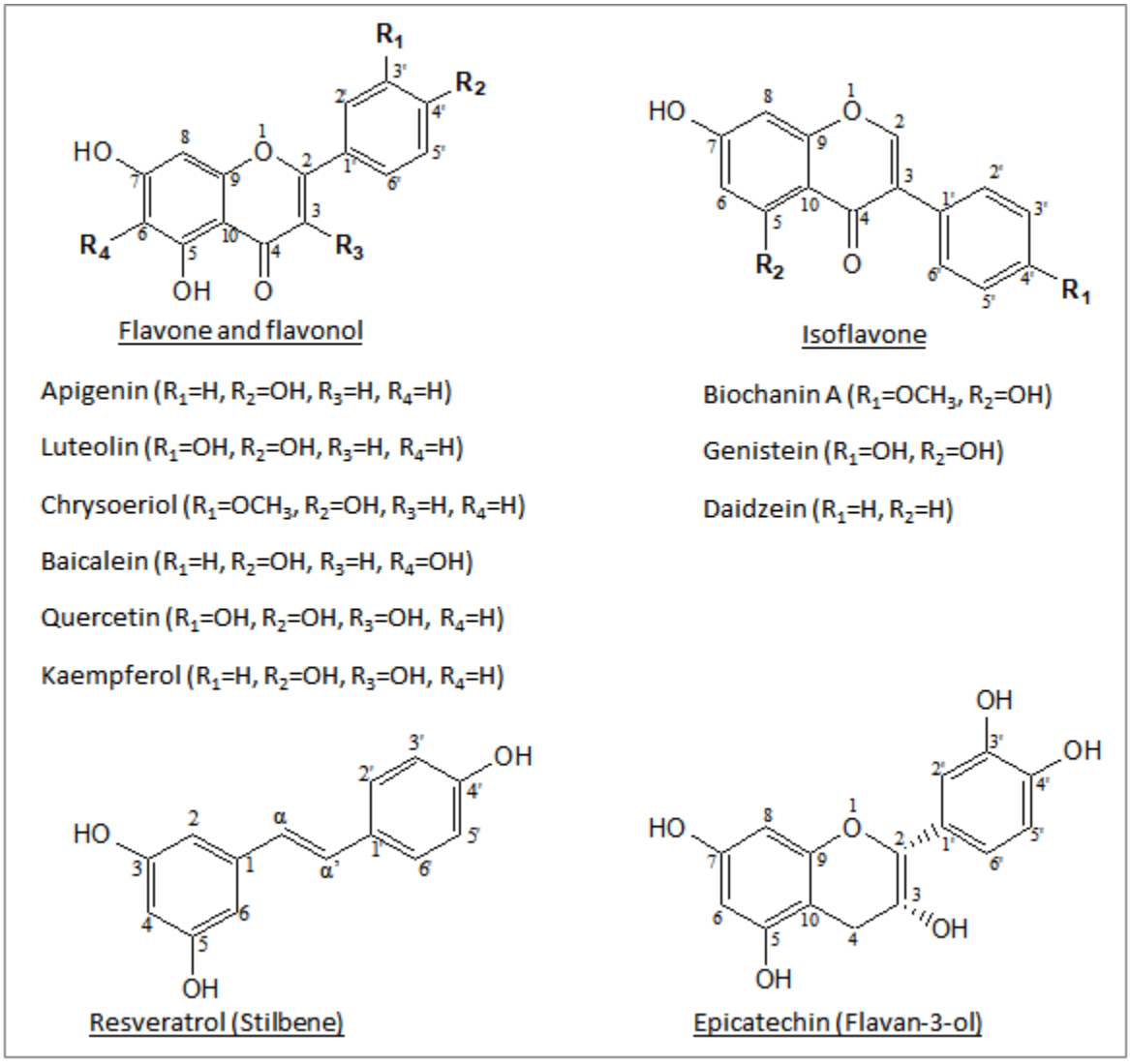
UGT84F9 is the major flavonoid UDP-glucuronosyltransferase in *Medicago truncatula* leaves[1]

Olubu A. Adiji[a], Maite L. Docampo-Palacios[a], Anislay Alvarez-Hernandez[a], Giulio M. Pasinetti[b], Xiaoqiang Wang[a] and Richard A. Dixon[a,2,3]

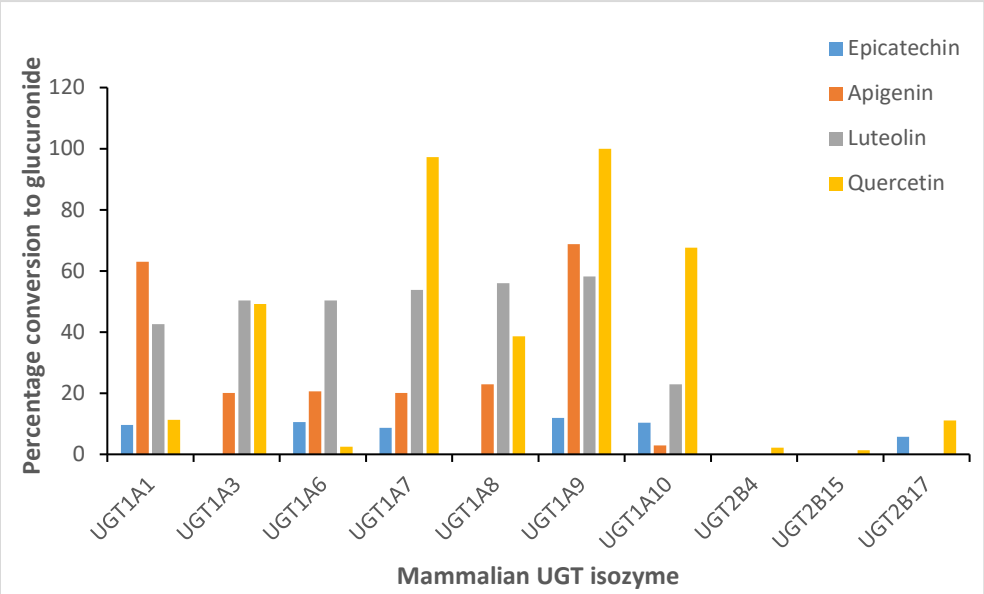
[a]BioDiscovery Institute and Department of Biological Sciences, University of North Texas, Denton, Texas, USA

[b]Department of Psychiatry, The Mount Sinai School of Medicine, New York, NY, USA

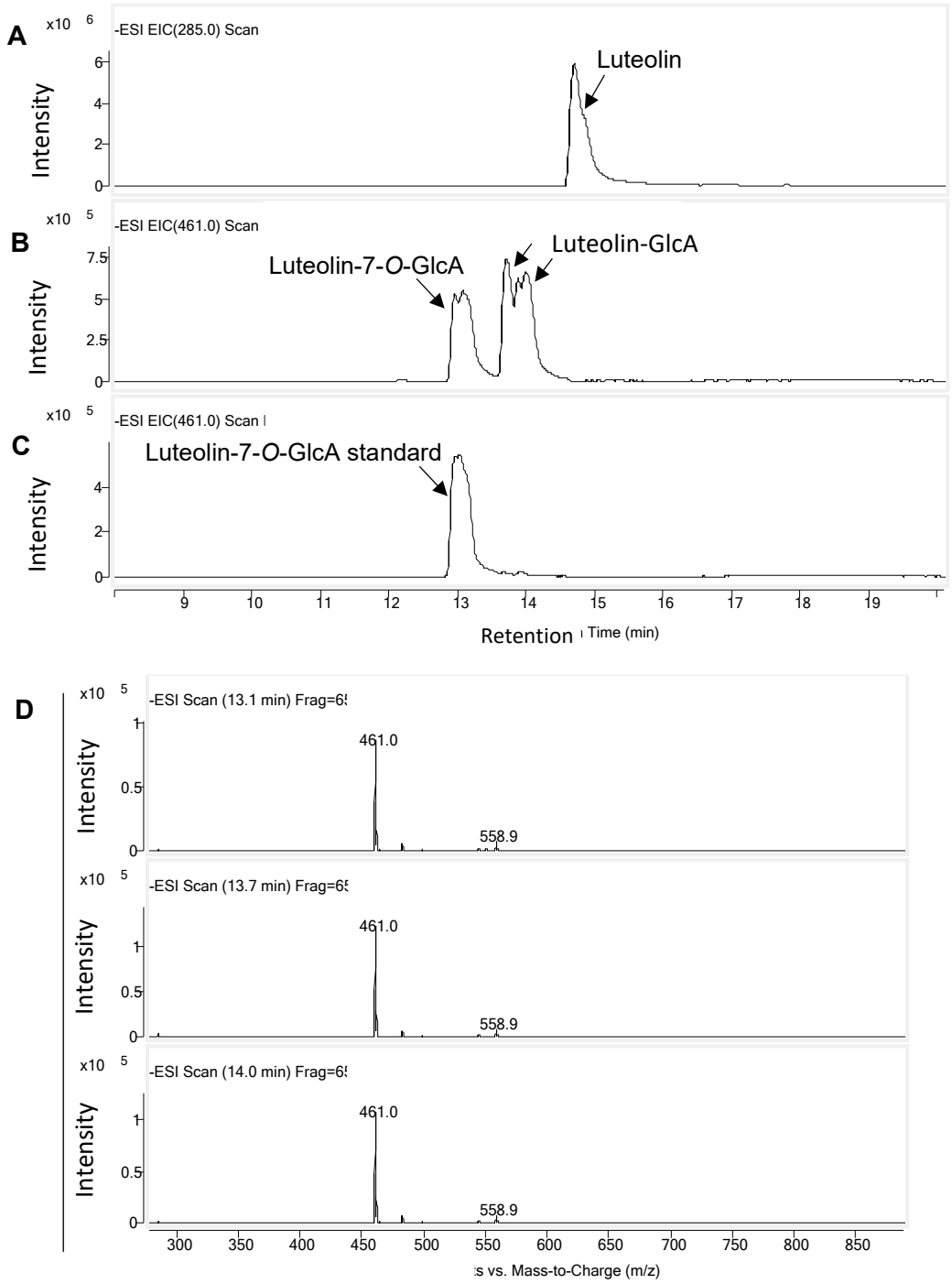
SUPPLEMENTAL DATA



Supplemental Figure S1. Chemical structures of the major (iso)flavonoid compounds tested as substrates for UG(A)Ts in the present work.

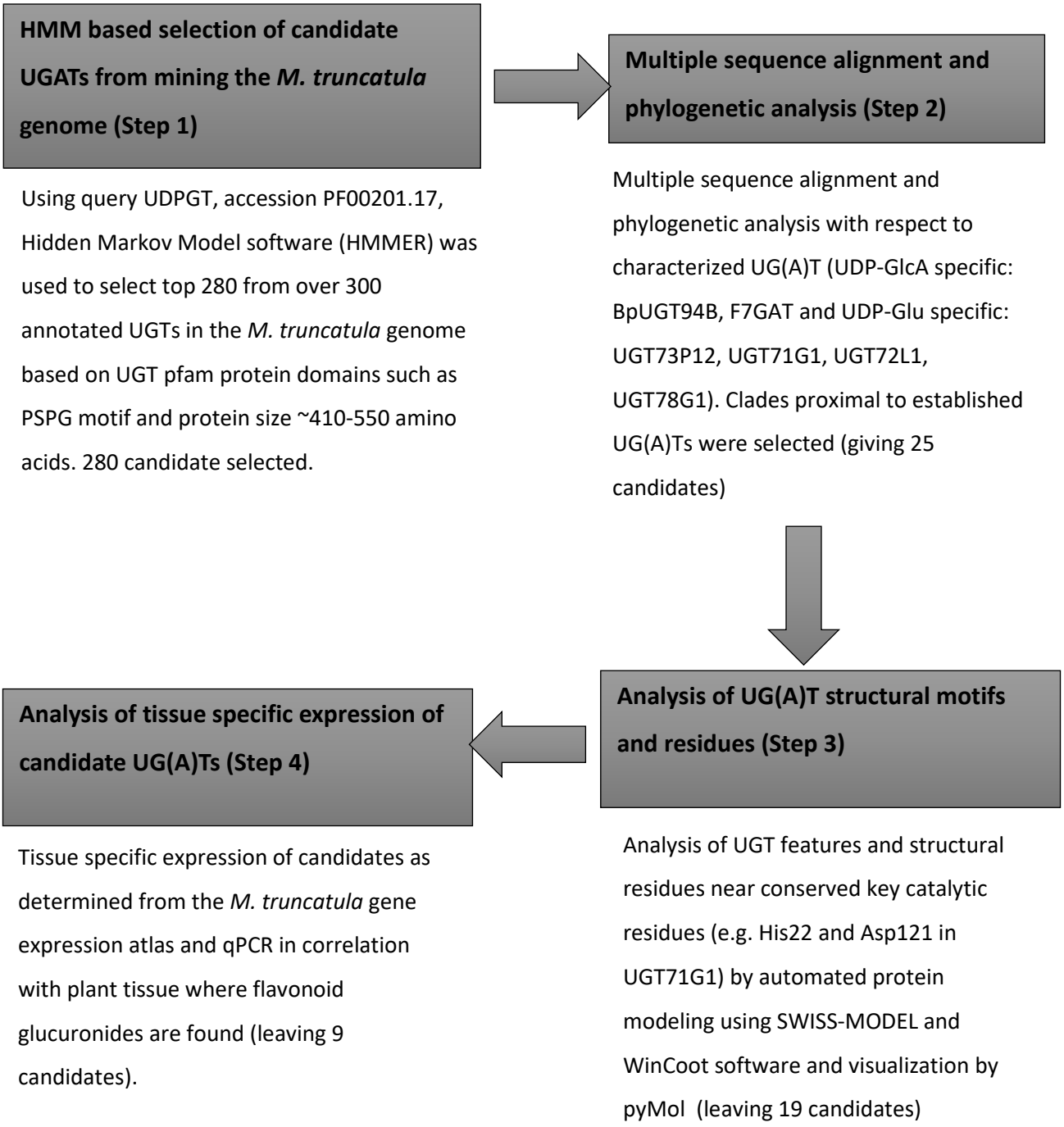


Supplemental Figure S2. Screening of mammalian UGATs for activity against flavonoid compounds. Mammalian UGATs were incubated with the acceptor substrate (0.75 mM) and UDP-GlcA (2 mM) for 2 h, and the percentage conversion to the glucuronidated products was measured by HPLC. Results are expressed relative to UGT1A9, where 100 % activity for quercetin = 2.79 nmol/min/mg protein. The screening assay was performed one time.

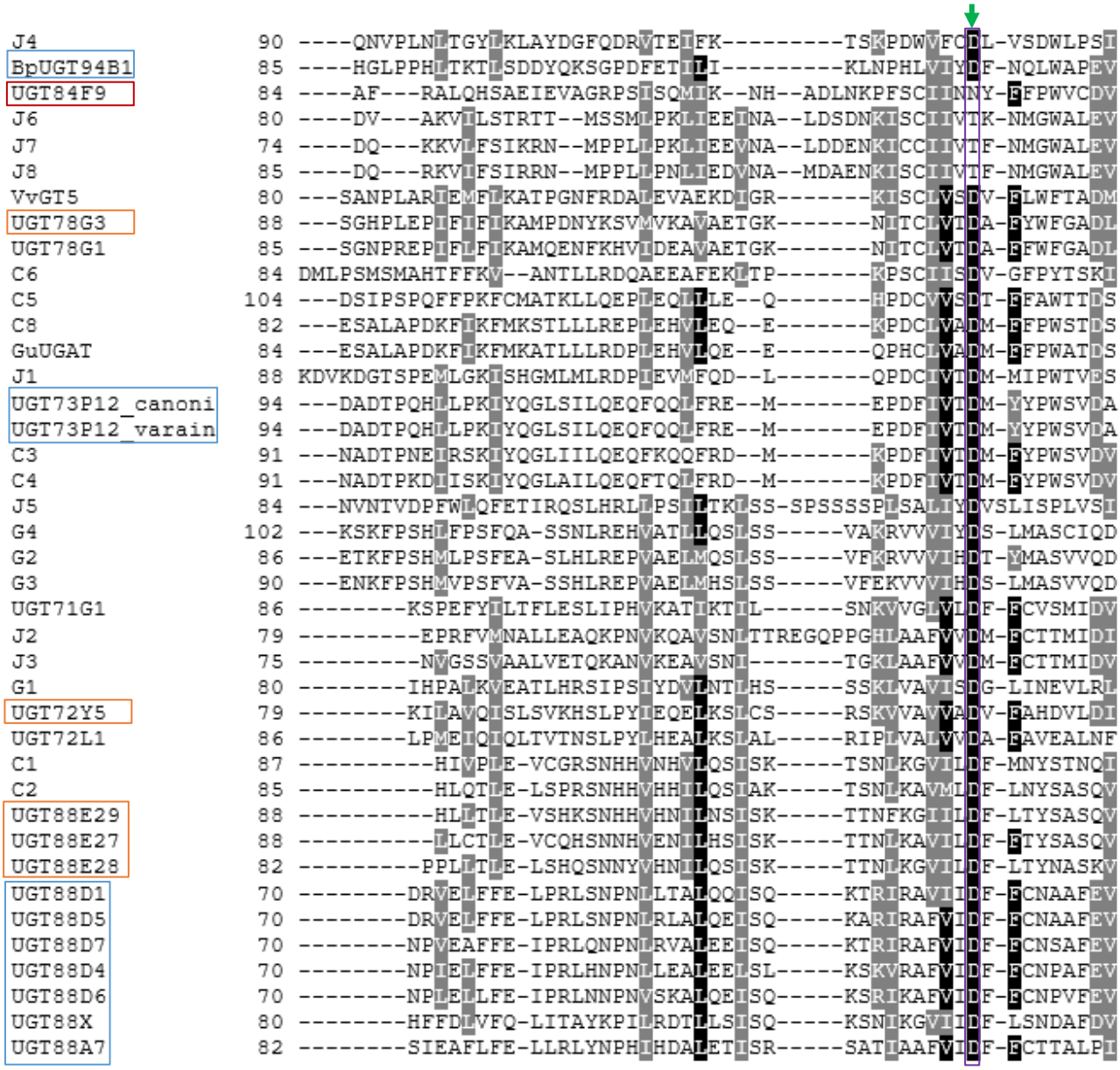


Supplemental Figure S3. LC-MS analysis of the reaction products from incubating UGT1A9 with luteolin as acceptor substrate and UDP-GlcA as sugar donor. A, LC-MS chromatogram of the acceptor substrate luteolin. B, LC-MS chromatogram of the glucuronidated products of luteolin resulting from incubating UGT1A9 with luteolin and UDP-GlcA. There are three possible

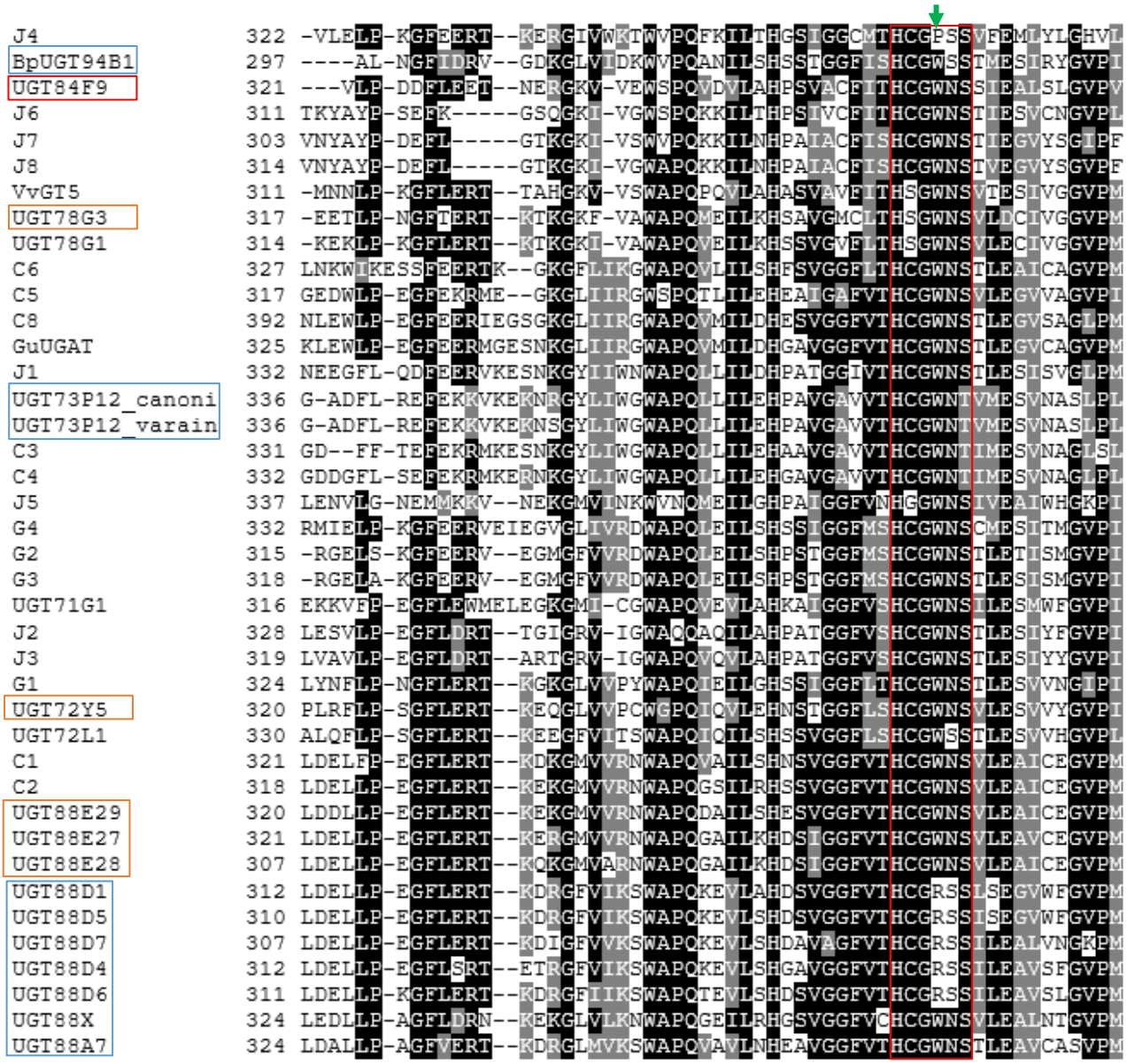
reaction product peaks resulting from glucuronidation at three of the five free -OH groups, including Lut-7-*O*-GlcA which co-eluted with the standard of this compound. C, LC-MS chromatogram of commercial standard of Lut-7-*O*-GlcA. D, Mass spectra of the glucuronidated products of luteolin generated by UGT1A9, showing the expected molecular mass of the parent ions (m/z = 461 in negative mode) for the three mono-glucuronidated products eluting at 13.1 min (Lut-7-*O*-GlcA), 13.7 min and 14 min.



Supplemental Figure S4. Outline of methods and criteria for selection of candidate *UG(A)T* genes from the genome of the model legume *M. truncatula*.

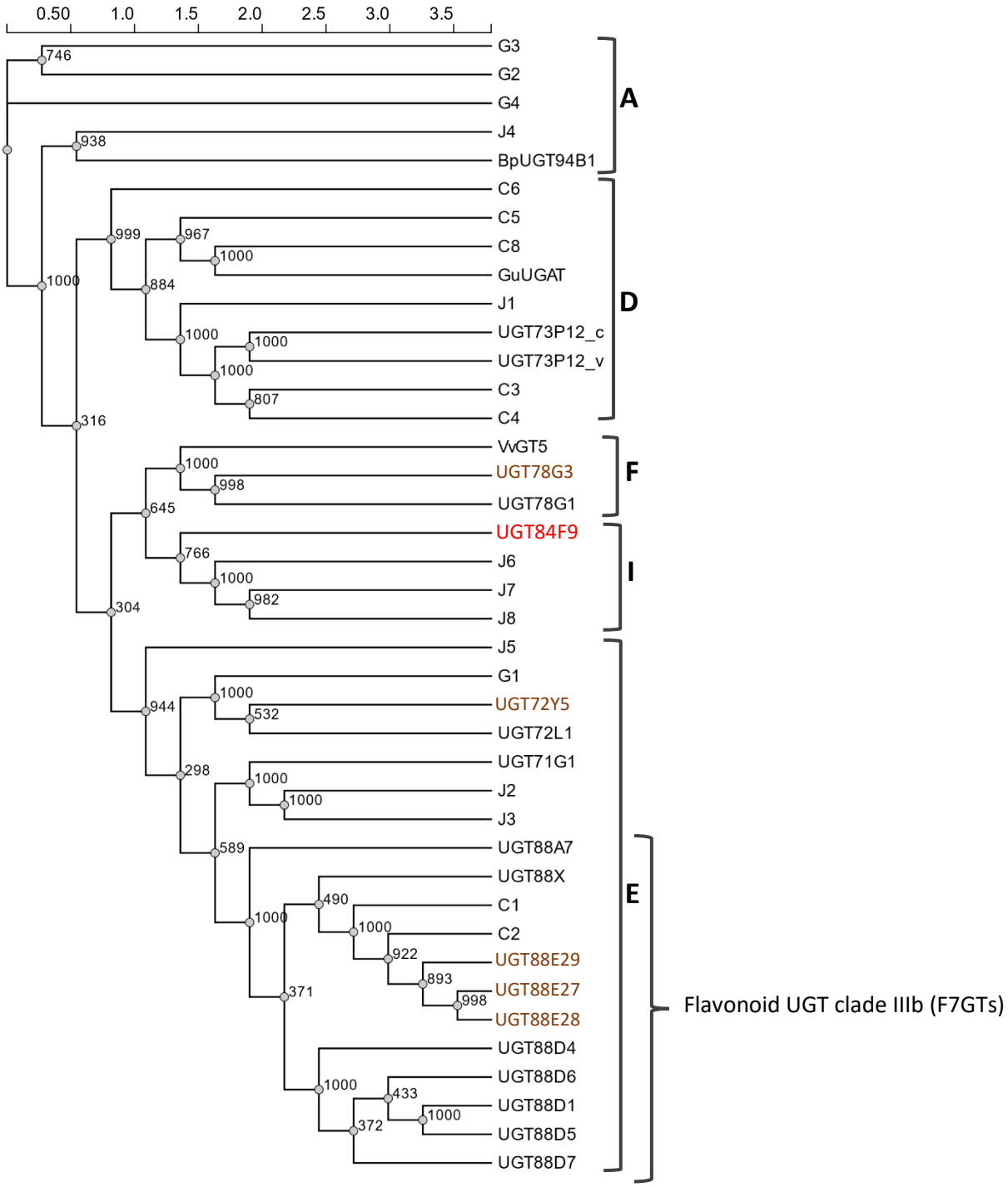


Supplemental Figure S5. Amino acid sequence alignment around the largely conserved aspartic acid (green arrow, Asp121 in UGT71G1) residue in plant UGTs believed to be involved in stabilizing the catalytic histidine (His 22 in UGT72G1) involved in deprotonation of the acceptor substrate during catalysis. This residue was found to be substituted in some of the candidate UGTs we identified. The origins and accession numbers of the sequences are given in the legend to Figure 1 and Table S2. Blue boxes in the left hand column indicated previously identified UDP-glucuronosyltransferases. Orange boxes and binary names (e.g. C3, J5) indicate *M. truncatula* UG(A)T candidates in this study.



Supplemental Figure S6. Multiple sequence alignment of the conserved PSPG motif located in the C-terminal domain of plant UGTs. The F7GATs (UGT88D-1, 5, 4, 6 & 7) from *Lamiales* species have an amino acid substitution (green arrow) in this region that is believed to contribute to sugar donor recognition. Arg351 in the UDP-GlcA dependent UGT88D7 replaces the tryptophan (e.g. Trp367 in UGT88A7) found to be highly conserved in UDP-Glu dependent UGTs. The origins and accession numbers of the sequences are given in the legend to Figure 1 and Table S2. Blue boxes in the left hand column indicated previously identified UDP-

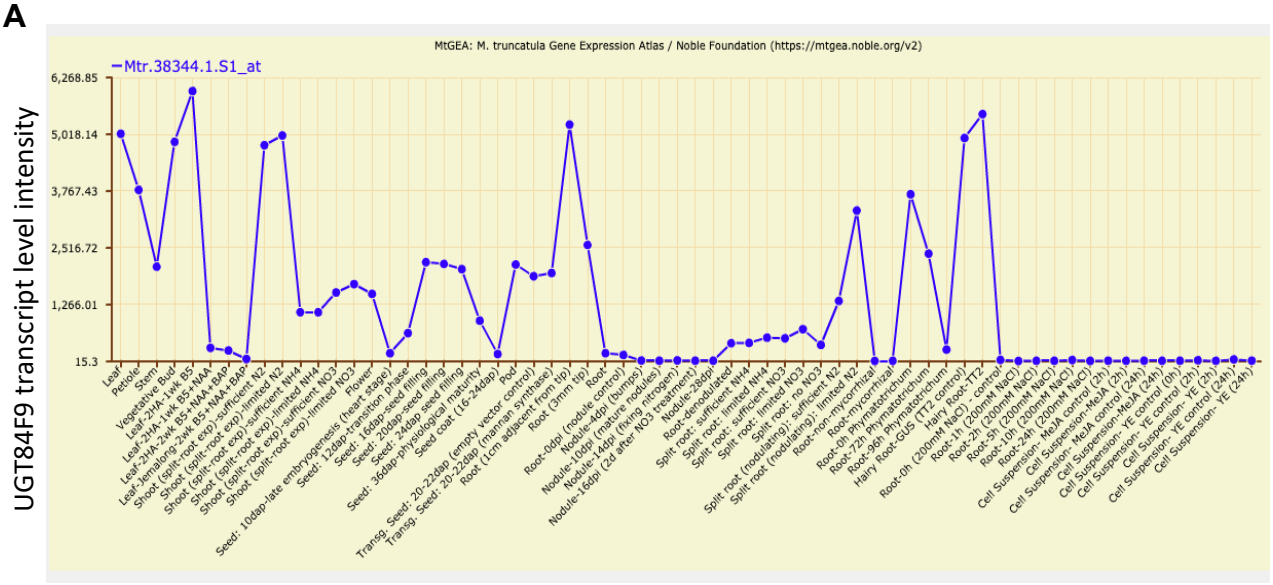
glucuronosyltransferases. Orange boxes and binary names (e.g. C3, J5) indicate *M. truncatula* UG(A)T candidates in this study.



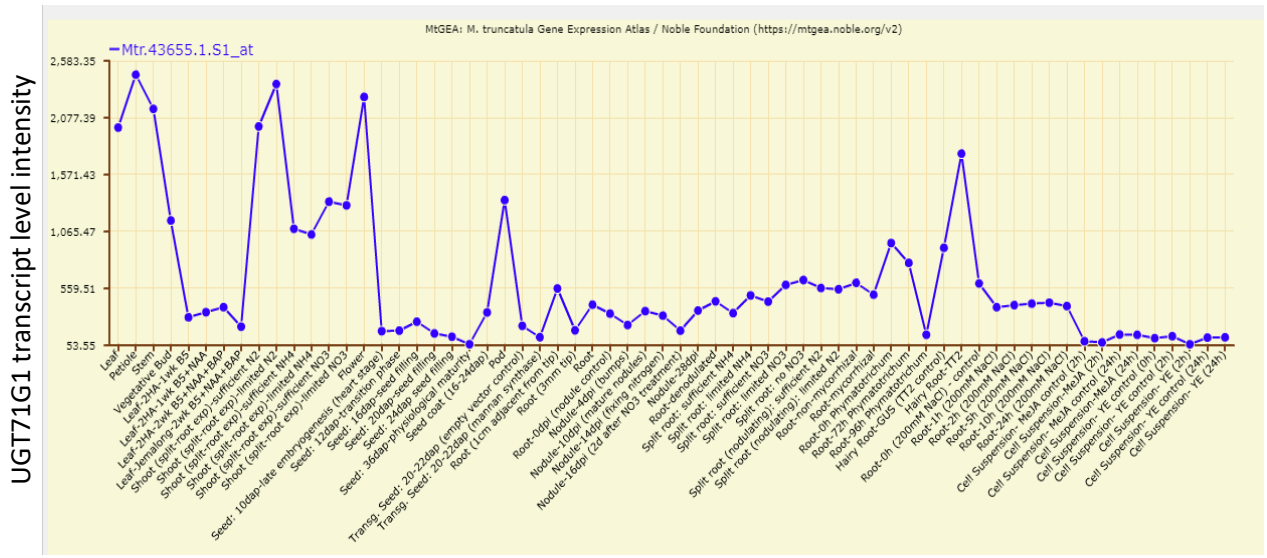
Supplemental Figure S7. Phylogenetic analysis of selected UG(A)T candidates. Candidate protein sequences from *M. truncatula* were aligned with previously characterized plant UG(A)Ts using multiple sequence comparison by Log-Expectation (MUSCLE). The resulting alignment was used to construct the maximum-Likelihood phylogeny tree by performing 1000

bootstrap replicates using PhyML 3.0 algorithm. The number for each node indicates the bootstrap support values for the node where 1000 represents maximum support for the node and the scale on top of the tree indicates the branch length. Length on the scale indicates evolutionary distance in substitutions per amino acid. The UGT clades are labelled on the tree. UGT84F9 (conformed UDP-glucuronosyltransferase) is marked in red; those Medicago UGATs marked in brown passed all selection criteria but were shown to utilized UDP-Glc rather than UDP-GlcA. The origins and accession numbers of the sequences are given in the legend to Figure 1 and Table S2.

A

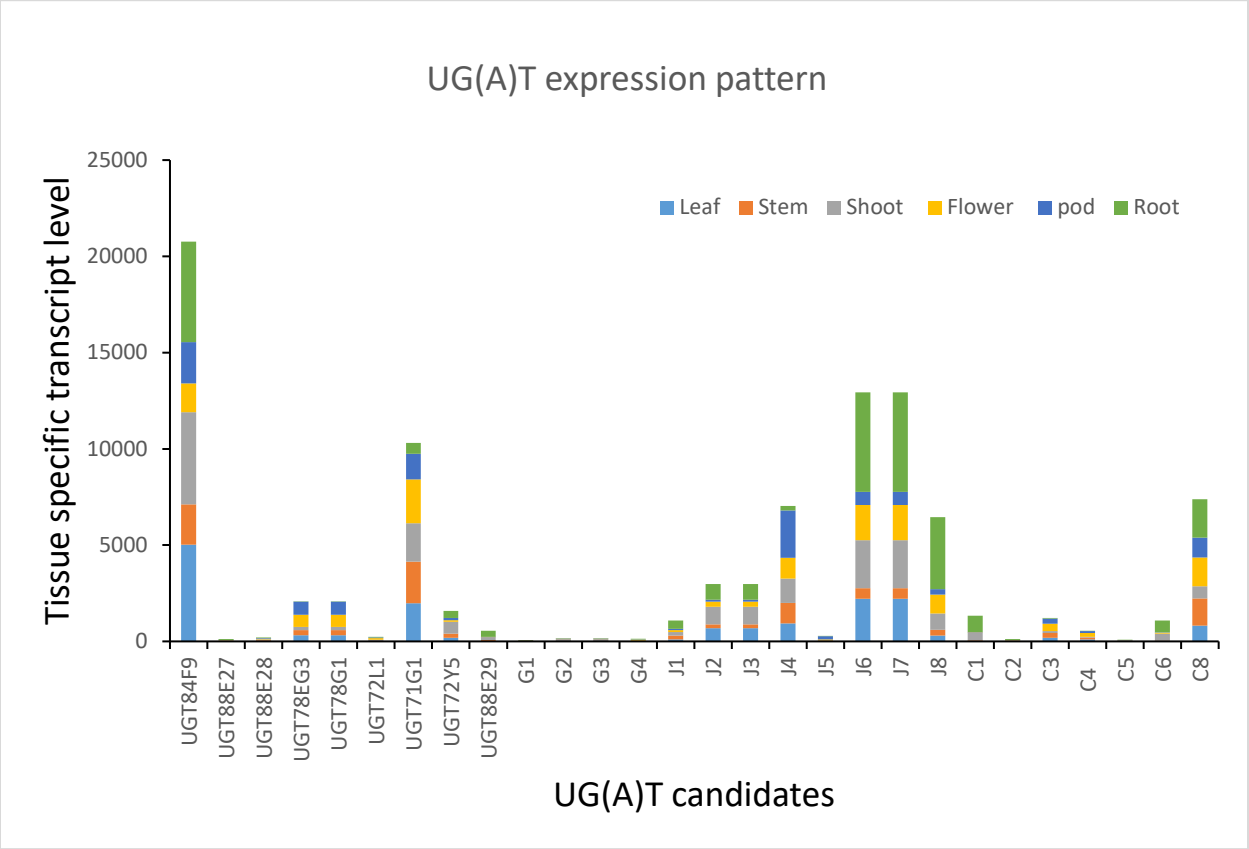


B

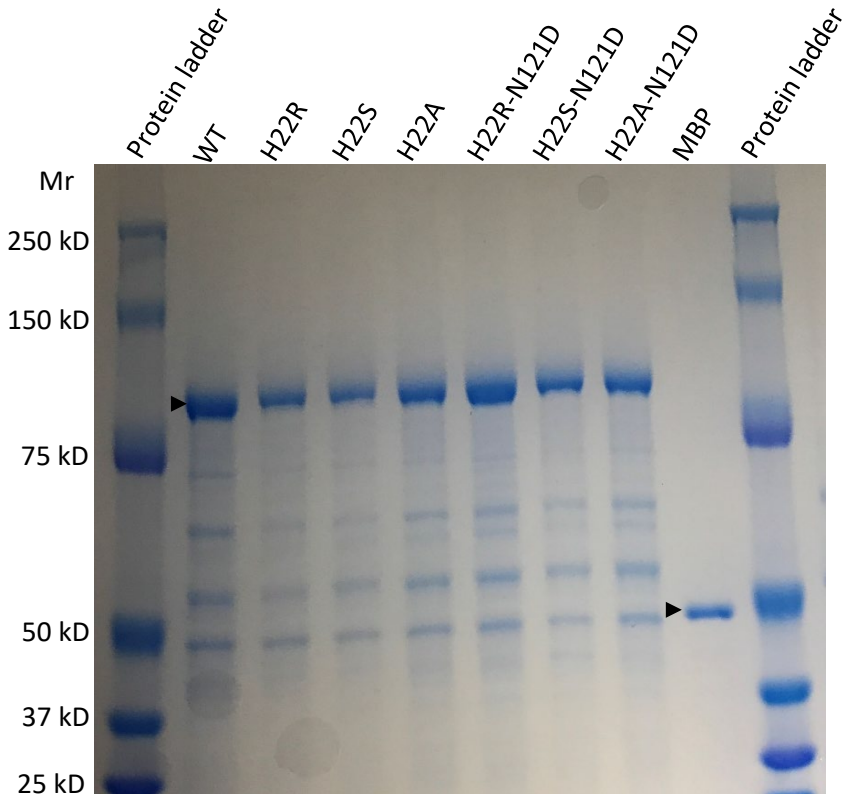


M. truncatula tissue or treatment conditions

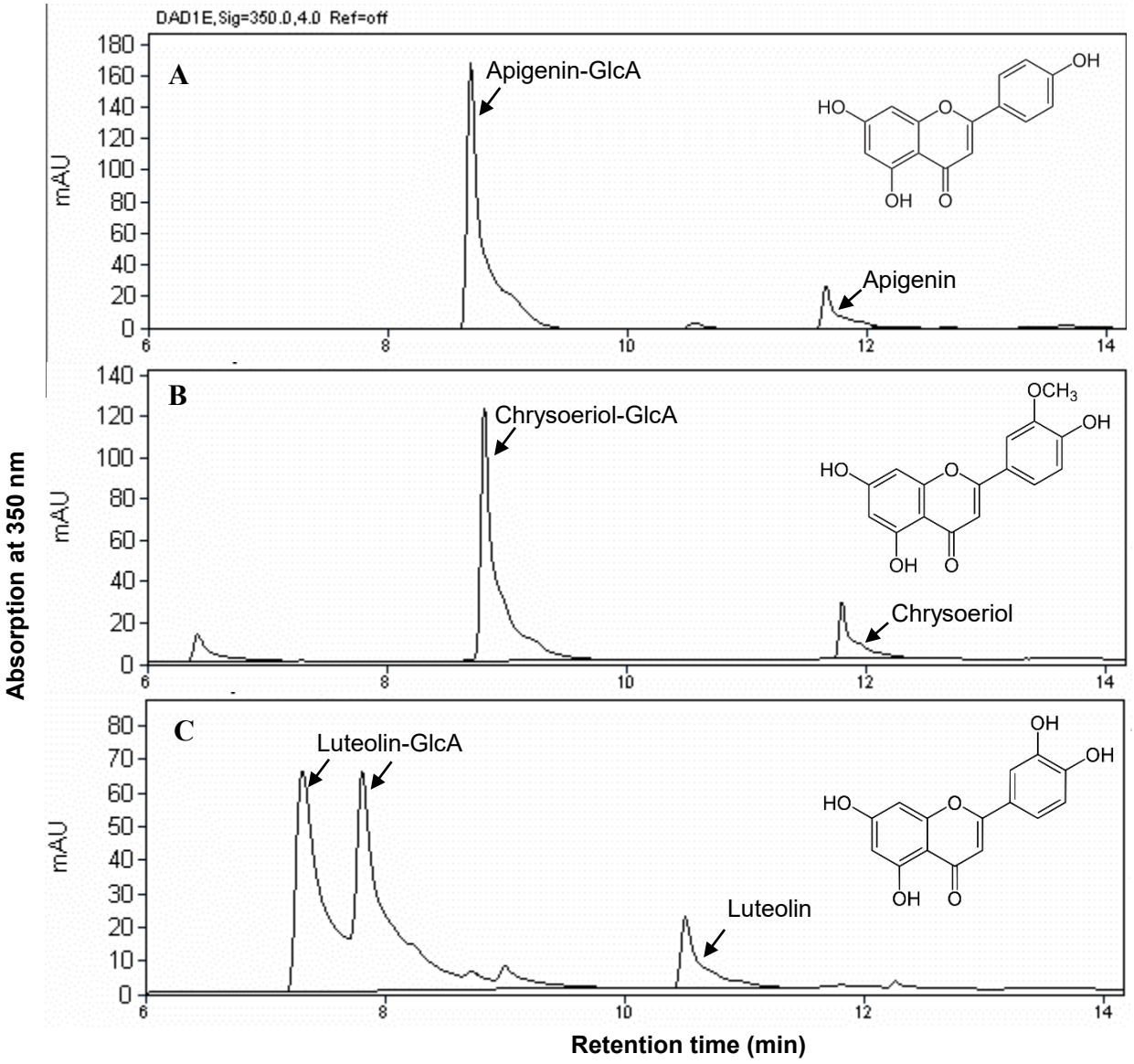
C



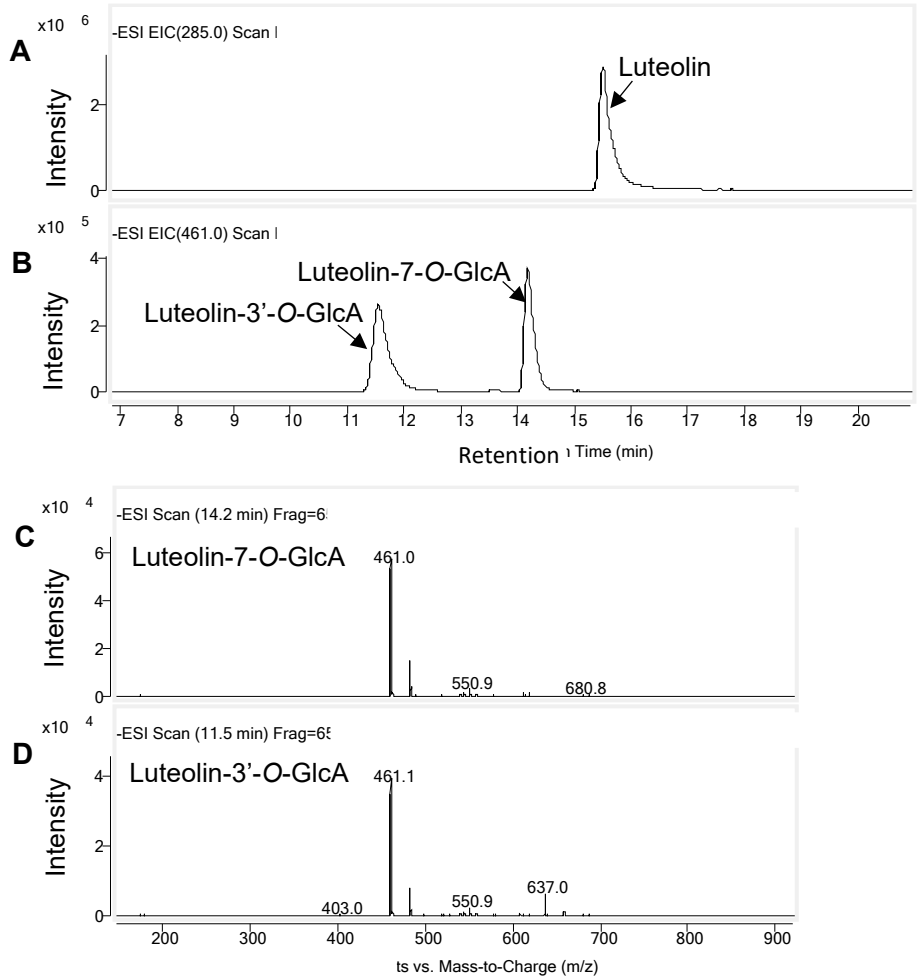
Supplemental Figure S8. Expression of UGAT candidates and previously characterized UGTs in *M. truncatula*. A, B, The transcript expression pattern of A, UGT84F9, and B, UGT71G1, in different tissues and under different conditions, retrieved from the *M. truncatula* Gene Expression Atlas (https://mtgea.noble.org/v3/). C, Tissue-specific expression of all UGAT candidates and previously characterized UGTs as retrieved from the Gene Expression Atlas.



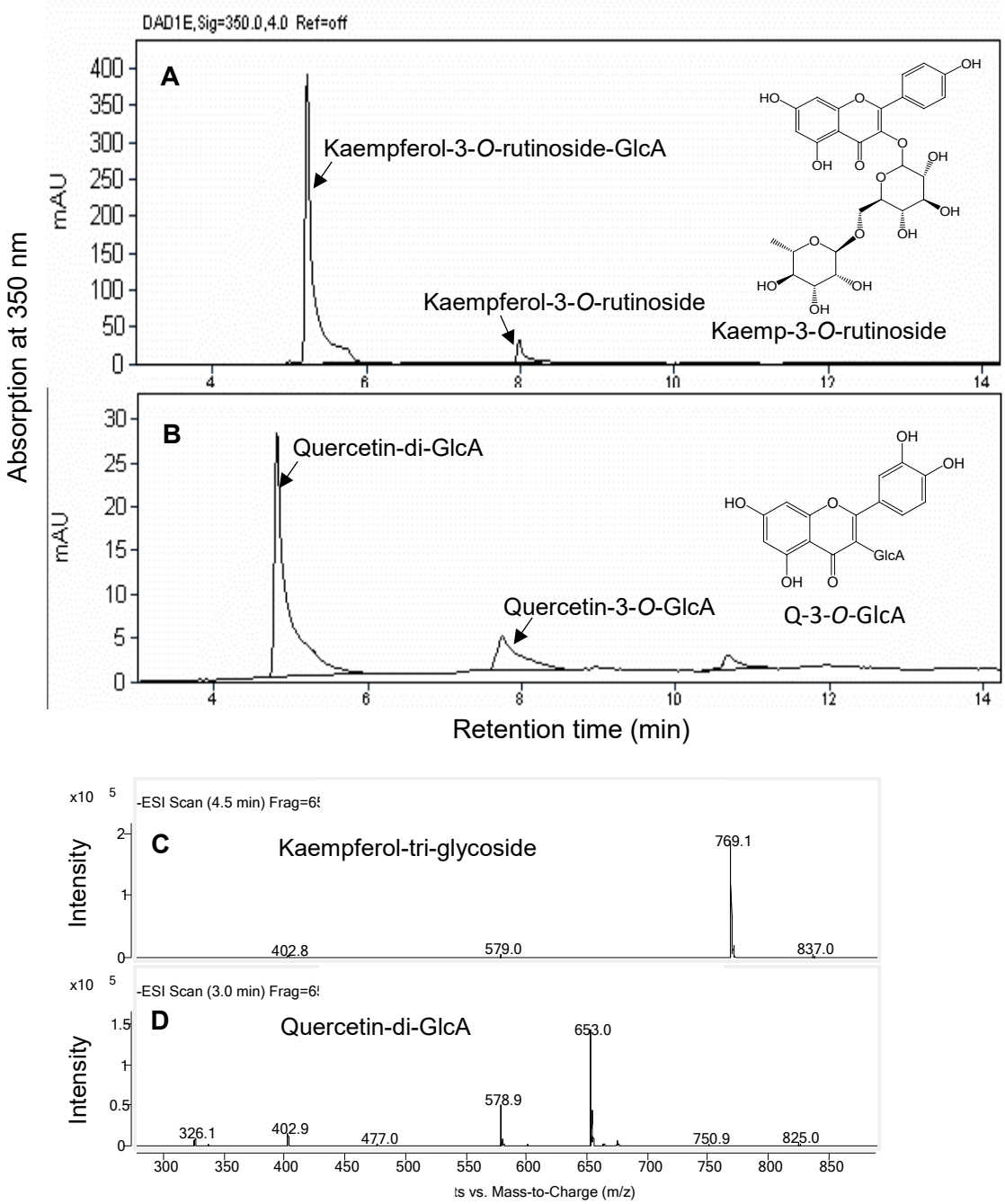
Supplemental Figure S9. SDS-PAGE of purified recombinant UGT84F9 and its site-directed mutants. UGT84F9 and mutant UGT84F9s are MBP-tagged with a total molecular weight of 95.44kD.



Supplemental Figure S10. HPLC-UV analysis of conjugated flavones generated by recombinant UGT84F9 with UDP-GlcA as sugar donor. A, Apigenin as acceptor. B, Chrysoeriol as acceptor. C, Luteolin as acceptor. Reactions with apigenin and chrysoeriol yielded one major glucuronidated product while the reaction with luteolin generated two glucuronidated products. Incubation was for 2 h.

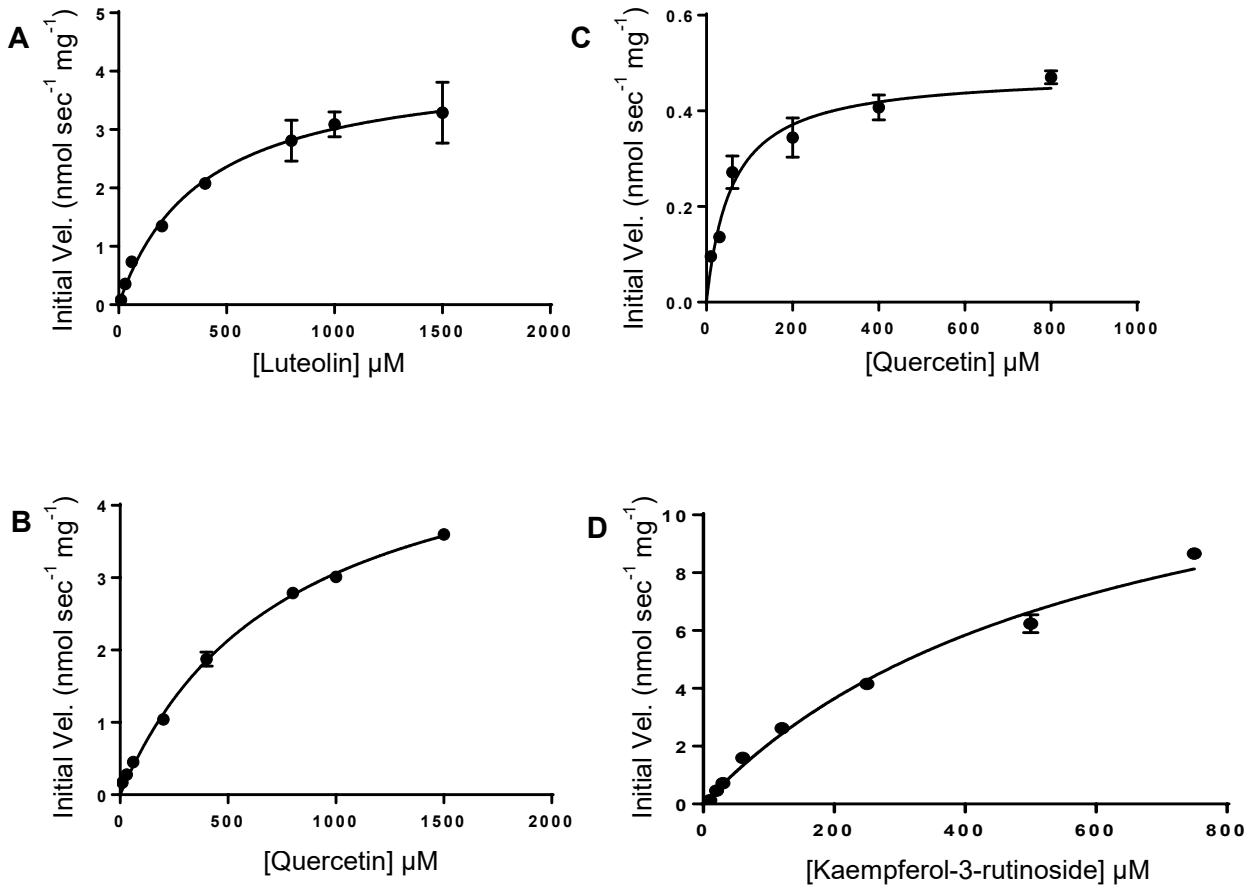


Supplemental Figure S11. LC-MS analysis of glucuronidated products of luteolin generated by recombinant UGT84F9 with UDP-GlcA as sugar donor. A, LC-MS peak for luteolin. B, LC-MS peak for luteolin-glucuronides eluting at 11.6 min and 14.2 min. C, D, Mass spectra showing the molecular ions of the two glucuronide products of luteolin (m/z = 461). The LC-MS was performed in negative ion mode.

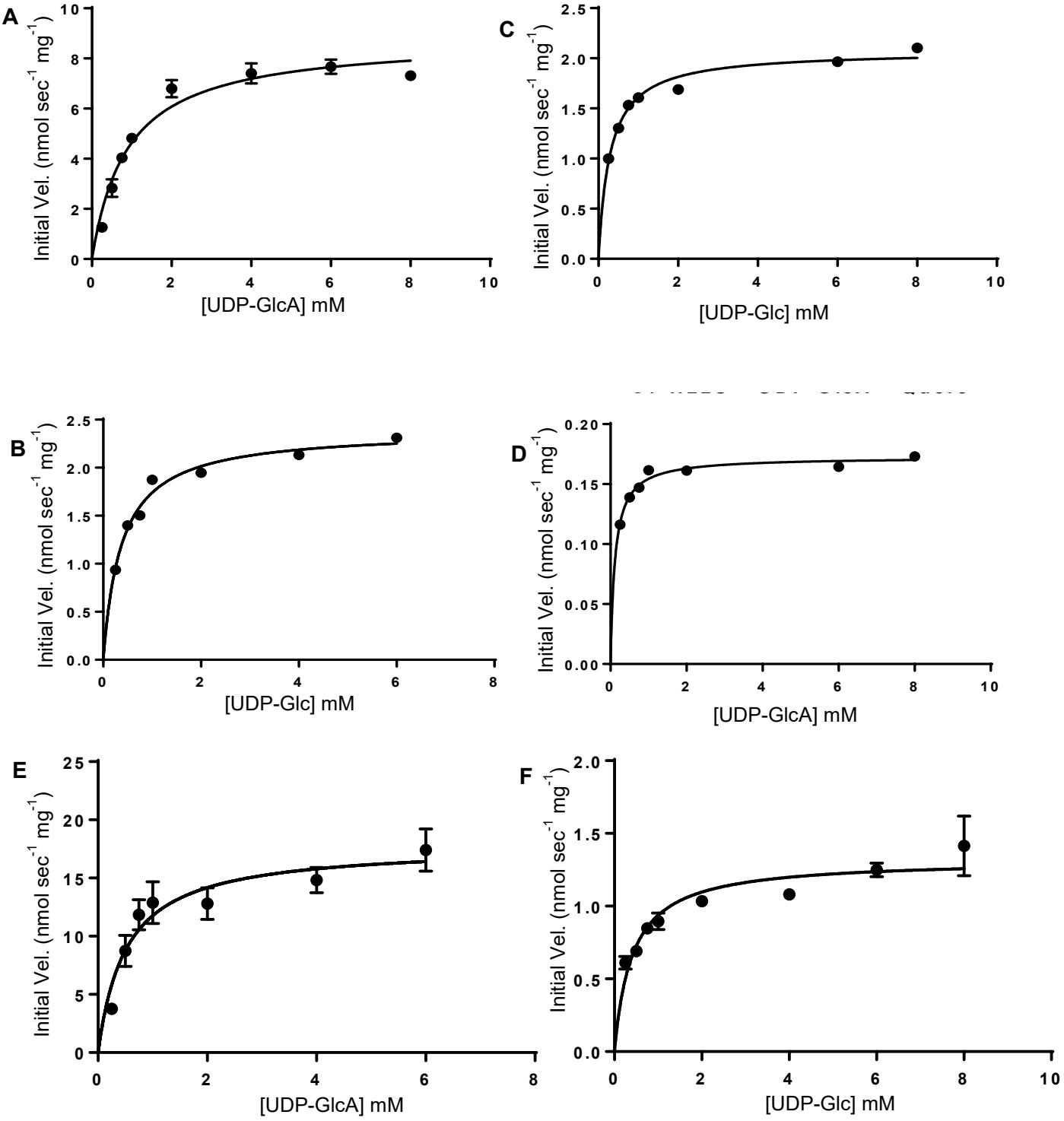


Supplemental Figure S12. HPLC-UV and LC-MS analysis of glucuronidated products of UGT84F9 activity with the di-glycoside kaempferol-3-*O*-rutinoside and the mono-glucuronide Q-3-*O*-GlcA as sugar acceptors and UDP-GlcA as sugar donor. Incubation was for 2 h. A, HPLC-UV profile of reaction products from kaempferol-3-*O*-rutinoside. B, HPLC-UV profile of reaction products from

Q-3-*O*-GlcA. C, The mass spectrum of the reaction product from kaempferol-3-*O*-rutinoside showing the molecular ion (m/z = 769.1) D, The mass spectrum of the reaction product from Q-3-*O*-GlcA showing the molecular ion (m/z = 653). The LC-MS was performed in negative mode

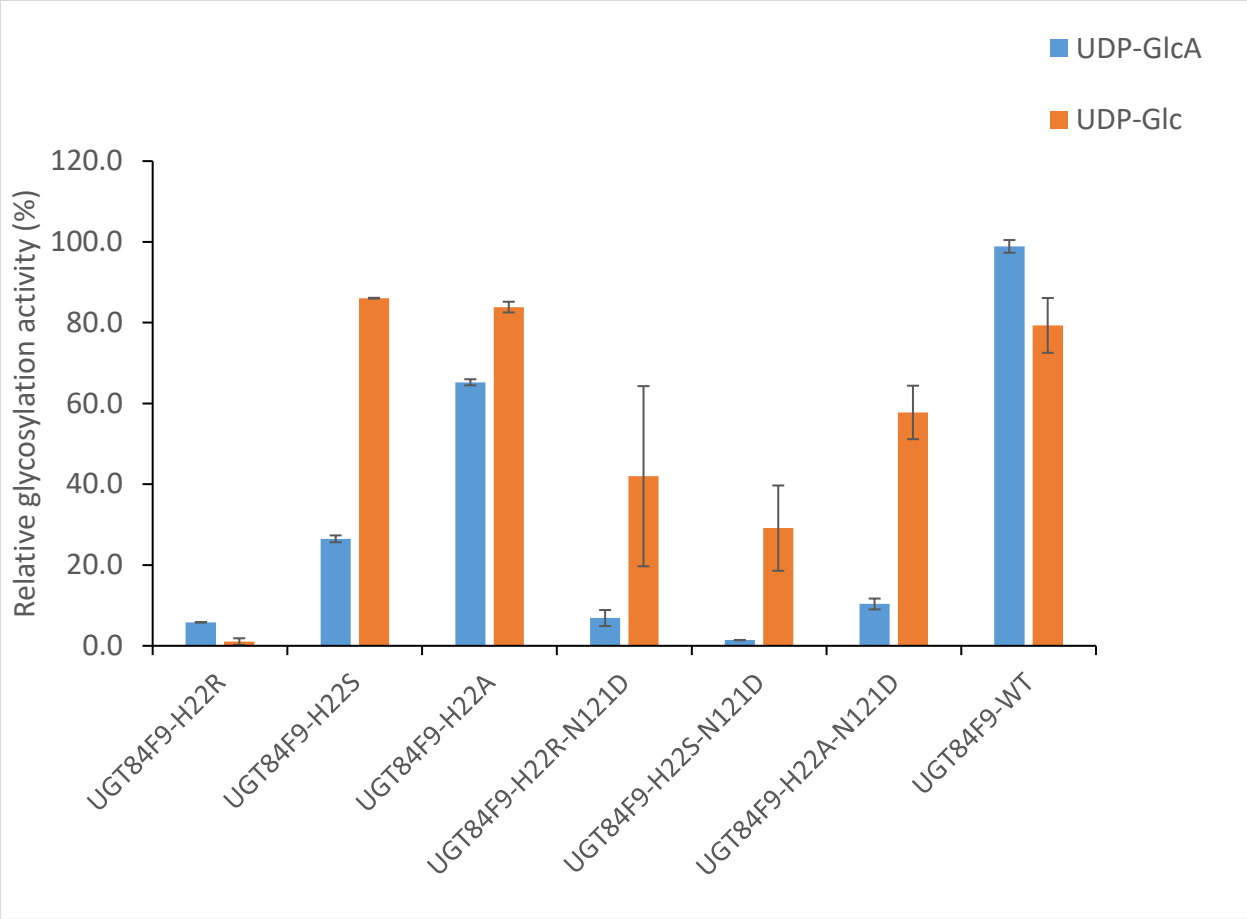


Supplemental Figure S13. Michaelis-Menten curves for activity of UGT84F9 towards acceptor substrates. Products were detected and measured by HPLC, and the apparent V_{max} and K_m values (Table S2) determined by fitting the initial velocity to the Michaelis-Menten equation by non-linear regression using GraphPad Prism 7.04. A, Luteolin using UDP-GlcA as sugar donor. B, Quercetin using UDP-GlcA as sugar donor. C, Quercetin using UDP-Glc as sugar donor. D, Kaempferol-3-rutinoside using UDP-GlcA as sugar donor. The analyses were performed with three technical replicates and error bars represent standard deviation.

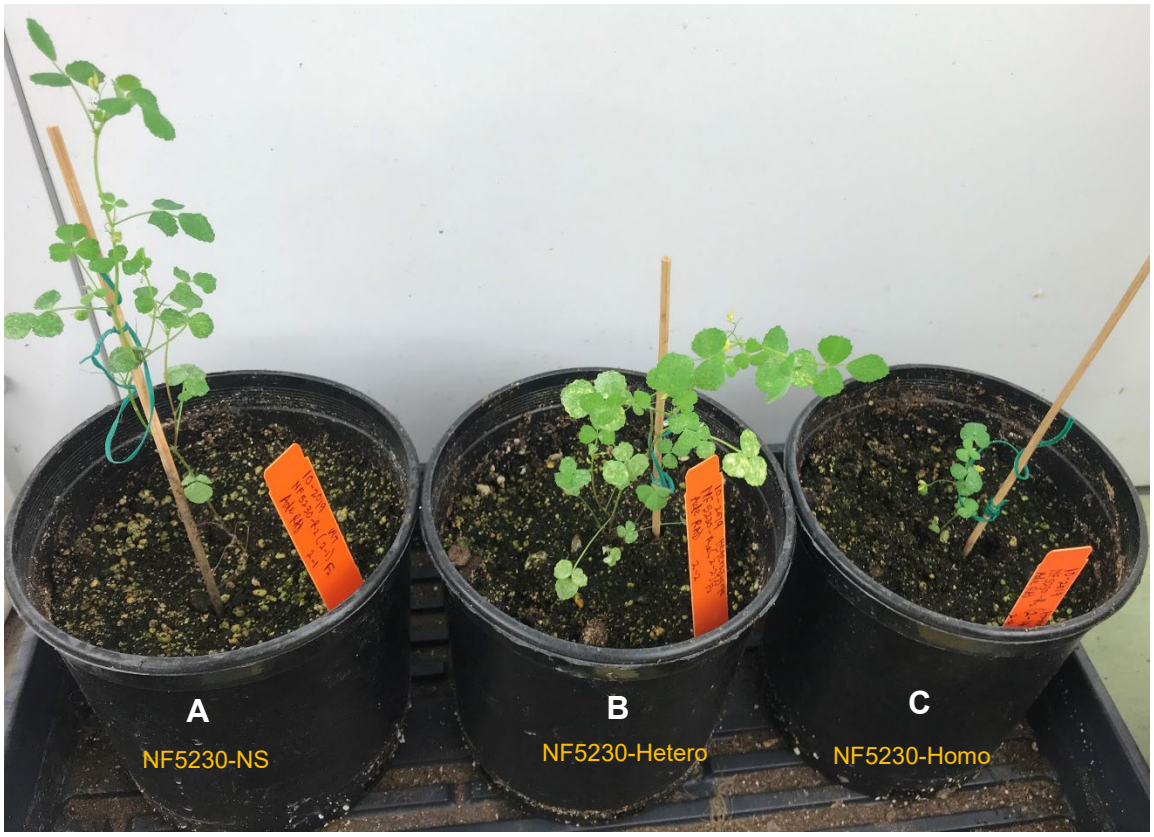


Supplemental Figure S14. Michaelis-Menten curves for wild-type and mutant UGT84F9 towards sugar donor substrates with quercetin and apigenin as acceptors. A, Kinetics of UGT84F9 towards UDP-GlcA using quercetin as acceptor substrate. B, Kinetics of UGT84F9 towards UDP-

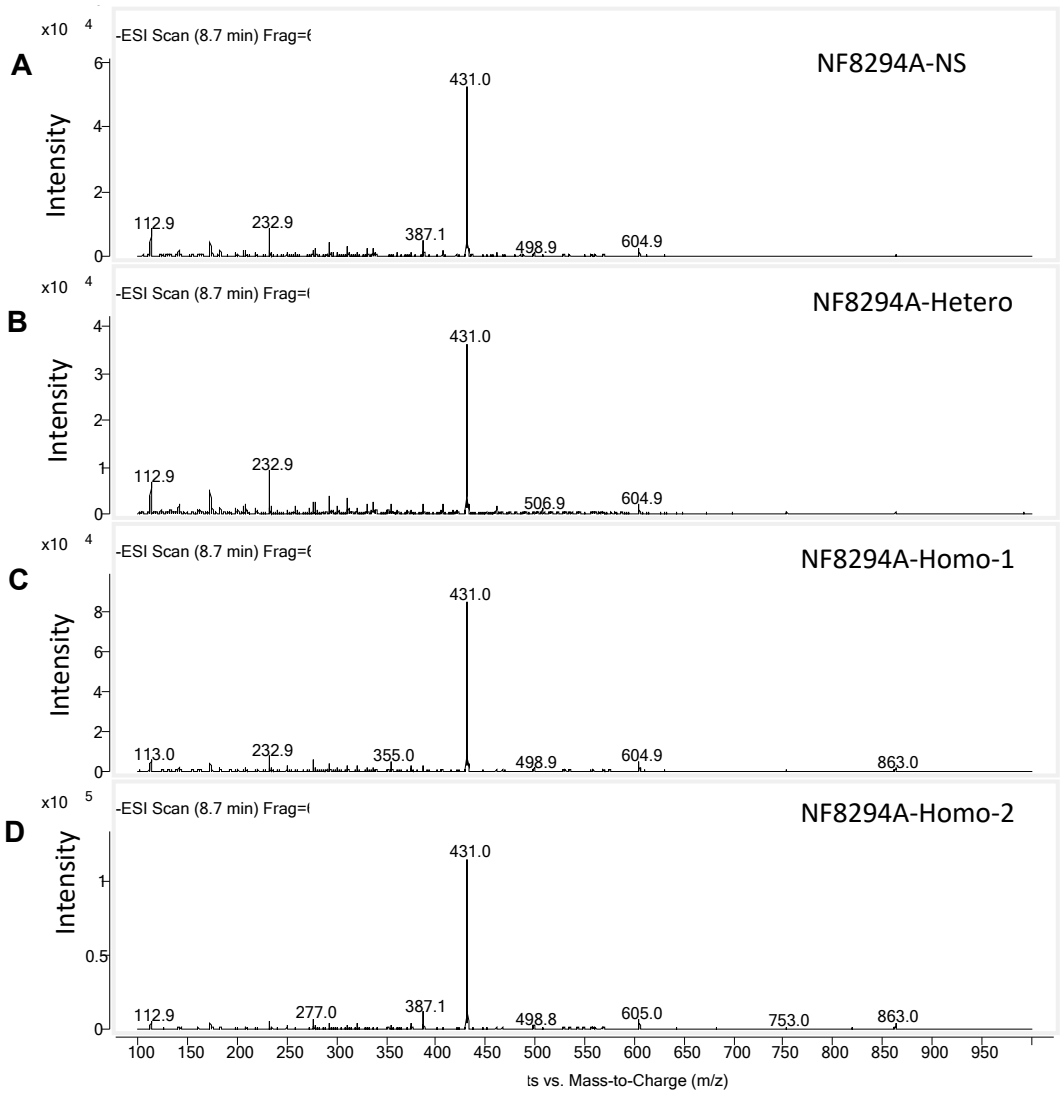
Glc using quercetin as acceptor substrate. C, Kinetics of UGT84F9-H22S towards UDP-Glc using quercetin as acceptor substrate. D, Kinetics of UGT84F9-H22S towards UDP-GlcA using quercetin as acceptor substrate E, Kinetics of UGT84F9 towards UDP-GlcA using apigenin as acceptor substrate. F, Kinetics of UGT84F9 towards UDP-Glc using apigenin as acceptor substrate. The analyses were performed with three technical replicates and error bars represent standard deviation.



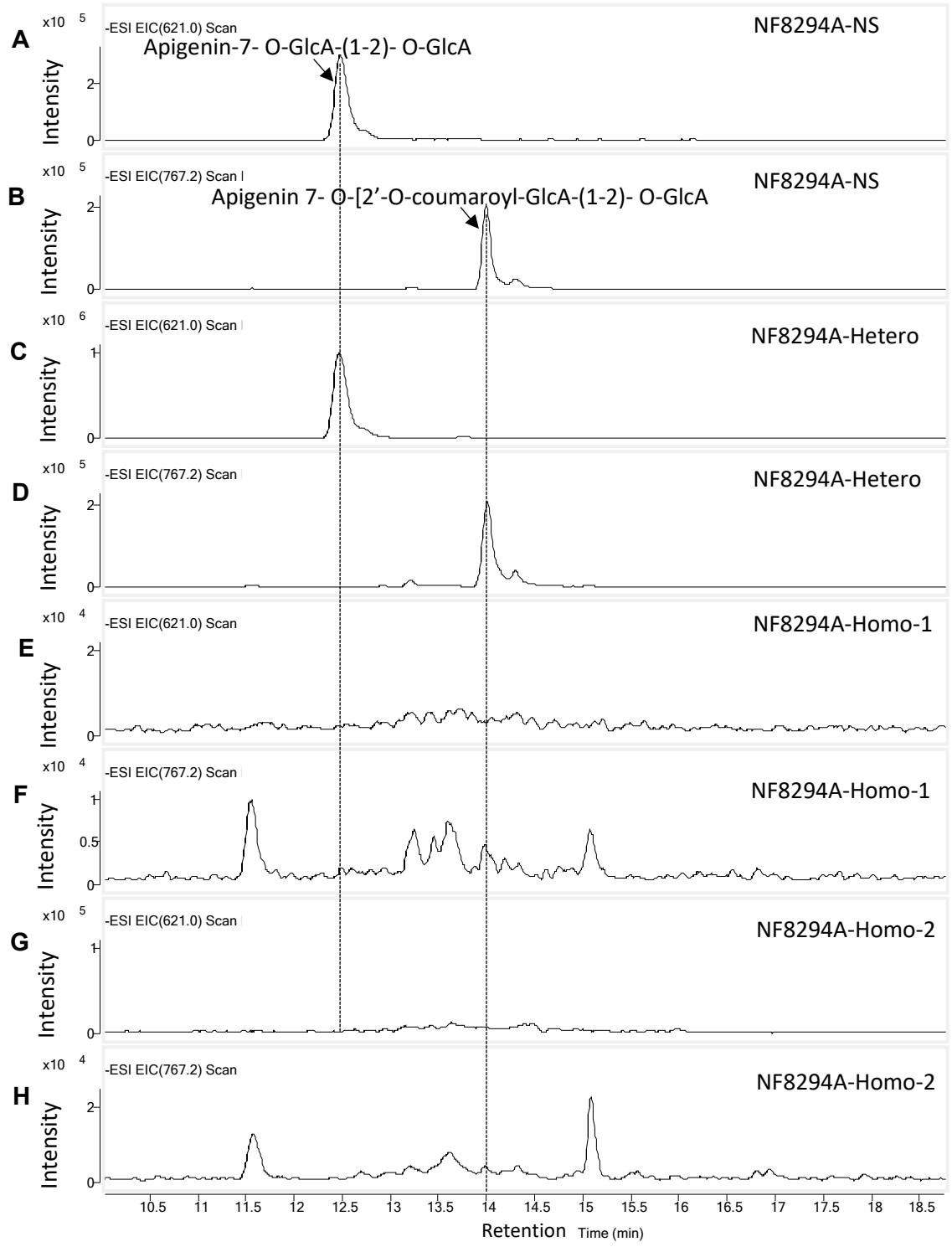
Supplemental Figure S15. Relative activity of UGT84F9 mutant and wild-type proteins towards UDP-GlcA and UDP-Glc. Quercetin was used as the acceptor substrate in all assays. The activities are expressed with respect to the most active enzyme; UGT84F9-H22S for UDP-Glc with specific activity = 33.038 µmol/min/mg of protein and wild-type UGT84F9 for UDP-GlcA with specific activity = 41.44 µmol/min/mg of protein. Relatively the same amount of mutant or wildtype protein was used following quantification by SDS PAGE resolution and densitometry as described in Materials and Methods. The enzyme assays were performed in duplicate.



Supplemental Figure S16. Five weeks-old plants of NF5230 mutant lines harboring a *Tnt1* insertion in the *UGT84F9* gene. The homozygous *UGT84F9 Tnt1* insertion line (C) shows smaller size compared to the corresponding heterozygote at the *UGT84F9* locus (B) and a corresponding NF5230 line that is null segregant wild type at the *UGT84F9* locus (A). PCR genotyping is shown in **Fig. 7B**.



Supplemental Figure S17. MS spectra of apigenin glucosides detected in extracts from *M. truncatula* leaf tissue of NF8294A null segregant, and heterozygous and homozygous knockout lines. All the lines analyzed accumulate apigenin glucoside at retention time 8.7 min. The MS spectrum for apigenin glucoside shows the molecular mass of the parent ion at $[M-H]^+ = 431$. The LC-MS was performed in negative mode.



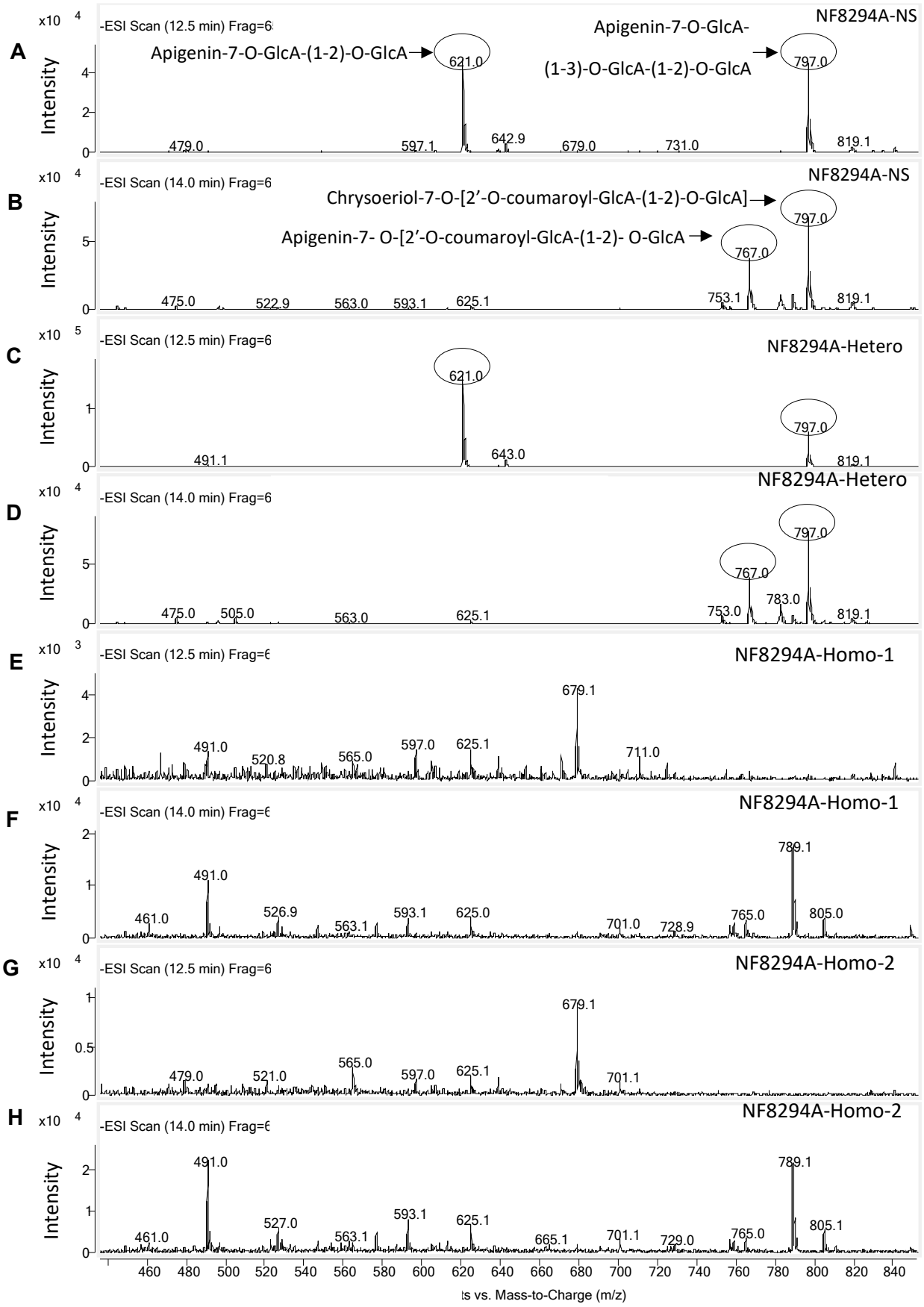
Supplemental Figure S18. LC-MS analysis of extracts from *M. truncatula* leaf tissue for detection of apigenin-7-*O*-GlcA-(1-2)-*O*-GlcA and apigenin-7-*O*-[2´-*O*-coumaroyl-GlcA-(1-2)-*O*-GlcA by scanning for their molecular ions of $[M-H]^+$ = 621 and $[M-H]^+$ = 767 respectively.

A,C, LC-MS chromatograms showing apigenin-7-*O*-GlcA-(1-2)-*O*-GlcA in the NF8294A- null segregant (NS), and NF8294A- heterozygote, respectively.

B, D, LC-MS chromatograms showing apigenin-7-*O*-[2´-*O*-coumaroyl-GlcA-(1-2)-*O*-GlcA in the NF8294A- null segregant (NS) and NF8294A- heterozygote, respectively.

Neither apigenin-7-*O*-GlcA-(1-2)-*O*-GlcA nor apigenin-7-*O*-[2´-*O*-coumaroyl-GlcA-(1-2)-*O*--GlcA accumulate in extracts from NF8294A- homozygote (E, F, G, H).

The LC-MS was performed in negative mode.


A
NF8294A-NS
-ESI Scan (12.5 min) Frag=6
Apigenin-7-O-GlcA-(1-2)-O-GlcA
621.0
Apigenin-7-O-GlcA-
(1-3)-O-GlcA-(1-2)-O-GlcA
797.0
479.0
597.1
642.9
679.0
731.0
819.1
B
NF8294A-NS
-ESI Scan (14.0 min) Frag=6
Chrysoeriol-7-O-[2'-O-coumaroyl-GlcA-(1-2)-O-GlcA]
797.0
Apigenin-7- O-[2'-O-coumaroyl-GlcA-(1-2)- O-GlcA
767.0
475.0
522.9
563.0
593.1
625.1
753.1
819.1
C
NF8294A-Hetero
-ESI Scan (12.5 min) Frag=6
621.0
797.0
491.1
643.0
819.1
D
NF8294A-Hetero
-ESI Scan (14.0 min) Frag=6
797.0
767.0
783.0
475.0
505.0
563.0
625.1
753.0
819.1
E
NF8294A-Homo-1
-ESI Scan (12.5 min) Frag=6
679.1
491.0
520.8
565.0
597.0
625.1
711.0
F
NF8294A-Homo-1
-ESI Scan (14.0 min) Frag=6
789.1
491.0
461.0
526.9
563.1
593.1
625.0
701.0
728.9
765.0
805.0
G
NF8294A-Homo-2
-ESI Scan (12.5 min) Frag=6
679.1
479.0
521.0
565.0
597.0
625.1
701.1
H
NF8294A-Homo-2
-ESI Scan (14.0 min) Frag=6
491.0
789.1
461.0
527.0
563.1
593.1
625.1
665.1
701.1
729.0
765.0
805.1
Intensity
460 480 500 520 540 560 580 600 620 640 660 680 700 720 740 760 780 800 820 840
s vs. Mass-to-Charge (m/z)

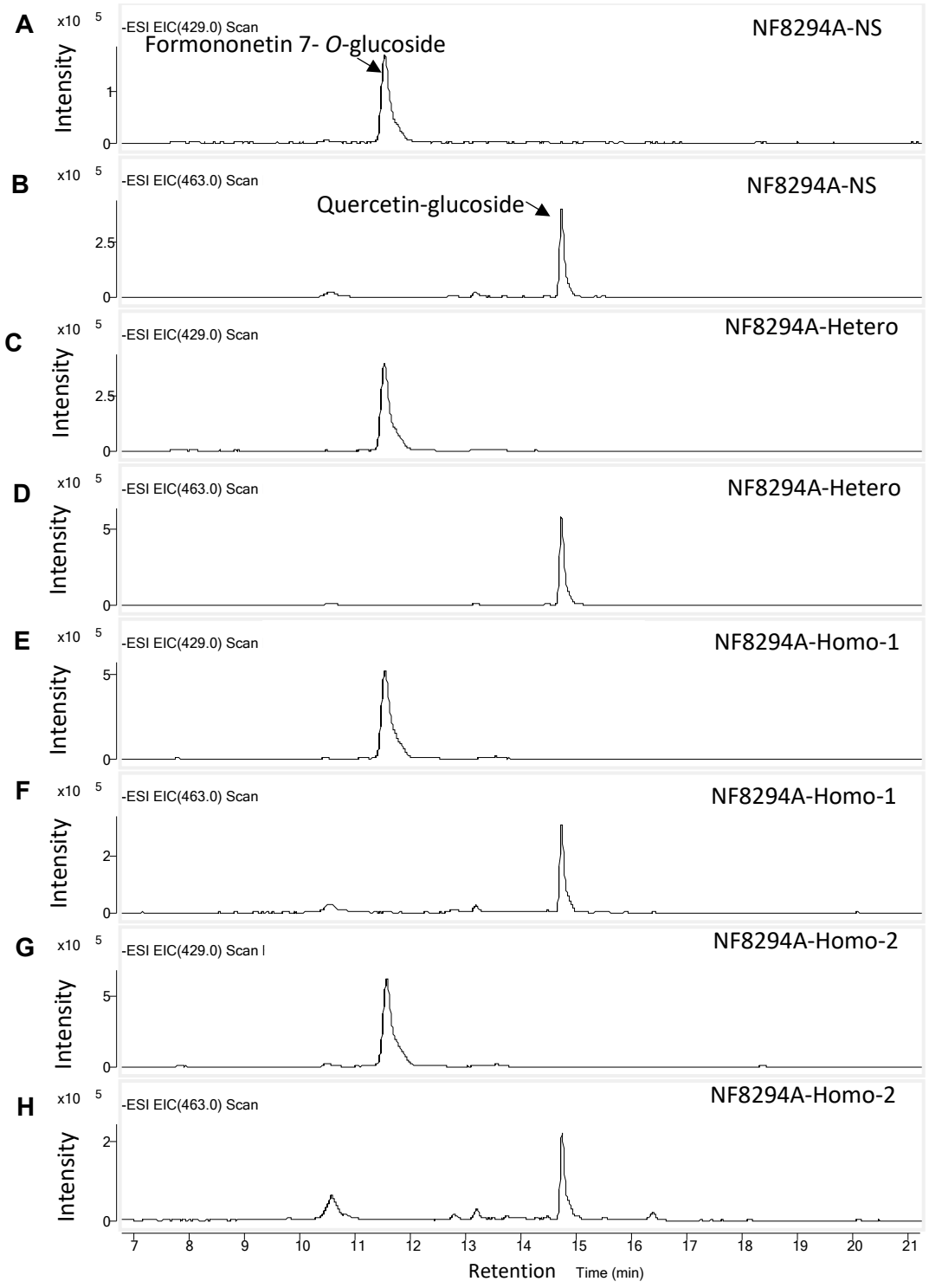
Supplemental Figure S19. Mass spectra of extracts from *M. truncatula* leaf tissue for detection of apigenin-7-*O*-GlcA-(1-2)-*O*-GlcA (m/z = 621), apigenin-7-*O*-GlcA-(1-3)-*O*-GlcA-(1-2)-*O*-GlcA (m/z = 797 at retention time 12.5 min), chrysoeriol-7-O-[2'-*O*-coumaroyl-GlcA-(1-2)-*O*-GlcA] (m/z = 797 at retention time 14 min), and apigenin-7-*O*-[2'-O-coumaroyl-GlcA-(1-2)-*O*-GlcA (m/z = 767).

A, C, The mass spectra for apigenin-7-*O*-GlcA-(1-2)-*O*-GlcA (m/z = 621) and apigenin-7-*O*-GlcA-(1-3)-*O*-GlcA-(1-2)-*O*-GlcA in extracts from NF8294A- null segregant (NS) and NF8294A- heterozygote, respectively.

B, D, The mass spectra for apigenin-7- *O*-[2'-O-coumaroyl-GlcA-(1-2)- *O*-GlcA (m/z = 767) and chrysoeriol-7-*O*-[2'-*O*-coumaroyl-GlcA-(1-2)-*O*-GlcA] (m/z = 797 at retention time 14 min) in NF8294A- null segregant (NS) and NF8294A- heterozygote, respectively.

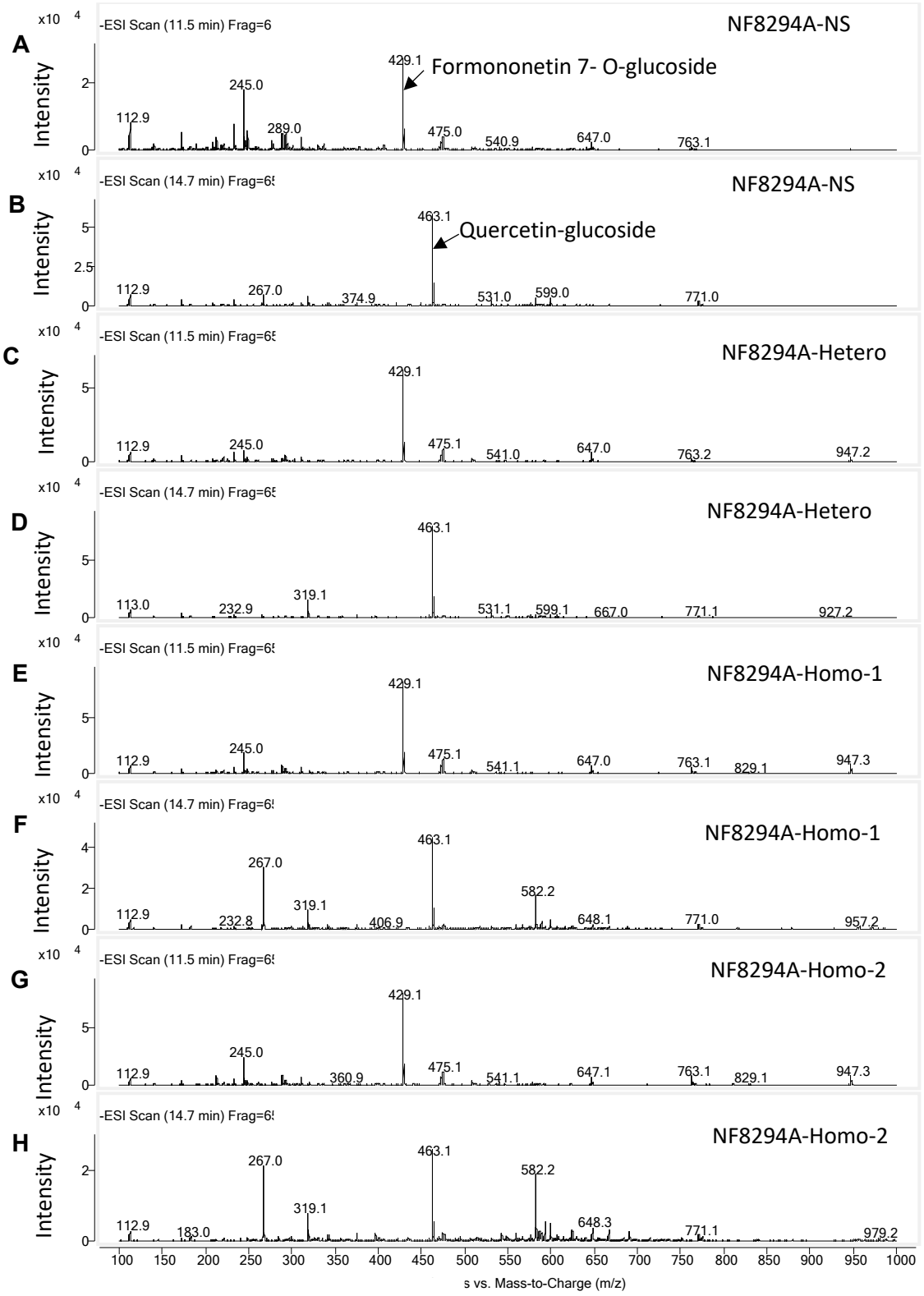
None of these flavonoid glucuronide metabolites was detected in NF8294A- homozygote extract (E, F, G, H).

The LC-MS was performed in negative mode.



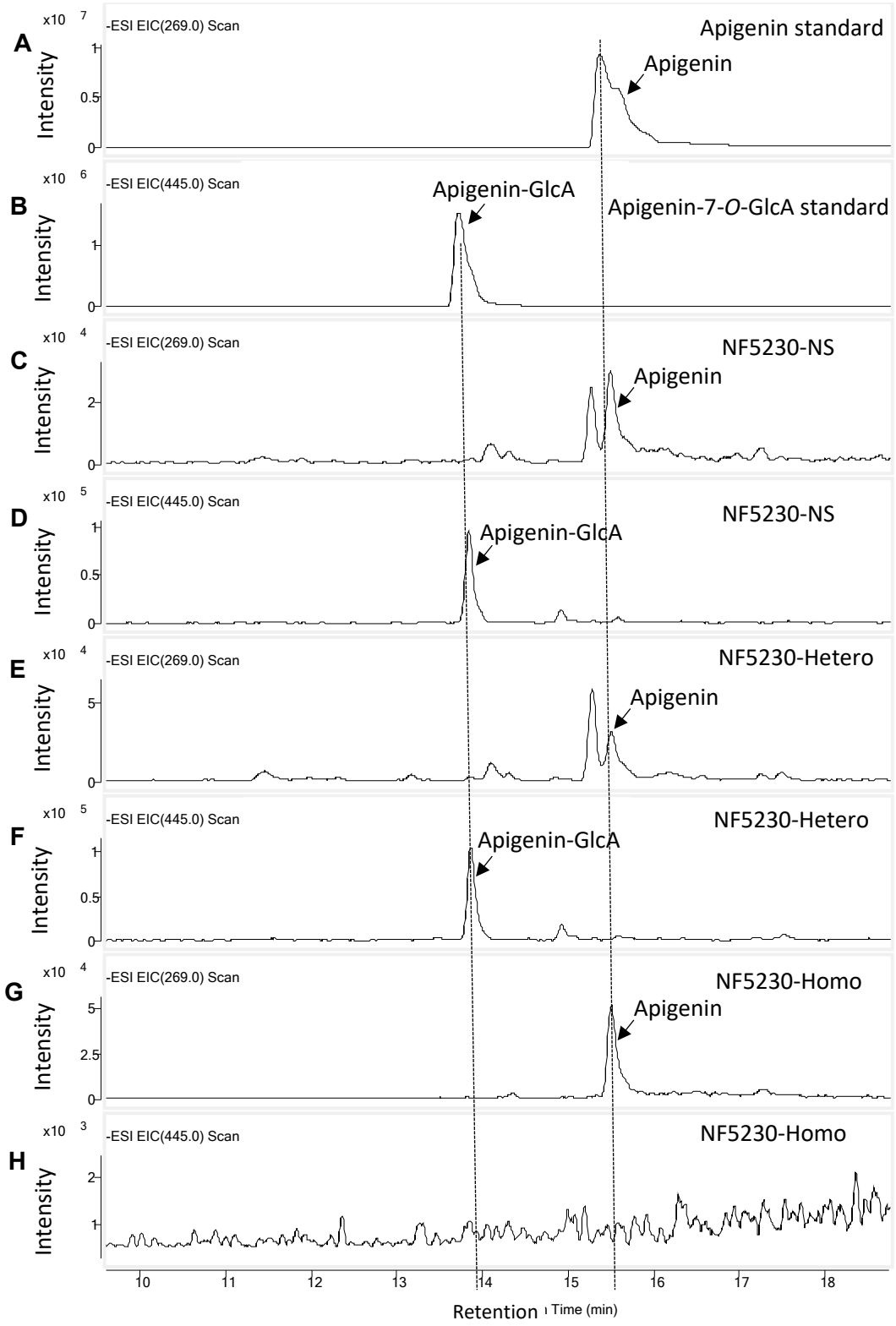
Supplemental Figure S20. LC-MS chromatograms of formononetin-7- *O*-glucoside and quercetin glucoside detected in extracts from *M. truncatula* leaf tissue of NF8294A *UGT84F9* null segregant (NS) (A, B), heterozygote (C, D) and homozygous knockout lines (E, F, G, H). All

the lines analyzed accumulate formononetin-7-*O*-glucoside and quercetin glucoside as detected by scanning for their molecular ions at $[M-H]^+$ = 429.1 and $[M-H]^+$ 463.1 respectively.



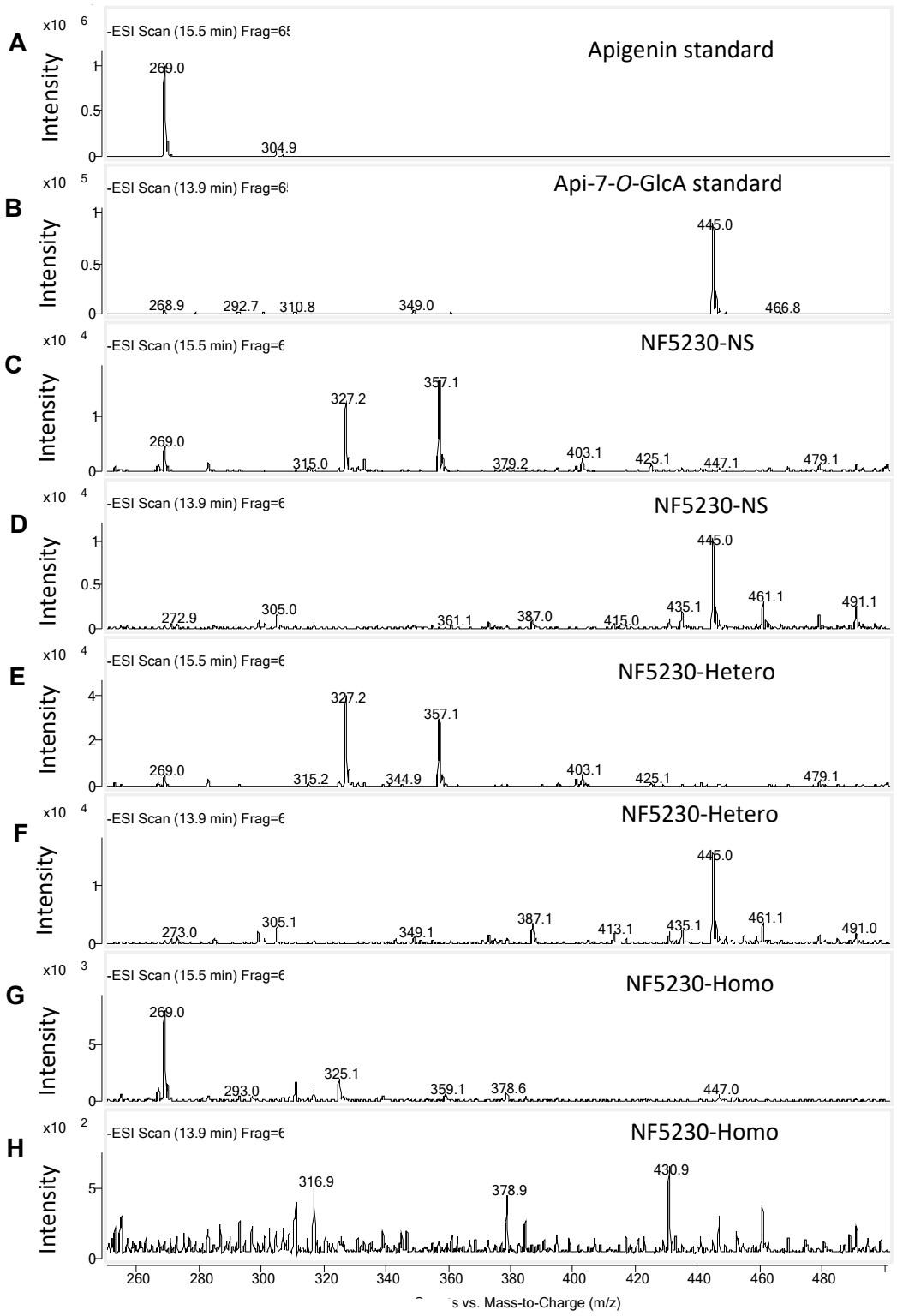
Supplemental Figure S21. Mass spectra corresponding to the chromatograms showing formononetin-7- *O*-glucoside and quercetin glucoside in **Figure S20**. All the lines analyzed

accumulate formononetin-7-*O*-glucoside (at retention time 11.5 min) (A, C, E, G) and quercetin glucoside (at retention time 14.7 min) (B, D, F, H) based on the presence of their respective molecular masses at $[M-H]^+$ = 429.1 and 463.1 respectively. The LC-MS was performed in negative mode.



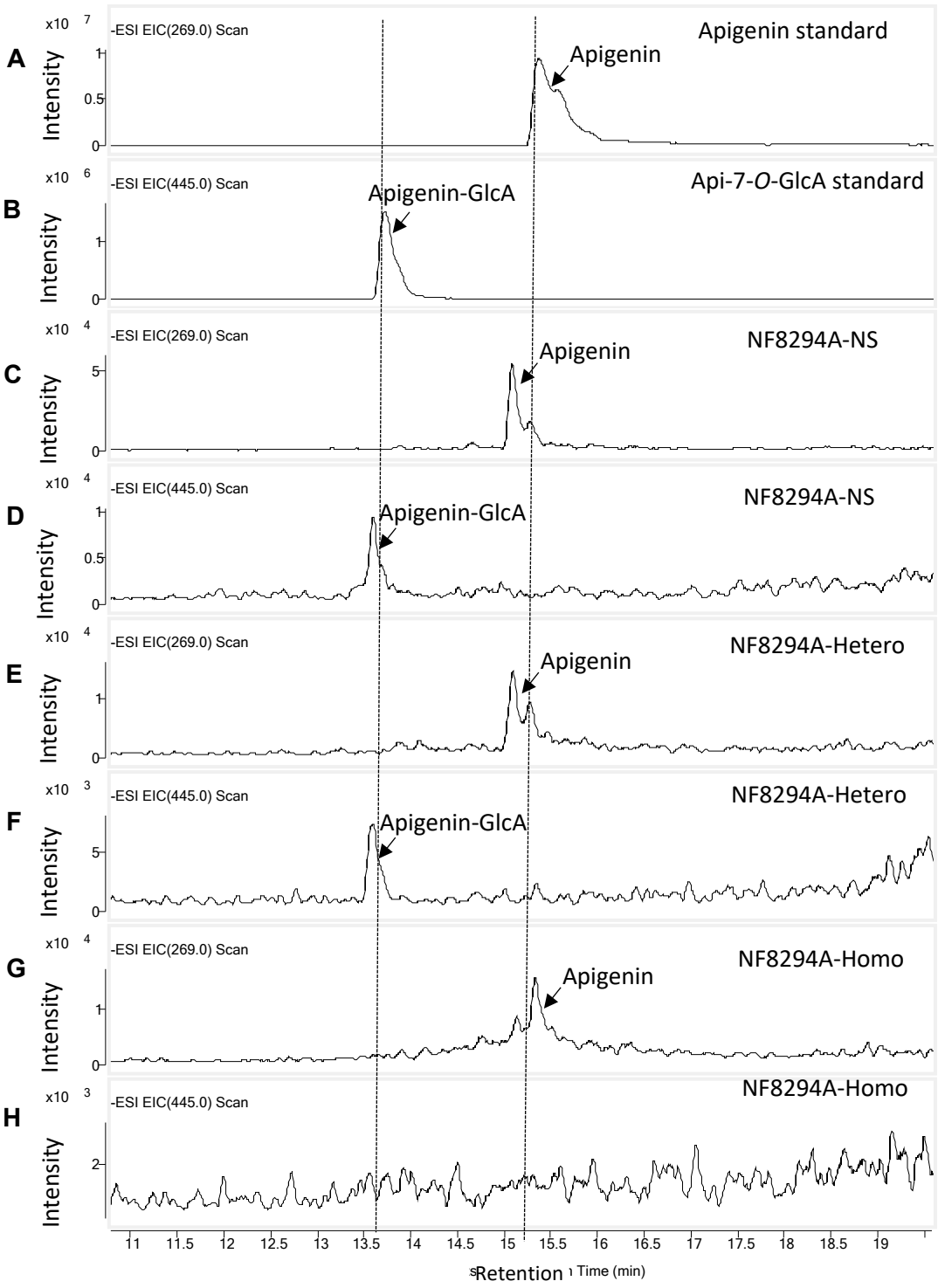
Supplemental Figure S22. Disappearance of apigenin glucuronide due to loss of function of UGT84F9 in the homozygote *ugt84f9 Tnt1* insertion in NF5230 mutant lines *M. truncatula*.

Figure shows MS analysis of flavonoid metabolites from leaf extracts of null segregant, heterozygote and homozygote *ugt84f9 Tnt1* insertion in NF5230 mutant of *M. truncatula*. In comparison with apigenin standard (m/z = 269) and apigenin-7-O-GlcA standard (m/z 445) (A, B), apigenin was detected in the extracted ion chromatogram from scanning the NF5230-null segregant, heterozygote and homozygote extracts for m/z = 269 (C, E, G), whereas apigenin-7-*O*-GlcA (m/z = 445) was detected in leaf extracts of null segregant and heterozygote (D, F) but not the homozygotic *ugt84f9 Tnt1* insertion in the NF5230 mutant (H). The LC-MS was performed in negative mode.



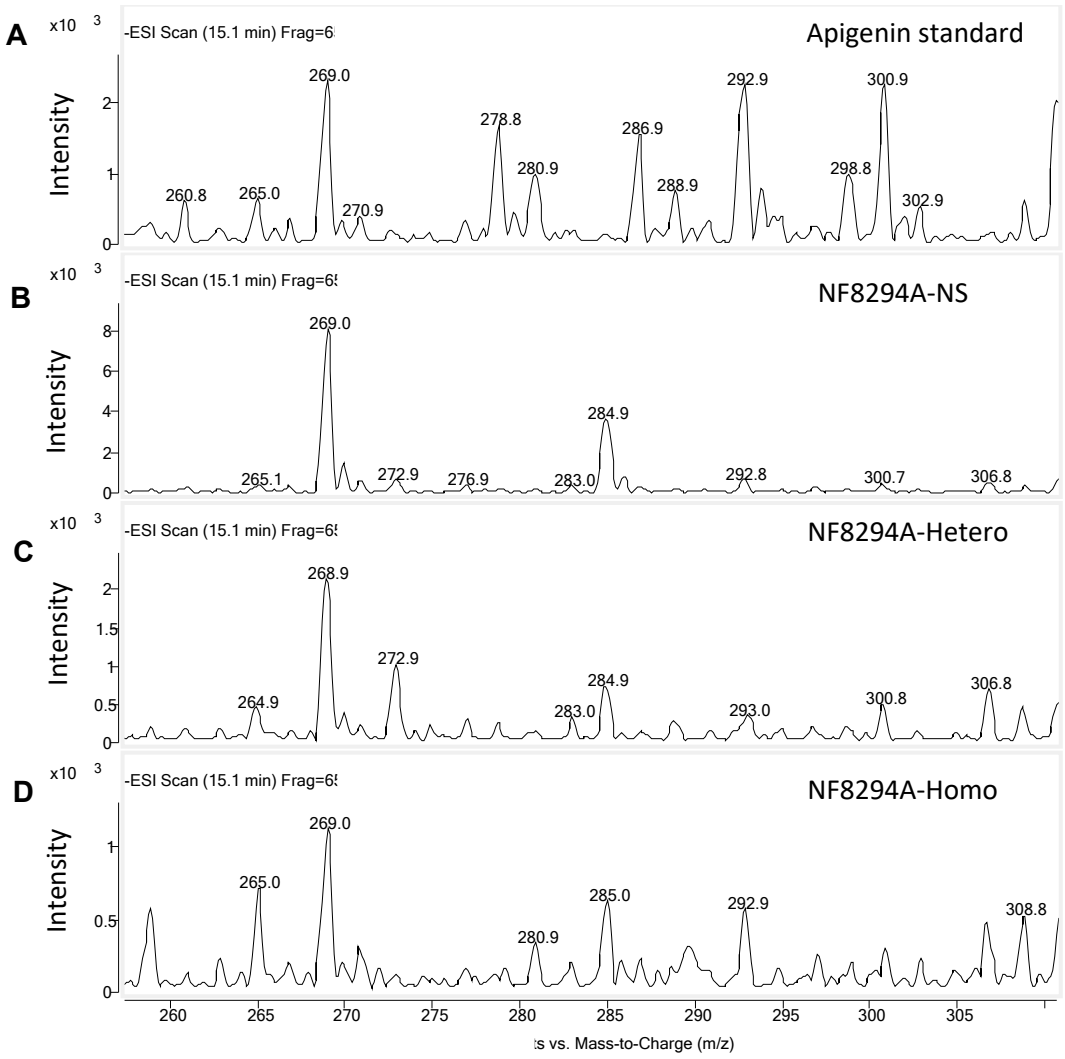
Supplemental Figure S23. Mass spectra corresponding to the metabolites in **Supplemental Figure S21**. The MS spectrum for apigenin, with molecular mass of the parent ion at $[M-H]^+$ = 269, was seen in all the samples (C, E, G); The spectrum of apigenin-*7-O*-GlcA, $[M-H]^+$ = 455, was present in the extract

from NF5230- null segregant and heterozygote (D, F) but was not detected in the extract from NF5230-homozygote (H). The LC-MS was performed in negative mode.

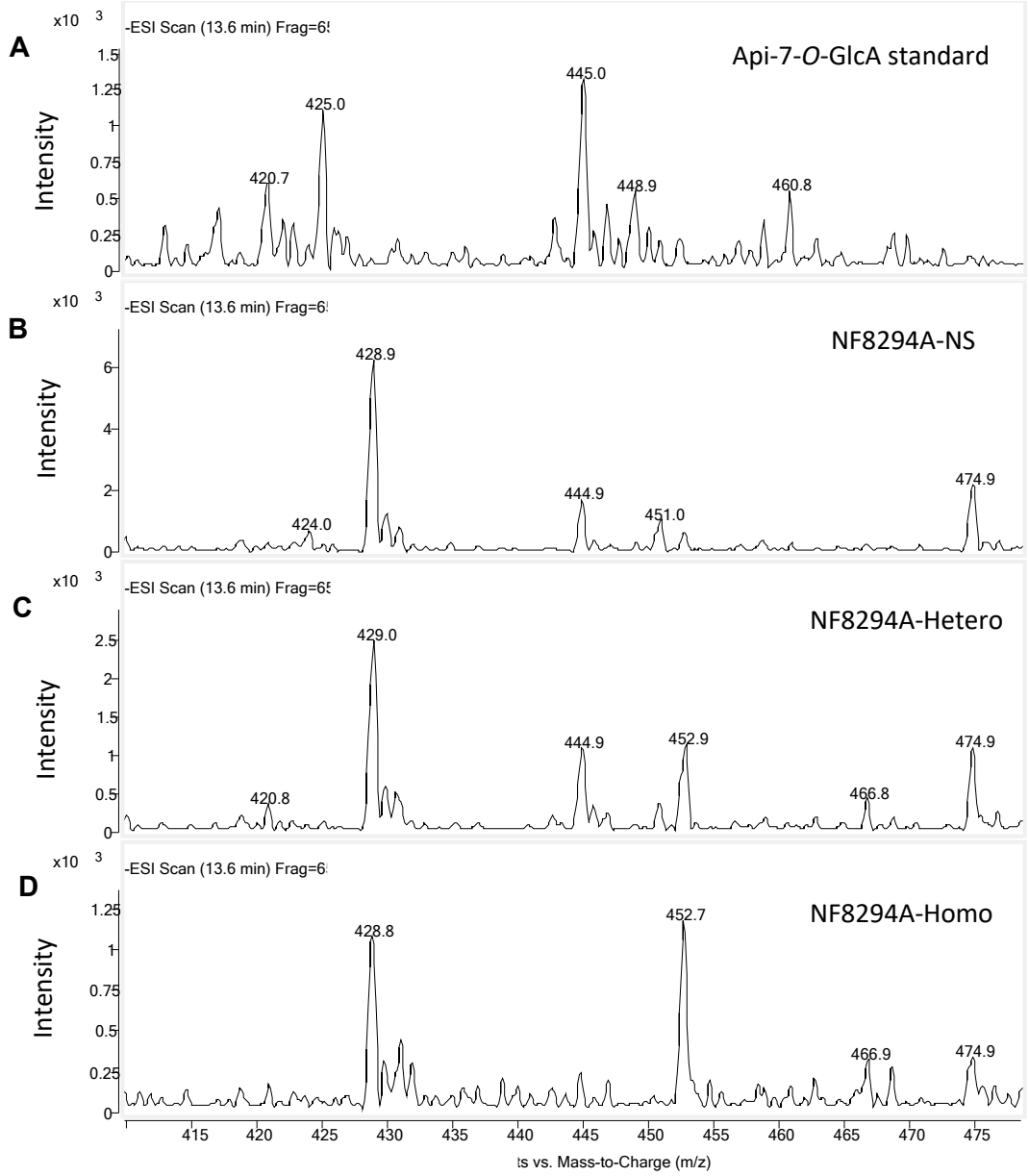


Supplemental Figure S24. LC-MS analysis of flavonoid metabolites from root extracts of NF8294A-null segregant, -heterozygote and -homozygote *ugt84f9 Tnt1* insertion mutant of *M. truncatula*. The disappearance of apigenin glucuronide was observed due to loss of function of UGT84F9 in the homozygote *ugt84f9 Tnt1* insertion in NF8294A mutant lines.

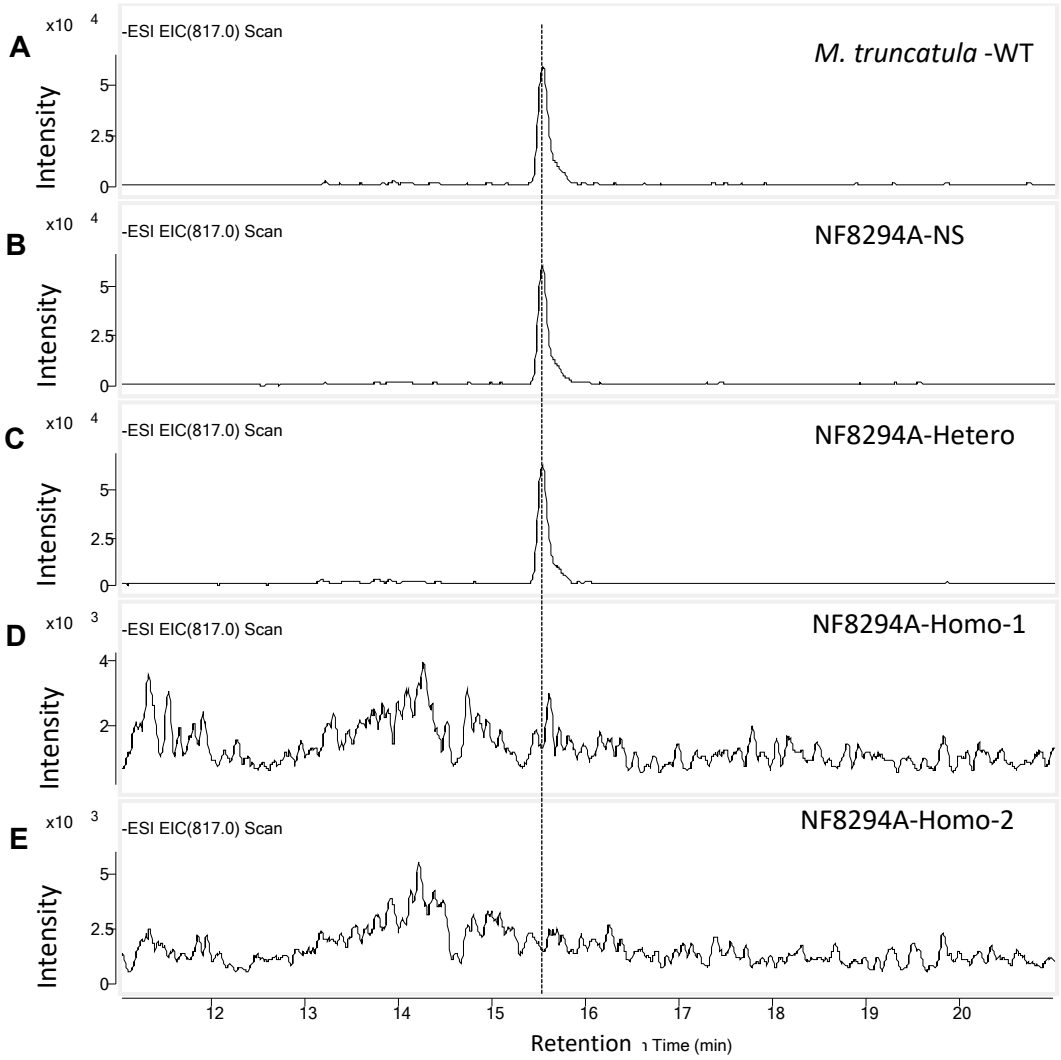
A, B, Apigenin standard (m/z = 269) and apigenin-7-O-GlcA standard (m/z 445). Apigenin was seen in the extracted ion chromatogram from scanning NF8294A-null segregant, -heterozygote and -homozygote root extracts (C, E, G), whereas apigenin-7-*O*-GlcA was detected in root extracts of NF8294A-null segregant and -heterozygote (D, F) but not in the homozygotic *ugt84f9 Tnt1* insertion in the NF8294A mutant (H). The LC-MS was performed in negative mode.



Supplemental Figure S25. Mass spectra corresponding to apigenin in **Supplemental Figure S24.** The molecular mass of the parent ion at $[M-H]^+ = 269$ (A), was seen in extracted ion chromatograms of root extracts from NF8294A-null segregant, -heterozygote and -homozygote (B, C, D). The LC-MS was performed in negative mode.

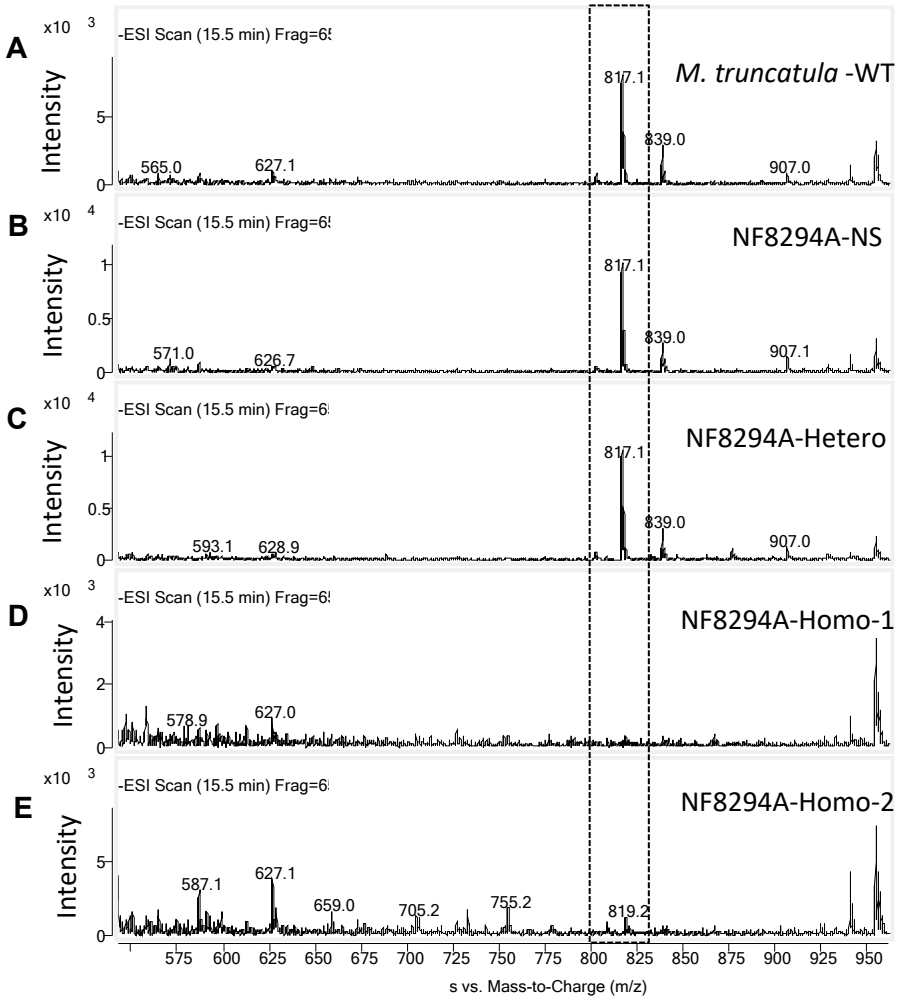


Supplemental Figure S26. Mass spectra corresponding to apigenin-*7-O*-GlcA in **Supplemental Figure S24**. The molecular mass of the parent ion at $[M-H]^+$ =445 (A), was seen in root extracts of NF8294A-null segregant and NF8294A -heterozygote (B, C), but was not detected in NF8294A-homozygote root extracts (**D**). The LC-MS was performed in negative mode.

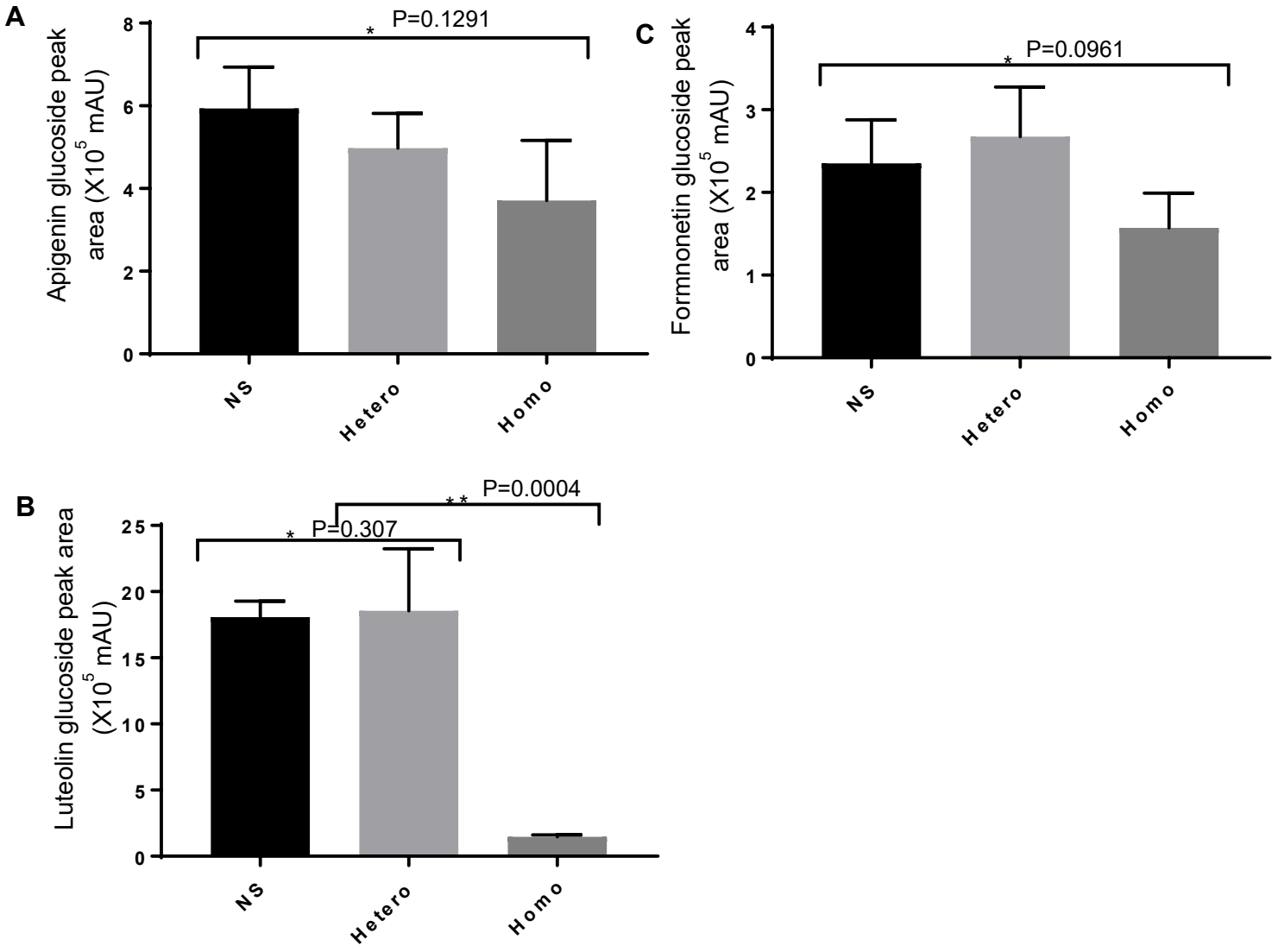


Supplemental Figure S27. Disappearance of laricitrin-3-*O*-glucopyranoside-5'-*O*-glucopyranosyl-7-*O*-glucoside due to loss of function of UGT84F9 in the homozygous *ugt84f9 Tnt1* insertion in NF8294A mutant lines of *M. truncatula*.

A-E, LC-MS analysis of laricitrin-3-*O*-glucopyranoside-5'-*O*-glucopyranosyl-7-*O*-glucoside (m/z = 817) in leaf extracts of *M. truncatula* WT, NF8294A -null segregant, -heterozygote and -homozygote *ugt84f9 Tnt1* insertion mutant. The target compound was detected in the extracted ion chromatograms from the *M. truncatula* WT, NF8294A -null segregant and -heterozygote leaf extracts (A, B, C) whereas laricitrin-3-*O*-glucopyranoside-5'-*O*-glucopyranosyl-7-*O*-glucoside was not detected in the homozygous *ugt84f9 Tnt1* insertion in the NF8294A mutant (D, E). The LC-MS was performed in negative mode.



Supplemental Figure S28. Mass spectra corresponding to laricitrin-3-*O*-glucopyranoside-5'-*O*-glucopyranosyl-7-*O*-glucoside in **Supplemental Figure S27.** Laricitrin-3-*O*-glucopyranoside-5'-*O*-glucopyranosyl-7-*O*-glucoside, with molecular mass of the parent ion at $[M-H]^+$ =817.1, was seen in leaf extracts of *M. truncatula* WT, NF8294A-null segregant and NF8294A -heterozygote (A, B, C), but was not detected in NF8294A-homozygote leaf extracts (**D, E**). The LC-MS was performed in negative mode.



Supplemental Figure S29: Levels of apigenin, luteolin and formononetin mono-glucosides in *ugt84f9 Tnt1* insertion lines. Flavonoid glucosides were determined by LC-MS analyses of extracts of the leaves of 4-week old *M. truncatula* NF8490A lines (NS, null-segregant; Hetero, heterozygote; Homo, homozygote) as described in Methods. The values represent the mean ± s.d. (n = 3 biological replicates). *P > 0.05 indicates statistically insignificant differences and **P < 0.05 indicates significant differences compared with the null-segregant in one-way ANOVA with Tukey test (GraphPad Prism 7.04). The experiment was repeated with a second set of plants of the same age, and essentially the same results were obtained.

Supplemental Table S1. Screening of selected flavonoid compounds and their derivatives for glucuronidation by mammalian UGATs.

Mammalian UGAT	Acceptor substrates	Position of conjugation on acceptor substrate or product formed [a]
UGT1A1	apigenin, luteolin, epicatechin, 4´-*O*-Me-Epi, quercetin, 3´ and 4´- *O*-Me-Q	7-OH (q, 3´mq, 4´mq, a, l) , 3'-OH (q, 4´mq, l), 3-OH (q, 3´mq), 5-OH (e, 4'me)
UGT1A3	apigenin, luteolin, quercetin, 4´-*O*-Me-Q	7-OH (q, 4´mq, a, l), 3´-OH (q, 4´mq, l)
UGT1A4	apigenin, luteolin, quercetin	7-OH (a, l, q), 3´-OH (l, q)
UGT1A6	apigenin, luteolin, epicatechin, 4´-*O*-Me-Epi, quercetin	7-OH (q, 4´mq, a, l), 3´-OH (q, 4´mq, l), 5-OH (e, 4´me)
UGT1A7	apigenin, luteolin, epicatechin, 4´-*O*-Me-Epi, quercetin	7-OH (q, 4´mq, a, l), 3´-OH (q, 4´mq, l), diglucuronide (4´mq), 5-OH (e, 4´me)
UGT1A8	apigenin, luteolin, quercetin, 4´-*O*-Me-Q	7-OH (q, 4´mq, a, l), 3´-OH (q, 4´mq, l)
UGT1A9	apigenin, luteolin, epicatechin, 4´-*O*-Me-Epi, quercetin, 3´ and 4´- *O*-Me-Q	7-OH (q, 4´mq, 3´mq ,a, l) , 3´-OH (q, 4'mq, l), 4´-OH (q, 3´mq , l), 3-OH (q, 3´mq), 5-OH (e, 4´me)
UGT1A10	apigenin, luteolin, epicatechin, 4´-*O*-Me-Epi, quercetin, 3´ and 4´- *O*-Me-Q,	7-OH (q, 3´mq , 4´mq, a, l), 3´-OH (q, 4´mq, l), 3-OH (q, 3´mq , 4´mq), 4´-OH (q, l), diglucuronide (4´mq), 5-OH (e, 4´me)

[a]Identities were based on co-chromatography with authentic standards where available, or are tentative assignments based on deductions or literature information. See Table S5 for details of sources of standards, and Docampo et al. (2017) for a summary of products reported in the literature. a, apigenin; l, luteolin; q, quercetin; 3´mq, 3´-O-methyl quercetin; 4´mq, 4´-O-methyl quercetin; e, epicatechin; 4´me, 4´-O-methyl epicatechin.

Supplemental Table S2. GenBank accession numbers of the UG(A)T sequences analyzed in the present work.

Organism name	Protein Sequence ID	Assigned annotation
Medicago truncatula	XP_013454613.1	G3
Medicago truncatula	XP_013466442	G2
Medicago truncatula	XP_013446783.2	G4
Medicago truncatula	XP_003619124.1	J4
Bellis perennis	Q5NTH0	BpUGT94B1
Medicago truncatula	AFK34364.1	C6
Medicago truncatula	XP_024637606	C5
Medicago truncatula	RHN62167.1	C8
Glycyrrhiza uralensis	ANJ03631	GuUGAT
Medicago truncatula	XP_003609043.1	J1
Glycyrrhiza uralensis	LC314779	UGT73P12 (canonical)
Glycyrrhiza uralensis	LC315805	UGT73P12 (variant)
Medicago truncatula	XP_013466939.1	C3
Medicago truncatula	ABI94026.1	C4
Vitis vinifera	BAI22846.1	VvGT5
Medicago truncatula	XP_003610166.3	UGT78G3
Medicago truncatula	XP_003610163.1	UGT78G1
Medicago truncatula	XP_013470035.1	UGT84F9
Medicago truncatula	XP_003621424.1	J6
Medicago truncatula	XP_003621425.2	J7
Medicago truncatula	XP_003621427.1	J8
Medicago truncatula	XP_003620191.1	J5
Medicago truncatula	XP_003618190.1	G1
Medicago truncatula	XP_003622620.2	UGT72Y5
Medicago truncatula	EU434684	UGT72L1
Medicago truncatula	XP_003615613.1	UGT71G1
Medicago truncatula	RHN57636.1	J2
Medicago truncatula	XP_003618094.1	J3
Perilla frutescens	BAG31949.1	UGT88A7
Erigeron breviscapus		UGT88X
Medicago truncatula	XP_013451680.1	C1
Medicago truncatula	XP_013451685.1	C2
Medicago truncatula	XP_024642419.1	UGT88E29
Medicago truncatula	XP_013451676.1	UGT88E27
Medicago truncatula	XP_013451675.1	UGT88E28
Antirrhinum majus	BAG31945	UGT88D4
Sesamum indicum	NP_001306616.1	UGT88D6
Scutellaria baicalensis	BAH19313.1	UGT88D1
Scutellaria laeteviolacea	BAG31946.1	UGT88D5
Perilla frutescens	BAG31948.1	UGT88D7

Supplemental Table S3. Kinetic parameters of recombinant UGT84F9 with selected flavonoid substrates in the presence of saturating concentrations of sugar donor (5mM UDP-Glc or UDP-GlcA).

Enzyme	Substrate	Sugar donor	K_m (µM)	K_{cat} (sec^{-1})	K_{cat}/K_m ($sec^{-1}mM^{-1}$)
UGT84F9	Kaempferol-3-*O*-rutinoside	UDP-GlcA	609.3 ± 103.5	1.41 ± 0.137	2.314
UGT84F9	Luteolin	UDP-GlcA	380.4 ± 65.25	0.397 ± 0.024	1.04
UGT84F9	Quercetin	UDP-Glu	58.65 ± 10.43	0.046 ± 0.003	0.782
UGT84F9	Quercetin	UDP-GlcA	768.2 ± 56.39	0.516 ± 0.018	0.672
UGT84F9	Apigenin	UDP-GlcA	772.6 ± 249.3	0.491 ± 0.072	0.635

Supplemental Table S4. Primer sets used in this study

Primer name	**Primers Sequence (5'-3') used for molecular cloning of UGT84F9**
UGT84F9Nde1-F	CAC CCA TAT GAC ATA CGA AGA TCC CAT TAA GC
UGT84F9EcoRV-R	GGG CCC GAT ATC TTA GAT GTT AAC ATT ATT AAT TAA TG
Primer name	**Primers Sequence (5'-3') used for mutagenesis of UGT84F9**
UGT84F9-H22R-F	GGA CAC ATA AAC CGT CTT GTT GGA CTA GG
UGT84F9-H22R-R	CCT AGT CCA ACA AGA CGG TTT ATG TGT CC
UGT84F9-H22S-F	GGA CAC ATA AAC AGT CTT GTT GGA CTA GG
UGT84F9-H22S-R	CCT AGT CCA ACA AGA CTG TTT ATG TGT CC
UGT84F9	GGA CAC ATA AAC GCT CTT GTT GGA CTA GG
UGT84F9	CCT AGT CCA ACA AGA GCG TTT ATG TGT CC
UGT84F9-N122D-F	CAT GTA TCA TAA ACG ATT ATT TTT TTC CAT GGG TTT G
UGT84F9-N122D-R	CAA ACC CAT GGA AAA AAA TAA TCG TTT ATG ATA CAT G
Primer name	**Primers Sequence (5'-3') used for qPCR**
qUGT84F9-F	TCC AGC ACA AGG ACA CAT AAA
qUGT84F9-R	CGT CTC TGT TGT GGT GAA GAT
qUGT78G3F	GAT GGA TTA CCA GAA GGG TAT GT
qUGT78G3R	CTG CCA CAG CTT TAA CCA TAA C
qUGT88E27F	GCA GAA CAA AGG CTG AAC AAG
qUGT88E27R	CTA ACT CGG TTC CAC TCA CAA A
qUGT88E28F	AGG ATC TCC ATA CGC CTC TT
qUGT88E28R	CCC ATC ACA TTG CCT CAT AGT
qTUB-F	TTT GCT CCT CTT ACA TCC CGT G
qTUB-R	GCA GCA CAC ATC ATG TTT TTG G
Primer name	**Primers Sequence (5'-3') used for analysis of UGT84F9 *Tnt1* transposon-insertion mutant lines**
UGT84F9tnt-F	GAG CAA GTA TTC TCT TGC TAT AGG TAT CTC C
UGT84F9tnt-R	GGA TAG TTT ACT AGA GTC CCG AAA GAG ACG
UGT84F9tnt-R_2	GGG CCC GAT ATC TTA GAT GTT AAC ATT ATT AAT TAA TG
TNT1-F	CCT TGT TGG ATT GGT AGC CAA CTT TGT TG
TNT1-R	TGT AGC ACC GAG ATA CGG TAA TTA ACA AGA

Supplemental Table S5. List of standards or methods used for identification of glucuronide products

Compound	**Standard or method used for identification**	**Source**
Quercetin-3-*O*-GlcA	Purchased standard	Sigma-Aldrich
Quercetin-3´-O-GlcA	Tentative, based on comparison to chemically synthesized 4´-O-Me-Q-3´-O-GlcA	(Docampo-Palacios et al 2020a)
7-*O*-glucuronides of apigenin, luteolin, quercetin	Enzymatically generated	Generated using F7GAT, UGT88D4 (Noguchi et al 2009)
4´-*O*-Me-quercetin-3´-*O*-GlcA	Chemically synthesized standard	(Docampo-Palacios at al 2020a)
3´-*O*-Me-quercetin-3-*O*-GlcA	Chemically synthesized standard	(Docampo-Palacios at al 2020a)
4´-*O*-Me-quercetin-7-*O*-GlcA	Chemically synthesized standard	(Docampo-Palacios at al 2020a)
Quercetin-4´-*O*-GlcA	Tentative, by chromatographic separation behavior	
Di-glucuronides	Tentative, by early chromatographic separation relative to monoglucuronide, plus m/z value from MS analysis	
Epicatechin-5-*O*-GlcA and 4´-*O*-Me-epicatechin-5-*O*-GlcA	Chemically synthesized standard	(Docampo-Palacios at al 2020b)