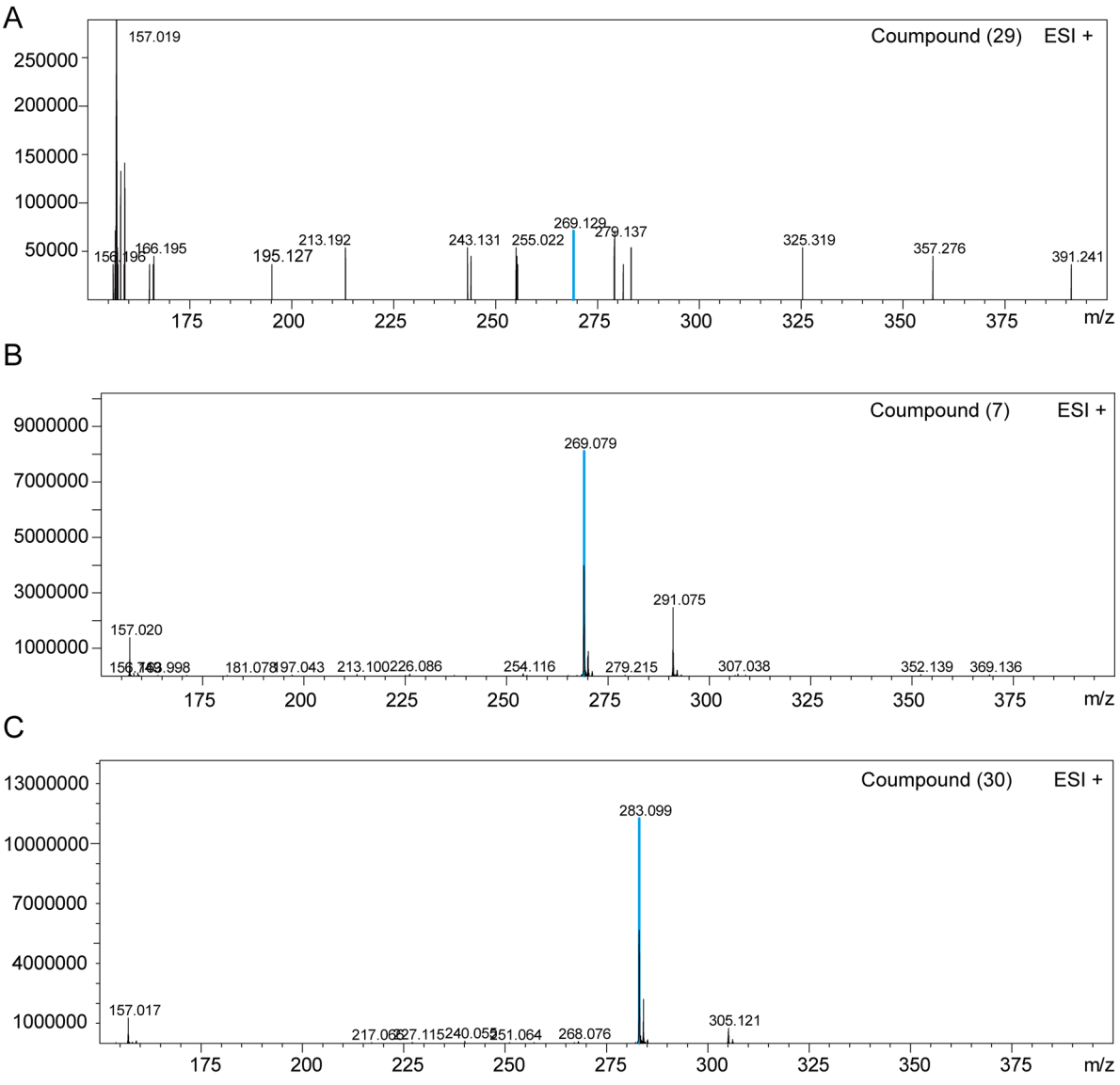
Supplementary Materials

Block-building with yeast to elucidate an artificial pathway for de novo biosynthesis of glabridin

Hanning Deng[1,2,3], Lu Cao[1,2], Shan Li[1,2], Changtai Zhang[1,2], Weizhu Zeng[1,2,3], Xinrui Zhao[1,2], Jingwen Zhou[1,2,3*]

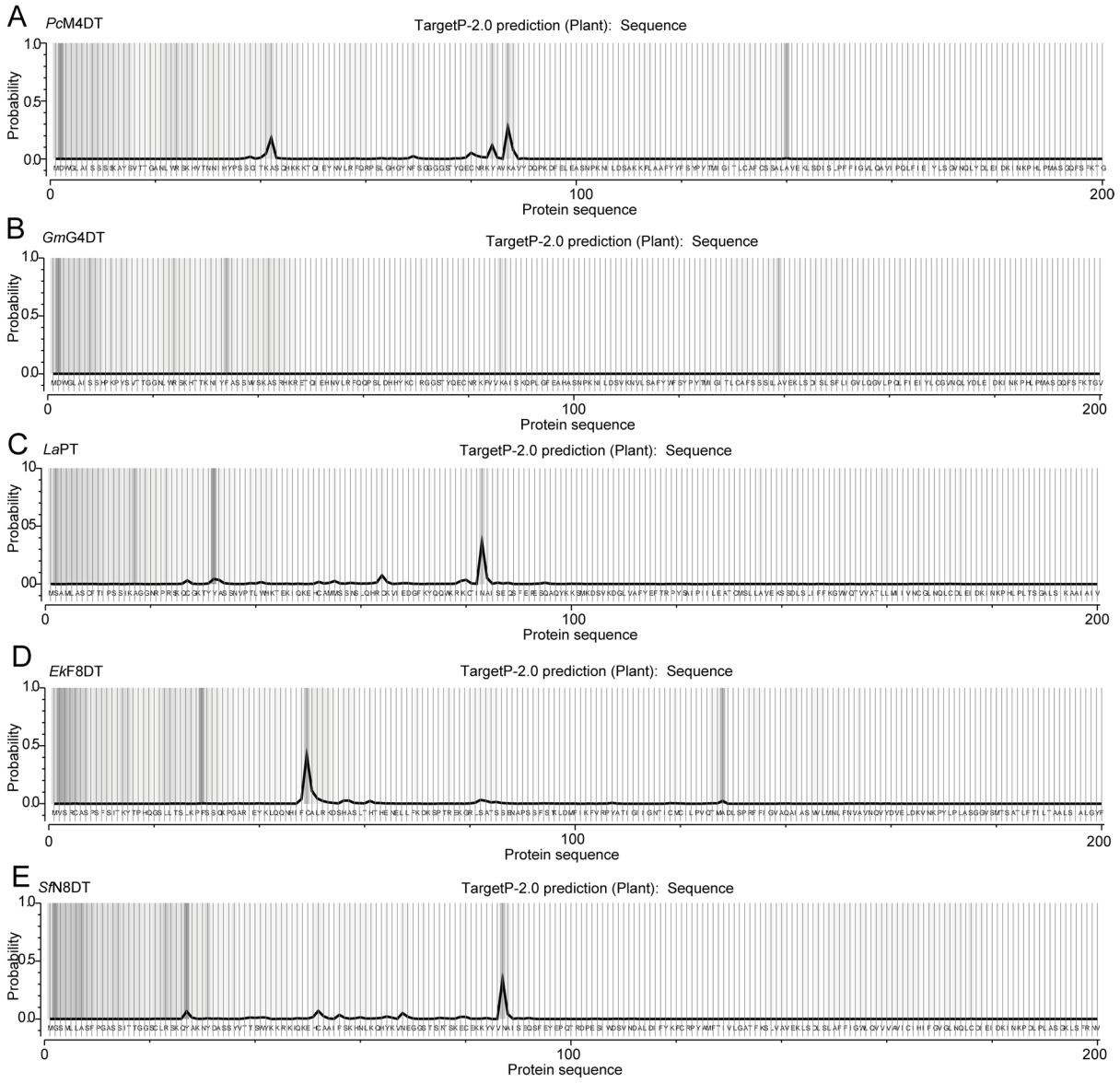
1 State Key Laboratory of Food Science and Resources, Jiangnan University, Wuxi 214122, China;
2 Science Center for Future Foods, School of Biotechnology, Jiangnan University, Wuxi 214122, China;
3 Jiangsu Province Engineering Research Center of Food Synthetic Biotechnology, Jiangnan University, Wuxi 214122, China.

* Correspondence:
Jingwen Zhou
E-mail: zhoujw1982@jiangnan.edu.cn
Phone: +86-510-85914371, Fax: +86-510-85914371



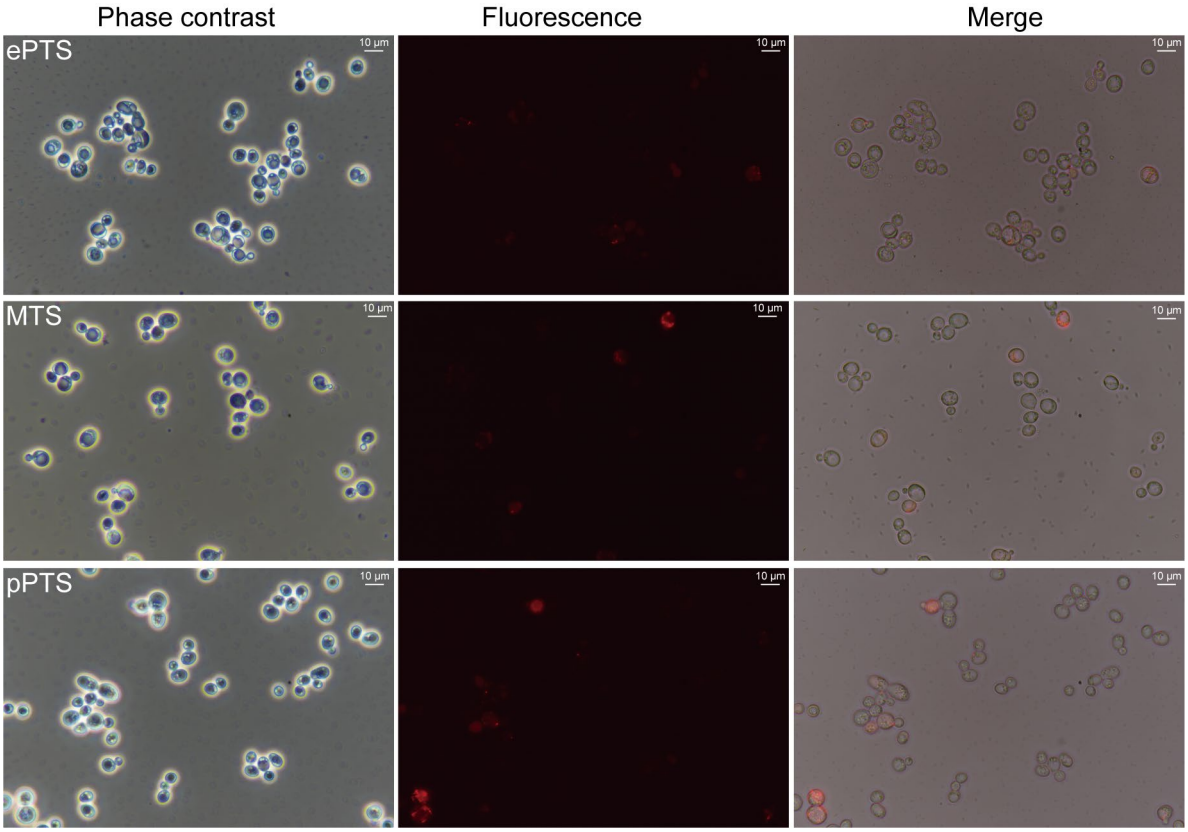
Supplementary figure 1 MS spectra (ESI^+) of different methylated products of daidzein catalyzed by *Pl*OMT9.

Panels (A) and (B) show mono-methylated products, while panel (C) represents the di-methylated product. The major peaks in all spectra correspond to the protonated molecular ions $[M+H]^+$, and the labeled ions are used as diagnostic signals for product identification and differentiation. The x-axis represents the mass-to-charge ratio (m/z), and the y-axis represents ion intensity.



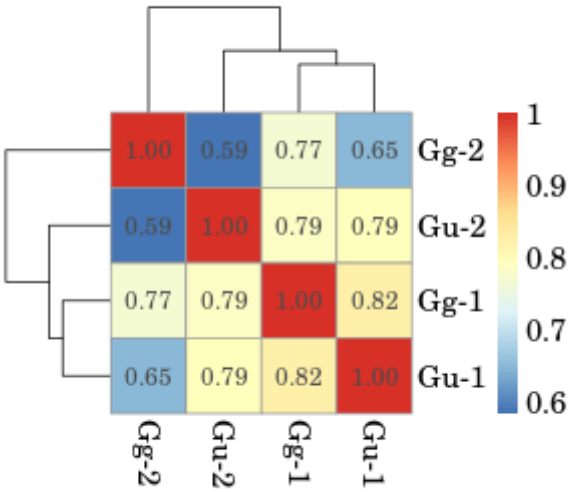
Supplementary figure 2 Signal peptide prediction of different prenyltransferases using TargetP-2.0 (Plant).

Signal peptide prediction profiles of different prenyltransferases analyzed using TargetP-2.0 (Plant). Panels (A–E) correspond to *Pc*M4DT, *Gm*G4DT, *La*PT, *Ek*F8DT, and *Sf*N8DT, respectively. The x-axis denotes the amino acid position along the protein sequence, and the y-axis indicates the predicted probability score. Peaks in the probability curves indicate regions with elevated signal peptide likelihood, providing insight into potential differences in subcellular targeting among these enzymes.



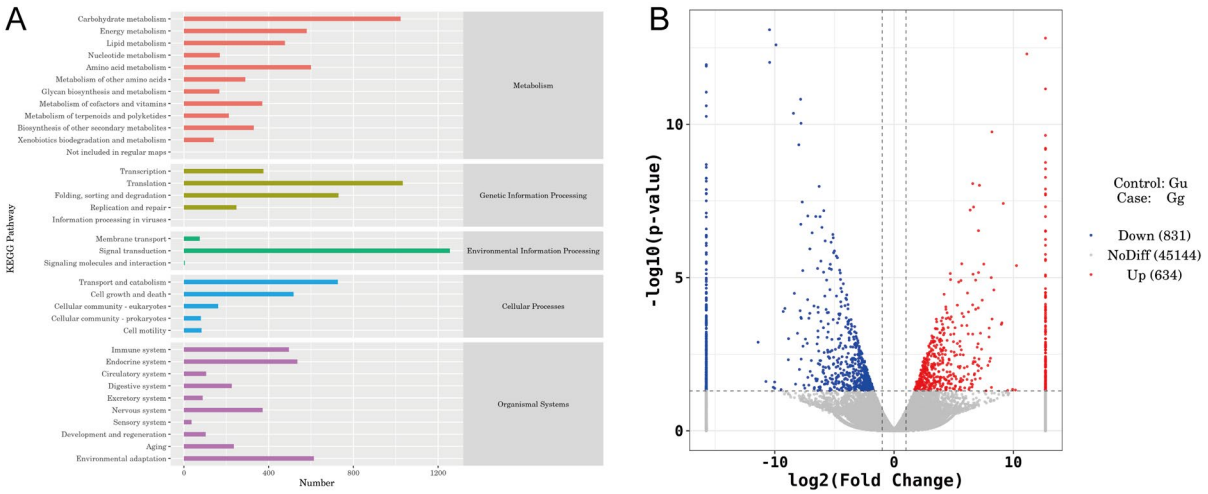
Supplementary figure 3 Fluorescence imaging of subcellular localization directed by different organelle-targeting signals.

Rows correspond to constructs fused with ePTS, MTS, and pPTS targeting signals, respectively. Columns show phase-contrast images (left), fluorescence images (middle), and merged images (right). The fluorescence signals indicate the intracellular distribution of the tagged proteins under the guidance of different targeting peptides, highlighting distinct organelle localization patterns. The experiment was repeated twice (n=2) with similar results, under the same conditions. Images were captured using a Nikon fluorescence microscope with a 100x oil objective.

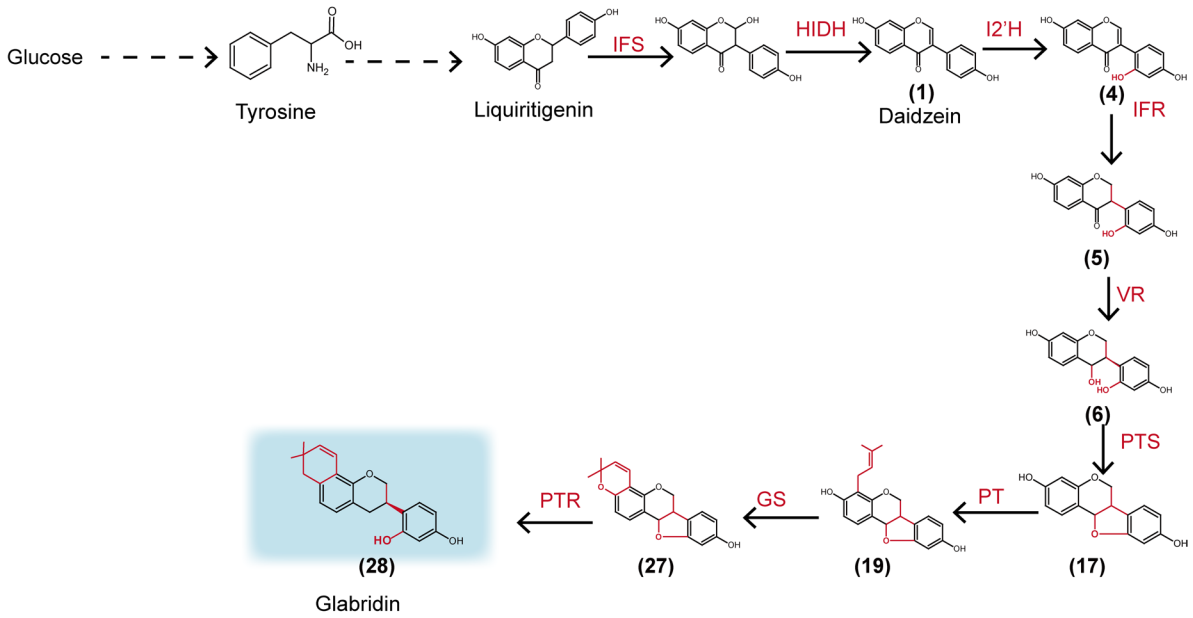


Supplementary figure 4 Correlation analysis of transcriptome samples.

The heatmap shows pairwise correlation coefficients among transcriptome samples, with values indicated by the color scale. Hierarchical clustering is applied to both axes to illustrate the similarity relationships between samples. High correlation coefficients indicate good reproducibility and consistency among biological replicates. Source data are provided as a Source Data file.

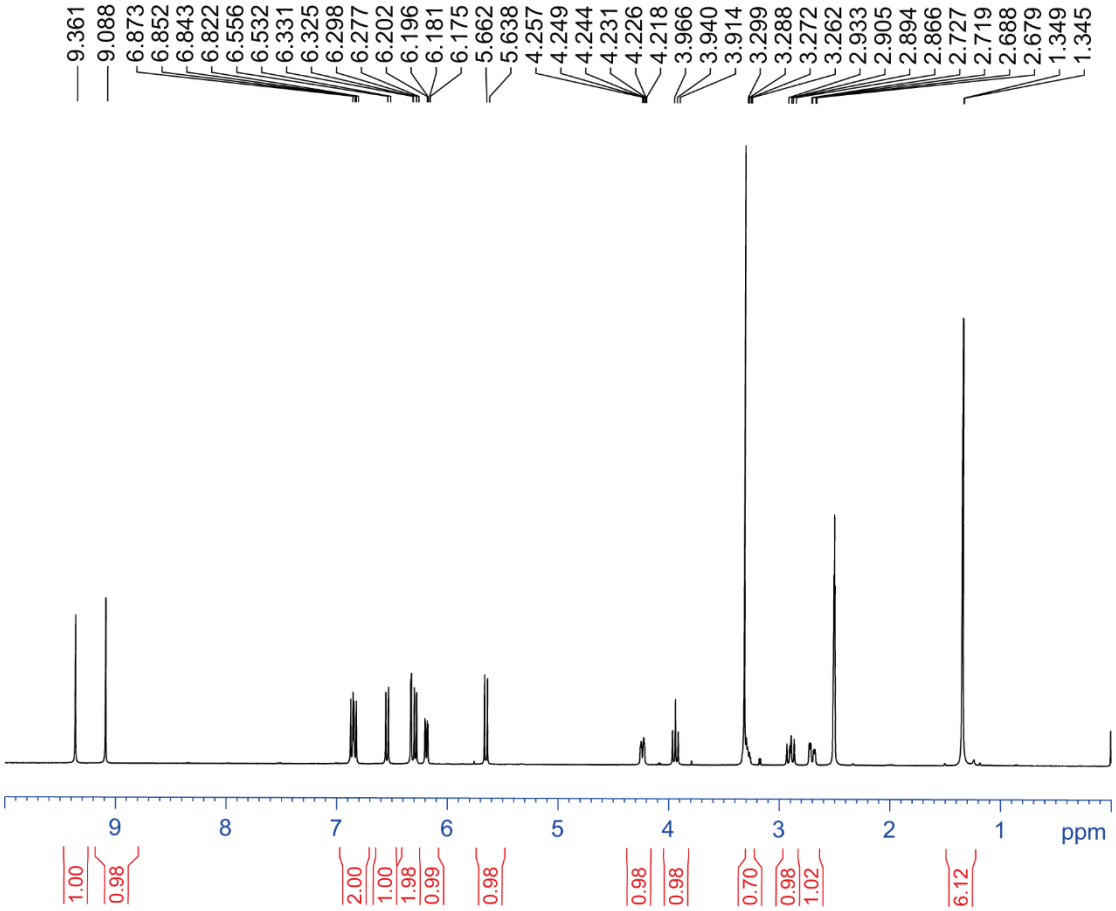


Supplementary figure 5 KEGG pathway annotation and differential gene expression analysis. (A) KEGG pathway annotation of annotated genes, classified into major functional categories including metabolism, genetic information processing, environmental information processing, cellular processes, organismal systems, and human diseases. The x-axis indicates the number of genes annotated to each pathway. (B) Volcano plot showing differentially expressed genes between the Cg and Gg groups. The x-axis represents $\log_2$(fold change), and the y-axis represents $-\log_{10}$(p-value). Red and blue dots indicate significantly upregulated and downregulated genes, respectively, while gray dots represent non-differentially expressed genes. Source data are provided as a Source Data file.



Supplementary figure 6 Final elucidated biosynthetic pathway of glabridin.

Schematic representation of the final elucidated biosynthetic pathway of glabridin. The pathway originates from glucose-derived tyrosine and proceeds through the isoflavonoid biosynthetic branch to form the key intermediate daidzein (1). Subsequent sequential reactions catalyzed by I2′H, IFR, VR, PTS, PT, GS, and PTR lead to the formation of multiple intermediates and ultimately yield glabridin (28). Enzymes involved in each step are indicated in red, compound numbers are shown in parentheses, and the final product glabridin is highlighted.

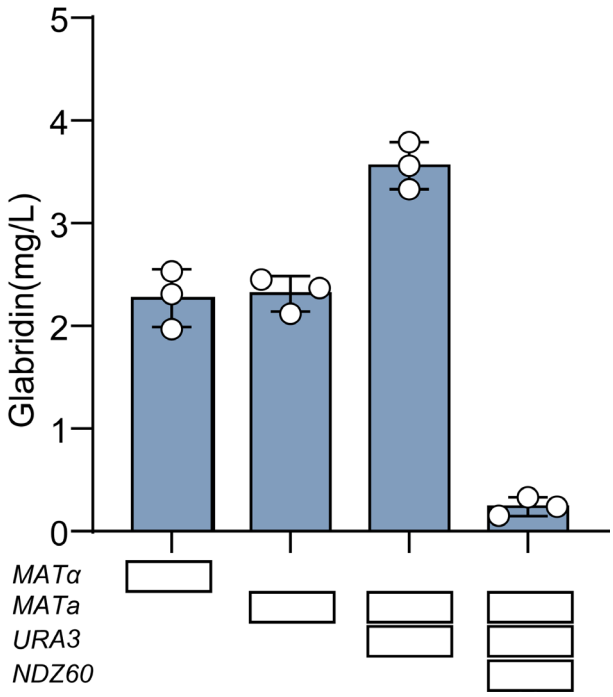


72

73 Supplementary figure 7 The ^{1}H-NMR of glabridin purified from culture broth.

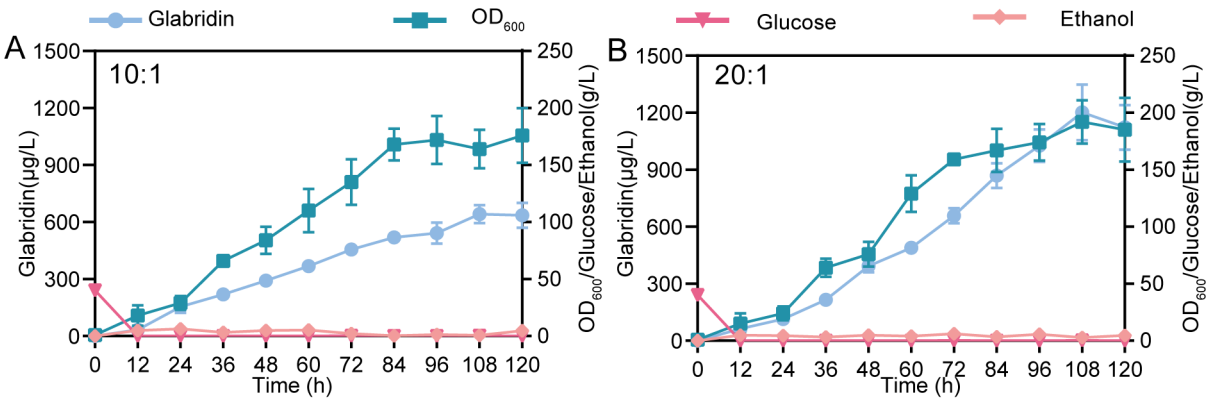
74 ^{1}H NMR (400 MHz, DMSO-d$_6$): δ 9.36 (s, 1H), 6.87–6.82 (m, 4H), 6.56–6.53 (m, 2H), 6.33 (s, 1H), 6.20–6.18 (m,

75 2H), 5.66 (s, 1H), 4.25–4.22 (m, 3H), 3.96–3.91 (m, 2H), 2.86–2.72 (m, 2H), 1.35 (s, 6H).



Supplementary figure 8 The biosynthesis of glabridin.

The effects of mating-type switching, URA3 tag complementation, and diploid strain construction on the synthesis of glabridin in the daidzein-to-glabridin biosynthetic module strain, where the mating-type switching and tag complementation were performed with 250 mg/L of daidzein supplementation. Data are presented as mean ± SD from n = 3 independent biological replicates. Source data are provided as a Source Data file.



83

84 Supplementary figure 9 Production of glabridin in 5 L bioreactor.

85 (A) Production of glabridin under a C/N ratio of 10:1; (B) Production of glabridin under a C/N ratio

86 of 20:1, with all other culture conditions kept identical. This comparison investigates how C/N ratio

87 modulation affects the biosynthetic efficiency of glabridin. Fermentations were performed in a 5-L

88 bioreactor, with two independent replicates (n=2). Source data are provided as a Source Data file.