Electronic Supplementary Material (ESI) for RSC Advances.

SUPPLEMENTARY INFORMATION

Two putative parallel pathways for naringenin biosynthesis in *Epimedium wushanense*

Table S1. Summary of RNA-seq data of *Epimedium wushanense*.

Sample	Raw Reads	Raw Bases	Clean Reads	Clean Bases
Ew_1	45422458	6858791158	44970122	6609887019
Ew_2	44379346	6701281246	43867346	6487050975
Ew_3	50598586	7640386486	50022334	7378743841

Table S2. List of enzymes identified from *E. wushanense* transcriptome.

Query	**Unigene**	**Length (aa)**	**Identity**
SbPAL (XP_021319560.1)	EwPAL	736	76%
Nt4CL2 (AAB18638.1)	Ew4CL1	569	76%
Nt4CL2 (AAB18638.1)	Ew4CL2	546	65%
AtCHS (AAB35812.1)	EwCHS1	393	86%
AtCHS (AAB35812.1)	EwCHS2	389	85%
AtCHS (AAB35812.1)	EwCHS3	393	78%
AtCHI (CAB94981.1)	EwCHI1	212	85%
AtCHI (CAB94981.1)	EwCHI2	271	82%
AtCHI (CAB94981.1)	EwCHI3	263	80%
HlCHIL2 (AVR53897.1)	EwCHIL1	247	23%
HlCHIL2 (AVR53897.1)	EwCHIL2	393	75%
AtCHI (CAB94981.1)	EwCHIL3	208	82%
AtCHI (CAB94981.1)	EwCHIL4	296	75%

Table S3. List of primers used in this study.

EwPAL-F	GTGCCGCGCGGCAGCCATATGGCGACCACCACCACCAC
EwPAL-R	TGCGGCCGCAAGCTTTTAGCAGATCGGCAGCGGCGCA
Ew4C1-F	GTGCCGCGCGGCAGCCATATGGAAACCCCGGCGGCTCCG
Ew4C1-R	GAGTGCGGCCGCAAGCTTTTAGTTCGGGATACCCGCCG
EwCHS1-F	GTGCCGCGCGGCAGCCATATGGTGACTGTCGAGGAA
EwCHS1-R	GAGTGCGGCCGCAAGCTTTCACTGAGTTGCAACA
EwCHS3-F	GTGCCGCGCGGCAGCCATATGGGCATGGTTAGCGTG
EwCHS3-R	GAGTGCGGCCGCAAGCTTCTCGAGCTTCGCCGCCACCG
EwCHI1-F	AGCAAATGGGTCGCGGATCCATGGCGAAAAGCAGCCTGCAGG
EwCHI1-R	CGGCCGCAAGCTTCTCGAGTTAAACCAGCAGTTCGCTAATACG
EwCHIL3-F	GTGCCGCGCGGCAGCCATATGGAAACCGAAATGGTT
EwCHIL3-R	GAGTGCGGCCGCAAGCTTCGCCAGCAGCGCACCCAG

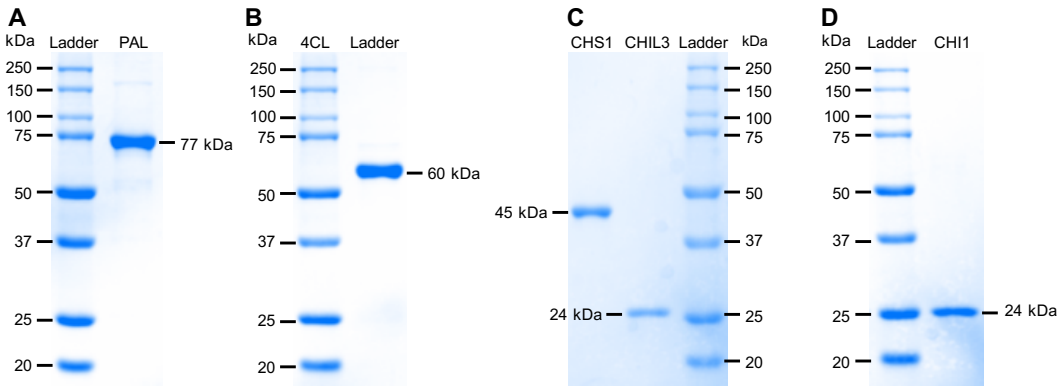

Figure S1. SDS-PAGE analysis of purified recombinant proteins. (A) EwPAL. **(B)** Ew4CL1. **(C)** EwCHS1 and EwCHIIL3. **(D)** EwCHI1.

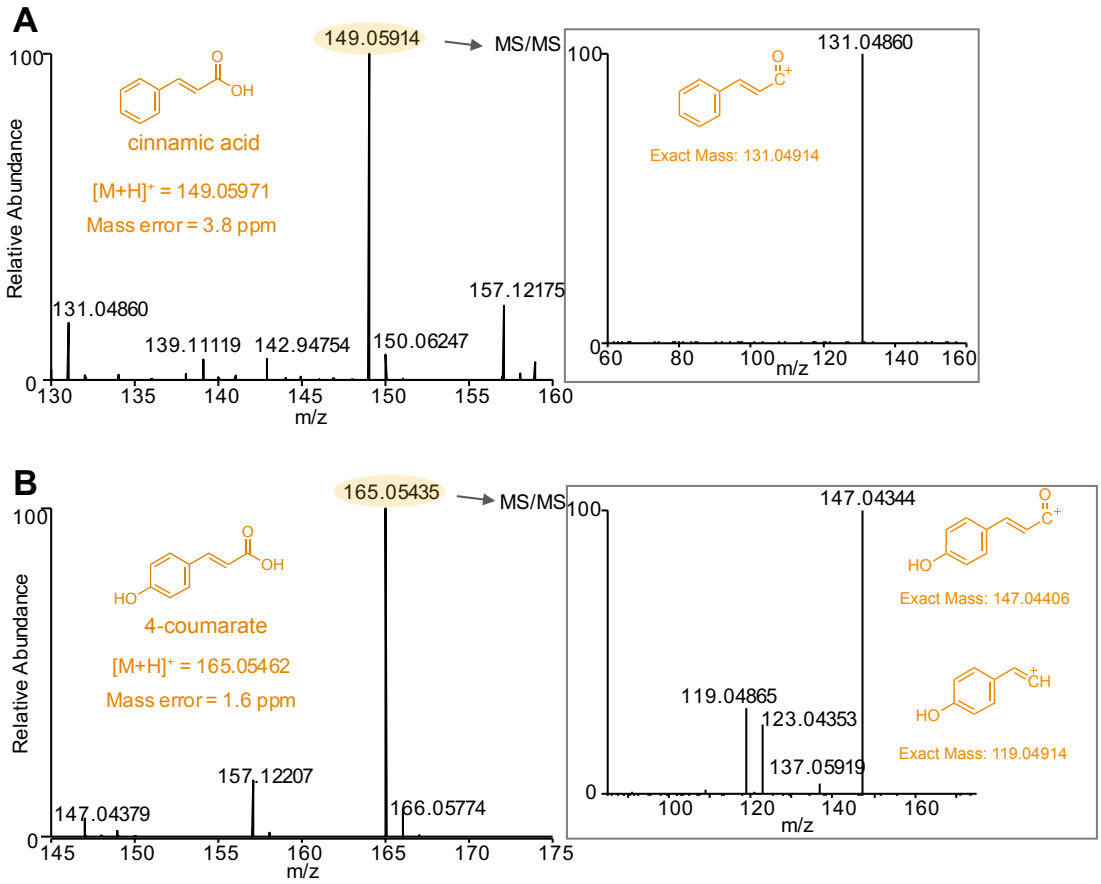

Figure S2. LC-MS analysis of the reaction products of EwPAL. (A) Mass spectrum analysis of EwPAL reaction product using L-phenylalanine as substrate, MS/MS analysis interpretation is shown on the right. **(B)** Mass spectrum analysis of EwPAL reaction product using L-tyrosine as substrate, MS/MS analysis interpretation is shown on the right.

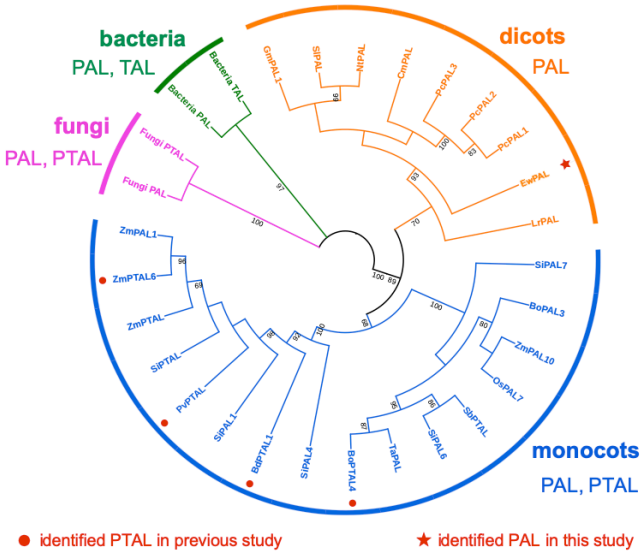

Figure S3. Phylogenetic tree analysis of EwPAL with characterized PALs, TALs and PTALs from plants, fungi and bacteria. PAL, phenylalanine ammonia-lyase, TAL, tyrosine ammonia-lyase, PTAL, bifunctional phenylalanine/tyrosine ammonia-lyase. Bootstrap values (based on 1,000 replicates) >50% are indicated for their corresponding edges.

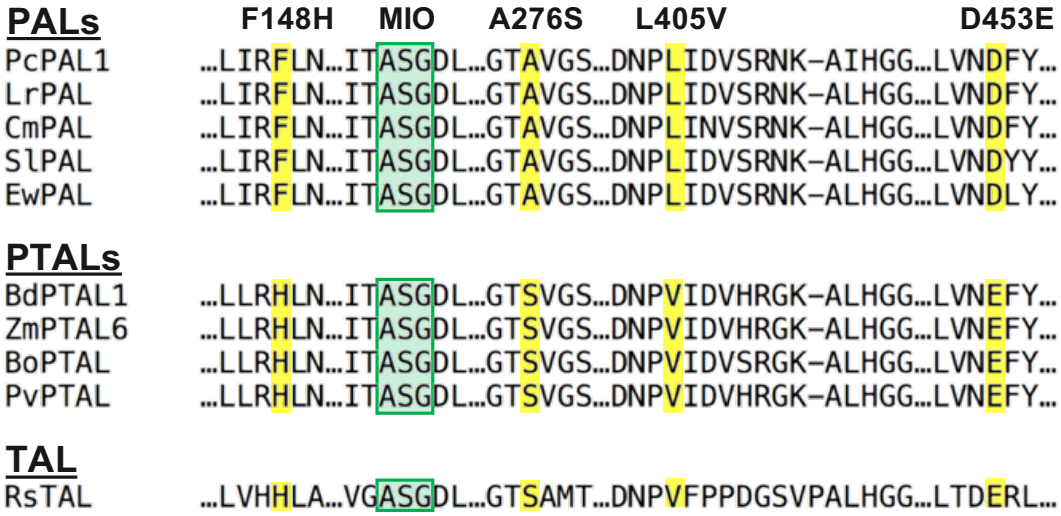

Figure S4. Multiple sequence alignment of EwPAL with representative PALs, PTALs and TAL. The three conserved residues [alanine-serine-glycine (ASG)] responsible for the co-factor 4-methylideneimidazole-5-one (MIO) generation are highlighted in green box. Residues involve in substrate selectivity between PAL/TAL (F148H, A276S, L405V and D453E) are marked in yellow.

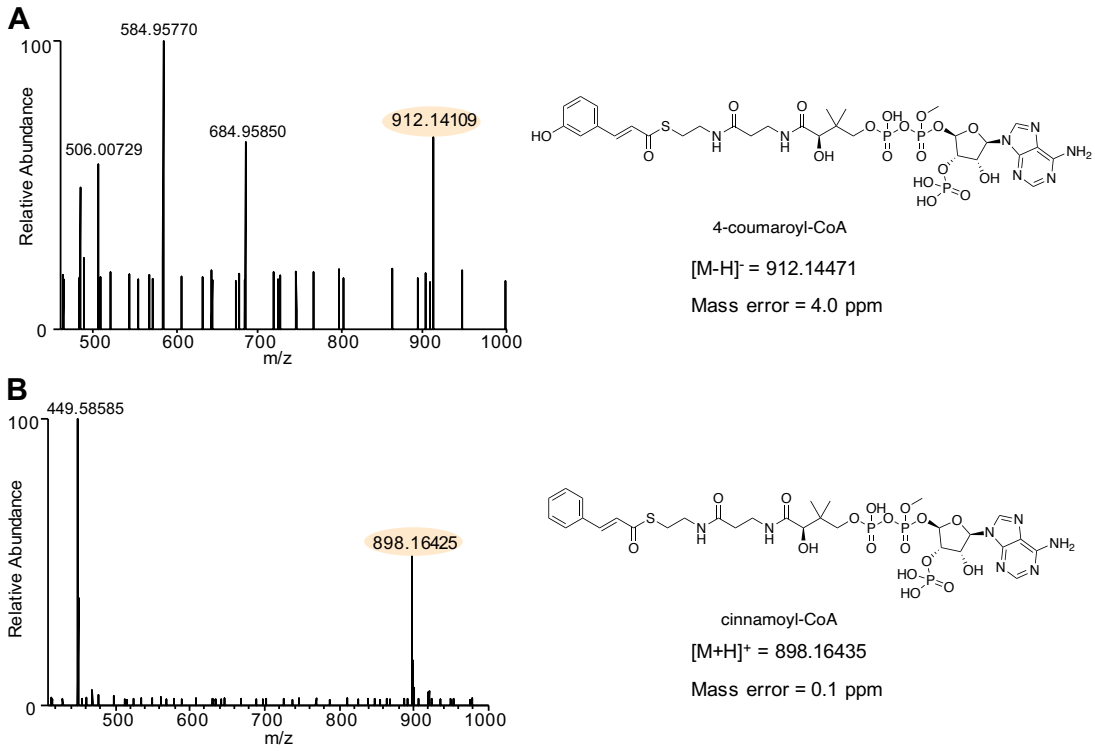

Figure S5. LC-MS analysis of the reaction product of Ew4CL1. (A) Mass spectrum analysis of Ew4CL1 reaction product using 4-coumarate as substrate, the $[M-H]^-$ m/z of the resulting product matches that of 4-coumaroyl-CoA. **(B)** Mass spectrum analysis of Ew4CL1 reaction product using cinnamic acid as substrate, the $[M+H]^+$ m/z of the resulting product matches that of cinnamoyl-CoA.

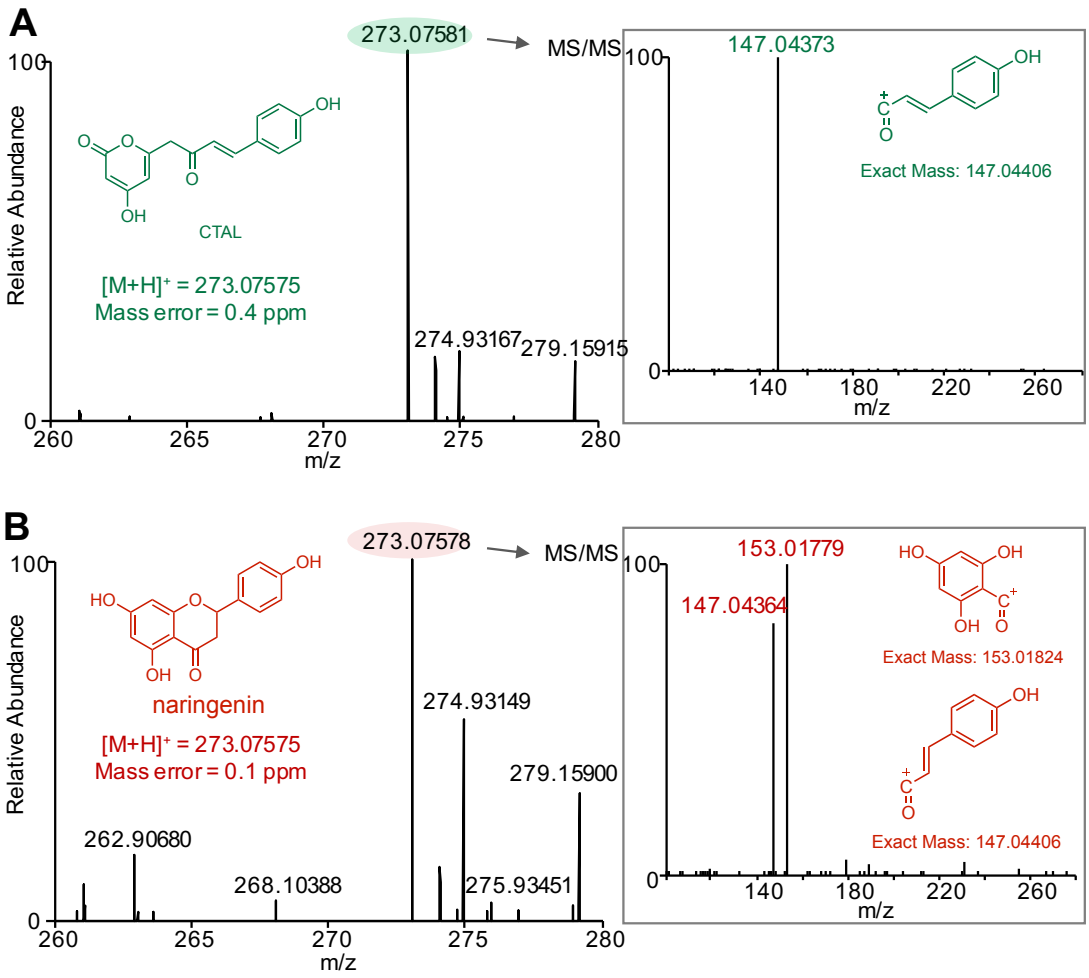

Figure S6. LC-MS analysis of the reaction products of EwCHS1 using 4-coumaroyl-CoA as substrate. (A) $[M+H]^+$ m/z and MS/MS spectrum of Product 1 matches that of CTAL**. (B)** $[M+H]^+$ m/z and MS/MS spectrum of Product 2 matches that of naringenin**.**

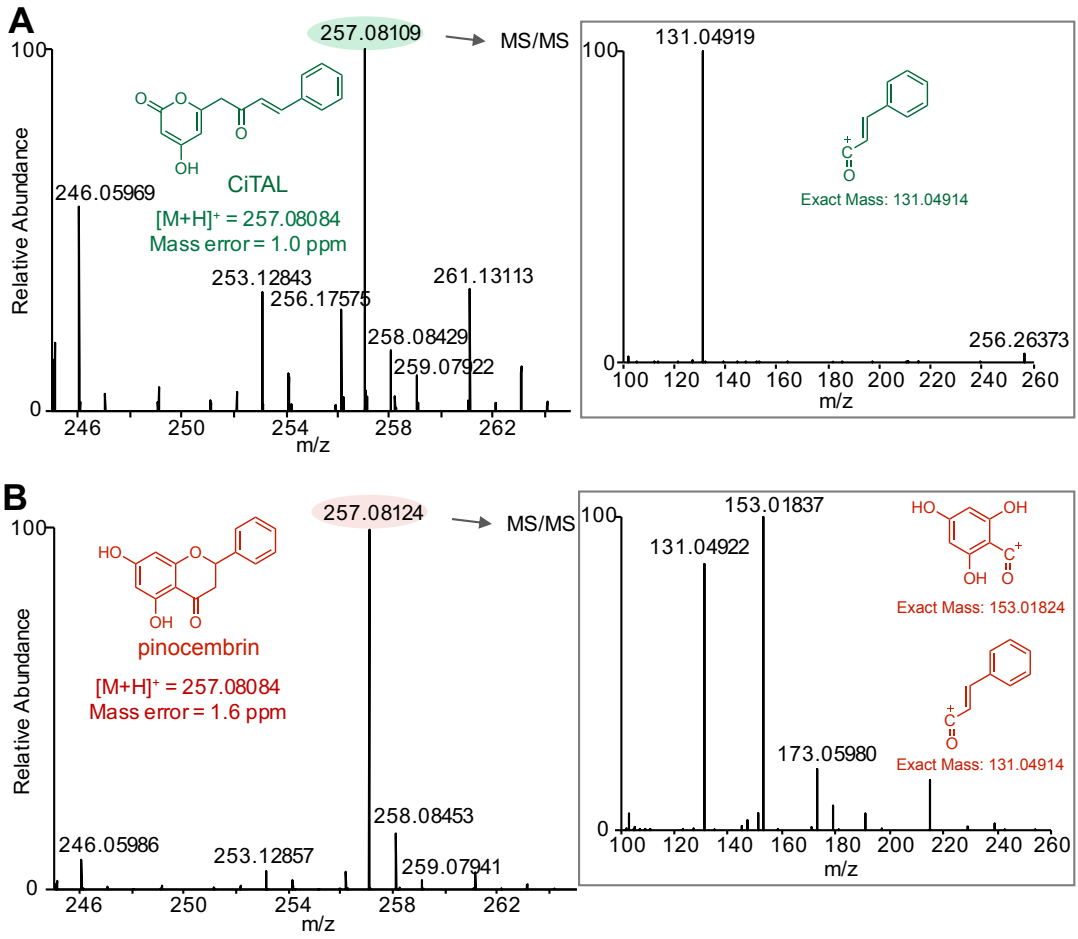

Figure S7. LC-MS analysis of the reaction products of EwCHS1 using cinnamoyl-CoA as substrate. **(A)** $[M+H]^+$ m/z and MS/MS spectrum of Product 1 matches that of CiTAL. **(B)** $[M+H]^+$ m/z and MS/MS spectrum of Product 2 matches that of pinocembrin.

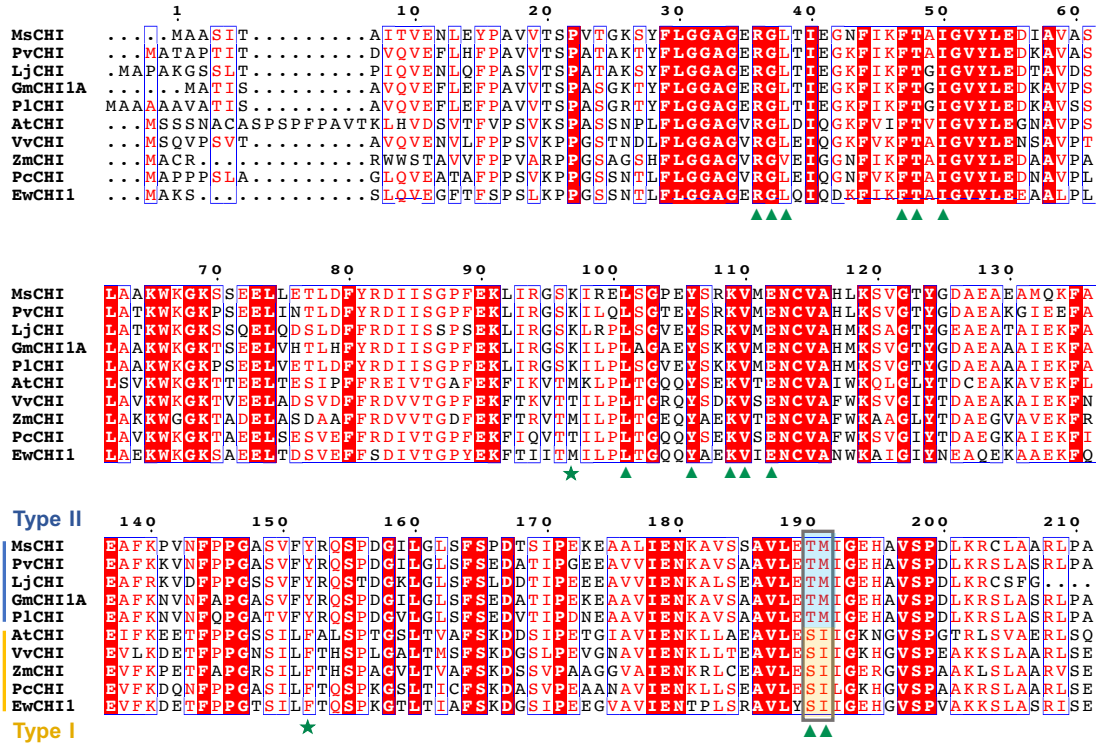

Figure S8. Multiple sequence alignment of EwCHI1 with other type I and type II CHIs from higher plants. Included are MsCHI (*Medicago sativa*, P28012), PvCHI (*Phaseolus vulgaris*, P14298), LjCHI (*Lotus japonicas*, CAD69022), GmCHI1A (*Glycine max*, AAT94358.1), PlCHI (*Pueraria lobate*, Q43056.1), AtCHI (*Arabidopsis thaliana*, P41088), VvCHI (*Vitis vinifera*, P51117). ZmCHI (*Zea mays*, Q08704.1) and PcCHI (*Pyrus communis*, A5HBK6.1). The multiple sequence alignment was built with MUSCLE and visualized using ESPript [1]. The apolar active site cleft constituting residues are indicated with green triangles, green stars indicate residues involved in hydrogen bond network in the active site. The grey box identifies residues proposed to determine substrate preference for isoliquiritigenin and naringenin chalcone.

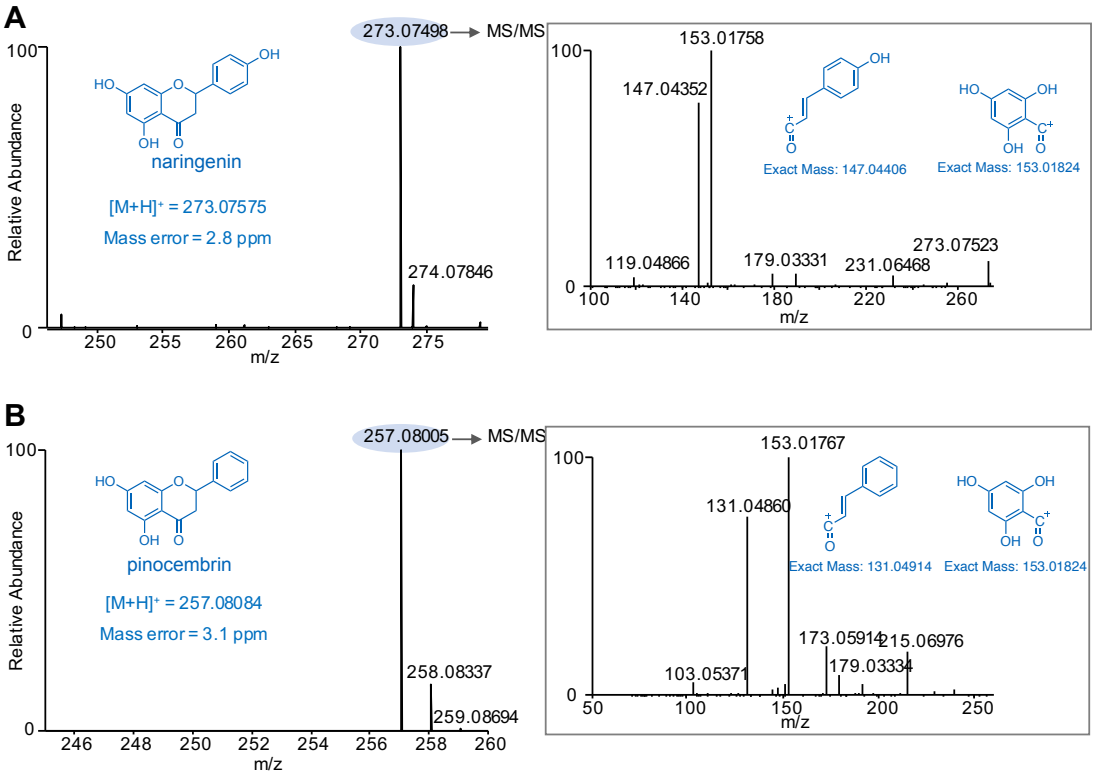

Figure S9. LC-MS analysis of the reaction product of EwCHI1. (A) Mass spectrum analysis of EwCHI1 reaction product using naringenin chalcone as substrate, the $[M+H]^+$ m/z of the resulting product matches that of naringenin. MS/MS analysis interpretation is shown on the right. **(B)** Mass spectrum analysis of EwCHI1 reaction product using pinocembrin chalcone as substrate, the $[M+H]^+$ m/z of the resulting product matches that of pinocembrin. MS/MS analysis interpretation is shown on the right.

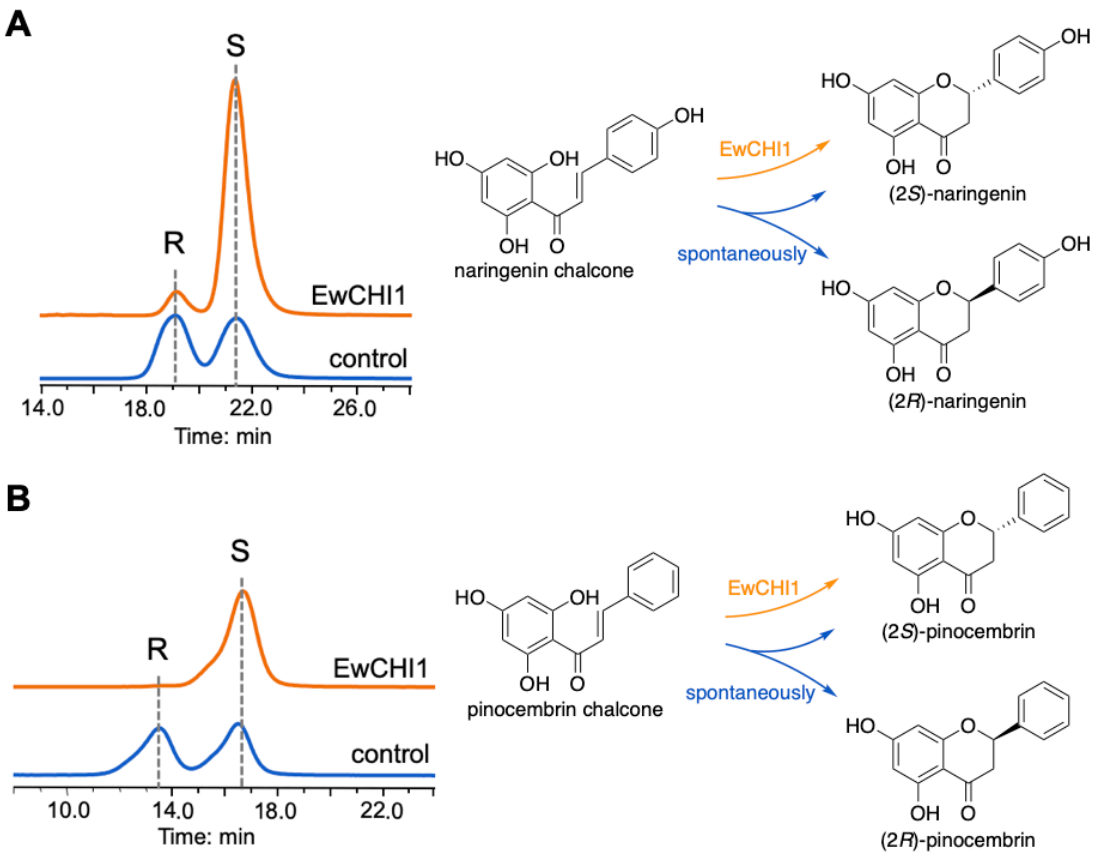

Figure S10. Chiral chromatography analysis of the EwCHI1 catalysed reactions. (A) Chiral chromatography analysis of the products of recombinant EwCHI1 using naringenin chalcone as substrate. R, (2*R*)-naringenin; S, (2*S*)-naringenin. Boiled EwCHI1 was used for negative control. **(B)** Chiral chromatography analysis of the reaction products of recombinant EwCHI1 using pinocembrin chalcone as substrate. R, (2*R*)-pinocembrin; S, (2*S*)-pinocembrin. The UV absorption was monitored at 280 nm.

References

1. Robert X, Gouet P, *Nucleic Acids Res.* , 2014, **42**, W320-W324.