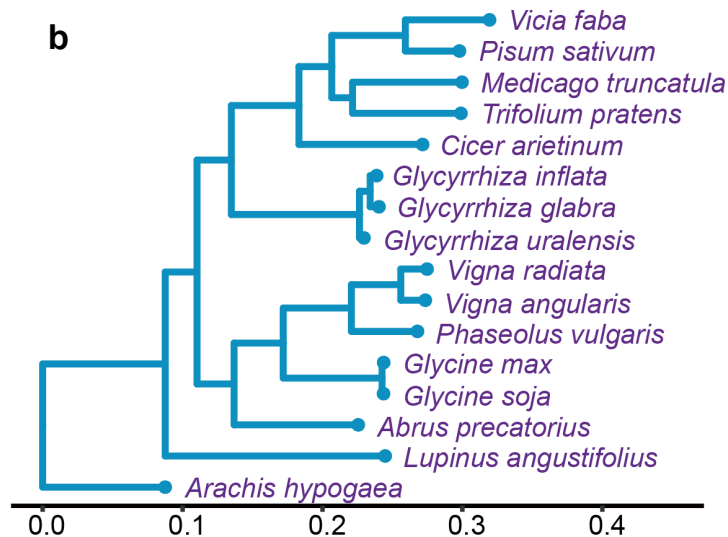
Page 21, line 34–35: *4'-O-methylglabridin and 4'-O-methylpreglabridin were custom-synthesized by Acme Biopharma Co., Ltd. (Shanghai, China).*

Page 24, line 24–25: *Pearson correlation analysis was performed using GgI2'H1 as the bait gene.*

Page 37, line 3–8: *a) Stepwise characterization of the GgI2'H1–GgIFR1–GgVR1–GgPTS1 module in yeast with 50 mg/L of the methylated substrate formononetin (5). b) Functional characterization of GgI2'H1–GgIFR1 combined with different vestitone reductases (VRs) for the conversion of methylated (5) and unmethylated (4) substrates. GgVR3 can convert 2'-hydroxydihydrodaidzein (17) to Demethylmedicarpin (12).*

1.7. Page 5, Line 23: In the phylogenetic tree in Extended Data Fig 2b, some taxa are only labeled at the genus level. Please uniformly label them to the species level, especially for *G. glabra*, *G. inflata*, and *G. uralensis*, which are of interest in this article.

The Latin names of the three *Glycyrrhiza* species have been added to Extended Data Fig. 2b to clearly indicate their phylogenetic relationships.



Extended Data Fig. 2b

1.8. Page 6, Line 1: "the overall transcriptomic differences among species were relatively minor", please note Extended Data Fig 3c.

The citation of Extended Data Fig. 3c has been corrected and repositioned for clarity. It now appears directly after the relevant sentence describing interspecies transcriptomic comparisons, ensuring accurate reference to the supporting data.

Page 6, line 33–Page 7, line 3: The clustering analysis revealed that the organ-specific transcriptome profiles group mainly into three clusters: 1) tap roots and horizontal roots, 2) stems and sprouts, 3) leaves. In contrast, the overall transcriptomic differences among species were relatively minor (Extended Data Fig. 3c), with variations observed only in certain specific genes.

1.9. Page 6, Line 3: Please supplement in the main text briefly how the candidate genes of 7 PTRs, 18 PTs, 39 OCs, and 6 DMTs were screened, and list the methods in detail in the Methods section.

We appreciate the reviewer's suggestion. A concise description has been added to the main text (Page 6, line 3–6), and a detailed workflow is now provided in the Methods section (Page 24, line 9–19). Briefly, we first used well-characterized enzymes from public databases as queries to search the *G. glabra* genome and compiled candidates based on functional annotation. Each enzyme class was then refined by phylogenetic analysis and domain/motif inspection: reductases were assigned into IFR/VR/PLR/PTR clades; PTs were enriched based on homology to known aromatic prenyltransferases with two conserved motifs; OCs were selected from BBE-like and P450 subfamilies associated with oxidative cyclization; and DMT candidates were drawn from dioxygenase and P450 implicated in O-demethylation. A summary of gene definitions and annotations is now provided in Table S1.

Page 6, line 3–6: Candidate genes were identified through a multi-layered screening strategy that combined sequence-based functional annotation and subsequent refinement using

phylogenetic clustering and conserved-motif analysis (Table S1). This integrative process yielded 7 PTRs, 18 PTs, 39 OCs, and 6 DMTs for downstream analysis.

Page 24, line 9–19: Candidate gene identification

*Candidate pathway genes were identified through a stepwise sequence-based workflow. First, functionally characterized enzymes reported for related reactions were compiled from public databases and used as queries for homology searches against the *G. glabra* genome. Reductase candidates were drawn from IFR/VR/PLR/PTR families and classified by phylogenetic analysis. Prenyltransferase candidates were filtered by homology to known aromatic PTs and the presence of conserved UbiA-type motifs (NQ--D---D, KD--D--GD). Oxidative cyclase candidates were selected from BBE-like proteins and P450 subfamilies (CYP71D, CYP82A, CYP736). O-demethylase candidates were obtained from Fe²⁺/2OG-dependent dioxygenases and P450 families implicated in O-demethylation (CYP80G, CYP82D, CYP71BE). Final gene sets are summarized in Table S1.*

1.10. Page 6, Line 5: "Notably, these genes are primarily distributed in open and active chromatin regions." What is the basis for this conclusion? Please provide auxiliary figures or tables to prove it.

We thank the reviewer for this comment. As it is not directly supported by experimental evidence and is not necessary to this study, we have removed this conclusion from the revised manuscript to maintain accuracy and focus on the storyline.

1.11. Page 6, Line 12: "This analysis prioritized candidate enzymes, including 1 PTR, 7 PTs, 2 OCs, and 2 DMTs." Please note "Fig. 2d" at the end.

We thank the reviewer for the careful reading. This result corresponds to the root-enriched cluster highlighted in red in Fig. 2c, which summarizes the transcriptional profiles of candidate genes across five organs of *G. glabra*. Within this root-expressed cluster, we identified 1 PTR, 7 PTs, 2 OCs, and 2 DMTs. To make this information clearer to readers, we have added Table S2 to list the detailed gene composition of this cluster and have updated the main text (Page 6, line 10–12) to cite Table S2 accordingly.

Page 6, line 10–12: The clustering revealed a root-enriched gene expression cluster, including 1 PTR, 9 PTs, 8 OCs, and 2 DMTs (Fig. 2c, Table S2).

1.12. Page 7, Line 17: Based on Fig 1a, I am curious whether reactions from 26 to 11, 18 to 14, and 19 to 15 were also attempted when detecting the catalytic activity of PTR. Whether these steps are valid affects the authenticity of the glabridin maze-like network. Please supplement this in the main text.

We appreciate the reviewer's insightful comment. Among the reactions mentioned, the conversion from 18 to 14 was indeed examined (Fig. 4b) by co-expressing PcM4DT and GgPTR2. In contrast, from 26 to 11 and from 19 to 15 could not be examined because the substrates 26 and 19 are not commercially available, and we found that GgOC1 cannot catalyze 25 to 26 or 18 to 19 (Fig. 4b). We agree that validation of these reactions would further enhance the authenticity of the proposed glabridin maze-like network. However, the inability to perform

these assays (from **26** to **11** and from **19** to **15**) is due to experimental constraints, rather than evidence against the existence of these branches. Importantly, both compounds **26** and **19** have been reported in *G. glabra* (Chem. Pharm. Bull. 1994, 42(5); J. Agric. Food Chem. 2007, 55(12)), supporting the biological plausibility of these routes. Thus, while not all branches could be experimentally validated, the network architecture remains consistent with known plant chemistry and available biological evidence. This clarification has been added to the revised main text (Page 16, line 28–32). We also note that in the updated Fig. 1a, compound **26** has been renumbered as **21**, and the text has been revised accordingly.

Page 16, line 28–32: We agree that validation of these reactions would further enhance the authenticity of the proposed glabridin maze-like network. However, the inability to perform these assays reflects experimental constraints rather than evidence against the existence of these branches. Nonetheless, the presence of key intermediates such as compounds 19 and 21 in G. glabra supports the biochemical plausibility of these branches^{46,47}.

1.13. Page 7, Line 22: "we further evaluated the remaining 6 PTR genes". Why are not all 6 PTR genes fully shown in Fig 3a? Do Chr06 2577684 and Chr06 2580801 represent GgPTR6 and GgPTR7?

We thank the reviewer for pointing this out. Chr06 2577684 and Chr06 2580801 really correspond to GgPTR6 and GgPTR7, respectively, and these have now been labelled in Fig. 3a in the revised manuscript for clarity. For consistency, we have also added the corresponding labels for the GgPT genes in Fig. 3d.

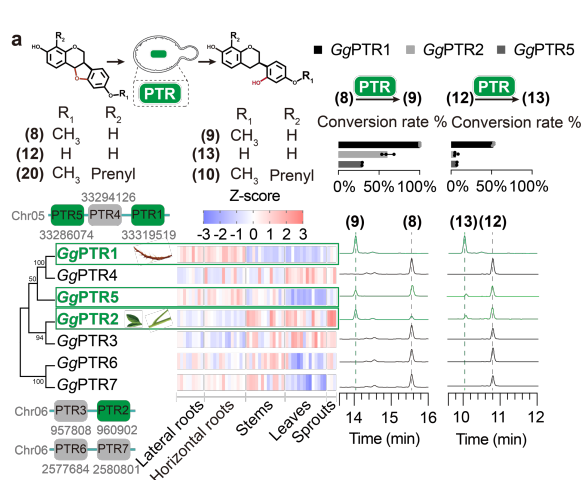


Fig. 3a

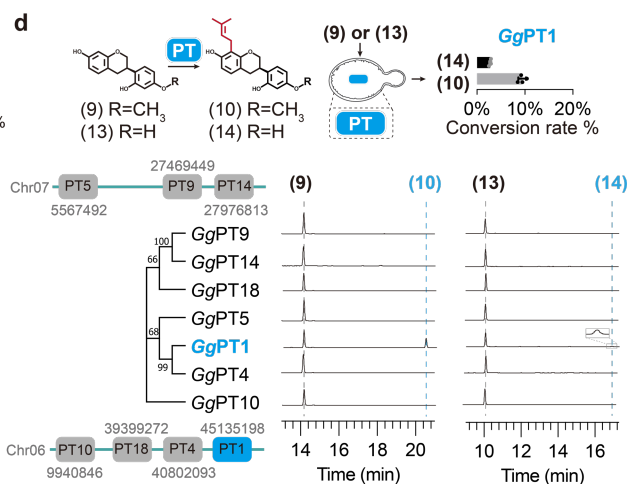


Fig. 3d

1.14. Page 7, Line 26: Change "Fig. 3b" to "Fig. 3a"; Page 7, Line 29: Change "Fig. 3a" to "Fig. 3b".

The figure references have been corrected accordingly, as following:

Page 7, line 29–31: with conversion rates of $58.8 \pm 7.6\%$ for medicarpin (8) and $4.7 \pm 2.5\%$ for demethylmedicarpin (12), respectively (Fig. 3a).

Page 7, line 32–34: GgPTR2 can convert licoagrocarpin (20) to 4'-O-methylpreglabridin (10), whereas GgPTR1 shows no activity, indicating selective recognition of prenylated substrates by GgPTR2 (Fig. 3b).

1.15. Page 8, Line 5: Since the catalytic activity of GgPTR2 from 25 to 10 was detected, why wasn't the reaction from 8 to 25 verified based on this when designing the activity test for candidate PT enzymes?

We thank the reviewer for pointing this out. The reaction from medicarpin (8) to licoagrocarpin (25) was tested, however, all 7 candidate PTs showed no detectable activity for this conversion. The complete results, together with the revised caption of Fig. S4, have been provided in the Supplementary Information (Supplementary information Page 5, line 2–7). The corresponding description has also been updated in the main text (Page 8, line 19–20).

Supplementary information Page 5, line 2–7: Supplementary Figure 4. Functional validation of GgPT homologs toward medicarpin. HPLC analysis of reactions catalyzed by GgPT1 and six GgPT homologs (GgPT4, GgPT5, GgPT9, GgPT10, GgPT14, GgPT18) with 20 mg/L medicarpin as the substrate. No detectable conversion or product formation was observed for any homolog, indicating that none of the tested GgPTs catalyze prenylation of medicarpin.

Page 8, line 19–20: Moreover, we tested the activity of all seven GgPTs toward compound 8 and found that none were capable of catalyzing its conversion to compound 20 (Fig. S4).

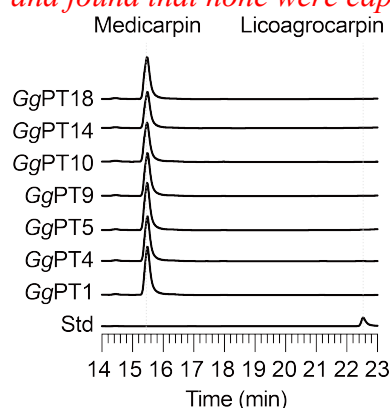


Fig. S4

1.16. Page 8, Line 9: "a reported PcM4DT can catalyze the production of 10 from 25". Considering the context, this should be corrected to "8 from 25".

We thank the reviewer for catching this error. The sentence has been corrected to "8 from 25" as suggested. We note that in the revised manuscript, the compound numbering in Fig. 1a has been updated, and the compound previously labelled as 25 is now numbered 20. The text has been revised accordingly to reflect this updated numbering (Page 8, line 20–22).

Page 8, line 20–22: In contrast, a previously reported enzyme, PcM4DT, was able to catalyze the reaction from 8 to 20²².

1.17. Page 8, Line 15: "the expression of OC is highly dependent on the transcriptional activity of the upstream genes(i.e. GgI2'H1)". Is there any additional evidence to support this? This conclusion cannot be drawn solely based on the high correlation from co-expression analysis.

We thank the reviewer for this insightful comment. We agree that co-expression analysis alone is insufficient to establish a direct regulatory relationship. Accordingly, we have revised the

statement in the manuscript (Page 8, line 27–30) to avoid implying causality and to more accurately reflect the correlation observed.

Page 8, line 27–30: Interestingly, the GgOC1 exhibits an asymmetric co-expression pattern (Table S4): the expression of OC shows a correlation with the transcriptional activity of the upstream genes (i.e. GgI2 H1), whereas the upstream genes show relatively weak association with OC expression.

1.18. Page 8, Line 19: "GgOC1 catalyzes the oxidative cyclization of 4'-O-methylpreglabridin (10) and preglabridin (14)". Why is it presented as 10 and 14 here, and in Fig 3e as well? Instead, why weren't the activities of OC catalyzing 10 or 14 detected separately?

We thank the reviewer for the helpful clarification. The label "(10)+(14)" in Fig. 3e was incorrect; the activities were tested separately. We have corrected the description to "(10) or (14)" in both the main text (Page 8, line 31–33) and the corresponding figure to accurately reflect the experiments.

Page 8, line 31–33: Indeed, functional assays of the two OCs demonstrated that only GgOC1 catalyzes the oxidative cyclization of 4'-O-methylpreglabridin (10) or preglabridin (14), yielding 4'-O-methylglabridin (11) or glabridin (15) (Fig. 3e).

Figure 3e: substrate label corrected to "(10) or (14)".

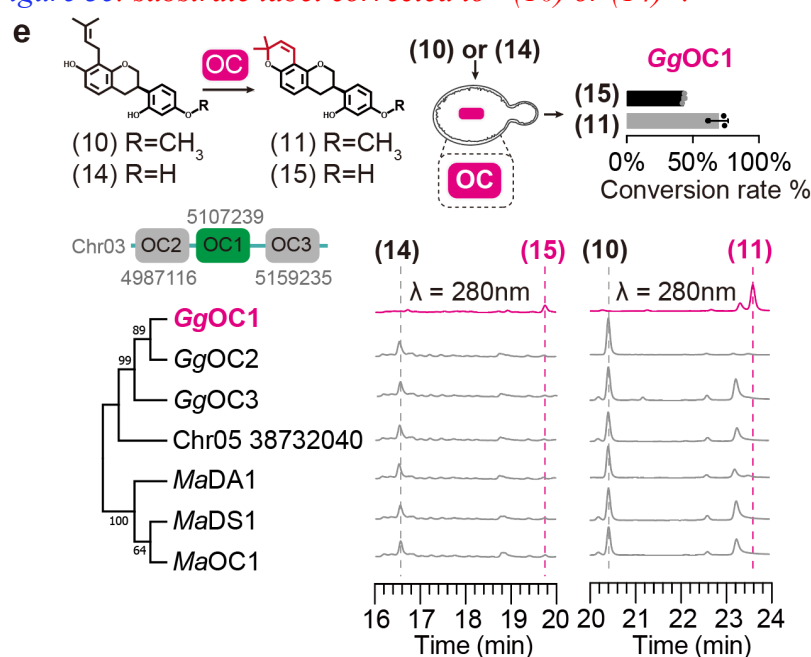


Fig. 3e

1.19. Page 9, Line 8: Change "Fig3h" to "Fig. 3h"

We thank the reviewer for pointing this out. The typographical error has been corrected to "Fig. 3h".

Page 9, line 33: (Extended Data Fig. 5, Fig. 3h).

1.20. Page 9, Line 10: The localization of GgPT1 and GgOC1 is only presented through mCherry fluorescence, which remains relatively vague. If classic organelle-specific marker genes (e.g., for the nucleus, mitochondria, or endoplasmic reticulum) are supplemented and co-localization experiments are performed with the mCherry-tagged GgPT1 and GgOC1, it will enable a more accurate determination of whether they reside within the nucleus or associate with other organelles. This would provide direct evidence for their “membrane-associated” status proposed in Fig. 3g and clarify the functional implications of their spatial distribution.

We thank the reviewer for this valuable suggestion. As described in Response 1.2, we have performed co-localization experiments using Sec61–mNeonGreen, a classical ER marker, to determine the subcellular localization of GgPT1 and GgOC1. The results confirmed their partial localization to the ER membrane (Fig. S8, *Page 9, line 34–Page 10, line 2*), providing direct evidence for their membrane association.

Page 9, line 34–Page 10, line 2: Furthermore, subcellular localization studies showed that GgDMT1 is cytoplasmic, whereas GgOC1 and GgPT1 co-localized with Sec61–mNeonGreen³⁶, indicating that both enzymes are localized to the endoplasmic reticulum (Fig. 3g, Fig. S8).

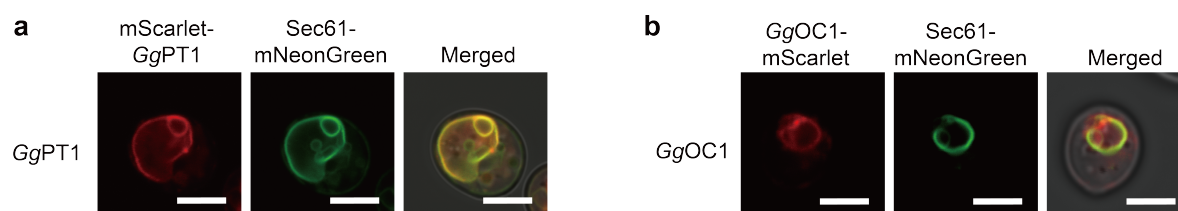


Fig. S8

1.21. Page 9, Line 20: Change "Extended Data Fig. 6c-e" to "Extended Data Fig. 6c-d".

We thank the reviewer for pointing this out. The reference has been corrected from “Extended Data Fig. 6c–e” to “Extended Data Fig. 6c–d”, as following:

Page 10, line 10–12: Molecular dynamics simulations revealed that 17 is more unstable in combination with VR compared to 7 (Extended Data Fig. 6c-d).

1.22. Page 10, Line 2: The newly identified PTRs, PT, OC, and DMT presented in the main text are all results of *in vivo* feeding experiments in yeast. To ensure the accuracy and reliability of the enzyme activity results, it is recommended to supplement *in vitro* enzyme activity experiments.

As noted in our response to point 1.3, We have now added *in vitro* validation experiments for newly characterized enzymes.

1.23. Page 11, Line 8: Change "Box plots in (i)" to "Box plots in (h)"

We thank the reviewer for pointing this out. The reference has been corrected from “Box plots in (i)” to “Box plots in (h)”, as following:

Page 12, line 7–9: Box plots in (h): lines are median, box limits are quartiles 1 and 3, whiskers are $1.5 \times$ interquartile range and points are outliers.

1.24. Page 12, Lines 1–11: It was mentioned above that NhPDA1 can convert 8 to 12 with a not-low conversion rate. Why does the pathway of 8 converting to 12 not appear in Fig 4d? How to explain this phenomenon? Similarly, GgDMT1 showed broad substrate recognition above and demonstrated good catalytic activity when using substances 9, 10, and 11 as substrates. Why does it only catalyze the reaction from 9 to 13 here?

We appreciate the reviewer's insightful questions. Fig. 4d-e were constructed based on the metabolites actually detected at the sampling time points, and only the dominant conversions are shown because competing branches do not accumulate detectable intermediates. Although only expression of NhPDA1 can convert 8 to 12 at a moderate rate, GgPTR1 rapidly catalyses the conversion of 8 to 9 in the assembled pathway, resulting in the absence of 12 at the sampling point shown in Fig. 4d. Similarly, while GgDMT1 can accept 9, 10, and 11 as substrates, under our assay conditions compound 9 was rapidly and almost completely converted to 13, and intermediates 10 and 11 were therefore not detected. To prevent confusion, we have clarified these details in the main text (Page 12, line 5–6; Page 12, line 12–13) and Fig.4 legends (Page 14, line 5–8).

Page 13, line 5–6: Although NhPDA1 can convert 8 to 12, this route was not observed because GgPTR1 rapidly converted 8 to 9.

Page 13, line 12–14: In this mode, 9 is consumed preferentially and intermediates 10, 11 and 12 do not accumulate to detectable levels.

Page 15, line 6–8: d-e) Metabolic dynamics of the ladder-like glabridin biosynthetic network incorporating with either (d) NhPDA1 or (e) GgDMT1. Only the dominant conversions are shown because competing branches do not accumulate detectable intermediates under these pathway conditions.

1.25. Page 12, Line 25: Change "Fig. 4h, S9b-c" to "Fig. 4h, Extended Data Fig. 9b-c"

The citation has been revised from "Fig. 4h, S9b-c" to "Fig. 4h, Extended Data Fig. 9b-c" in the revised manuscript:

Page 13, line 25–27: Interestingly, this model also indicates that shifting the demethylation step close to the end of the pathway could substantially improve glabridin production (Fig. 4h, Extended Data Fig. 9b-c)

1.26. Page 13, Line 2: Change "Fig. 4i, Extended Data Fig. d-e" to "Fig. 4i, Extended Data Fig. 10d-e".

The citation has been corrected from "Fig. 4i, Extended Data Fig. d-e" to "Fig. 4i, Extended Data Fig. 10d-e".

Page 14, line 1–3: Finally, this fully engineered yeast strain enabled de novo production of glabridin in shake-flask fermentation, reaching a final titer of 0.5 mg/L (Fig. 4i, Extended Data Fig. 10d-e).

1.27. Page 22, Line 25: Italicize "*G. glabra*".

The species name *G. glabra* has been italicized throughout the revised manuscript accordingly.

Page 23, line 33–34: Transcriptome sequencing of various G. glabra organs was performed by BGI Genomics Co., Ltd. (Shenzhen, China).

1.28. Page 23, Line 12: Change "Table 5" to "Table S5"; Page 27, Line 21: Change "Table 8" to "Table S8"; Page 31, Line 3: Change "Table 9" to "Table S9".

We thank the reviewer for the helpful suggestion. As requested, the references originally pointing to “Table 5,” “Table 8,” and “Table 9” have been updated to their corresponding Supplementary tables. During the revision, the Supplementary tables were renumbered for consistency, and the final labels in the revised manuscript are Table S7 (Page 24, line 31), Table S10 (Page 30, line 27), and Table S11 (Page 35, line 3), respectively.

Page 24, line 31: A complete list of plasmids used is available in Table S7.

Page 30, line 27: Km values were listed in the Table S10.

Page 35, line 3: The sequences used in this analysis are listed in Table S11.

1.29. Page 27, Line 6: Change "100-mL" to "100 mL".

The unit expression has been revised from “100-mL” to “100 mL” in the revised manuscript (Page 30, line 11–12)

Page 30, line 11–12: Subsequently, 1 mL of each culture was transferred into 20 mL fresh YPD medium in 100 mL shake flasks.

1.30. Page 27, Line 28: Change "24 hr" to "24 h".

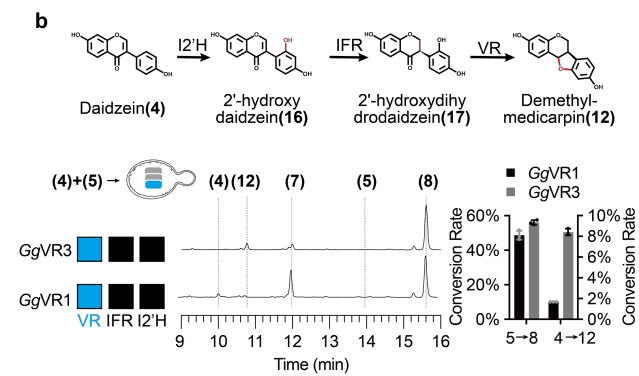
The expression “24 hr” has been revised to “24 h” in the revised manuscript (Page 27, Line 28).

Page 31, line 2–4: Enzyme concentrations were fixed based on biomass measurements at 24 h post-inoculation, coinciding with substrate supplementation timing in experimental protocols.

1.31. Page 33, Line 10: "GgVR1 can convert 2'-hydroxydihydrodrodaidzein (17) to Demethylmedicarpin (12)". Considering the context and Extended Data Fig. 6b, this should be corrected to "GgVR3".

We thank the reviewer for pointing out this issue. In fact, both GgVR1 and GgVR3 are able to catalyse the conversion of 2'-hydroxydihydrodrodaidzein (17) to demethylmedicarpin (12), with GgVR3 showing a substantially higher conversion rate. In the original figure, the activity of GgVR1 appeared less visible due to the single-axis scaling. We have now added dual y-axes to clearly display the conversion rates from 4 to 12 catalyzed by GgVR1 and GgVR3. The text has been revised accordingly in the main manuscript (Page 37, line 7–8).

Page 37, line 7–8: Both GgVR1 and GgVR3 can convert 2'-hydroxydihydrodrodaidzein (17) to Demethylmedicarpin (12) without PTS, which is similar with previous reported THIS..



Extended Data Fig. 6b

Referee #2 (Remarks to the Author):

This work examined enzymes capable of modifying isoflavan scaffolds. Authors reconstructed a network in yeast solidified that the metabolic redundancy and interconnectivity confers unexpected robustness and enables higher production efficiencies compared to a single-route design.

We thank Reviewer #2 for their careful reading and valuable comments. Below, we respond to each point and outline the corresponding changes to the manuscript.

Major concerns comprise

2.1. Can the biosynthetic pathway of glabridin inferred from the metabolic pathway search algorithm based on known enzymatic reactions and the retrosynthetic algorithm based on reaction rules represent the actual biosynthetic pathway present in licorice?

We appreciate this important question. In principle, the algorithm generates an extensive reaction network that is expected to encompass the native biosynthetic pathways operating in licorice, unless licorice employs enzymes catalysing reaction types not previously documented. Our algorithm utilizes a comprehensive database that includes nearly all known types of enzyme-catalysed reactions in nature. Thus, barring computational constraints, the method is capable of enumerating all feasible pathways based on known reaction types. In this study, we integrated multiple approaches—including metabolite profiling of licorice, computational pathway prediction, genomic analysis, and enzyme screening—to identify a plausible glabridin biosynthetic route, which was subsequently validated in *Saccharomyces cerevisiae*. Nevertheless, it should be noted that the computationally inferred pathway does not necessarily represent the exclusive or native route in licorice, and the possibility of alternative biosynthetic pathways for glabridin in the plant cannot be ruled out. We have now clarified this limitation in the main text (Page 3, line 14–17, Page 16, line 17–20)

Page 3, line 14–17: The reaction-rule-based pathway search is designed to enumerate all biochemically plausible routes supported by known enzyme reaction types, including both pathways that could represent the biosynthetic route in licorice and others that licorice may not actually use.

Page 16, line 17–20: Besides, our pathway search relies on known types of enzyme reactions, the inferred routes reflect reasonable biosynthetic possibilities rather than a definitive reconstruction of the native pathway in licorice. Licorice may employ enzymes or reaction chemistries that have not yet been described, and such steps would not be captured by the computational model.

2.2. Since its isolation from licorice in 1976, has no researcher explored the biosynthetic pathway of glabridin using isotope tracing methods?

We thank the reviewer for this insightful question. To the best of our knowledge, after carefully retrieving the available literature, no studies have investigated the biosynthetic pathway of glabridin using isotope tracing methods.

Page 2, line 29–31: Microbial biosynthesis represents a promising and sustainable alternative; however, although glabridin was first isolated from licorice in 1976, its native biosynthetic

pathway in *G. glabra* remains poorly resolved, and no isotope-tracing study has been reported to date.

2.3. Fig. 1a shows that most intermediates were not detected. Can the synthetic pathway be simplified for clearer presentation?

We appreciate the reviewer's helpful suggestion. In the revised manuscript, Fig. 1a has been simplified for clearer presentation by retaining only intermediates known to be endogenous to licorice (Page 2, line 17–24), based on our results and previously validated reports. The compound numbering has been updated accordingly.

Page 2, line 17–24: To focus on routes relevant to licorice, the network is further refined by retaining only intermediates detected in licorice (Fig. 1a). Using this curated compound–conversion network, we constructed a directed graph and performed a systematic depth-first traversal, which identified 13 potential biosynthetic pathways leading to the production of glabridin, involving L-phenylalanine (1) as precursor and 21 intermediate metabolites (Fig. 1a, Extended Data Fig. 1d). Among these, 9 compounds are located at pathway branch nodes, and over 69% of the predicted pathways include more than four branch nodes.

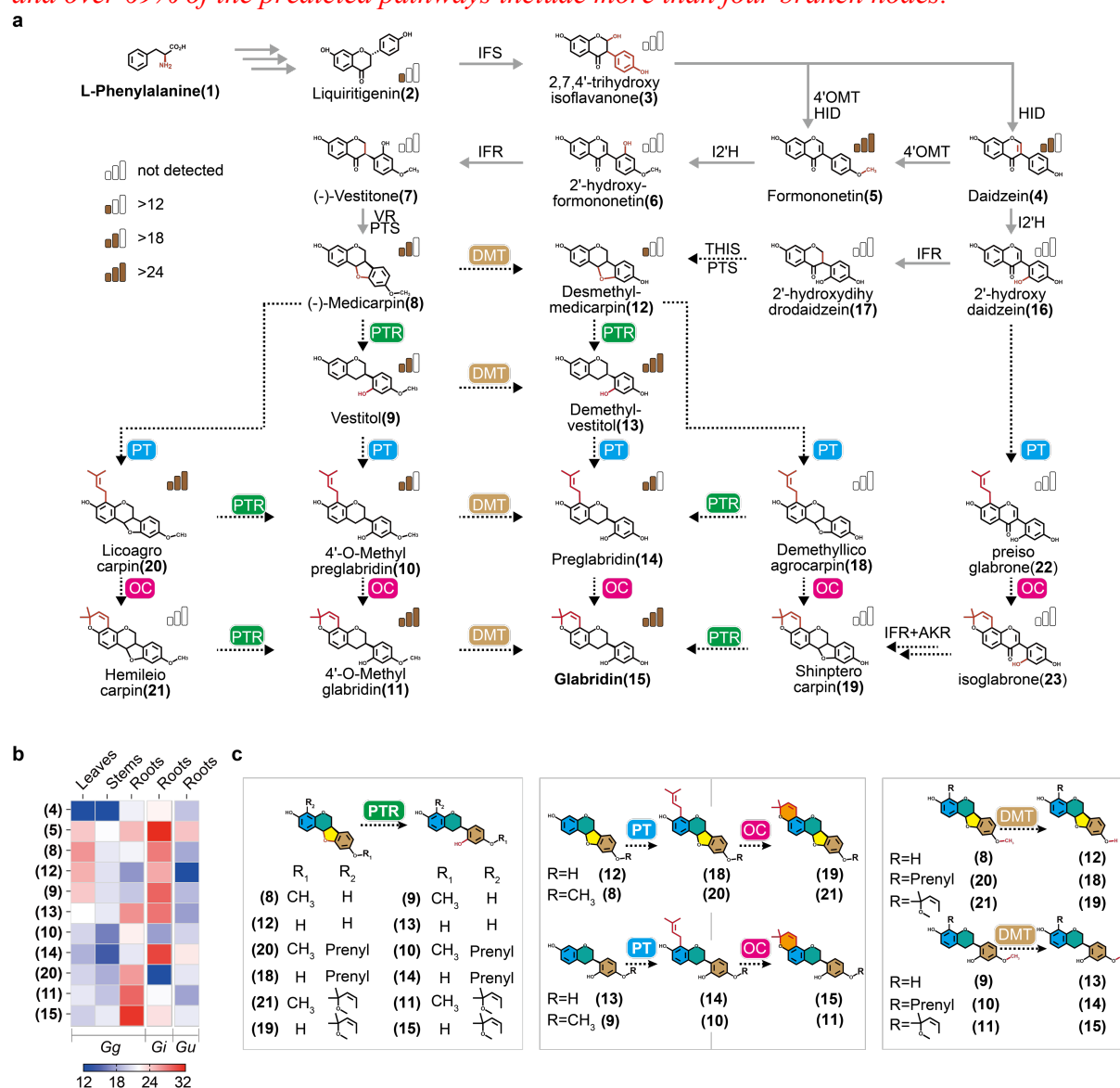
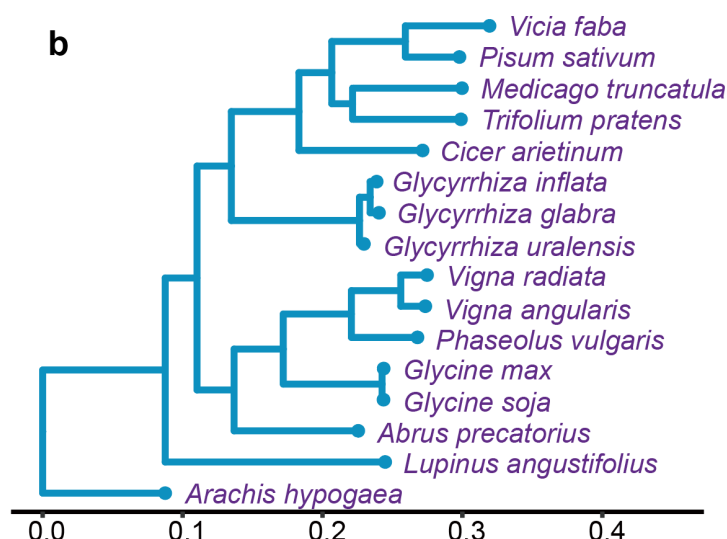


Fig. 1

2.4. Extended Data Fig. 2b does not explicitly state the Latin names of the three species, only indicating their genus names, making it impossible to discern their phylogenetic relationships

We appreciate the reviewer's helpful comment. The Latin names of the three *Glycyrrhiza* species have been added to Extended Data Fig. 2b to clearly indicate their phylogenetic relationships.



Extended Data Fig. 2b

2.5. Fig. 2d-e do not correspond to the results and figures. It is recommended to improve them in the results section. Additionally, Fig. 2e is redundant and should be deleted or placed in the appendix.

We thank the reviewer for the helpful suggestion. Fig. 2e has been removed as recommended, and the descriptions for Fig. 2d in the Results section have been revised to ensure full consistency between the text and the figure (Page 6, line 10–14).

Page 6, line 10–14: The clustering revealed a root-enriched gene expression cluster, including 1 PTR, 9 PTs, 8 OCs, and 2 DMTs (Fig. 2c, Table S2). To refine candidate genes, we used a previous reported key gene GgI2 H120 as the bait for co-expression analysis. This analysis prioritized candidate enzymes, including 1 PTR, 7 PTs, 2 OCs, and 2 DMTs(Fig. 2d).

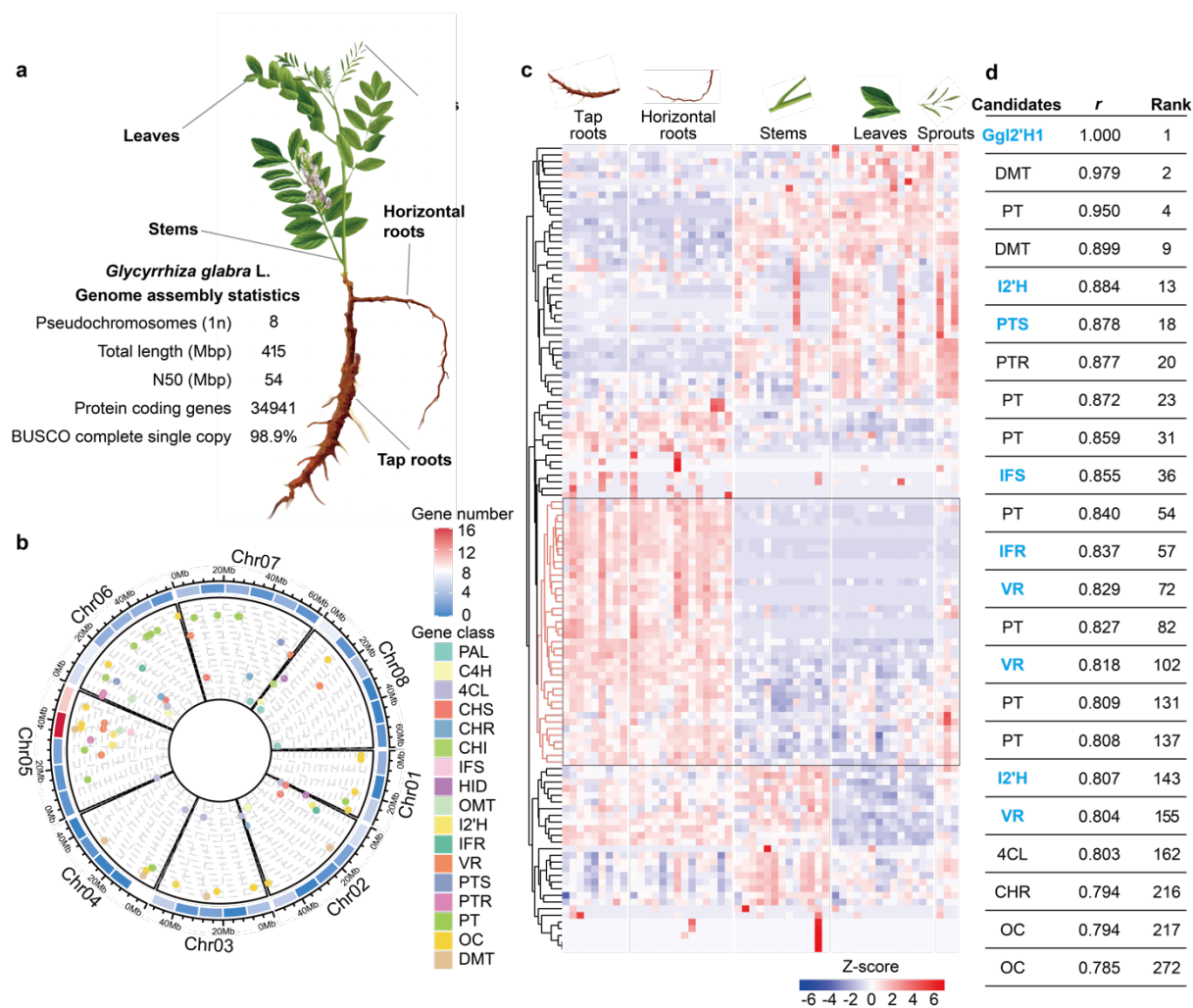


Fig. 2

2.6. The authors compared the enzymatic activities of GgPTR1 and GgPTR5, observing significantly lower activity in GgPTR5 without further mechanistic explanation. It is suggested to perform functional rescue mutations on GgPTR5 to determine whether its activity can be restored to that of GgPTR1.

We appreciate the reviewer's constructive suggestion. We performed functional rescue mutagenesis on GgPTR5 based on the six amino-acid differences identified within the substrate-binding pocket compared to GgPTR1 (Fig. S1). Among the introduced mutations, V264A notably enhanced GgPTR5 activity, increasing it by 27.5% to a total of 37.5%. However, none of the mutations fully restored the catalytic efficiency to the level of GgPTR1. To further investigate the underlying mechanism, we conducted molecular dynamics simulations, which showed that medicarpin (**8**) binds stably within the GgPTR1 but displays unstable binding in GgPTR5. These results indicate that the reduced activity of GgPTR5 is not attributable to a single residue but rather reflects broader structural differences that affect substrate positioning and pocket stability. A summary of sequence differences, mutational effects, and key structural features has been added to Fig. S1, and the corresponding description has been incorporated into the revised main text (Page 7, line 21–28) and in the caption of Fig. S1 (Supplementary information Page 2, line 4–15).

Page 7, line 21–28: Among them, the homolog GgPTR5 showed detectable activity but remained much weaker than GgPTR1. Because of its sequence similarity to GgPTR1, we attempted to improve GgPTR5 by introducing amino-acid substitutions in its substrate-binding pocket. Although these mutations produced slight increases in activity, none were able to bring GgPTR5 close to the activity of GgPTR1 (Fig. S1a-b). Molecular dynamics simulations further revealed unstable substrate binding in GgPTR5, suggesting that its low activity arises from overall structural limitations rather than defects at individual residues (Fig. S1c-d).

Supplementary information Page 2, line 4–15: Supplementary Figure 1. Functional and structural analysis of GgPTR. a) Structural comparison of GgPTR1 and GgPTR5 highlighting the six divergent amino-acid sites surrounding the substrate-binding pocket. b) Conversion rates of GgPTR1 and GgPTR5 mutants generated by reciprocal residue swapping at positions 128, 131, 132, 160, 263 and 264, using 20 mg/L medicarpin (8) as the substrate to quantify the conversion to vestitol (9). Notably, the V264A mutation increased GgPTR5 activity by 27.5% but did not restore activity to the level of GgPTR1. Data are mean $\pm$ s.d. from three biological replicates. c-d) Molecular dynamics simulations of GgPTR1 (c) and GgPTR5 (d) showing backbone RMSD trajectories over 200 ns across three independent replicates, where the black line represents the averaged RMSD, illustrating stable binding of medicarpin in GgPTR1 versus unstable binding in GgPTR5.

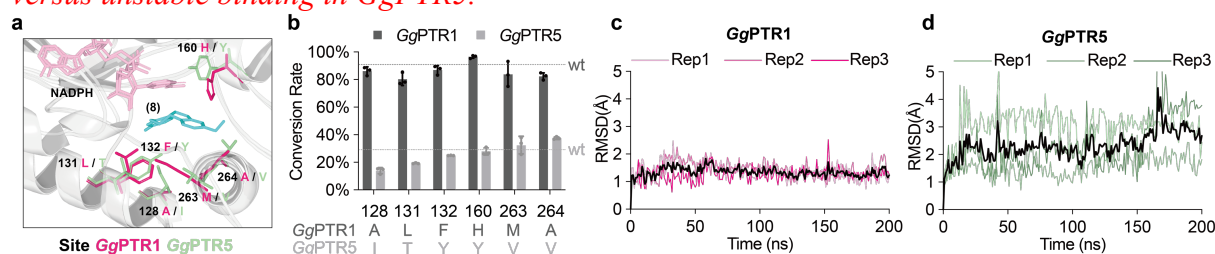


Fig. S1

2.7. It is recommended to further analyze the specific reasons why GgPT1 is active while the other six GgPTs are not. This can be done through amino acid sequence alignment and site-directed mutagenesis.

We thank the reviewer for this valuable suggestion. Following the recommendation, we conducted a multiple sequence align analysis of GgPT1 and the six non-functional GgPTs. These GgPTs differ from GgPT1 by more than 50 amino acids, indicating substantial divergence in their catalytic regions (Fig. S3a-b). Further inspection identified seven positions where all six non-functional GgPTs share conserved residues that are uniquely substituted in GgPT1. We mutated each of these residues in GgPT1 to the conserved amino acids found in the other GgPTs. Among them, the K242Q mutation markedly reduced the catalytic activity of GgPT1 (Fig. S3c), suggesting that Lys242 is a key determinant of its prenylation activity toward isoflavan substrates. Together, these results imply that the loss of activity in the other six GgPTs is attributable to both the absence of this critical residue and their extensive sequence divergence relative to GgPT1. The alignment, mutational data, and discussion have been incorporated into the revised manuscript (Page 8, line 16–19) and Fig. S3 (Supplementary information Page 4, line 2–13).

Page 8, line 16–19: Amino-acid comparison revealed that GgPT1 contains seven unique residues absent in the other GgPTs, and mutational analysis demonstrated that K242 is essential for its catalytic activity (Fig. S3).

Supplementary information Page 4, line 2–13: Supplementary Figure 3. Functional and structural analysis of GgPT1. a) Multiple sequence alignment of GgPT1 with six non-functional PT homologs showing conserved motifs and the seven amino-acid residues uniquely present in GgPT1. b) Summary of the total amino-acid differences between GgPT1 and each PT candidate, highlighting the extensive sequence divergence (>50 residues) of the six homologs relative to GgPT1. c) Conversion rates of GgPT1 point mutants in which each residue was mutated to the consensus amino acid shared by the six PT homologs at positions where GgPT1 uniquely differs, using 20 mg/L vestitol (9) as the substrate to measure the conversion to 4'-O-methylpreglabridin (10). Data represent mean $\pm$ s.d. from three biological replicates, and the marked loss of activity in the K242Q mutant identifies Lys242 as a key determinant of GgPT1 catalytic function.

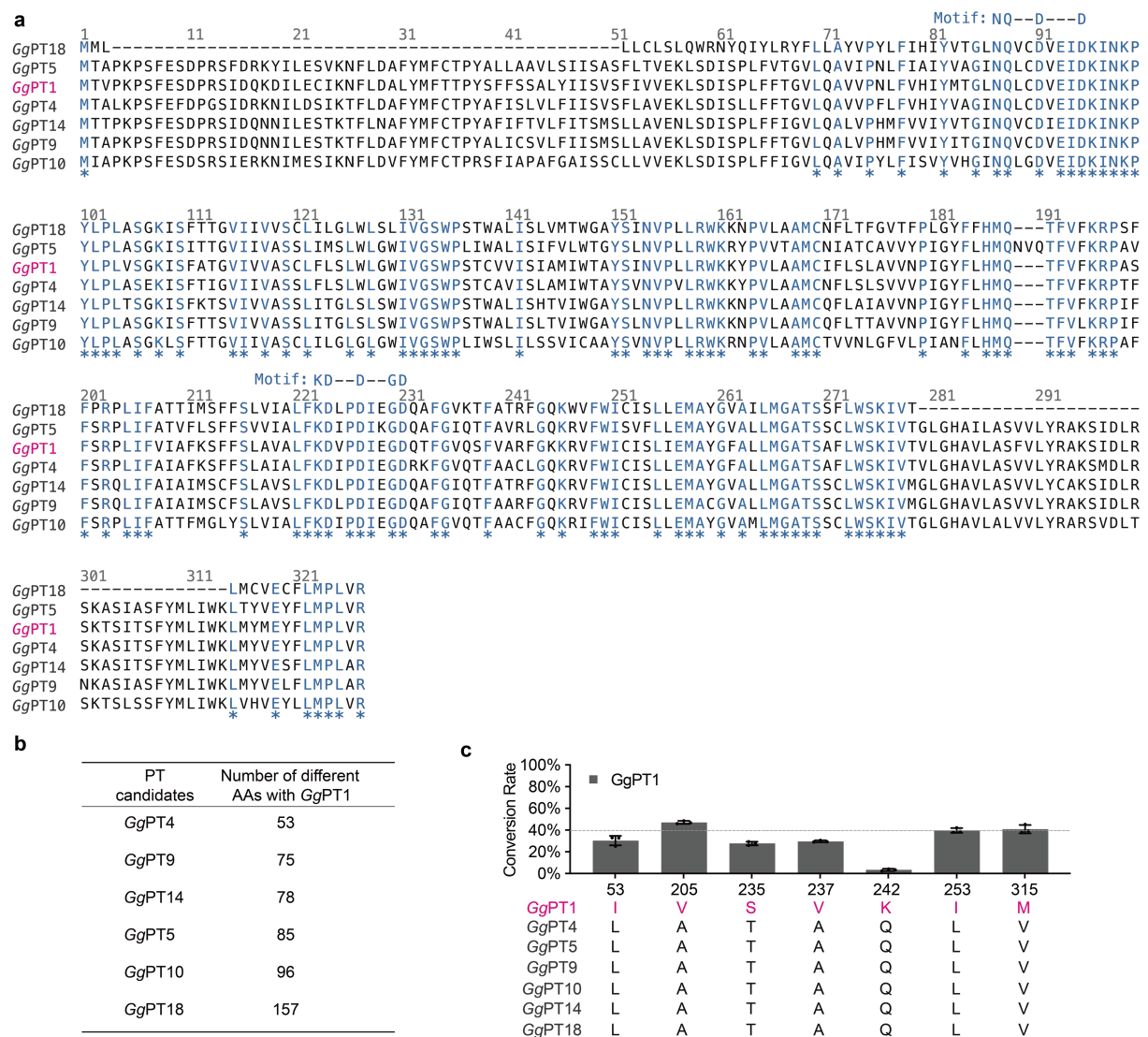


Fig. S3

2.8. The test showed that GgOC3 had no corresponding activity, but this was not described in the results section. Further analysis was then conducted on the differences in sequence and active sites between GgOC1, GgOC2, and GgOC3 to further explain the reason for the lack of activity.

We thank the reviewer for this valuable comment. In the revised manuscript, we have added a description of the experimental result showing that GgOC2 and GgOC3 exhibited no detectable catalytic activity. To investigate the basis of this loss of function, we performed a comparative sequence analysis of GgOC1, GgOC2, and GgOC3. GgOC2 and GgOC3 differed from GgOC1 by more than 30 amino acids overall, and by over 15 residues within the active-site region (Fig. S5a–c). Further inspection identified seven positions at which GgOC2 and GgOC3 both carry conserved residues, whereas GgOC1 contains unique substitutions. Mutational analysis showed that introducing these conserved residues into GgOC1 substantially impaired its activity: the I293F mutation nearly abolished catalytic function, and T295R markedly reduced activity (Fig. S5d). These findings suggest that these positions are critical for OC function. Combined with the extensive sequence divergence across GgOC2 and GgOC3, these results provide a mechanistic explanation for why GgOC2 and GgOC3 lack detectable oxidative cyclization activity. The relevant data and discussion have been added to the revised manuscript (Page 8, line 33–Page 9, line 4) and Fig. S5 (Supplementary information Page 6, line 1–14).

Page 8, line 33–Page 9, line 4: In addition, we tested GgOC3 and several other BBE-like enzymes and found that they were also inactive. To investigate the cause of their inactivity, we performed mutational analysis, which showed that residues I293 and T295 of GgOC1 are essential for catalysis, explaining the loss of activity in both GgOC2 and GgOC3 (Fig. S5)..

Supplementary information Page 6, line 1–14: Supplementary Figure 5. Sequence comparison and functional characterization of GgOC mutations. a) Multiple sequence alignment of GgOC1 with its two homologs GgOC2 and GgOC3 showing conserved regions and the extensive sequence divergence (>30 residues overall) in the two inactive homologs. b) Structural mapping of the active-site region highlighting more than 15 amino-acid differences between GgOC1 and the two homologs. c) Summary of full-length and active-site differences between GgOC1 and the two homologs. d) Conversion rates of GgOC1 variants generated by substituting residues at the seven positions where both GgOC2 and GgOC3 carry conserved amino acids, using 10 mg/L 4'-O-methylpreglabridin (10) as the substrate to measure the conversion to 4'-O-methylglabridin (11). Data represent mean $\pm$ s.d. from three biological replicates; the I293F mutation nearly abolished activity and T295R markedly reduced activity, identifying these residues as critical determinants of oxidative cyclization.



We thank the reviewer for this insightful comment. In the revised manuscript, we have now included the full experimental results, which show that GgDMT2 exhibited no detectable O-demethylation activity toward any of the five tested methylated substrates (**5**, **8**, **9**, **10**, and **11**) (Fig. S6). To explore the possible reasons for this difference, we compared the amino-acid sequences of GgDMT1 and GgDMT2 (Fig. S7a), which revealed 138 amino-acid differences in total. Docking analysis revealed that the substrate binds in an inverted orientation in GgDMT2, preventing correct positioning for catalysis (Fig. S7b). We further examined the eight amino-acid differences surrounding the predicted catalytic pocket and introduced the corresponding substitutions into GgDMT1 (Fig. S7c). Notably, the N303R mutation abolished GgDMT1 activity, likely due to steric hindrance caused by the bulky arginine side chain, which blocks proper substrate binding (Fig. S7d). Overall, the unfavorable 303R substitution and substantial sequence differences together underlie the inactivity of GgDMT2. The results (Page 9, line 18–22), Fig. S6 (Supplementary information Page 6, line 17–21) and Fig. S7 (Supplementary information Page 7, line 2–13) have been added to the revised manuscript.

Page 9, line 18–22: Besides, GgDMT2 showed no detectable activity toward any of the tested methylated substrates (Fig. S6). Docking analyses indicated that GgDMT2 misorients the substrate in the active site, impairing productive catalysis. Likewise, changing GgDMT1 at residue N303 (N303R), which affects substrate positioning, also abolishes its activity (Fig. S7).
 Supplementary information Page 6, line 17–21: Supplementary Figure 6. Characterization of GgDMT2. HPLC analysis showing that GgDMT2 exhibits no detectable O-demethylation activity toward any of the five methylated substrates formononetin (5), medicarpin (8), vestitol (9), 4'-O-methylpreglabridin (10), and 4'-O-methylgalbridin (11). Substrate-feeding assays were performed using 20 mg L⁻¹ of each compound.

Supplementary information Page 7, line 2–13: Supplementary Figure 7. Sequence comparison and functional analysis of GgDMT mutations. a) Sequence alignment of GgDMT1 and GgDMT2 showing extensive sequence divergence, with 138 amino-acid differences identified in total. b) Docking analysis of substrate 4'-O-methylgalbridin (11) showing that it binds in an inverted orientation in GgDMT2 compared with GgDMT1. c) Conversion rates of GgDMT1 and GgDMT2 variants generated by substituting the eight divergent residues surrounding the predicted catalytic pocket, using 20 mg/L 4'-O-methylgalbridin (11) as the substrate to measure the conversion to galbridin (15). Notably, the N303R substitution in GgDMT1 abolishes catalytic activity. Data represent mean $\pm$ s.d. from three biological replicates. d) Structural view showing that Arg303 introduces steric hindrance near the active center, blocking proper substrate accommodation and explaining the loss of activity.

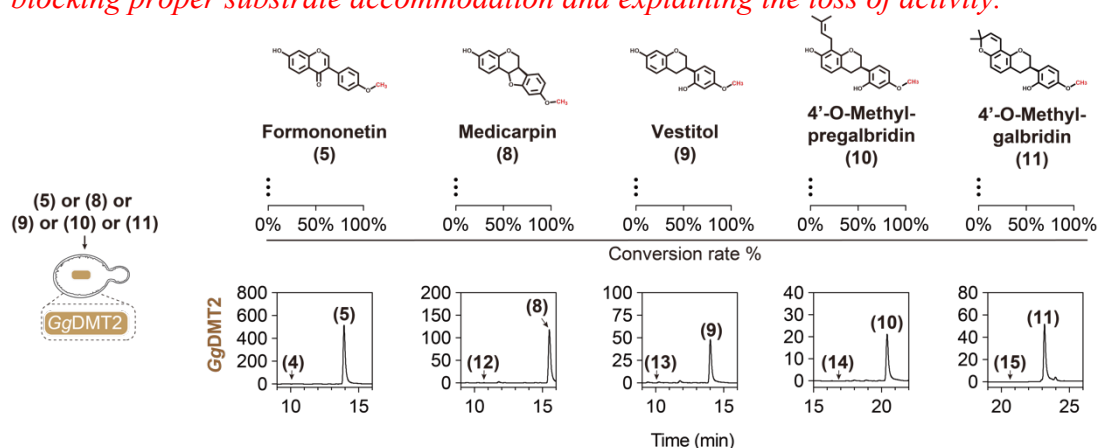


Fig. S6

Referee #3 (Remarks to the Author):

The authors have successfully elucidated the biosynthetic process of glabridin, and this work does provide a blueprint for glabridin production. However, the narrative of a "maze-like" metabolic network appears unfounded. Theoretically, there may be multiple potential pathways for producing the target compound, but in nature—particularly in plants where such compounds are enriched—evolution typically favors a specific, dominant route. Thus, the core premise of this study is questionable.

Moreover, the research only identifies several enzymes involved in the conversion from compound 8 to glabridin, and the yield of glabridin remains excessively low. As such, the study made a unreal story and lacks genuine novelty. Given these limitations, this work does not meet the criteria for publication in the journal.

We thank Reviewer #3 for the candid comments. We have addressed the core concerns related to the novelty, relevance, and logic of our multi-route pathway interpretation.

3.1. The narrative of a ‘maze-like’ metabolic network appears unfounded. Theoretically, there may be multiple potential pathways for producing the target compound, but in nature—particularly in plants where such compounds are enriched—evolution typically favors a specific, dominant route. Thus, the core premise of this study is questionable.

We appreciate the reviewer’s thoughtful comment. Our use of the term “maze-like” was not intended to suggest randomness or the absence of a dominant biosynthetic route. Rather, it reflects the presence of several partially overlapping tailoring branches that arise from enzyme promiscuity and pathway interconnectivity. Our functional assays and sequence analyses support the existence of parallel oxidative and methylation routes, which together form a ladder-like, multi-route network capable of channeling intermediates through more than one viable sequence of tailoring reactions.

While plants often favor a predominant metabolic route, it is well established that specialized-metabolism pathways can incorporate auxiliary or conditionally active branches that enhance robustness, buffer flux imbalances, or fine-tune metabolite profiles. The interconnected architecture we describe is therefore consistent with known principles of plant metabolic organization, rather than contradictory to them.

To improve clarity, we have revised the wording in the Results and Discussion (Page 3, line 14–22; Page 16, lines 26–32) and simplified the network depiction in Fig. 1 to highlight that the proposed network is based on experimentally supported alternative routes, not arbitrary pathway complexity.

Page 3, line 14–22: The reaction-rule-based pathway search is designed to enumerate all biochemically plausible routes supported by known enzyme reaction types, including both pathways that could represent the biosynthetic route in licorice and others that licorice may not actually use. To focus on routes relevant to licorice, the network is further refined by retaining only intermediates detected in licorice (Fig. 1a). Using this curated compound–conversion network, we constructed a directed graph and performed a systematic depth-first traversal, which identified 13 potential biosynthetic pathways leading to the production of glabridin, involving L-phenylalanine (1) as precursor and 21 intermediate metabolites (Fig. 1a, Extended Data Fig. 1d).

Page 16, line 26–32: Moreover, although we established two major biosynthetic routes, certain branches of the network (Fig. 1, Extended Data Fig. 1d) remain incomplete due to missing or unidentified enzymes. We agree that validation of these reactions would further enhance the authenticity of the proposed glabridin maze-like network. However, the inability to perform these assays reflects experimental constraints rather than evidence against the existence of these branches. Nonetheless, the presence of key intermediates such as compounds 19 and 25 in Glycyrrhiza supports the biochemical plausibility of these branches.

3.2. Moreover, the research only identifies several enzymes involved in the conversion from compound 8 to glabridin, and the yield of glabridin remains excessively low. As such, the study made a unreal story and lacks genuine novelty. Given these limitations, this work does not meet the criteria for publication in the journal.

We thank the reviewer for this thoughtful comment. Although our functional characterization focuses on the segment from intermediate 8 to glabridin, this portion of the pathway encompasses several key transformations that had not been previously elucidated. In particular, we identified four previously unreported classes of tailoring enzymes that together complete the biosynthetic route to glabridin. Notably, the discovery of a plant demethylation step acting on O-methylated isoflavonoids is, to our knowledge, unprecedented. These findings, which substantially advance the biochemical understanding of glabridin formation, are now emphasized in the revised Discussion (Page 15, lines 20–27).

Beyond this segment, we systematically identified all candidate enzymes involved in the full biosynthetic pathway, including the early scaffold-forming steps. This comprehensive pathway reconstruction enabled the first *de novo* biosynthesis of glabridin from glucose in *Saccharomyces cerevisiae*. To our knowledge, no previous study has established the complete pathway or achieved microbial total biosynthesis of this compound. We believe this represents significant conceptual and methodological novelty.

Regarding the reviewer's concern about the current glabridin titer, we agree that yield improvement is an important next step. As with many newly reconstructed plant specialized-metabolism pathways, optimization of enzyme expression, subcellular localization, cofactor supply, and flux balance will be required to enhance production. These aspects are well-recognized challenges in the field of metabolic engineering and lie beyond the scope of the present study. Importantly, the successful installation of the full biosynthetic route establishes a robust baseline chassis host and a necessary foundation for subsequent strain-engineering efforts.

Together, the identification of previously unknown enzyme classes, elucidation of a complete plant biosynthetic route and achievement of total biosynthesis in yeast represent substantial advances that we believe meet the novelty and significance standards of the journal.

Page 15, line 20–27: In this study, we systematically mapped the maze-like biosynthetic network of glabridin and successfully reconstructed it in S. cerevisiae, achieving de novo production from glucose. To our knowledge, this work provides the first complete biosynthetic route to glabridin and its total microbial biosynthesis. Within this pathway, we uncovered several previously uncharacterized transformations, including a plant demethylation step

acting on O-methylated isoflavonoids that has not been reported previously. By assembling key tailoring enzymes, we revealed a ladder-like network architecture mediated by a set of promiscuous enzymes that enhances pathway robustness through multiple interconnected routes.