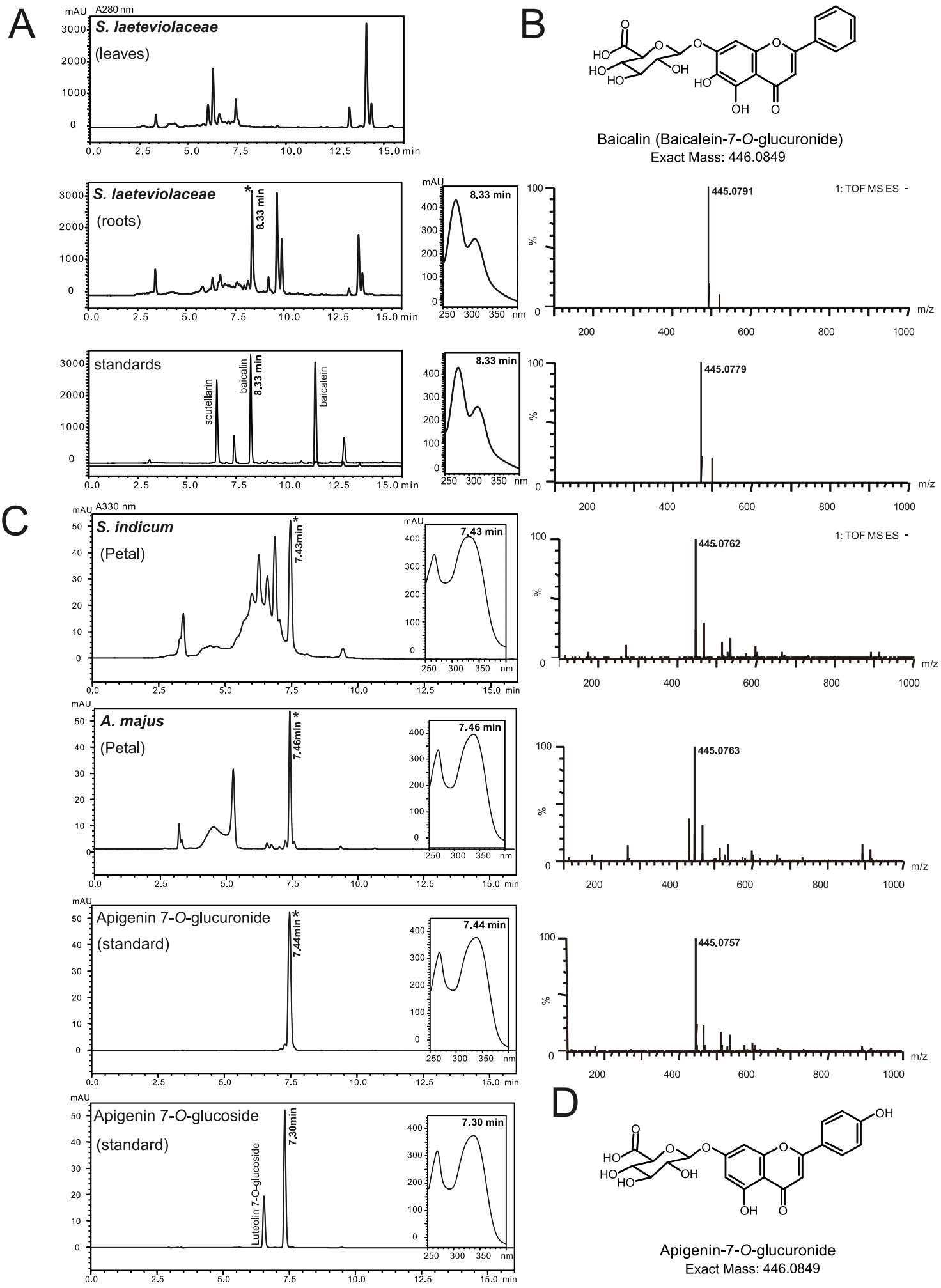
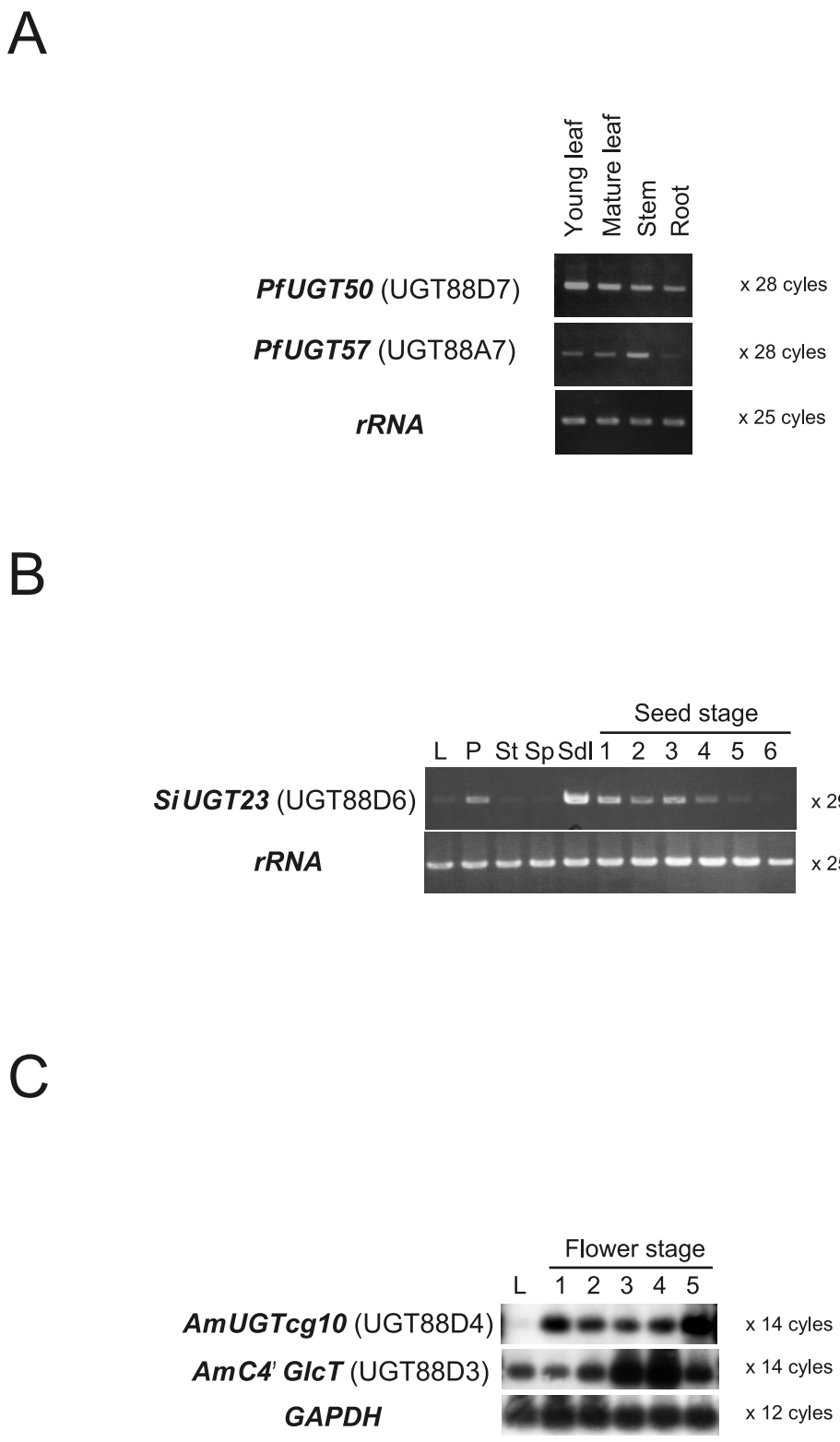
Supplemental Data. Noguchi et al. (2009). Local Differentiation of Sugar-donor Specificity of Flavonoid Glycosyltransferase in Lamiales.



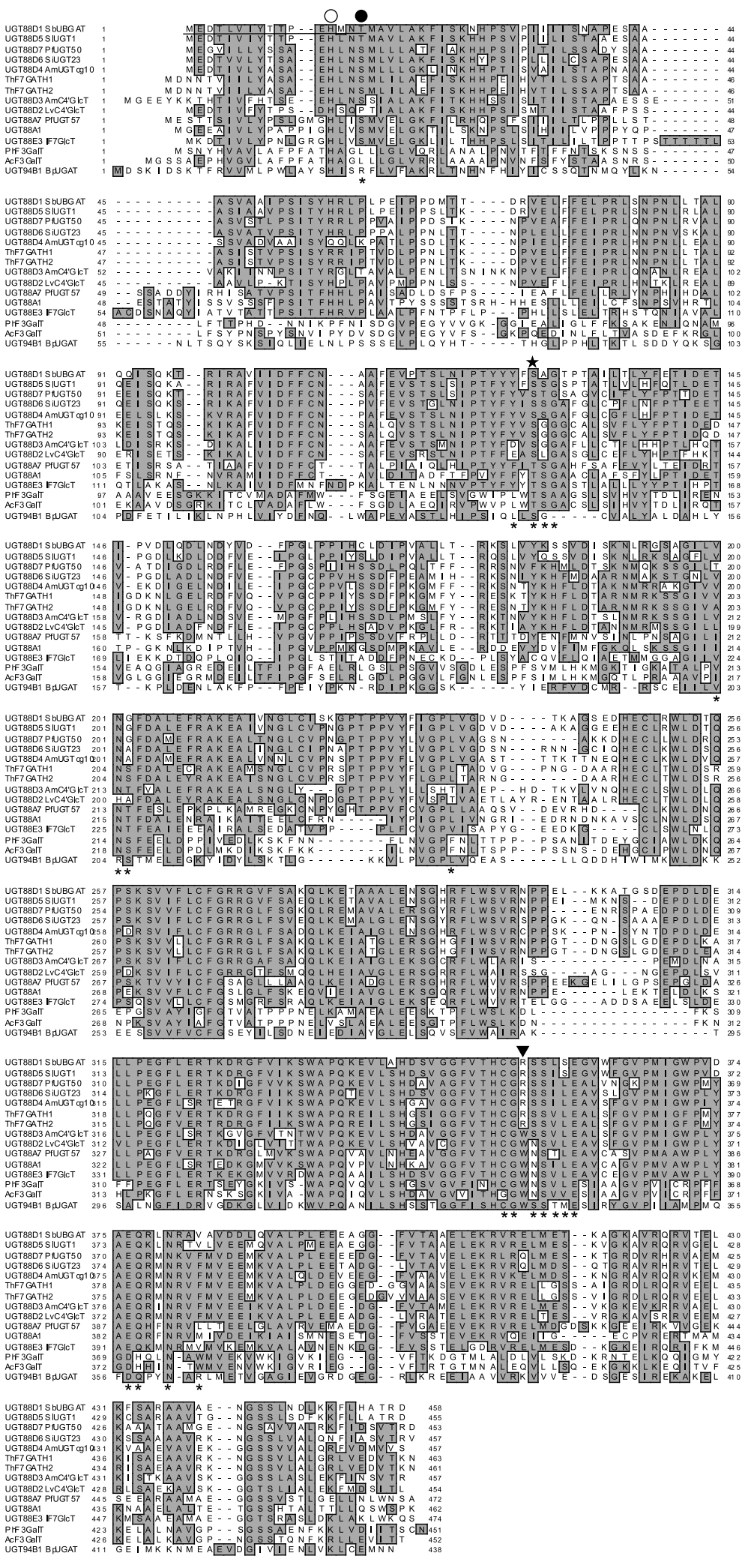
Supplemental Figure 1. Metabolite profiles of Lamiales plants.

(A) HPLC chromatograms of the extracts from leaves (top) and roots (middle) of *Scutellaria laeteviolaceae*, and authentic standards (bottom) at A280 nm . The baicalin in the roots is marked by asterisk. UV and MS spectra of baicalin from the roots of *S. laeteviolaceae* and an authentic standard are shown on the right of each chromatogram. (B) Chemical structure of baicalin. (C) HPLC chromatogram of petal extracts from *S. indicum* and *A. majus*, and authentic standards of apigenin 7-*O*-glucuronide and apigenin 7-*O*-glucoside at A330 nm (from top to bottom). Asterisk indicates apigenin 7-*O*-glucuronide, whose UV spectrum is shown in the inset. The mass spectrum of the marked peak is shown on the right of the chromatogram. (D) Chemical structure of apigenin 7-*O*-glucuronide.



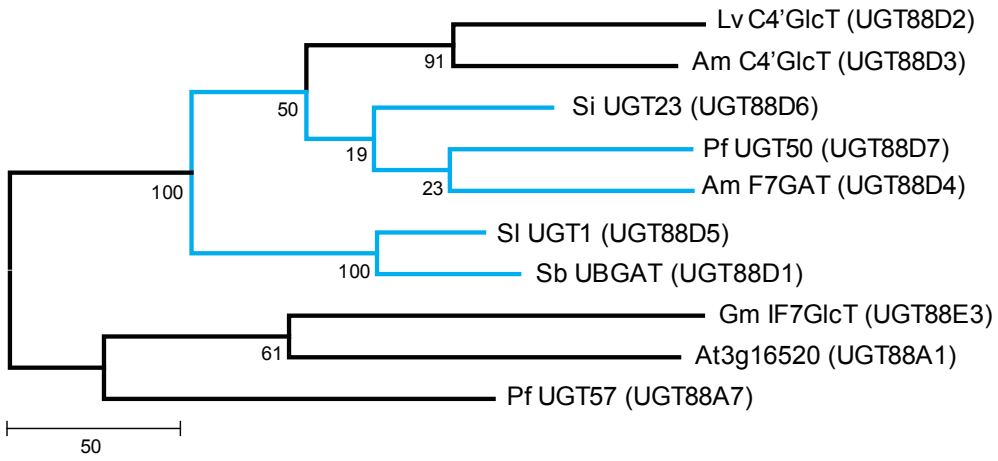
Supplemental Figure 2, Expression profile of three Lamiales F7GAT genes

(A) Reverse transcription-polymerase chain reaction (RT-PCR) analysis of Pf UGT50/UGT88D7 and Pf UGT57/UGT88A7 in organs of *P. frutescens* cultivar Aochirimen (B) RT-PCR analysis of Si UGT23/UGT88D6 in organs of *S. indicum* cultivar Masekin (C) RT-PCR analysis of Am UGTcg10/UGT88D4 and Am C4' GlcT/UGT88D3 in organs of *A. majus* cultivar Butterfly yellow. L; leaves. These results were replicated by repeated experiments (n=2).



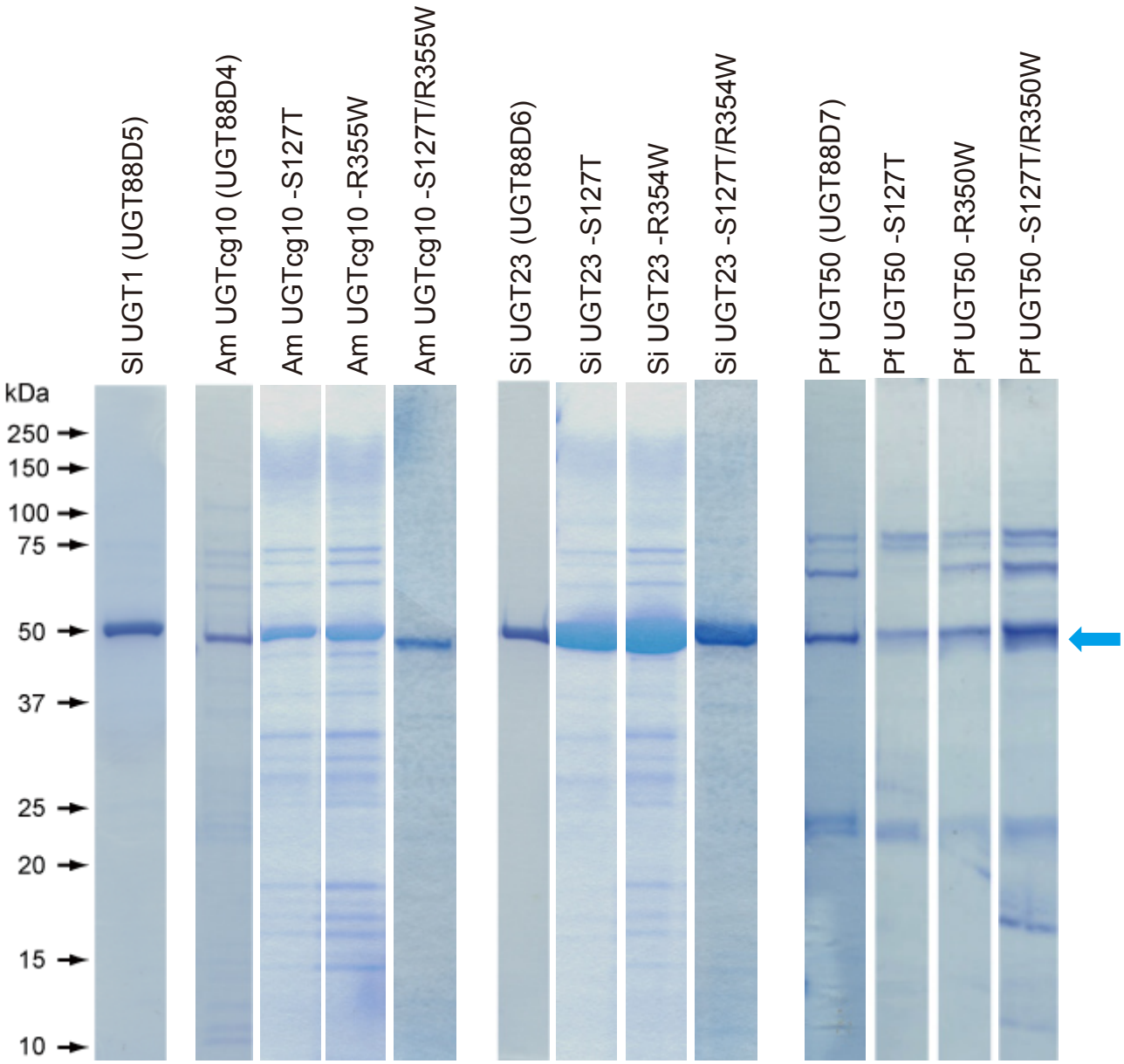
Supplemental Figure 3, Amino acid alignment of UGT88-related enzymes

The multiple alignment was performed using a CLUSTAL-W program packaged in MACVECTOR 7.2.2 software (Accelrys, San Diego)(Thompson et al., 1994. *Nucleic Acid Res*. 22, 4673-4680). The empty circle, black circle, star, and triangle indicate the position of His20 in Vv GT1, Arg25 in UGT94B1, Ser127 in UGT88D7, and Arg350 in UGT88D7, respectively. Other amino acid residues predicted to be in the vicinity of Arg350 (within 5 Å) of UGT88D7 are marked by an asterisk beneath the multiple alignment. The region of Sb UBGAT extended by 5'RACE is underlined.



Supplemental Figure 4 Evolutionary relationships of cluster IIIb of flavonoid UGTs

The evolutionary history was inferred using the Maximum Parsimony (MP) method (Eck, R. V. and Dayhoff, M. O. 1966. Atlas of Protein Sequence and Structure. National Biomedical Research Foundation, Silver Springs, Maryland). The consistency index is (0.775033), the retention index is (0.551813), and the composite index is 0.460635 (0.427673) for all sites and parsimony-informative sites (in parentheses). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches (Felsenstein, J. 1985. *Evolution* **39**:783-791). The MP tree was obtained using the Close-Neighbor-Interchange algorithm (Nei, M. and Kumar, S. 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York.3, pg. 128) with search level 3 in which the initial trees were obtained with the random addition of sequences (10 replicates). The tree is drawn to scale, with branch lengths calculated using the average pathway method and are in the units of the number of changes over the whole sequence. All positions containing gaps and missing data were eliminated from the dataset (Complete Deletion option). There were a total of 410 positions in the final dataset, out of which 210 were parsimony informative.
Blue line indicates the lineages of Lamiales F7GATs. Phylogenetic analyses were conducted using MEGA4 software (Tamura, K., et al., 2007. *Mol. Biol. Evol.* **24**:1596-1599).



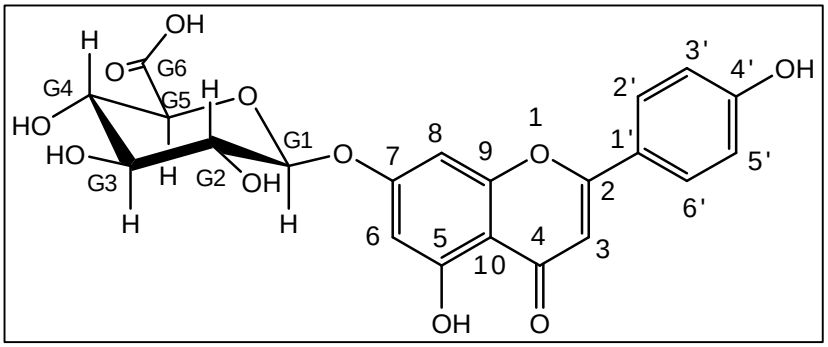
Supplemental Figure 5. SDS-PAGE after affinity purification of His-Tagged Lamiales F7GAT proteins expressed in *E. coli.*

Four His-Tagged Lamiales F7GAT proteins and their mutants were eluted with 200 mM imidazole, stained with Coomassie Blue, and detected as a major band close to the 50 kDa size expected from the molecular weight calculation (arrow). Using bovine serum albumin as a standard, the protein concentration was determined by the intensity of the protein band on the SDS-polyacrylamide gel.

Supplemental Table 1, NMR of apigenin 7-*O*-glucuronide

(in DMSO-*d*6)

#	^{1}H δ		*J* Hz	^{13}C
1				
2				164.21
3	6.87	s		103.05
4				181.91
5				161.09
6	6.47	d	2	99.26
7				162.39
8	6.86	d	2	94.55
9				156.86
10				105.36
1'				120.92
2'	7.96	d	9	128.52
3'	6.95	d	9	115.91
4'				161.28
5'	6.95	d	9	115.91
6'	7.96	d	9	128.52
G1	5.28	d	7	99.04
G2	3.41	brt	9	72.66
G3	3.33	brt	9	75.53
G4	3.33	brt	9	71.13
G5	4.05	d	9.5	75.30
G6				169.94



^{1}H and ^{13}C NMR chemical shifts and structure of the apigenin 7-*O*-monoglucuronide that was used as the standard in the enzyme assays

Supplemental Table 2, Kinetics of Lamiales F7GATs.

Pf UGT50/UGT88D7

Acceptor	K_m(μM)	k_{cat}(s^{-1})	k_{cat}/K_m (S^{-1} . mM^{-1})
Apigenin	11±3	8.8±0.5	800
Baicalein	49±11	4.3±0.3	88
Scutellarein	120±59	19±3	150
Kaempferol	100±24	4.0±0.4	38
Quercetin	130±100	5.9±2.1	47
Naringenin	680±26	6.6±0.2	10
UDP-GA	36±4		120

Am UGTcg10/UGT88D4

Acceptor	K_m(μM)	k_{cat}(s^{-1})	k_{cat}/K_m (S^{-1} . mM^{-1})
Apigenin	560±450	17±8	30
Baicalein	52±20	3.6±0.4	70
Scutellarein	46±27	29±8	640
Kaempferol	6.7±1.9	11±1	1600
Quercetin	210±140	3.4±1.1	16
Naringenin	100±3	17±0.2	170
UDP-GA	70±15		51

Si UGT23/UGT88D6

Acceptor	K_m(μM)	k_{cat}(s^{-1})	k_{cat}/K_m (S^{-1} . mM^{-1})
Apigenin	21±25	17±7	800
Baicalein	21±16	9.2±2.0	450
Scutellarein	78±30	24±4	300
Kaempferol	130±10	10±1	79
Quercetin	160±80	7.3±1.7	47
Naringenin	100±10	3.2±0.1	31
UDP-GA	16±3		570

Sl UGT1/UGT88D5

Acceptor	K_m(μM)	k_{cat}(s^{-1})	k_{cat}/K_m (S^{-1} . mM^{-1})
Apigenin	65±26	0.46±0.07	7.1
Baicalein	93±24	10±1	110
Scutellarein	100±50	23±4	220
Kaempferol	130±50	6.9±1.0	53
Naringenin	130±3	3.8±0.02	12
UDP-GA	36±3		280

These kinetic parameters were determined using each purified protein. As described in the Methods section, kinetic parameters were calculated by fitting the initial velocity data to the Michaelis-Menten equation by nonlinear regression. Baicalein was used as sugar-acceptor for the *K*m value on UDP-GA.

Supplemental Table 3, Sugar-donor specificity of Si UGT23, Am UGTcg10 and their mutants.

		Relative activity (%)		
Name	**UGT**	**UDP-Glc**	**UDP-Gal**	**UDP-GA**
Si UGT23	UGT88D6 (WT)	1	0	100
	UGT88D6 (S127T)	1	0	100
	UGT88D6 (R354W)	100	5	4
	UGT88D6 (S127T/R354W)	100	4	1
Am UGTcg10	UGT88D4 (WT)	0	0	100
	UGT88D4 (S127T)	0	1	100
	UGT88D4 (R355W)	100	6	24
	UGT88D4 (S127T/R355W)	100	6	15

The glycosylation activity of each enzyme on three types of sugar-donor (UDP-Glucose, UGT-Galactose, and UDP-Glucuronic acid) was tested. Baicalein was used as sugar-acceptor for the evaluation of sugar-donor specificity. Products were quantified based on an absorbance peak at 350 nm. The highest activity in the three UDP-sugars is set at 100%.

Supplemental Table 4, Primer sequences

Primer name	UGT	Primer Sequence (5'- -3')	Underlined site
Sl UGT-Fw	UGT88D5	AAACATATGGCGGTGCTGGCGAAGTTC	
Sl UGT-Rv	UGT88D5	TTTTGATCATTAATCCCGAGTGGCGTGAAG	
GR-Sl UGT-Rv	UGT88D5	TGGGAGGCAAACCAGGGATCTCGACAA	
Sl UGT-nest-Rv	UGT88D5	AATCATCCAAATCTTTAAGGT	
GR-Sl UGT-Fw	UGT88D5	AGAAGGGGTGTGTTCTCCGCTGAGCAA	
Sl UGT-nest-Fw	UGT88D5	GAACAGCGGTCACAGATTTCT	
Pf 50-Fw	UGT88D7	AAACATATGGAAGGCGTCATACTTC	NdeI
Pf 50-Rv	UGT88D7	TTTTGATCATTAATCACGAGTTACGGAATC	BclI
Pf 57-Fw	UGT88A7	AGGATCCGATGGAAAGTACCACTAGTATAG	BamHI
Pf 57-Rv	UGT88A7	TGGATCCTTAGGCAGAGTTCCACAAG	BamHI
Am UGTcg10-Fw	UGT88D4	AAACATATGGAGGACACTATCGTTCTC	NdeI
Am UGTcg10-Rv	UGT88D4	TTGGATCCTTAAGAAACCACCATATCAAC	BamHI
Sl UGT1-Fw	UGT88D5	AAACATATGGAGGACACGATTGTTATC	NdeI
Sl UGT1-Rv	UGT88D5	TTCATATGTCAATCCCTCGTGGCCAGAAG	NdeI
Si UGT23-Fw	UGT88D6	CACCATATGGAAGACACCGTTGTCCTCTA	NdeI
Si UGT23-Rv	UGT88D6	GGATCCTAACATCACTCAAACCCGAGTCA	BamHI
Pf UGT2-Fw	UGT73A7	AAACATATGGTGACAATGGCGATAAACG	NdeI
Pf UGT2-Rv	UGT73A7	TTTGGATCCTTACTGGACTACAATTTCAG	BamHI
Pf UGT31-Fw	UGT73A13	AAACATATGAAACAGCTACACATCGTTC	NdeI
Pf UGT31-Rv	UGT73A13	TTTGGATCCCTAGTTTATGTTATGTTTGGTG	BamHI
Am UGT21-Fw	UGT73A9	AAACATATGGGAAAACTTCACATTGCC	NdeI
Am UGT21-Rv	UGT73A9	TTTGGATCCCTAGTTTAAGTCTTGTTTTTC	BamHI
Am UGT36-Fw	UGT73E2	AAAGGATCCGATGGCCATTCATGAACAAAAAC	BamHI
Am UGT36-Rv	UGT73E2	TTTGGATCCTCATGTATTTATAACTGTAACACC	BamHI
Am UGT38-Fw	UGT73N1	AAACATATGGCCTTTCAAATTCAACC	NdeI
Am UGT38-Rv	UGT73N1	TTTGGATCCTTACACATCCCTCGCTACAC	BamHI
At UGT88A1-Fw	UGT88A1	CACCCATATGGGTGAAGAAGCTATAGTT	NdeI

At UGT88A1-Rv	UGT88A1	GGATCCTCACTTTGGGCTCCACGACTG	BamHI
Pf 50-S127T-Fw	UGT88D7	TTCTACGTCACCACCGGTTCC	**Thr**
Pf 50-S127T-Rv	UGT88D7	GGAACCGGTGGTGACGTAGAA	**Thr**
Pf 50-R350W-Fw	UGT88D7	ACTCACTGTGGGTGGAGCTCGATT	**Trp**
Pf 50-R350W-Rv	UGT88D7	AATCGAGCTCCACCCACAGTGAGT	**Trp**
Pf 57-T139S-Fw	UGT88A7	TACTTCATCAGCTCCGGGGCTCA	**Ser**
Pf 57-T139S-Rv	UGT88A7	TGAGCCCCGGAGCTGATGAAGTA	**Ser**
Pf 57-W367R-Fw	UGT88A7	ACTCACTGCGGGCGGAACTCGACT	**Arg**
Pf 57-W367R-Rv	UGT88A7	AGTCGAGTTCCGCCCGCAGTGAGT	**Arg**
Gm IF7GlcT-T150S-Fw	UGT88E3	ACTTCTACTACAGTTCTGGCGCCT	**Ser**
Gm IF7GlcT-T150S-Rv	UGT88E3	AGGCGCCAGAACTGTAGTAGAAGT	**Ser**
Gm IF7GlcT-W371R-Fw	UGT88E3	TCACTGCGGTAGGAACTCGGTGTT	**Arg**
Gm IF7GlcT-W371R-Rv	UGT88E3	AACACCGAGTTCCTACCGCAGTGA	**Arg**
Si UGT23-S127T-Fw	UGT88D6	TACTTCTACATCACCTCCGGCGCGTTT	**Thr**
Si UGT23-S127T-Rv	UGT88D6	AAACGCGCCGGAGGTGATGTAGAAGTA	**Thr**
Si UGT23-R354W-Fw	UGT88D6	CACTGCGGTTGGAGCTCAATTCTGGAA	**Trp**
Si UGT23-R354R-Rv	UGT88D6	TTCCAGAATTGAGCTCCAACCGCAGTG	**Trp**
Am UGTcg10-S127T-Fw	UGT88D4	TTCTATGTCACCAGCGGCGCGTTT	**Thr**
Am UGTcg10-S127T-Rv	UGT88D4	AAACGCGCCGCTGGTGACATAGAA	**Thr**
Am UGTcg10-R355W-Fw	UGT88D4	ACGCACTGTGGGTGGAGTTCGATA	**Trp**
Am UGTcg10-R355R-Rv	UGT88D4	TATCGAACTCCACCCACAGTGCGT	**Trp**
Pf rRNA-Fw		TTTGGTGTGTATGGGCATGTT	
Pf rRNA-Rv		TTGCGCGCCTGCTGCCTTCCTT	
GR-Sb 7GAT-Rv	UGT88D1	GCCTGCCGATCCGCGCAGGTTCTTCGAAAT	
Sb 7GAT-nest-Rv	UGT88D1	AACAAG ACTCTTACGCGTGAGCAA	
Am UGT88-Fw2		ATACCCACTTACTTCTATGTCAGC	
Am UGT88-Rv3		ACCCTATTCATCCTCTGCTCC	
GR-Th 7GATH-Fw		AACGGCTTGTGCGTGCCCCGGTCTCCTACT	
GR-Th 7GATH-Rv		AGTAGGAGACCGGGGCACGCACAAGCCGTT	
Th F7GATH-nest-Fw		AGACAATGGGTCCTTGGGTGA	

Th F7GATH-nest-Rv	GAGTAAATTGGCGGGCAACCT