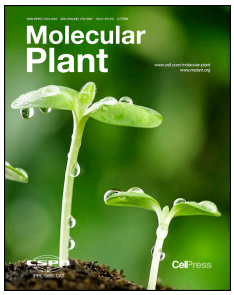
Journal Pre-proof



Steroidal scaffold decorations in *Solanum* alkaloid biosynthesis

Rosalind Lucier, Mohamed O. Kamileen, Yoko Nakamura, Sofiia Serediuk, Ranjit Barbole, Jens Wurlitzer, Maritta Kunert, Sarah Heinicke, Sarah E. O'Connor, Prashant D. Sonawane

PII: S1674-2052(24)00193-X

DOI: https://doi.org/10.1016/j.molp.2024.06.013

Reference: MOLP 1752

To appear in: *MOLECULAR PLANT*

Received Date: 7 January 2024

Revised Date: 10 June 2024

Accepted Date: 25 June 2024

Please cite this article as: Lucier R., Kamileen M.O., Nakamura Y., Serediuk S., Barbole R., Wurlitzer J., Kunert M., Heinicke S., O'Connor S.E., and Sonawane P.D. (2024). Steroidal scaffold decorations in *Solanum* alkaloid biosynthesis. Mol. Plant. doi: https://doi.org/10.1016/j.molp.2024.06.013.

Steroidal scaffold decorations in *Solanum* alkaloid biosynthesis

Rosalind Lucier[1], Mohamed O. Kamileen[1], Yoko Nakamura[1,2], Sofiia Serediuk[1], Ranjit Barbole[3], Jens Wurlitzer[1], Maritta Kunert[1], Sarah Heinicke[1], Sarah E. O'Connor[1*], Prashant D. Sonawane[1*]

[1]Department of Natural Product Biosynthesis, Max Planck Institute for Chemical Ecology, Jena 07745, Germany

[2]Research Group Biosynthesis and NMR, Max Planck Institute for Chemical Ecology, Jena 07745, Germany

[3]Biochemical Sciences Division, CSIR-National Chemical Laboratory, Pune, 411008, Maharashtra, India

*Corresponding authors: oconnor@ice.mpg.de

psonawane@ice.mpg.de

Running title: Steroidal alkaloid decorations in *Solanum* plants

Short Summary: This study uncovers the biosynthetic pathways of the most widespread SGA scaffold, solasodine and downstream three main bioactive SGAs, *α*-solasonine, *α*-solamargine and malonyl-solamargine in *Solanum nigrum* (black nightshade) and *S. melongena* (eggplant), and also enables a platform for engineering the sustainable production of high value, bioactive steroidal molecules using synthetic biology and metabolic engineering applications.

Abstract

Steroidal glycoalkaloids (SGAs) are specialized metabolites produced by hundreds of *Solanum* species, including important vegetable crops such as tomato, potato and eggplant. Though SGAs are better known for their role in defence in plants and 'anti-nutritional' effects (e.g., toxicity and bitterness) to humans, many of these molecules have documented anti-cancer, anti-microbial, anti-inflammatory, anti-viral and anti-pyretic activities. Among these, *α*-solasonine and *α*-solamargine isolated from black nightshade (*Solanum nigrum*), are reported to have potent anti-tumor, anti-proliferative and anti-inflammatory activities. Notably, *α*-solasonine and *α*-solamargine, along with the core steroidal aglycone solasodine are the most widespread SGAs produced among the *Solanum* plants. However, it is still unknown how plants synthesize these bioactive steroidal molecules. Through comparative metabolomic-transcriptome guided approach, biosynthetic logic, combinatorial expression in *Nicotiana benthamiana* and functional recombinant enzyme assays, here we report the discovery of 12 enzymes from *S. nigrum* that converts the staring cholesterol precursor to solasodine aglycone, and the downstream *α*-solasonine, *α*-solamargine and malonyl-solamargine SGA products. We further identified 6 enzymes from cultivated eggplant that catalyse the production of *α*-solasonine, *α*-solamargine and malonyl-solamargine SGAs from solasodine aglycone, via glycosylation and atypical malonylation decorations. Our work provides the gene tool box and platform for engineering the production of high value, steroidal bioactive molecules in heterologous hosts using synthetic biology.

Keywords: Steroidal glycoalkaloids, *Solanum*, biosynthetic pathway, *α*-solamargine, malonyl-solamargine, specialized metabolites

Introduction

Steroidal glycoalkaloids (SGAs) are nitrogen containing specialized metabolites produced by hundreds of wild and cultivated species of the genus *Solanum*, including agriculturally important food crops such as tomato (*Solanum lycopersicum*), potato (*Solanum tuberosum*), and eggplant (*Solanum melongena*) (Friedman, 2002; Friedman, 2006; Cárdenas et al., 2015; Zhao et al., 2021; Sonawane et al., 2020). SGAs act as chemical defenses against a broad range of plant pathogens, pests and herbivores (Friedman, 2002; Friedman, 2006; Zhao et al., 2021; Sonawane et al., 2018). For humans, some SGAs are considered as anti-nutritional factors in the diet due to their toxicity and bitterness (e.g., α-solanine and α-chaconine in potato); however several *Solanum* SGAs are known for their anti-cancer, anti-microbial, anti-inflammatory, anti-viral and anti-pyretic activities (Milner et al., 2011; Friedman, 2015; Winkiel, 2022; Delbrouck et al., 2023). Some renowned examples of bioactive SGAs are α-tomatine in tomato, α-solanine and α-chaconine in potato and α-solasonine and α-solamargine in eggplant (Zhao et al., 2021; Milner et al., 2011; Sinani and Eltayeb, 2017). In recent years, *Solanum nigrum* Linn., a wild member of the *Solanum* genus has captured much attention due to its remarkable anti-tumor and anti-proliferative bioactivities against range of different cancer cells (e.g., liver, cervical, lung, breast, colon etc.) displayed mainly by α-solamargine and α-solasonine, two SGAs isolated from green fruits (Sinani and Eltayeb, 2017; Ding et al., 2012; Shiu et al., 2007; Ding et al., 2013; Gu et al., 2018; Kalalinia and Karimi-Sani, 2017; Liang et al., 2022). Notably, *S. nigrum*, commonly known as black nightshade has been part of traditional Chinese and Indian medicine for thousands of years as a treatment of fever, kidney and urinary infections, skin diseases etc. (Jabamalairaj et al., 2019; Chen et al., 2022).

A typical SGA structure contains two components, a steroidal alkaloid aglycone and glycoside residues attached to C-3 hydroxyl of aglycone. The biosynthesis of dehydrotomatidine, the first steroidal alkaloid aglycone in tomato and potato, through action of several GLYCOALKALOID METABOLISM (GAME) enzymes from the starting precursor cholesterol has been well studied (Itkin et al., 2013;

Umemoto et al., 2016; Nakayasu et al., 2017; Nakayasu et al., 2021) (Figure 1, Supplemental Figure 1). Subsequent removal of the C-5,6 double bond from dehydrotomatidine by 3βHSD (3β-hydroxysteroid dehydrogenase/3-Ketosteroid reductase, Lee et al., 2019), also known as GAME25 (Sonawane et al., 2018) and 5α reductase2 (5αR2) (Akiyama et al., 2019) results in the formation of tomatidine, the main steroidal aglycone in tomato (Figure 1). Finally, both dehydrotomatidine and tomatidine are decorated by four UDP-glycosyltransferases (UGTs); GAME1, GAME2, GAME17 and GAME18 to form dehydrotomatine and *α*-tomatine, respectively (Itkin et al., 2011; Itkin et al., 2013**)** (Figure 1, Supplemental Figure 1). In potato, dehydrotomatidine is hypothesized to generate glycosylated intermediates, *α*-solamarine and *β*-solamarine, which are further converted to potato specific SGAs, *α*-solanine and *α*-chaconine, respectively (Akiyama et al., 2021) (Figure 1, Supplemental Figure 1). In contrast, solasodine is the main steroidal alkaloid aglycone in black nightshade (*S. nigrum*) and eggplant (*S. melongena*), which is further glycosylated to produce diverse SGA structures such as *α*-solasonine, *α*-solamargine and malonyl-solamargine (Zhao et al., 2021; Sinani and Eltayeb, 2017**)** (Figure 1). Notably, solasodine is the most widespread steroidal aglycone, present in almost 200 out of 350 *Solanum* species reported to produce SGAs (Sinani and Eltayeb, 2017; Eich, 2008**)**. Despite the long and rich histories that demonstrate the therapeutic potential of *α*-solasonine and *α*-solamargine SGAs, their biosynthetic pathways starting from cholesterol remain to be identified in any *Solanum* species (Figure 1).

Here we report the elucidation of *α*-solasonine, *α*-solamargine and malonyl-solamargine biosynthesis in wild *S. nigrum* (black nightshade) and cultivated *S. melongena* (eggplant) plants. We identified 12 enzymes including six GAMEs, five UGTs and one malonyltransferase that are together involved in biosynthesis of solasodine aglycone scaffold and its downstream bioactive SGAs (e.g., *α*-solamargine) from common precursor cholesterol. Our work paves the way for metabolic engineering

and production of potential therapeutic SGA repertoire using synthetic biology approaches in microbial or plant host platforms.

Results and Discussion

Biosynthesis of the solasodine aglycone scaffold

Comparative profiling of metabolites from various tissues, coupled with analysis of transcriptome data obtained from the same tissues serves as a powerful approach to identify candidate genes in a biosynthetic pathway. Both unripe green and ripe purple fruits (commonly known as berries) of *S. nigrum* produce substantial levels of SGAs (Ding et al., 2013; Gu et al., 2018; Zhao et al., 2023). To guide our study of SGA biosynthesis in *S. nigrum*, we first employed ultra high performance liquid chromatography coupled to quadrupole time-of-flight mass spectrometry (UHPLC-qTOF-MS) and examined the SGA content in green and purple berries. We observed high accumulation of *α*-solasonine and *α*-solamargine in green berries of *S. nigrum* (Figure 2A), consistent with previous reports (Zhao et al., 2023; Bednarz et al., 2019). Additionally, we also noted the high accumulation of SGA in green berries with the mass that corresponded to putative malonyl-solamargine (*m/z* 954.5038; $C_{48}H_{78}NO_{18}$ $[M+H]^+$), for which structural assignment was supported by tandem mass spectrometry (MS/MS or MS^2) fragmentation (Figure 2A, Supplemental Figure 2). It seemed logical that malonyl-solamargine could form from *α*-solamargine through a single malonylation reaction. Notably, MS/MS analysis comparison between putative malonyl-solamargine and *α*-solamargine standard led us to propose that this malonyl group was attached to one of the rhamnose sugars (Rha I or Rha II) present in the structure (Supplemental Figure 2), though the exact position remained unverified. We isolated this compound from green berries of *S. nigrum* using analytical scale HPLC, and confirmed its structure as malonyl-solamargine by NMR spectroscopy, which revealed that the malonyl group is unambiguously attached to the C-4 position of the Rha II sugar (Supplemental Table 1, 2 and Supplemental Figures 3-12 for NMR spectra). The

assignment of aglycone portion of this compound was identical to the aglycone assignment of *α*-solamargine standard, carried out by 1D and 2D NMR spectroscopy experiments (Supplemental Figures 13-24 for NMR spectra). In addition, SELTCOSY (Selective TOCSY) experiments were performed to clarify the assignment of sugar moieties of this compound (Supplemental Figure 25 and 26). The isolated malonyl-solamargine contain solasodine aglycone and the C-3 sugar chain consisting of one D-glucose and two L-rhamnose (Rha I and Rha II) units, similarly as that of *α*-solamargine, but in addition has a malonyl moiety at C-4 of Rha II sugar (Supplemental Figure 3 and 13, see structures). Thus, the malonylation pattern observed here in case of steroidal glycoalkaloids is unusual as numerous plant secondary/specialized metabolites such as anthocyanin-, flavonoid-, diterpenoid- and triterpene-glycosides are often malonylated with the malonyl group typically attached to glucose moiety present in glycosides (Li et al., 2018; Luo et al., 2007; Taguchi et al., 2010; Ahmed et al., 2017). We were also able to detect *α*-solasonine, *α*-solamargine and malonyl-solamargine in purple ripe berries of *S. nigrum*, but their levels were very low compared to green berries (Figure 2A). Guided by these metabolic profiling results, we next generated the transcriptomic data from green (unripe) and purple (ripe) berries of *S. nigrum* to identify corresponding SGA biosynthetic genes.

In tomato and potato SGA biosynthesis, the starting precursor cholesterol is proposed to be converted to the first core steroidal alkaloid aglycone, dehydrotomatidine (*22S*, *25S*), via series of hydroxylation, oxidation and transamination reactions catalysed by a set of five GAME enzymes; GAME6 (CYP72A188), GAME8 (CYP72A208), GAME11 (dioxygenase), GAME4 (CYP88D) and GAME12 (aminotransferase) (Figure 1, Supplemental Figure 1). The main SGAs in *S. nigrum* green berries, *α*-solasonine, *α*-solamargine and malonyl-solamargine share the stereoisomer solasodine (*22R*, *25R*) as the core steroidal alkaloid aglycone scaffold in their structures (Figure 1). We envisaged that the pathway to solasodine would occur in an analogous fashion as the established pathway for dehydrotomatidine (Itkin et al., 2013; Umemoto et al., 2016; Nakayasu et al., 2017; Nakayasu et al.,

2021), with similar modification steps on the steroid backbone, and corresponding GAME enzyme orthologues would catalyze the formation of solasodine from cholesterol precursor in *S. nigrum*. Indeed, we identified orthologues of biosynthetic GAME genes (GAME6, GAME8, GAME11, GAME4, and GAME12) in our *S. nigrum* transcriptome sharing 70 to 93 % amino acid sequence identity to known GAME genes from tomato and potato (Supplemental Figure 27). Notably, all five candidate GAME genes were preferentially expressed in green berries (Figure 2B), consistent with the profile of SGAs (e.g. α-solasonine) that accumulate in the same tissue (Figure 2A).

Using gene specific primers, we next cloned each GAME coding sequence from cDNA prepared from green berries of *S. nigrum*, and expressed these genes simultaneously in *Nicotiana benthamiana* leaves using *Agrobacterium tumefaciens*-mediated transient expression. *N. benthamiana* produces substantial amounts of cholesterol (Sonawane et al., 2017), the starting precursor for SGA biosynthesis, and thus well suited to test the activity of biosynthetic GAME enzymes involved in solasodine aglycone formation. Co-expression of all five GAME genes, *GAME6*, *GAME8*, *GAME11*, *GAME4* and *GAME12* in *N. benthamiana* did not result in the production of expected solasodine aglycone. Notably, however when we included the gene *GAME15*, annotated as cellulose synthase like protein, together with identified five *GAME* genes in transient expression experiments, solasodine formation was observed in infiltrated *N. benthamiana* leaves (Figure 2C). Earlier reports suggested the association of GAME15, a cellulose synthase like protein from tomato and potato in SGA metabolism because of its strong co-expression with known SGA biosynthetic GAME (e.g., GAME4, GAME6 etc.) genes as well as its presence as a part of SGA biosynthetic metabolic gene cluster (chromosome 7) together with other GAME genes (e.g. *GAME6*, *GAME11*) in tomato and potato (Itkin et al., 2013; Sonawane et al., 2020). Notably, we identified a clear orthologue of GAME15 in *S. nigrum* that shares 90% homology (at amino acid level) with tomato and potato GAME15 proteins (Supplementary Figure 27). Moreover, in our transcriptome data, GAME15 is predominantly expressed in the green berries, resembling the expression

144 profile of upstream SGA biosynthetic GAME genes (Figure 2B). Altogether, our results showed that six
145 *S. nigrum* enzymes, GAME6, GAME8, GAME11, GAME4, GAME12 and GAME15 mediate
146 solasodine aglycone formation from the cholesterol precursor. We also observed an additional product
147 (marked as * in Figure 2C) in transient expression experiments showing the same mass as that of
148 solasodine (*m/z* 414.30) in selected ion monitoring (SIM) mode of LC-Triple Quadrupole (TQ)-MS
149 analysis. Further MS/MS analysis on LC-qTOF-MS revealed that the actual mass of this peak is *m/z*
150 412.3577 ($C_{28}H_{46}NO$ $[M+H]^+$), different than solasodine produced in the same experiment (*m/z*
151 414.3372; $C_{27}H_{44}NO_2$ $[M+H]^+$) (Supplemental Figure 28). Since very low amounts of this peak were
152 produced, we were not able to structurally characterize it further.
153
154 **Four glycosyltransferases are involved in *α*-solasonine and *α*-solamargine biosynthesis from**
155 **solasodine**
156 Having identified the core biosynthetic GAME enzymes that produce solasodine from cholesterol, we
157 next focused on identifying the enzymes that decorate the solasodine scaffold. The major *S. nigrum*
158 SGAs, *α*-solasonine and *α*-solamargine are glycosylated at the hydroxyl of the C-3 position of solasodine
159 by solatriose (D-galactose, D-glucose and L-rhamnose) and chacotriose (one D-glucose and two L-
160 rhamnose) moieties respectively (for chemical structures see Figure 2D). UDP-dependent
161 glycosyltransferases (UGTs), members of glycosyltransferase 1 (GT1) family, typically carry out
162 glycosylation steps in plant specialized metabolic pathways (Louveau and Osbourn, 2019). Therefore,
163 we speculated that decoration of six sugar moieties (three in each case) in *α*-solasonine and *α*-
164 solamargine biosynthesis would likely be carried out by several UGT enzymes. The order in which these
165 glycosylation reactions would occur in the biosynthetic pathway was unknown (Figure 1).
166 To identify candidate UGT genes involved in *α*-solasonine and *α*-solamargine biosynthesis, we
167 mined the generated *S. nigrum* transcriptome data and initially selected ten candidate UGT genes that

were highly expressed in green unripe berries compared to purple ripe berries (Supplemental Figure 29). Testing all these candidates by *in vitro* enzyme assays was challenging since the order of biosynthesis is unknown, and the glycosylated pathway intermediates (e.g. solasodine-galactoside, solasodine-glucoside etc.) required as substrates in activity assays are not accessible due to lack of availability. To functionally characterize the UGTs rapidly, we cloned all 10 UGT candidates and screened them using a stepwise-pooled approach of pathway reconstitution via transient gene expression in *N. benthamiana* (Pool 1-3, Figure 2D) as outlined by Christ *et al.* (2019). In Pool 1, all 10 UGT candidates were co-infiltrated into *N. benthamiana* leaves together with the six 'solasodine' producing *GAME* genes (*GAME6*, *GAME8*, *GAME11*, *GAME4*, *GAME12* and *GAME15*) and monitored for SGA product formation, specifically *α*-solasonine and *α*-solamargine. LC-MS analysis of leaf extracts revealed that transient expression of these 10 UGTs (Pool 1) indeed resulted in accumulation of both *α*-solasonine and *α*-solamargine in infiltrated *N. bemthamiana* leaves suggesting that the UGT enzymes required for *α*-solasonine and *α*-solamargine biosynthesis are likely present within pool 1 group (Figure 2D). Next, UGT candidates were removed individually, one at a time from the pool to identify the specific UGT enzymes critical for *α*-solasonine and *α*-solamargine product formation (Pool 2, Figure 2D and Supplemental Figure 30). Pool 2 screening revealed that individual removal of three UGTs; *UGT73L14*, *UGT73DU3* and *UGT93N4* abolished *α*-solasonine formation (Figure 2D, Supplemental Figure 30). Notably, removal of *UGT73DU3* also resulted in loss of *α*-solamargine formation, suggesting that UGT73DU3 could be a common enzyme involved in both *α*-solasonine and *α*-solamargine biosynthesis (Figure 2D, Supplemental Figure 30). Nevertheless, another UGT candidate, *UGT93M3* was also found to be crucial for *α*-solamargine production as its removal resulted in no *α*-solamargine formation (Pool 2, Figure 2D and Supplemental Figure 30). Finally, specific UGT candidates from Pool 2 results were co-expressed together with 'solasodine' pathway genes (Pool 3, Figure 2D). Transient co-expression of *UGT73L14*, *UGT73DU3* and *UGT93N4* with solasodine pathway genes clearly generated *α*-solasonine,

while *UGT73DU3* and *UGT93M3* in combination with solasodine pathway genes resulted in *α*-solamargine production in *N. benthamiana* (Pool 3, Figure 2D). Thus, using this pooled-screen approach, we successfully identified four UGT enzymes that are capable of catalyzing *α*-solasonine and *α*-solamargine formation in *N. benthamiana*. To further demonstrate the glycosylation capabilities of these enzymes directly on native substrate, we first expressed *UGT73L14*, *UGT73DU3* and *UGT93N4* genes transiently in *N. benthamiana* leaves followed by infiltration of solasodine aglycone. LC-MS analysis of leaf extracts clearly showed the final SGA product, *α*-solasonine (Figure 2E and 2F). Similarly, transient co-expression of *UGT73DU3* and *UGT93M3* in *N. benthamiana* leaves supplemented with solasodine generated *α*-solamargine (Figure 2E and 2G). In summary, four specific UGT enzymes identified here enable the conversion of solasodine steroidal aglycone to either *α*-solasonine or *α*-solamargine SGAs.

UGT73L14, UGT73DU3 and UGT93N4 act sequentially on solasodine to produce *α*-solasonine

The structure of *α*-solasonine comprises two components; solasodine aglycone and solatriose sugar moieties at C-3 consisting of D-galactose, D-glucose and L-rhamnose (Figure 1). Having successfully reconstituted the *α*-solasonine pathway in *N. benthamiana* with three UGT enzymes; UGT73L14, UGT73DU3 and UGT93N4 discovered here, we next sought to investigate their glycosylation capacities as well as to deconvolute the order of glycosylation in biosynthetic pathway. To address this, we characterized these UGTs using both *in vitro* (biochemical assays) and *in planta* (combinatorial expression in *N. benthamiana*) approaches.

The first sugar residue attached to C-3 position in *α*-solasonine is D-galactose suggesting that formation of solasodine-galactoside, named here as *γ*-solasonine, from solasodine aglycone is the first logical step in *α*-solasonine biosynthetic pathway (Figure 3A). We next expressed UGT73L14, UGT73DU3 and UGT93N4 proteins separately in *Escherichia coli*. Each purified UGT enzyme then assayed *in vitro* using solasodine as a substrate. LC-MS analysis of the assay mixtures revealed that only

UGT73L14 exhibited UDP-galactosyltransferase activity and generated *γ*-solasonine (solasodine-galactoside, *m/z* 576.3902) when incubated with solasodine and UDP-Galactose (UDP-Gal) as the sugar donor (Figure 3, *in vitro* assays). We also observed the accumulation of *γ*-solasonine in leaves of *N. benthamiana* infiltrated with UGT73L14 and solasodine (Figure 3, *in planta* assays). Interestingly, when solasodine was fed to *N. benthamiana* leaves expressing empty vector (control), two compounds with the masses similar as that of *γ*-solasonine, but different retention times were detected, indicating non-specific glycosylated metabolites production via activity of the host UGT enzymes **(Figure 3A,** peak A and B**)**. As no *γ*-solasonine formation was observed in these control samples, it appears that endogenous UGT enzymes from *N. benthamiana* are capable of decorating solasodine, presumably via addition of different hexose (e.g., glucose) than galactose to produce glycosylated solasodine derivatives (i.e. peak A and B). Our results show that UGT73L14 catalyzes the first galactosylation step in *α*-solasonine biosynthesis. The identification of peak A and B is discussed in the next section of *α*-solamargine biosynthesis.

The next reasonable step in biosynthesis is glycosylation of *γ*-solasonine, and either D-glucose or L-rhamnose can be added to γ-solasonine at this stage. Having UGT73L14 activity in hand, we next screened UGT73DU3 and UGT93N4 enzymes for their ability to glycosylate *γ*-solasonine using *in vitro* coupled enzyme assays. Incubation of UGT73L14 and UGT93N4 recombinant enzymes with solasodine substrate and necessary UDP-sugar donors did not generate any di-glycosylated assay products (γ-solasonine-glucoside or *γ*-solasonine-rhamnoside) (Supplemental Figure 31). This shows that UGT93N4 is not involved in the glycosylation of *γ*-solasonine. Conversely, an assay reaction containing UGT73L14, UGT73DU3, solasodine, UDP-Gal and UDP-Rha produced a di-glycosylated product; γ-solasonine-rhamnoside, named here *β*-solasonine (*m/z* 722.4489) (Figure 3, *in vitro* assays). When UDP-Rha was substituted with UDP-Glc as sugar donor in the UGT73L14 and UGT73DU3 coupled assays, no glucosylated product (*γ*-solasonine-glucoside) was detected (Supplemental Figure 31), suggesting the

preference of UGT73DU3 enzyme for UDP-Rha compared to other UDP-sugar donors. The same *β*-solasonine peak was detected when solasodine was infiltrated into *N. benthamiana* leaves that transiently co-expressed UGT73L14 and UGT73DU3 (Figure 3, *in planta* assays). Thus, UGT73DU3 encodes a rhamnosyltransferase catalyzing the second glycosylation step in *α*-solasonine biosynthetic pathway. From the above results, it is now clear that the last step in the biosynthetic pathway is glucosylation of *β*-solasonine to produce *α*-solasonine, likely catalyzed by UGT93N4, which is predicted to encode a glucosyltransferase. Indeed, when recombinant UGT73L14, UGT73DU3 and UGT93N4 enzymes were incubated with solasodine and all required UDP-sugar donors (UDP-Gal, UDP-Rha and UDP-Glc) in a single tube, we observed the accumulation of the final glycosylated SGA product, *α*-solasonine (Figure 3, *in vitro* assays). The *α*-solasonine peak observed here was identical to authentic standard as well as to the one detected in agroinfiltrated leaves of *N. benthamiana* transiently transformed with UGT73L14, UGT73DU3 and UGT93N4 and further supplemented with solasodine substrate (Figure 3, *in planta* assays).

Order of glycosylation in *α*-solamargine biosynthesis

α-solamargine share the same steroidal aglycone, solasodine as that of *α*-solasonine, but glycosylated with chacotriose sugar moieties at C-3 consisting of one D-glucose and two L-rhamnose (Rha I and Rha II) units (Figure 2E, see structure). A biosynthetic proposal based on chemical logic suggests that solasodine is first decorated with glucose to form *γ*-solamargine (solasodine-glucoside). The next biosynthetic step involves attachment of first rhamnose sugar to *γ*-solamargine, and that either Rha I (1→2 linkage, route 1; Supplemental Figure 32) or Rha II (1→4 linkage, route 2; Supplemental Figure 32) can be added at this step to form β_1-solamargine (*γ*-solamargine-Rha I) or β_2-solamargine (*γ*-solamargine-Rha II), respectively. Finally, last step in pathway would be *α*-solamargine formation from either β_1- or β_2- solamargine via second rhamnose addition (Supplemental Figure 32). Therefore, three

UGT enzyme activities (one UDP-glucosyltransferase and two UDP-rhamnosyltransferase) are required to complete the biosynthesis of *α*-solamargine from solasodine aglycone. Our stepwise-pooled approach identified not three, but two UGTs, UGT73DU3 and UGT93M3 that are necessary and sufficient to produce *α*-solamargine in *N. benthamiana* suggesting that additional UGTs are not required in this heterologous host for *α*-solamargine reconstitution (Figure 2D and 2G). However, no conversion of solasodine to *α*-solamargine was observed when recombinant UGT73DU3 and UGT93M3 enzymes were assayed together with appropriate UDP-sugar donors (UDP-Glc and UDP-Rha) (Supplemental Figure 32). Therefore, an additional, third UGT activity is certainly required in the biosynthetic pathway of *α*-solamargine starting from solasodine. These results also clearly indicate that an endogenous *N. benthamiana* UGT enzyme is involved in *α*-solamargine production in addition to the identified UGT73DU3 and UGT93M3 enzymes during *in planta* reconstitution assays.

UGT73DU3 acts as common enzyme in both the *α*-solasonine and *α*-solamargine pathways (Figure 2D, 2F, 2G). In the *α*-solasonine pathway, UGT73DU3 adds the rhamnose sugar (1→2 linkage) to *γ*-solasonine via 1,2-rhamnosyltransferase activity (Figure 3). Thus, we anticipated that UGT73DU3 would exhibit a similar 1,2 rhamnosyltransferase activity in *α*-solamargine biosynthesis by adding the Rha I sugar (1 → 2 linkage) to *γ*-solamargine (Supplemental Figure 32 for proposed pathway). This means that the remaining UGT93M3 enzyme could possibly catalyze either addition of glucose to solasodine (first step) or addition of Rha II sugar (1→4 linkage) on mono- or di- glycosylated solasodine (Supplemental Figure 32). However, UGT93M3 did not exhibit UDP-glucosyltransferase activity when incubated with solasodine and UDP-Glc (Supplemental Figure 32), which led us to believe that UGT93M3 encoded 1,4-rhamnosyltransferase activity. Thus, it appears that UDP-glucosyltransferase activity responsible for conversion of solasodine to *γ*-solamargine, an initial step in the pathway is indeed missing in our analysis. In other words, endogenous *N. benthamiana* UDP-glucosyltransferase is able to utilize solasodine to produce γ-solamargine, which is transformed to *α*-solamargine by further action of

UGT73DU3 and UGT93M3 enzymes in heterologous host *N. benthamiana* (Supplemental Figure 33, see pathway scheme). At this point, despite additional UGT candidates screening, we were unable to identify a glucosyltransferase that glucosylate solasodine to γ-solamargine.

Our *in planta* and *in vitro* assays demonstrated that UGT73L14 catalyzes first step in *α*-solasonine pathway by adding the galactose sugar on solasodine aglycone to form *γ*-solasonine (Figure 3). We next tested whether UGT73L14 could utilize UDP-Glucose as an alternative sugar donor in enzymatic reactions, and indeed recombinant UGT73L14 enzyme generated γ-solamargine (*m/z* 576.3901) in presence of solasodine and UDP-Glc (Supplemental Figure 33, *in vitro* assays). Notably, UGT73L14 only produces γ-solasonine (solasodine-galactoside) and not γ-solamargine (solasodine-glucoside) when expressed transiently in *N. benthamiana* (Supplemental Figure 33, see also Figure 3, *in planta* assays). This suggests that UGT73L14 preferred UDP-Gal as a specific sugar donor over UDP-Glc *in planta*. *In vitro* enzyme kinetics assay experiments also supported this observation where UGT73L14 displayed favored specificity toward UDP-Gal over UDP-Glc (k_{cat}/K_m = 149.91 min^{-1} μM^{-1} for UDP-Gal compared with k_{cat}/K_m = 66.61 min^{-1} μM^{-1} for UDP-Glc) (Supplemental Figure 34). Moreover, accumulation of solasodine glycosides (peak A and B) were observed in control *N. benthamiana* plants (transiently expressing empty vector + solasodine fed) due to endogenous activity of host UGTs (Supplemental Figure 33, see also Figure 3). Interestingly, peak A detected had a same retention time and mass spectrum as that of *γ*-solamargine generated *in vitro* (Supplemental Figure 33). Hence, we confirm the identity of peak A as *γ*-solamargine. On the other hand, peak B detected in control *N. benthamiana* leaves had an actual mass *m/z* 738.4431, consistent with addition of one additional glucose to *γ*-solamargine (Supplemental Figure 35A). Thus, peak B was putatively assigned as *γ*-solamargine-glucoside based on MS/MS fragmentation analysis (Supplemental Figure 35A). We speculate that another endogenous UGT enzyme from *N. benthamiana* is involved in the glycosylation of *γ*-solamargine to form *γ*-solamargine-glucoside (Supplemental Figure 35B).

Addition of first rhamnose sugar, either Rha I or Rha II to γ-solamargine is the next step in the biosynthetic pathway. Hypothetically, either UGT73DU3 (1,2 rhamnosyltransferase) or UGT93M3 (1,4 rhamnosyltransferase) enzymes can catalyze this step. To clarify this biosynthetic step, *in vitro* activity assays were set up using recombinant UGT73L14, UGT73DU3 and UGT93M3 enzymes. Incubation of UGT73L14 and UGT73DU3 enzymes with solasodine as a substrate in presence of UDP-Glc and UDP-Rha resulted in a di-glycosylated product (*m/z* 722.4484), β_1-solamargine (γ-solamargine-Rha I) with a mass consistent with addition of rhamnose (Supplemental Figure 33, *in vitro* assays**)**. Notably, combination of UGT73L14 and UGT93M3 enzymes also generated a new product with the same mass (*m/z* 722.4484), β_2-solamargine (γ-solamargine-Rha II) corresponding to addition of rhamnose, but at different retention time (Supplemental Figure 33, *in vitro* assays). The same β_1-solamargine and β_2-solamargine metabolites (with similar retention time and mass spectrum) were produced with concomitant reduction in γ-solamargine and its downstream derivative γ-solamargine-glucoside levels, by infiltrating solasodine into *N. benthamiana* leaves that transiently expressed UGT73DU3 and UGT93M3, respectively (Supplemental Figure 33, see *in planta* assays and Supplemental Figure 35C). Thus, both UGT73DU3 and UGT93M3 enzymes demonstrate the ability to utilize γ-solamargine as a substrate to produce corresponding rhamnosylated derivatives. However, we clearly detected α-solamargine production when UGT73DU3 and UGT93M3 were co-expressed together in *N. benthamiana* leaves (supplemented with solasodine substrate) (Supplemental Figure 33, *in planta* assays) as also shown in earlier sections (Figure 2F), suggesting that either of the enzyme is capable of accepting the reaction product generated by other as a substrate in the biosynthetic pathway. This is well supported by consumption of both β_1-solamargine and β_2-solamargine in transiently co-expressing UGT73DU3 and UGT93M3 as compared to the levels produced when expressed them individually in *N. benthamiana* (Supplemental Figure 36). To accomplish *in vitro* reconstruction, we added three recombinant UGT enzymes, UGT73L14, UGT73DU3 and UGT93M3 along with solasodine and

required UDP-sugar donors (UDP-Glc and UDP-Rha) in a single tube that eventually resulted in accumulation of *α*-solamargine, the expected final SGA product (Supplemental Figure 33, *in vitro* assays).

UGT73L17 catalyzes glucosylation of solasodine to form *γ*-solamargine and completes *α*-solamargine biosynthetic pathway

Although UGT73L14 and *N. benthamiana* endogenous UGT activities served as an alternative for the first biosynthetic step in *α*-solamargine pathway (Supplemental Figure 33), we still continued to search for remaining glucosyltransferase enzyme responsible for glucosylation of solasodine to *γ*-solamargine. After screening many potential candidates in *S. nigrum*, we revisited the SGA biosynthetic pathways in different *Solanum* species and focused on UGT enzymes that are known to decorate steroidal alkaloid substrates. We noticed that the major potato SGAs, *α*-solanine and *α*-chaconine share a similar glycosylation pattern compared to *α*-solasonine and *α*-solamargine, respectively (Supplemental Figure 37). Moreover, three UGT enzymes (SGT1, SGT2 and SGT3; steroidal alkaloid glycosyltransferase) involved in the conversion of solanidine aglycone to either *α*-solanine or *α*-chaconine have been reported (Moehs et al., 1997; McCue et al., 2005; McCue et al., 2006; McCue et al., 2007) (Supplemental Figure 37). Among these UGTs, UDP-glucose:solanidine glucosyltransferase, also known as SGT2, is responsible for glucosylation of solanidine leading to formation of *γ*-chaconine in potato (Supplemental Figure 37). Indeed, a BLAST search using SGT2 as the query against the *S. nigrum* transcriptome resulted in identification of one similar protein, *SnGT-21* (termed here UGT73L16; 81% amino acid sequence identity) that was previously selected as a potential UGT candidate and tested in *N. benthamiana* using a pooled screen approach (Supplemental Figures 29 and 30). As the absence of UGT73L16 (SnGT-21) did not affect the *α*-solamargine production in *N. benthamiana* earlier (Supplemental Figure 30), we did not consider this candidate for further functional characterization. In

light of the sequence similarity to SGT2, we retested this enzyme by *in vitro* and *in planta* assays. Recombinant UGT73L16 (SnGT-21) enzyme did not show any glucosyltransferase activity when incubated with solasodine and UDP-Glc (Supplementary Figure 38). Moreover, α-solamargine levels remain unaffected upon transient expression of UGT73DU3 and UGT93M3 genes into *N. benthamiana* with or without UGT73L16 (Supplementary Figure 38). Interestingly, while analyzing the Sanger sequencing results of UGT73L16 coding regions amplified from *S. nigrum* green fruit cDNA, we noticed a distinct UGT73L16 isoform having a complete open reading frame and sharing 89% amino acid sequence identity to the characterized UGT73L16. We named this newly obtained UGT isoform as UGT73L17 based on UGT nomenclature. We next expressed UGT73L17 in *E. coli* cells and the purified enzyme was assayed for glucosyltransferase activity using solasodine as a substrate. LC-MS analysis of the assay mixtures clearly showed that UGT73L17 exhibited UDP-glucosyltransferase activity and generated γ-solmargine (*m/z* 576.3902) when incubated with solasodine and UDP-Glucose (UDP-Glc) (Figure 4, *in vitro* assays). UGT73L17 showed clear specificity for UPD-Glc, and did not show any activity when UPD-Gal was used in assay reaction. Moreover, higher accumulation of γ-solmargine was observed when UGT73L17 was infiltrated in *N. benthamiana* plants compared to control (transiently expressing empty vector + solasodine) plants (Figure 4, *in planta* assays). This confirms that UGT73L17 catalyzes the first glucosylation step in α-solamargine biosynthesis. Having UGT73L17 activity in hand, we further successfully reconstituted the α-solamargine biosynthetic pathway *in vitro* together with UGT73DU3 and UGT93M3 enzymes (Figure 4, *in vitro* assays). Moreover, transient expression of UGT73L17 together with UGT73DU3 and/or UGT93M3 genes in solasodine fed *N. benthamiana* displayed much higher levels of pathway intermediates (e.g., β_1- and β_2- solamargine) as well as final product α-solamargine, as compared to levels obtained in case of assay combinations lacking UGT73L17 (Figure 4, *in planta* assays).

A novel malonyltransferase activity converts *α*-solamargine to malonyl-solamargine in SGA biosynthetic pathway

α-Solasonine, *α*-solamargine and malonyl-solamargine are the major SGAs present in green unripe berries of *S. nigrum* (Figure 2A). Having uncovered UGT enzymes that are involved in the *α*-solasonine and *α*-solamargine pathways, our next goal was to identify the missing biosynthetic step responsible for malonyl-solamargine formation. Biosynthetic logic suggests that a single malonyl group addition to *α*-solamargine would generate malonyl-solamargine**.** The addition of a malonyl group to the glucose moiety of secondary/specialized metabolites (e.g. anthocyanin, flavonoids etc.) has been well documented (Li et al., 2018; Luo et al., 2007; Taguchi et al., 2010; Ahmed et al., 2017). The decoration of these metabolites via malonylation not only increases their stability and solubility, but can also enhance the biological activity of these compounds, improving the anti-herbivore defense of the producer plants (Li et al., 2018). NMR based structural assignment of the malonyl-solamargine isolated from *S. nigrum* green berries revealed that the malonyl group is not attached to glucose sugar, but instead to C-4 hydroxyl of the Rha II sugar, which is much rarer derivatization pattern (Supplemental Figure 3). Malonyltransferases, members of the versatile BAHD acyltransferase enzyme family, typically catalyze malonylation steps in plant specialized metabolic pathways (D'Auria, 2006; Moghe et al., 2023). Therefore, we identified eight BAHD candidate genes in our generated *S. nigrum* transcriptome that were highly expressed in green berries compared to purple berries (Supplemental Figure 39). Each BAHD candidate was cloned, and tested through transient co-expression in *N. benthamiana* with *α*-solamargine pathway UGTs (UGT73L17, UGT73DU3 and UGT93M3) followed by co-infiltration of solasodine as starting substrate. Transient expression of one BAHD candidate (SnMAT) with UGT73L17, UGT73DU3 and UGT93M3 genes in *N. benthamiana* leaves generated malonyl-solamargine, along with concomitant reduction in the levels of the *α*-solamargine precursor (Figure 5A, left panel). Moreover, infiltration of *α*-solamargine instead of upstream solasodine aglycone as a

substrate into *N. benthamiana* plants expressing only SnMAT clearly resulted in malonyl-solamargine production (Figure 5A, right panel). Finally, recombinant SnMAT protein expressed in *E. coli* efficiently converted *α*-solamargine to malonyl-solamargine in the presence of malonyl-CoA (Figure 5B). Importantly, recombinant SnMAT enzyme did not exhibit any malonyltransferase activity when assayed with *α*-solasonine substrate, an SGA that lacks the rhamnose sugar. We isolated the malonyl-solamargine product generated enzymatically using analytical scale HPLC (Supplemental Figure 40), and confirmed the structural assignment by NMR (Supplemental Table 3 and 4). Altogether, our in vivo and *in vitro* results confirmed that SnMAT catalyzes the regiospecific malonylation on Rha II sugar of *α*-solamargine to produce malonyl-solamargine (see pathway scheme, Figure 5C).

Elucidation of SGA biosynthetic pathway in cultivated eggplant (*S. melongena*)

Up to this point, we uncovered five UGT and one malonyltransferase enzymes that decorate the solasodine aglycone scaffold to establish the biosynthetic route for *α*-solasonine, *α*-solamargine and malonyl-solamargine SGAs, found in green berries of *S. nigrum*, a wild *Solanum* species. This discovery of solasodine tailoring enzymes prompted us to track the SGA biosynthetic pathway in eggplant (*S. melongena*), one of the most important vegetable crop worldwide, and the only cultivated *Solanum* species renowned for accumulating *α*-solasonine and *α*-solamargine SGAs mainly in developing fruits (Lelario et al., 2019; Sanchez-Mata et al., 2010). Analysis of SGA content by LC-MS revealed that young leaves of cultivated eggplant share a similar SGA profile compared to the profile observed in *S. nigrum* green berries, with predominant accumulation of *α*-solasonine, *α*-solamargine and malonyl-solamargine (Figure 6A). We hypothesized that cultivated eggplant should contain homologues of 5 UGTs (UGT73L14, UGT73L17, UGT73DU3, UGT93N4 and UGT93M3) and malonyltransferase (MAT) identified from *S. nigrum* that are required for biosynthesis of *α*-solasonine, *α*-solamargine and malonyl-solamargine, and that these genes would be preferentially expressed in leaves. A BLAST search

against the cultivated eggplant genome was conducted using the characterized UGTs and MAT from *S. nigrum*. A clear homologue for each UGT gene (sharing at least 80% amino acid identity to corresponding *S. nigrum* UGT) was readily identified in the cultivated eggplant genome (Figure 6B). Based on UGT nomenclature, hereafter we refer to them as UGT73L15, UGT73L19, UGT73DU2, UGT93N3 and UGT93M2. This analysis also revealed a single malonyltransferase hit (SmMAT) in cultivated eggplant with 75% amino acid sequence identity to SnMAT (Figure 6B). All five UGTs and SmMAT coding sequences were successfully cloned from cDNA prepared from the leaves of cultivated eggplant.

To test the function of these candidates in eggplant SGA biosynthesis, we used *Agrobacterium*-mediated transient expression in leaves of *N. benthamiana*. We first infiltrated *N. benthamiana* leaves with *Agrobacterium* harboring the respective candidate genes (UGTs or SmMAT), and three days later, infiltrated the same leaves with solasodine substrate. Metabolic profiling of the leaf extracts by LC-MS revealed that transient co-expression of UGT73L15, UGT73DU2 and UGT93N3 led to the production of α-solasonine (Figure 6C), whereas co-expression of UGT73L19, UGT73DU2 and UGT93M2 resulted in α-solamargine formation (Figure 6D), similarly as observed in case of *S. nigrum* UGTs (Figure 2E and 2F). Moreover, we clearly detected malonyl-solamargine in case of leaves infiltrated with malonyltransferase (SmMAT), UGT73L19, UGT73DU2 and UGT93M2, and solasodine substrate (Figure 6E and Figure 6F left panel). Expression of SmMAT alone in *N. benthamiana* leaves followed by infiltration of precursor substrate α-solamargine also generated malonyl-solamargine, the expected SGA product (Figure 6E and Figure 6F right panel). When SmMAT was replaced with its homolog from potato (StMAT, sharing 82% amino acid sequence identity) in *N. benthamiana* assays transiently expressing α-solamargine pathway UGTs either from *S. nigrum* (UGT73L17, UGT73DU3 and UGT93M3) or cultivated eggplant (UGT73L19, UGT73DU2 and UGT93M2), de novo malonyl-solamargine formation was noted (Figure 6G). These results confirm the role of UGT73L15, UGT73L19,

456 UGT73DU2, UGT93N3, UGT93M2 and SmMAT enzymes in the production of main SGAs in
457 cultivated eggplant. The characterized biosynthetic pathways of *α*-solasonine, *α*-solamargine and
458 malonyl-solamargine SGAs from black nightshade (*S. nigrum*) and cultivated eggplant (*S. melongena*)
459 are presented in Figure 7.
460
461 **Generating diverse SGA profiles by combinatorial expression of tailoring enzymes *in planta***
462 The discovery of five UGTs and malonyltransferase from both *S. nigrum* and cultivated eggplant convert
463 solasodine aglycone (*22R, 25R*) to *α*-solasonine, *α*-solamargine and malonyl-solamargine in the
464 heterologous host *N. benthamiana*. It is well known that plant UGTs are highly specific for their UDP-
465 sugar donors, but can act promiscuously on array of acceptor substrates (Louveau and Osbourn, 2017).
466 Therefore, we tested the capacity of these UGTs and the malonyltransferase to decorate
467 dehydrotomatidine and tomatidine, different stereoisomers of steroidal alkaloid aglycone scaffolds (*22S,*
468 *25S*) that are produced in tomato. Both dehydrotomatidine and tomatidine are similar in structure except
469 the presence (dehydrotomatidine) or absence (tomatidine) of double bond at C-5,6 position (Figure 1 for
470 structures**)**. Hence, instead of solasodine, tomatidine was infiltrated as a substrate into *N. benthamiana*
471 leaves transiently co-expressing the appropriate combination of pathway genes leading to specific SGA
472 product formation. We noted that the commercial tomatidine standard contained small amounts of
473 dehydrotomatidine as an impurity. Co-expression of either UGT73L14, UGT73DU3 and UGT93N4 (*S.*
474 *nigrum*) or UGT73L15, UGT73DU2 and UGT93N3 (cultivated eggplant) in *N. benthamiana* leaves
475 resulted in two new glycosylated products; *α*-solamarine, an unsaturated SGA (*m/z* 884.4980,
476 $C_{45}H_{74}NO_{16}$ $[M+H]^+$) and dihydro-*α*-solamarine (*m/z* 886.5132, $C_{45}H_{76}NO_{16}$ $[M+H]^+$), a saturated SGA
477 (Supplemental Figure 41). Furthermore, transient expression of either UGT73L17, UGT73DU3 and
478 UGT93M3 (*S. nigrum*) or UGT73L19, UGT73DU2 and UGT93M2 (cultivated eggplant) in *N.*
479 *benthamiana* leaves supplemented with tomatidine also resulted in two new glycosylated products; first,

480 *β*-solamarine, an unsaturated SGA (*m/z* 868.5038, $C_{45}H_{74}NO_{15}$ $[M+H]^+$) and dihydro-*β*-solamarine (*m/z*
481 870.5193, $C_{45}H_{76}NO_{15}$ $[M+H]^+$), a saturated SGA (Supplemental Figure 42). Addition of
482 malonyltransferase (SnMAT or SmMAT) to UGT73DU3 /UGT93M3 or UGT73DU2/UGT93M2
483 combinations generated a new malonylated product (saturated), malonyl-dihydro-*β*-solamarine (*m/z*
484 956.5186, $C_{48}H_{78}NO_{18}$ $[M+H]^+$) in *N. benthamiana* leaves (Supplemental Figure 43). These SGA
485 metabolites were putatively identified based on accurate mass-derived elemental composition and
486 MS/MS fragmentation analysis. In potato (*S. tuberosum*), earlier studies suggested that the main SGAs,
487 *α*-solanine and *α*-chaconine biosynthesis proceeds through solanidine aglycone (Moehs et al., 1997;
488 McCue et al., 2005; McCue et al., 2006; McCue et al., 2007) (see Supplemental Figure 37 for pathway
489 scheme). Furthermore, a recent study showed that *α*-solamarine and *β*-solamarine are the key
490 glycosylated intermediates in the *α*-solanine and *α*-chaconine biosynthetic pathway that are derived from
491 unsaturated dehydrotomatidine aglycone through the action of yet unreported UGT enzymes (Akiyama
492 et al., 2021) (see Figure 1 for pathway scheme). Notably, transient expression of either *S. nigrum*
493 [UGT73L14, UGT73L17, UGT73DU3, UGT93N4 and UGT93M3] or cultivated eggplant [UGT73L15,
494 UGT73L19, UGT73DU2, UGT93N3 and UGT93M2] UGTs in *N. benthamiana* leaves supplemented
495 with solanidine aglycone did not result in the formation of *α*-solanine or *α*-chaconine or any pathway
496 intermediates (Supplemental Figure 44). This strongly suggests that *S. nigrum* and eggplant UGTs act
497 specifically on spirosolane type aglycones (e.g., dehydrotomatidine, tomatidine) but not on solanidane
498 type ones (e.g., solanidine) to generate potato specific SGAs (e.g., *α*-solamarine and *β*-solamarine) *in*
499 *planta* (Supplemental Figure 44). We did not find the presence of saturated dihydro-*α*-solamarine,
500 dihydro-*β*-solamarine and malonyl-dihydro-*β*-solamarine in any of the SGAs producing *Solanum*
501 species. Thus, our results show that *S. nigrum* and cultivated eggplant decorating enzymes (UGTs and
502 MAT) generate a range of yet unreported, possibly 'new to nature', saturated SGAs (Supplemental
503 Figures 41-43).

In tomato SGA biosynthesis, the main steroidal aglycone tomatidine is suggested to get converted into *α*-tomatine through sequential action of GAME1 (galactosyltransferase), GAME17 and GAME18 (glucosyltransferases) and GAME2 (xylosyltransferase) (Itkin et al., 2011; Itkin et al., 2013). Interestingly, transient expression of GAME1, GAME17, GAME18 and GAME2 in *N. benthamiana* leaves supplemented with tomatidine did not generate *α*-tomatine, but *β1*-tomatine, the penultimate intermediate in the pathway (Supplemental Figure 45). In addition, we were able to see accumulation of proposed pathway intermediates (e.g., *γ*-tomatine) when respective tomato GAME enzymes were infiltrated in sequential manner as predicted earlier by Itkin et al. (2013). Thus, the order of the reactions catalyzed by GAME1, GAME17 and GAME18 seems correct in *α*-tomatine biosynthesis, but the involvement of GAME2 enzyme catalyzing the last step of pathway (conversion of *β1*-tomatine to *α*-tomatine) remains unclear (Supplemental Figure 45). We next tested the capacity of these UGT enzymes (GAME1, GAME17 and GAME18) to glycosylate various steroidal aglycones produced mainly by *S. nigrum*/eggplant (e.g., solasodine) and potato (e.g., solanidine). Indeed, transient co-expression of tomato GAME UGTs (GAME1, GAME18, GAME17) in *N. benthamiana* leaves supplemented with solasodine or solanidine resulted in the formation of new glycosylated SGA products that we assigned as γ-solasonine-di-glucoside or γ-solanine-di-glucoside, respectively (Supplemental Figure 45). Moreover, co-infiltration of respective tomato GAME enzymes according to pathway reaction order generated glycosylated solasodine and solanidine SGA derivatives (Supplemental Figure 45) in *N. benthamiana* infiltrated leaves. These results show that unlike *S. nigrum* and eggplant, tomato GAME UGTs act on both spirosolane (e.g. solasodine, tomatidine) and solanidane (e.g. solanidine) type steroidal aglycones to generate diverse SGAs profile.

Potato steroidal alkaloid glycosyltransferases (SGTs) also act on both spirosolane- and solanidane type-aglycones

We continued our combinatorial expression approach in *N. benthamiana* by assaying the activity of potato SGTs against different steroidal aglycones. As mentioned earlier, *α*-solanine and *α*-chaconine, both derived from the solanidine aglycone, are the major SGAs in potato. To date, three glycosyltransferases, also known as SGTs have been identified in the biosynthesis of these SGAs (Supplemental Figure 37). SGT1, a UDP-galactose:solanidine galactosyltransferase is responsible for the first step of *α*-solanine biosynthesis and attaches the initial galactose to the solanidine aglycone (McCue et al., 2005). SGT2, a UDP-glucose:solanidine glucosyltransferase catalyzes the first step in *α*-chaconine pathway by adding a glucose moiety to solanidine (McCue et al., 2006). Finally, SGT3, a UDP-rhamnose:β-steroidal glycoalkaloid rhamnosyltransferase acts as a common enzyme and catalyzes the terminal step of both *α*-chaconine and *α*-solanine biosynthesis (McCue et al., 2007). Notably, two biosynthetic UGT enzymes, one glucosyltransferase and one rhamnosyltransferase are still missing in the *α*-solanine and *α*-chaconine pathways, respectively (Supplemental Figure 37). This was further supported in our *in planta* assays where no *α*-solanine or *α*-chaconine formation was observed when SGT1/SGT3 or SGT2/SGT3 gene combinations were transiently expressed in *N. benthamiana* (supplemented with solanidine) (Supplemental Figures 46A and 47A). Having knowledge of similar glycosylation pattern profiles between *S. nigrum* and potato SGAs and complete set of UGTs from *S. nigrum* in hand, we attempted to reconstruct potato SGA biosynthesis *in planta* by using the *S. nigrum* UGT93N4 (glucosyltransferase) and UGT93M3 (rhamnosyltransferase) genes together with known potato SGT genes. Metabolic profiling of the leaf extracts by LC-MS revealed that transient co-expression of potato SGT1 and SGT3 with UGT93N4 (*S. nigrum*) led to the production of *α*-solanine (Supplemental Figure 46A), whereas co-expression of SGT2, SGT3 and UGT93M2 (*S. nigrum*) resulted in *α*-chaconine formation (Supplemental Figure 47A). Furthermore, transient expression of either SGT1, SGT3 and UGT93N4 or SGT2, SGT3 and UGT93M2 in *N. benthamiana* leaves supplemented with solasodine aglycone clearly generated *α*-solasonine (Supplemental Figure 46B) or *α*-solamargine

(Supplemental Figure 47B), respectively. When tomatidine was infiltrated as a substrate in the *N. benthamiana* leaves expressing above gene combinations, 'new to nature', saturated dihydro-*α*-solamarine and dihydro-*β*-solamarine SGAs formation was observed (Supplemental Figures 46C and 47C). Thus, like tomato GAME UGTs, potato SGTs are also able to act on spirosolane- and solanidane type steroidal aglycones.

Phylogenetic analysis of UGTs involved in SGA biosynthetic pathways

The diverse glycosylation capacity of UGTs involved in SGA biosynthesis is also reflected in the phylogenetic analysis (Supplemental Figure 48). Based on the type of UGT activity, *S. nigrum* and eggplant UGTs are clearly separated into five clades. UGT73L14 (*S. nigrum*), UGT73L15 (eggplant), GAME1 (tomato) and SGT1 (potato) exhibiting UDP-galactosyltransferase activity are present in clade I (Supplemental Figure 48). All of these UGTs act on spirosolane type (e.g., solasodine, tomatidine) steroidal aglycones, but only tomato GAME1 and potato SGT1 act on solanidane type (e.g., solanidine) aglycones (Supplemental Figures 45, 46 and 49). Despite sharing 80 to 86 % homology (amino acid level) with potato SGT1 (Supplemental Figure 50), UGT73L14 (*S. nigrum*) and UGT73L15 (eggplant) showed minimal or no activity on solanidine when tested transiently in *N. benthamiana* (*in planta*) assays (Supplemental Figure 49). This was also the case with clade 2 proteins that consist of homologous UGT73L17 (*S. nigrum*), UGT73L19 (eggplant) and SGT2 (potato) displaying UDP-glucosyltransferase activity on solasodine and tomatidine aglycones (Supplemental Figures 47, 48 and 50). Unlike SGT2, both UGT73L17 (*S. nigrum*) and UGT73L19 (eggplant) did not show any glucosyltransferase activity on solanidine infiltrated *N. benthamiana* leaves (Supplemental Figure 49). The inability of *S. nigrum* and eggplant UGT73L subfamily enzymes to utilize solanidine aglycone explains the absence of *α*-solanine and *α*-chaconine formation in our *in planta* reconstitution experiments (Supplemental Figure 44). UGT73DU (clade 3), UGT93N (clade 4) and UGT93M (clade 5) proteins from *S. nigrum* and

575 eggplant possessing rhamnosyltransferase (1→2), glucosyltransferase and rhamnosyltransferase (1→ 4)
576 activities in SGA biosynthesis form distinct clades in the phylogeny (Supplemental Figure 48). Notably,
577 there are clear orthologues of UGT93N (clade 4) and UGT93M (clade 5) in potato sharing 80 to 85%
578 homology with *S. nigrum* and eggplant UGT proteins (Supplemental Figure 50). As mentioned earlier,
579 the potato SGA biosynthetic pathway is still not completely elucidated and requires one
580 rhamnosyltransferase and glucosyltransferase that remained to be identified till date. Notably, addition
581 of *S. nigrum* UGT93N4 or UGT93M3 enzymes to potato SGT combinations resulted in heterologous
582 production of potato specific SGAs *in planta* (Supplemental Figures 46 and 47). Therefore, potato
583 UGT93N and UGT93M orthologues could be promising candidate enzymes for catalyzing the missing
584 steps in potato SGA biosynthetic pathway.
585
586 In summary, using a combination of metabolomics, transcriptomics, chemical logic, *in planta* and *in*
587 *vitro* enzyme characterization, we have been able to establish a complete biosynthetic pathway of *α*-
588 solasonine, *α*-solamargine and malonyl-solamargine, plant-derived bioactive steroidal molecules with
589 potential physiological functions in wild *S. nigrum* and cultivated *S. melongena* plants (Figure 7). Our
590 discoveries expand the understanding of how solasodine and its downstream diverse SGAs are produced
591 in hundreds of *Solanum* plant species. The successful reconstitution of classical SGA biosynthetic
592 pathways and generation of 'new to nature SGAs' via combinatorial expression in heterologous host
593 such as *N. benthamiana* and *E. coli* provides the feasibility to produce bioactive SGAs in sustainable
594 hosts through metabolic engineering and synthetic biology applications.
595
596
597
598

Methods

Plant materials

Black nightshade (*S. nigrum* Linn.), tomato (*S. lycopersicum cv.* Micro Tom) cultivated potato (*S. tuberosum* cv. Annabelle) and cultivated eggplant (*S. melongena* cv. DR2) plants were grown in a climate-controlled greenhouse at 24 °C during the day and 18 °C during night, with natural light. *Nicotiana benthamiana* plants were grown in a growth room maintained at 23 ± 2 °C with 16-h day/8-h night light regime.

Analytical standards

Analytical standards including solanidine, tomatidine (contains dehydrotomatidine as impurity), solasodine, α-solasonine, α-solamargine were purchased from Sigma-Aldrich Chemie GmbH, Germany and unless stated otherwise, were dissolved in methanol to a concentration of 1 mg ml^{-1}. UDP-Glc and UDP-Gal sugar donors were purchased from Sigma-Aldrich, while UDP-Rha was purchased from PeptaNova GmbH, Germany.

Targeted profiling of SGAs by LC-MS

Preparation of extracts and SGAs profiling of *S. nigrum* (unripe green and ripe purple berries) and *S. melongena* (young leaves and roots) tissues were much performed as described earlier (Sonawane et al., 2018; Sonawane et al., 2023). Briefly, 100 mg of frozen powder tissue was extracted with 80% methanol and 0.1% formic acid, vortexed for 1 min, then sonicated for 15 min at room temperature. Finally, the extracts were centrifuged for 20 min at 20,000 × *g* and filtered through 0.22 μm filters. Except pooled-screen approach experiment, LC-MS analysis for all other sample sets was done as described here: samples were analyzed using a Thermo Scientific UltiMate 3000 RS ultra-high performance liquid chromatography (UHPLC) system coupled to an Impact II UHR-Q-ToF (Ultra-High Resolution Quadrupole-Time-of-Flight) mass spectrometer (Bruker Daltonics) with the standard (43 min., positive mode) run conditions as follows: 5% B for 1 min; 5% B to 28% B in 22 min; then changing to 100% B

in 14 min and further at 100% B for 3 min, and finally returned to the initial conditions (5% phase B) within 0.5 min. The column was equilibrated with 5% B for another 2.5 min before next injection. Separation of metabolites was performed on an Acquity Premier BEH VanGuard FIT C18 column (100 mm x 2.1 mm, 1.7 µm particles, 130 Å) by Waters (Milford, Massachusetts, United States)The mobile phase consisted of 0.1% formic acid in water (phase A) and acetonitrile (phase B). The flow rate was 0.3 ml min^{-1}, and the column temperature was kept at 35°C. Mass spectrometry was performed in positive electrospray ionization mode (capillary voltage = 3500 V; end plate offset = 500 V; nebulizer pressure = 2.0 bar; drying gas: nitrogen at 250 °C and 10 L min^{-1}). Mass spectrometry data was recorded at 12 Hz ranging from 100 to 1300 *m/z* in auto MS/MS mode with an active exclusion window of 0.2 min. Fragmentation was triggered on an absolute threshold of 400 counts and restricted to a total cycle time range of 0.5 s, with dynamic collision energy (20-60 eV). To calibrate MS spectrum recording, each run was initiated with the direct source infusion of a sodium formate-isopropanol calibration solution (using external syringe pump at 0.18 ml h^{-1}). The initial 1 min of the chromatographic gradient was directed towards the waste. SGAs were identified by comparing the retention times and mass spectra of authentic standards analyzed on the same instrument. When the corresponding standards were not available, metabolites were putatively identified based on accurate mass derived elemental composition and MS/MS fragmentation pattern. For pooled screen experiment, samples were analyzed using an UltiMate 3000 RS UHPLC system (Thermo Fisher Scientific, Germering, Germany) coupled to an EVOQ Elite™ Triple Quadrupole mass spectrometer (Bruker Daltonics, Bremen, Germany). Separation of analytes was performed on an Acquity Premier BEH VanGuard FIT C18 column (100 mm x 2.1 mm, 1.7 µm particles, 130 Å) by Waters (Milford, Massachusetts, United States). Water containing 0.1% formic acid and acetonitrile were used as mobile phases A and B, respectively, with a flow rate of 0.3 ml min^{-1}. The column temperature was kept at 35°C. The following gradient was used for analysis: 5 % B for 1 min; 5 % to 28 % B in 22 min; 28 % to 100 % B in 14 min, 100 % B for 3 min, returning to the

initial conditions (5 % phase B) within 0.5 min and a final re-equilibration to 5 % B for another 2.5 min before next injection. The mass spectrometer was operated in positive ionization mode. The EVOQ source parameters were as follows: heated ESI spray voltage (+) 4000 V; cone gas flow 20 arbitrary units at 350°C; probe gas flow 45 arbitrary units at 400°C; nebulizer gas flow 55 arbitrary units; and exhaust gas on. The analysis was performed in selected ion monitoring (SIM), operating the Q1 mass analyzer under unit resolution (0.7 Da FWHM). In each analytical run, solasodine *m/z* 414.30; *α*-solasonine *m/z* 884.50; *α*-solamargine *m/z* 868.50 and malonyl-solamargine *m/z* 954.50 were recorded. The EVOQ SIM chromatograms were analysed using Data Review version 8.2.1 of the MS workstation software (Bruker Daltonics, Bremen, Germany).

Transcriptome analysis

Total RNA from unripe green and purple ripe berries of *S. nigrum* was extracted using the RNeasy Mini Kit (Qiagen). Samples were submitted to BGI (https://www.bgi.com/) for preparing mRNA libraries and further RNA-seq analysis (PE 2x150, ~40M reads per sample) according to the company's standard protocols. De Novo transcriptome assemblies were generated from cleaned, trimmed reads using Trinity (Grabherr et al., 2011). Transdecoder (https://github.com/TransDecoder/TransDecoder) was used to identify candidate-coding regions within transcript sequences. Functional annotation was then performed with seven functional databases (NR, NT, GO, KOG, KEGG, SwissProt and InterPro). Gene expression measured by fragments per kilobase of transcript per million mapped reads was calculated using RSEM (Li and Dewey, 2011).

A. tumefaciens* mediated transient expression in *N. benthamiana

Genes of interest (e.g., GAMEs, UGTs, and MATs) were amplified by PCR using cDNA prepared from RNA isolated from either green berries (*S. nigrum*) or young leaves (*S. melongena*). SGT (SGT1-3) and malonyltransferase genes from cultivated potato (StMAT) was amplified using potato leaf cDNA template, while tomato GAME UGTs (GAME1, GAME17 and GAME18) coding sequences were

amplified using tomato leaf cDNA. Nucleotide and amino acid sequences of all GAMEs, UGTs and MATs characterized (from *S. nigrum*, eggplant, tomato and potato) in this study are provided in Supplemental Data 1. The resulting amplicons were cloned into binary 3Ω1 destination vector using Goldenbraid cloning (Sarrion-Perdigones et al., 2013) and transformed into *A. tumefaciens* (GV3101) by electroporation. It is important to mention that *S. nigrum* UGT73L16 and UGT73L17 gene sequences were obtained and confirmed by Sanger sequencing of several 3Ω1 clones. Single colonies with each target construct were inoculated in liquid LB medium supplemented with antibiotics (200 μg ml^{-1} spectinomycin, 50 μg ml^{-1} gentamicin and 25 μg ml^{-1} rifampicin) and incubated overnight at 28 °C with shaking (200 rpm). Overnight grown cultures were centrifuged at 2,000 × *g* for 20 min and cell pellets were washed once with 5 ml of infiltration buffer [50 mM MES buffer (pH 5.6), 10 mM $MgCl_2$, 150 μM acetosyringone]. Finally, each pellet was resuspended in 10 ml of infiltration buffer and incubated at room temperature for 2 h. For combinatorial infiltrations, optical density (OD_{600}) for each strain was set at 0.2. *Agrobacterium* suspensions were infiltrated into 4–6-week-old *N. benthamiana* leaves. After 5 days, infiltrated leaves were harvested for further LC-MS based SGA analysis. In case of exogenous substrate feeding, steroidal aglycones (solasodine or tomatidine or solanidine) or SGA (*α*-solamargine) substrates (20 μg ml^{-1}) were injected to the infiltrated leaves 3 days post infiltration. After 48 h, leaves were collected for further metabolites analysis. Biological replicates consisted of several leaves collected from different infiltrated plants. Leaves infiltrated with empty vector was used as control. Sample preparation (except for the ones that produce solasodine, Figure 2C) and LC-MS analysis for SGAs was carried out on UHPLC-qTOF-MS as described above. Due to non-polar nature of the solasodine steroidal aglycone, its recovery from *N. benthamiana* leaf samples (infiltrated with six GAME genes) using methanolic extraction method is rather poor. Therefore, we adopted different extraction method for these samples as follows: powdered leaf tissues (100 mg) was first saponified at 70 °C for 2 h in 0.6 ml of 20% KOH (w/v) in 50% ethanol. Samples were mixed every 30 min during this procedure. Upon cooling

695 to room temperature, samples were extracted three times with 0.5 ml of hexane. The combined hexane
696 phases were evaporated to dryness using a gentle stream of nitrogen and resuspended in 150 µl of
697 ethanol, and injected on LC-Triple Quadrupole (TQ)-MS and analyzed using SIM mode as described
698 above.

699 **Heterologous expression in *E. coli* and *in vitro* enzyme assays**

700 UGT candidates, UGT73L14, UGT73L16, UGT73L17, UGT73DU3, UGT93M3, UGT93N4 and a
701 malonyltransferase (SnMAT) genes from *S. nigrum* were cloned separately into the pOPINF (*HindIII* /
702 *KpnI* digested) vector and expressed in *E. coli* BL21 (DE3) cells. Briefly, single colonies with target
703 gene construct were grown in LB medium at 37 °C, 250 rpm, overnight. 1% of the seed culture was used
704 to inoculate 100 mL 2 x YT medium with ampicillin (100 µg ml^{-1}) and the cultures were grown further
705 at 37°C, 250 rpm shaking. When cultures reached OD_{600} = 0.6, protein expression was induced with 200
706 µM of IPTG at 15 °C, for 20 h. Bacterial cells were harvested by centrifugation (3000 x g, 4°C, 20 min),
707 resuspended in 10 mL lysis buffer (50 mM TRIS-HCl, 50 mM glycine, 5% v/v glycerol, 0.5 M NaCl, 20
708 mM imidazole, pH 8) with 0.2 g L^{-1} lysozyme and SIGMAFAST protease inhibitor tablet (Sigma-
709 Aldrich), and lysed by sonication for 4 min (3 s on, 2 s off cycle) on ice. Cell debris were removed by
710 centrifugation at 35,000 × g at 4°C for 20 min and each soluble protein was purified using Ni-NTA
711 agarose beads (Qiagen) according to the manufacturer's instructions. Finally, purified proteins were
712 eluted with 250 mM imidazole in buffer containing 50 mM NaH_2PO_4 (pH 7.5) and 150 mM NaCl.
713 Standard UGT *in vitro* enzyme assay was performed in 100 µL reaction mix containing 5µg solasodine
714 substrate, 100 mM potassium phosphate buffer (pH 7.5), UDP-sugar donor (2 mM) and purified UGT
715 enzyme (2 µg). The reaction mixture was incubated at 30 °C for 3 h. The reaction was stopped by the
716 addition of 250 µL methanol containing 0.1% formic acid. After centrifugation for 20 min at 20,000 × *g*
717 and filtering through 0.22 µm filters, analysis of enzyme assay products (glycosylated SGAs) was
718 performed on UHPLC-qTOF-MS using standard (43 min.) run conditions as above. While performing

719 the coupled *in vitro* assays, 2 µg of each purified enzyme and an appropriate UDP-sugar donor was
720 included in the reaction mixtures. The recombinant SnMAT (malonyltransferase) *in vitro* assay was
721 performed in 50 mM potassium phosphate buffer (pH 7.5) containing 5µg *α*-solamargine substrate,
722 malonyl CoA (200 µM) and purified SmMAT protein (5 µg). The reaction mixture was incubated at 30
723 °C for 2 h. After incubation, the reaction was mixed with 250 µL methanol containing 0.1% formic acid,
724 extracted, and analyzed by UHPLC-qTOF-MS, as described above. Protein extracts obtained from empty
725 pOPINF vector-transformed *E. coli* BL21 (DE3) cells were used in control reactions.

726 For kinetic analysis of UGT73L14, we tested two different UDP-sugar donors (UDP-Gal and UDP-Glc).
727 Briefly, 120 µM solasodine was used as the acceptor substrate, with variable concentrations of UDP-
728 sugar donors (0-1500 µM). Each reaction (100 µl) included 5 µl of recombinant UGT73L14 (0.6 µg
729 μl^{-1}) and other standard assay components as described above. The reactions were run at 30 ^{0}C, and
730 stopped by the addition of two volumes of methanol with 0.1% formic acid. Samples were centrifuged
731 for 15 minutes at 20,000 × *g*, supernatant was placed in a LC vial. All reactions were done in triplicate
732 and assay products were analyzed by LC-MS using standard (43 min.) gradient. Kinetics parameters
733 (K_m, k_{cat}, V_{max}) were calculated using the non-linear regression model in GraphPad Prism 8.0 software.

734 **Isolation of Malonyl-solamargine from unripe green berries of *S. nigrum*.**

735 30 g of green berries were collected and used as starting material to isolate malonyl-solamargine. Briefly,
736 berries were ground in liquid nitrogen to a fine powder and extracted with 50 ml of methanol at room
737 temperature for 3 hours with constant stirring. The resulting methanol extract was first filtered through
738 miracloth (Merk Millipore) and then through a filter paper (MN 615, Macherey-Nagel). The filtered
739 methanolic extract was evaporated to dryness using a rotary evaporator. The dried extract was
740 reconstituted in 10 ml of 50% methanol (aqueous) and applied on to a conditioned Chromabond-HLB
741 SPE (solid phase extraction) cartridge (60 µm, 6 mL 200 mg, Macherey-Nagel) for further processing.
742 The SPE cartridge was first conditioned by the addition of 6 ml of methanol followed by 6 ml of water.

743 The extract (10 ml) was then applied on to the SPE cartridge, washed three times with 6 ml of 30%
744 methanol (aqueous) and cartridge was dried using vacuum. Finally, the extract was eluted with 6 ml of
745 methanol and evaporated to dryness for analytical-HPLC separation.

746 **Enzymatic workup for the production of malonyl-solamargine from recombinant SnMAT.**

747 *E. coli* expressed and purified recombinant SnMAT enzyme was used for scaling up malonyl-
748 solamargine production. A 5 ml one-pot reaction was set up at final concentrations of 0.2 mg ml^{-1} enzyme
749 (SnMAT), 0.25 mM α-solamargine and 0.4 mM malonyl-CoA in 100 mM TRIS buffer at pH 7.5 and
750 incubated 16 hours at 30 °C with gentle stirring. The reaction was quenched with the addition of 5 ml
751 methanol and vortexed vigorously. The quenched reaction was filtered through a 0.22 µm PTFE syringe
752 filter and the filtrate was evaporated to dryness using a rotary evaporator. The dried enzymatic workup
753 was reconstituted in 50% methanol (aqueous) and processed through Chromabond-HLB SPE cartridge
754 as described above.

755 **Analytical HPLC method for the purification of malonyl-solamargine.**

756 Malonyl-solamargine from green berries and enzymatic reaction were purified using high-performance
757 liquid chromatography (HPLC). An Agilent infinity II 1260 HPLC instrument paired with an auto
758 sampler, diode array detector (DAD), and analytical-scale fraction collector was used for compound
759 detection and isolation. Reversed-phase (C18) chromatography was performed using a Phenomenex
760 Kinetex XB-C18 column (5.0 µm, 100 Å 100 x 2.1 mm) with water + 0.1% formic acid (phase A) and
761 ACN (phase B) as mobile phases. A flow rate of 0.5 ml min^{-1} was used while the column was maintained
762 at 35 °C. Chromatographic separation was performed at 15% B for 4 min, followed by a linear gradient
763 from 15% to 50% B in 16 min, 90% B for 5 min, 15% B for 8 min. In case of green berries extract, the
764 sample was diluted to 1 mg ml^{-1} with 50% methanol and filtered using a 0.22 µm PTFE syringe filter.
765 The diluted extract was placed in the auto sampler and 20 uL injections were performed and fractions
766 were collected by monitoring the UV 208 nm corresponding to Malonyl-solamargine absorbance.

Fractions were tested by LC-MS to confirm identity of malonyl-solamargine and verified factions were pooled for evaporation. Evaporation was performed using a rotary evaporator to dryness and the sample was submitted for NMR analysis.

NMR methods

NMR measurements were carried out on a 500 MHz Bruker Avance III HD spectrometer and a 700 MHz Bruker Avance III HD spectrometer (Bruker Biospin GmbH, Rheinstetten, Germany), equipped with a TCI cryoprobe using standard pulse sequences as implemented in Bruker Topspin ver. 3.6.1. (Bruker Biospin GmbH, Rheinstetten, Germany). Chemical shifts were referenced to the residual solvent signals of pyridine-d_5 (δ_H 8.74/δ_C 150.35). The assignment of α-solamargine was carried out by 1D and 2D NMR spectroscopy (^{1}H, DEPTQ, COSY, HSQC, HMBC, HSQC-TOCSY, and ROESY; Supplemental Figures 13-24). The stereochemistry of C22 and C25 were confirmed as 22α and 25*R* from a comparison with the literature (Gu et al., 2018; Yamashita et al., 1990). The aglycone spectra of malonyl-solamargine (isolated from green berries, GB) and that of malonyl-solamargine (from enzyme assay, EA) were identical with the aglycone spectrum of α-solamargine (Supplemental Table 1 and 3). The HMBC spectra of malonyl-solamargine (GB) suggested that the malonyl-group was attached to Rha II at position 4.

Phylogenetic analysis

UGT sequences from tomato, potato, *S. nigrum* and eggplant were obtained either by PCR cloning and Sanger sequencing (in this study) or obtained from NCBI and public databases (from Sol Genomics Network server). The Maximum Likelihood tree was inferred in MEGAX (Kumar et al., 2018) using 1000 bootstrap replications with the following parameters: Poisson model, discrete gamma distribution (five categories), and partial deletion. Evolutionary distances are in units of a number of amino acid substitutions per site. All UGT amino acid sequences used in the phylogenetic analysis are provided in Supplemental Data 1.

References:

Ahmed, M., Li, P., Wang, J., Rehman, N. U. & Zhao J. (2017). Isoflavone malonyltransferases GmIMat1 and GmIMat3 differently modify isoflavone glucosides in soybean (*Glycine max*) under various stresses. Front. Plant Sci. **8:**735.

Akiyama, R. *et al.* (2019). Characterization of steroid 5α-reductase involved in α-tomatine biosynthesis in tomatoes. Plant Biotechnol. (Tokyo) **36:**253-263.

Akiyama, R. *et al.* (2021). The biosynthetic pathway of potato solanidanes diverged from that of spirosolanes due to evolution of a dioxygenase. Nat. Commun. **12:**1300.

Bednarz, H., Roloff, N. & Niehaus, K. (2019). Mass spectrometry imaging of the spatial and temporal localization of alkaloids in nightshades. J. Agric. Food Chem. **67:**13470-13477.

Cárdenas, P. D. *et al.* (2015). The bitter side of the nightshades: Genomics drives discovery in Solanaceae steroidal alkaloid metabolism. Phytochemistry **113:**24-32.

Chen, X., Dai, X., Liu, Y., Yang, Y., Yuan, L., He, X. & Gong, G. (2022). *Solanum nigrum* Linn.: An Insight into Current Research on Traditional Uses, Phytochemistry, and Pharmacology. Front. Pharmacol. **13:**918071.

Christ, B. *et al.* (2019). Repeated evolution of cytochrome P450-mediated spiroketal steroid biosynthesis in plants. Nat. Commun. **10:**3206.

D'Auria, J. (2006). Acyltransferases in plants: a good time to be BAHD. Curr Opin. Plant. Biol. **9:**331-340.

Delbrouck, J. A., Desgagne, M., Comeau, C., Bouarab, K., Malouin F. & Boudreault Pierre-Luc. (2023). The therapeutic value of *Solanum* steroidal (glyco)alkaloids: a 10-year comprehensive review. Molecules **28:**4957.

Ding, X., Zhu, F. S., Li, G. Y., & Gao, S. G. (2012). Purification, antitumour and induction of apoptosis in human hepatoma SMMC-7721 cells by solamargine from *Solanum nigrum* L. Food Chem. **139:**599-604.

Ding, X., Zhu, F., Yang, Y. & Li, M. (2013). Purification, antitumour activity in vitro of steroidal glycoalkaloids from black nightshade (*Solanum nigrum* L.) Food Chem. **141:**1181-1186.

Eich, E. (2008). Solanaceae and convolvulaceae–specialized metabolites: biosynthesis chemotaxonomy biological and economic significance: a handbook (Springer, Berlin).

Friedman, M. (2002). Tomato glycoalkaloids: Role in the plant and in the diet. J. Agric. Food Chem. **50:** 5751-5780.

Friedman, M. (2006). Potato glycoalkaloids and metabolites: Roles in the plant and in the diet. J. Agric. Food Chem. **54:**8655-8681.

Friedman, M. (2015). Chemistry and anticarcinogenic mechanisms of glycoalkaloids produced by eggplants, potatoes, and tomatoes. J. Agric. Food Chem. **63:**3323-3337.

Grabherr, M. G. *et al.* (2011). A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. **29:**644-652.

Gu, Xin-Yue *et al.* (2018). Bioactive steroidal alkaloids from the fruits of *Solanum nigrum*. Phytochemistry **147:**125-131.

Itkin, M. *et al.* (2011). GLYCOALKALOID METABOLISM1 is required for steroidal alkaloid glycosylation and prevention of phytotoxicity in tomato. Plant Cell **23:**4507-4525.

Itkin, M. *et al.* (2013). Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. Science **341:**175-179.

Jabamalairaj, A., Priatama, R.A., Heo, J. & Park S. J. (2019). Medicinal metabolites with common biosynthetic pathways in *Solanum nigrum*. Plant Biotechnol. Rep. **13:**315-327.

Kalalinia, F. & Karimi-Sani, I. (2017). Anticancer properties of solamargine: A systemic review. Phytother. Res. **31:**858-870.

Kumar, S., Stecher, G., Li, M., Knayz, K. & Tamura, K. (2018). MEGAX: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. **35:**1547-1549.

Lee, H. *et al*. (2019). Identification of a 3β-hydroxysteroid dehydrogenase/ 3-ketosteroid reductase involved in α-tomatine biosynthesis in tomato. Plant Cell Physiol. **6:**1304-1315.

Li, B. & Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics **12:**323.

Li, J. *et al.* (2018). The decoration of specialized metabolites influences stylar development. eLife **7:** e38611.

Liang, X. *et al.* (2022). Solasonine inhibits panvreatic cancer progression with involvement of ferroptosis induction. Front. Oncol. **12:**834729.

Lelario, F., Maria, S., Rivelli, A., Russo, D., Milella, L., Bufo, S. & Scrano, L. (2019). A complete survey of glycoalkaloids using LC-FTICR-MS and IRMPD in a commercial variety and a local landrace

of eggplant (*Solanum melongena* L.) and their anticholinesterase and antioxidant activities. Toxins **11:**230.

Louveau, T. & Osbourn, A. (2019). The sweet side of plant-specialized metabolism. Cold Spring Harb. Perspect. Biol. **11:**a034744.

Luo, J. *et al.* (2007). Convergent evolution in the BAHD family of acyl transferases: identification and characterization of anthocyanin acyl transferases from *Arabidopsis thaliana*. Plant J. **50:**678-695.

McCue, K.F. *et al.* (2005). Metabolic compensation of steroidal glycoalkaloid biosynthesis in transgenic potato tubers: using reverse genetics to confirm the in vivo enzyme function of a steroidal alkaloid galactosyltransferase. Plant Sci. **168:**267-273

McCue, K.F. *et al.* (2006). The primary in vivo steroidal alkaloid glucosyltransferase from potato. Phytochemistry **67:**1590-1597.

McCue, K.F. *et al.* (2007). Potato glycosterol rhamnosyltransferase, the terminal step in triose side-chain biosynthesis. Phytochemistry **68:**327334.

Milner, S. E. *et al.* (2011). Bioactivities of glycoalkaloids and their aglycones from *Solanum* species. J. Agric. Food Chem. **59:**3454-3484.

Moehs, C.P., Allen, P.V., Friedman, M. & Belknap, W.R. (1997). Cloning and expression of solanidine UDP-glucose glucosyltransferase from potato. Plant J. **11:** 227-236.

Moghe, G., Kruse, L.H., Petersen, M., Scossa, F., Fernie, A. R., Gaquerel, E. & D'Auria, J. C. (2023). BAHD company: the ever-expanding roles of the BAHD acyltransferase gene family in plants. Annu. Rev. Plant Biol. **74:**165-194.

Nakayasu, M. *et al.* (2017). A dioxygenase catalyzes steroid 16α-hydroxylation in steroidal glycoalkaloid biosynthesis. Plant Physiol. **175:**120–133.

Nakayasu, M. *et al.* (2021). Characterization of C-26 aminotransferase, indispensable for steroidal glycoalkaloid biosynthesis. Plant J. **108:**81–92.

Sanchez-Mata, M.C., Yokoyama, W.E., Hong, Y.J. & Prohens, J. (2010). α-solasonine and α-solamargine contents of gboma (*Solanum macrocarpon* L.) and scarlet (*Solanum aethiopicum* L.) eggplants. J. Agric. Food Chem. **58:**5502-5508.

Sarrion-Perdigones, A. *et al.* (2013). GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. Plant Physiology. **162**:1618-1631.

Sinani Al, S.S.S. & Eltayeb, E. A. (2017). The steroidal glycoalkaloids solamargine and solasonine in *Solanum* plants. South African Journal of Botany **112:**253-269.

Shiu, L. Y., Chang, L. C., Liang, C. H., Huang, Y. S., Sheu, H. M., & Kuo, K. W. (2007). Solamargine induces apoptosis and sensitizes breast cancer cells to cisplatin. Food Chem. Toxicol. **45:**2155–2164.

Sonawane, P. D. *et al.* (2017). Plant cholesterol biosynthetic pathway overlaps with phytosterol metabolism. Nat. Plants **3:**16205.

Sonawane, P. D. *et al.* (2018). Short-chain dehydrogenase/reductase governs steroidal specialized metabolites structural diversity and toxicity in the genus *Solanum.* Proc. Natl Acad. Sci. USA **115:** E5419-5428.

888 Sonawane, P.D., Jozwiak, A., Panda, S. & Aharoni, A. (2020). 'Hijacking' core metabolism: a new
889 panache for the evolution of steroidal glycoalkaloids structural diversity. Curr Opin. Plant Biol. **55:**118–
890 128.

891 Sonawane, P. D. *et al.* (2023). A BAHD-type acyltransferase concludes the biosynthetic pathway of non-
892 bitter glycoalkaloids in ripe tomato fruit. Nat. Commun **14:**4540.

893 Taguchi, G. *et al.* (2010). Malonylation is a key reaction in the metabolism of xenobiotic phenolic
894 glucosides in Arabidopsis and tobacco. Plant J. **63:**1031-1041.

895 Umemoto, N. *et al.* (2016). Two cytochrome P450 monooxygenases catalyze early hydroxylation steps
896 in the potato steroid glycoalkaloid biosynthetic pathway. Plant Physiol. **171:**2458–2467.

897 Winkiel, J. M., Chowanski S. & Slocinska M. (2022). Anticancer activity of glycoalkaloids from
898 Solanum plants: A review. Front. Pharmacol. **13:**979451.

899 Yamashita, T., Fujimura, N., Yahara, S., Nohara, T., Kawanobu, S. Fujida, K. (1990). Structures of three
900 new steroidal alkaloid glycosides, solaverines I, II and III from *Solanum toxicarium* and *S.*
901 *Verbascifolium*. Chemical and Pharmaceutical Bulletin **38:**827-829.

902 Zhao, D-K., Zhao, Yi., Chen, Sui-Yen. & Kennelly E. J. (2021). Solanum steroidal glycoalkaloids:
903 structural diversity, biological activities, and biosynthesis. Nat. Prod. Rep. **38:**1423-1444.

904 Zhao, S., Yan, T., Hunag, X. & Zhang, Y. (2023). Analysis of steroidal glycoalkaloids and their
905 metabolites in *Solanum nigrum* fruits based on liquid chromatography-tandem mass spectrometry and
906 molecular networking. J. Sep. Sci. **46:**202200804.

907

Data availability

Data supporting the findings of this work are available within the paper and its Supplemental Information files. *S. nigrum* RNA-seq data associated with this manuscript have been deposited into the NCBI Sequence Read Archive with BioProject ID XXXXX. Correspondence and requests for materials should be addressed to P.D.S. or S.E.O'C.

Author Contributions

R.L. performed the research and wrote the paper. M.O.K. and S.S. isolated malonyl-solamargine from *S. nigrum* green berries and scale up recombinant enzyme assays. Y.N. performed the NMR experiments and structural assignments. R.B. and J.W. assisted in candidate genes cloning for *in planta* assays. M.K. and S.H. assisted with metabolomics data analysis and operated the LC-MS. S.E.O'C and P.D.S. designed the research and wrote the paper.

Acknowledgements

R.L. gratefully acknowledges financial support by the Fulbright U.S. Student Program, which is sponsored by The U.S. Department of State and the German-American Fulbright Commission. We thank Prof. M. Court for UGT nomenclature assignment. *N. benthamiana*, green unripe and purple ripe *S. nigrum* berries used in Main and Supplementary Figures were generated in BioRender. We kindly acknowledge the Max Planck Society and the European Research Council (788301) for funding.

Competing interests

The authors declare no competing interests.

Figure Legends

Figure 1: Steroidal Glycoalkaloids (SGAs) biosynthetic pathway in cultivated (tomato, potato and eggplant) and wild (*Solanum nigrum*) *Solanum* species. The hypothesized biosynthetic pathway for *α*-solasonine, *α*-solamargine and malonylsolamargine in *S. nigrum* and eggplant starts with cholesterol and proceeds via the solasodine aglycone (22*R*, 25*R* spirosolane). Biosynthetic steps specific to tomato, potato and *S. nigrum*/eggplant are shown by light green, orange and gold shades, respectively. Common steps between tomato and potato pathway are shown in accent blue. Known biosynthetic enzymes (solid arrows) are marked in black. The uncharacterized steps (dashed arrows) and corresponding hypothesized enzymes are shown in blue. See Supplemental Figure 1 for more detailed tomato and potato SGA biosynthetic pathway. GAME: GLYCOALKALOID METABOLISM; Glc: Glucose; Gal: Galactose; Xyl: Xylose; Rha: Rhamnose; Mal: Malonyl; DPS: Dioxygenase for potato solanidane synthesis; 3βHSD: 3β-hydroxysteroid dehydrogenase/3-Ketosteroid reductase; 5αR2: 5α reductase2; UGT: UDP-glycosyltransferase.

Figure 2: Discovery of *α*-solasonine and *α*-solamargine biosynthesis pathway in black nightshade (*S. nigrum*). (A) Profiling of *α*-solasonine, *α*-solamargine and malonyl-solamargine in unripe green and ripe purple berries of *S. nigrum*. **(B)** Expression of candidate *GAME* genes in green and ripe berries of *S. nigrum* (RNA-seq expression data). Normalized FPKM (Fragments Per Kilobase of transcript per Million mapped reads) values were used to infer the expression profile. (**C**) Transient expression of six *S. nigrum GAME* genes (termed as solasodine genes) in *N. benthamiana* resulted in solasodine steroidal aglycone production. Aligned selected ion monitoring (SIM) chromatograms from LC-MS analysis are presented. The asterisk shows the presence of a peak with the same mass (*m/z* 414.30) as that of solasodine, but at a different retention time. See supplemental Figure 28 for MS/MS spectra of this new

compound (*). Leaves infiltrated with empty vector were used as control. Simplified scheme of solasodine biosynthesis from cholesterol in *S. nigrum* is presented. **(D)** Three step pooled-screen approach (pool 1-3) coupled with transient gene expression system in *N. benthamiana* led to identification of four UGT enzymes, UGT73L14, UGT73DU3, UGT93M3, UGT93N4 from *S. nigrum* catalyzing *α*-solasonine and *α*-solamargine biosynthesis. 10 UGT candidates from *S. nigrum* were initially selected for this stepwise screening approach. Aligned SIM chromatograms are shown. **(E)** A simplified pathway scheme of main SGAs, *α*-solasonine and *α*-solamargine from solasodine aglycone in *S. nigrum* based on pooled screens results. **(F, G)** Aligned extracted ion chromatograms (EICs) showing the accumulation of *α*-solasonine (F) and *α*-solamargine (G) in *N. benthamiana* leaves transiently expressing gene combinations UGT73L14/UGT73DU3/UGT93N4 (F) and UGT73DU3/UGT93M3 (G), respectively, with and without infiltration of solasodine substrate. Leaves infiltrated with empty vector were used as control. LC-MS was used for targeted SGAs analysis. *m/z*, mass to charge.

Figure 3: Three distinct UGT activities enable step-by-step conversion of solasodine to *α*-solasonine. Aligned LC-MS chromatograms (EICs) presenting the assay products (glycosylated SGAs) produced in *N. benthamiana* (*in planta*) and *E. coli* (*in vitro*) after expression of indicated UGT enzymes from starting solasodine aglycone precursor. Assay products from the *in planta* assays had same mass and retention time with those produced in the *in vitro* assays. Control (*in planta* assays): solasodine infiltrated *N. benthamiana* leaves transiently expressing empty 3Ω1 vector. Control (*in vitro* assays): assay reaction with solasodine and protein extracts of empty pOPINF vector-transformed *E. coli* cells. MS/MS fragmentation spectrum for glycosylated products in the *α*-solasonine biosynthetic pathway and *α*-solasonine standard are shown. Non-specific glycosylated metabolites, peak A and B are produced in control *N. benthamiana* samples due to activity of host UGT enzymes (see Supplemental Figures 33 and 35 for more details).

978 **Figure 4: UGT73L17, UGT73DU3 and UGT93M3 activities convert solasodine aglycone to *α*-**
979 **solamargine *in planta* and *in vitro*.** Aligned extracted ion chromatograms (EICs) from LC-MS
980 presenting the assay products (glycosylated SGAs) produced in *N. benthamiana* (*in planta*) and *E. coli*
981 (*in vitro*) after expression of indicated UGT enzymes from starting solasodine aglycone precursor. Assay
982 products from the *in planta* assays had same mass and retention time with those produced in the *in vitro*
983 assays. Control (*in planta* assays): solasodine infiltrated *N. benthamiana* leaves transiently expressing
984 empty 3Ω1 vector. Control (*in vitro* assays): assay reaction with solasodine and protein extracts of empty
985 pOPINF vector-transformed *E. coli* cells. MS/MS fragmentation spectrum for glycosylated
986 intermediates in *α*-solamargine pathway are shown. Mass to charge (*m/z*) is shown for assay products.
987 Dashed green arrows represent alternative route for *α*-solamargine formation. See supplemental Figure
988 35 for Peak B characterization observed in control *in planta* assays.

989 **Figure 5: A novel malonyltransferase catalyzes malonylation of *α*-solamargine to malonyl-**
990 **solamargine. (A)** Extracted ion chromatograms (EICs) of malonyl-solamargine from *N. benthamiana*
991 leaf extracts expressing *SnMAT* (*S. nigrum* malonyltransferase) together with upstream
992 UGT73L17/UGT73DU3/UGT93M3 genes (left panel) or alone itself (right panel), with infiltration of
993 solasodine (left panel) or *α*-solamargine (right panel) substrates, respectively. *De novo* production of
994 malonyl-solamargine was observed in *N. benthamiana* leaves upon transient expression of SnMAT. **(B)**
995 Malonylation of *α*-solamargine by the recombinant SnMAT enzyme produced in *E. coli* cells.
996 Recombinant SnMAT catalyzes conversion of *α*-solamargine to malonyl-solamargine in presence of
997 malonyl-CoA as acyl donor. Control reactions were performed with *α*-solamargine substrate, but either
998 without recombinant enzyme or without malonyl-CoA. Aligned ion chromatograms were obtained by
999 LC-MS analysis. **(C)** Summary of the biosynthetic pathway for malonyl-solamargine from solasodine in
1000 *S. nigrum*.

Figure 6: Elucidation of *α*-solasonine, *α*-solamargine and malonyl-solamargine biosynthesis in cultivated eggplant (*S. melongena*). (A) LC-MS based profiling of *α*-solasonine, *α*-solamargine and malonyl-solamargine in young leaves and roots of cultivated eggplant. **(B)** Amino acid sequence identity matrix of UGT and malonyltransferase enzymes from cultivated eggplant compared with its homologues from *S. nigrum*. **(C, D)** Transient expression of eggplant UGT73L15/UGT73DU2/UGT93N3 (C) and UGT73L19/UGT73DU2/UGT93M2 (D) gene combinations in *N. benthamiana* (solasodine fed to infiltrated leaves) produced *α*-solasonine and *α*-solamargine SGAs, respectively. Extracted ion chromatograms from LC-MS are shown. Leaves infiltrated with empty vector were used as control. **(E)** Summary of biosynthetic pathway of malonyl-solamargine from solasodine in cultivated eggplant. **(F)** Aligned extracted ion chromatograms (LC-MS) of malonyl-solamargine from *N. benthamiana* leaf extracts after transient expression of various gene combinations. Co-expression of malonyltransferase from eggplant (SmMAT) together with UGT73L19/UGT73DU2/UGT93M2 in *N. benthamiana* converts infiltrated solasodine to malonyl-solamargine (left panel). Transient expression of SmMAT alone in *N. benthamiana* leaves fed with *α*-solamargine, immediate precursor also produces malonyl-solamargine (right panel). **(G)** Co-expression of potato malonyltransferase (StMAT) either with UGT73L17/UGT73DU3/UGT93M3 (*S. nigrum*) or UGT73L19/UGT73DU2/UGT93M2 (cultivated eggplant) gene set also resulted in malonyl-solamargine formation in *N. benthamiana* leaves (infiltrated with solasodine substrate). Extracted ion chromatograms from LC-MS analysis are shown. *m/z*: mass to charge.

Figure 7: The complete SGA biosynthetic pathway in wild *S. nigrum* and cultivated eggplant (*S. melongena*). Twelve biosynthetic enzymes including six core GAMEs, five UGTs and one malonyl-transferase identified from *S. nigrum* (solid green arrows) directs the biosynthesis of solasodine steroidal aglycone and its downstream three main SGAs namely, *α*-solasonine, *α*-solamargine and malonyl-solamargine from cholesterol precursor. Six genes including five UGTs and one malonyltransferase

identified from cultivated eggplant (*S. melongena*) enable the conversion of solasodine aglycone to the similar SGA repertoire (solid purple arrows) as that of *S. nigrum* mentioned above. Altogether eighteen enzymes were functionally characterized by *in planta* (transient expression in *N. benthamiana*) and/or by *in vitro* (recombinant enzyme activity) assay approach. Dashed green and purple arrows represent alternative route for *α*-solamargine biosynthesis in *S. nigrum* and cultivated eggplant, respectively. Malonyl group in malonyl-solamargine is marked in red. GAME: GLYCOALKALOID METABOLISM; Glc: Glucose; Gal: Galactose; Rha: Rhamnose; UGT: UDP-glycosyltransferase; SnMAT: *S. nigrum* malonyltransferase; SmMAT: *S. melongena* malonyltransferase.

Figure 1

Common tomato/potato pathway

Tomato specific pathway

Potato specific pathway

***Solanum nigrum* and *S. melongena* (cultivated eggplant) specific pathway**

GAME6
GAME8
GAME11

GAME4
GAME12

GAME enzymes ?

UGTs ?

3βHSD (GAME25)
5αR2

GAME1
GAME2
GAME17
GAME18

DPS

UGTs ?

UGTs ?

Malonyl
Transferase?

Dehydrotomatidine

Cholesterol

Solasodine

Tomatidine

α-solamarine

α-solanine

α-solasonine

α-solamargine

Malonyl-solamargine

α-tomatine

Figure 2

A

Green berries
Ripe berries
Malonyl-solamargine standard
α-solamargine standard
α-solasonine standard
20 21 22 23 24 25
Time (min)

B

	Green berries	Ripe berries
GAME6	75.52	1.11
GAME8	197.07	8.58
GAME11	3093.14	116.19
GAME4	178.81	0.15
GAME12	865.19	0.08
GAME15	130.32	1.03

3000
2000
1000
0

C

SIM *m/z* 414.30
Solasodine standard
Solasodine genes [GAME6/GAME8/GAME11/GAME4/GAME12+GAME15]
GAME6/GAME8/GAME11/GAME4/GAME12
Control
20 23 26 29
Time (min)

Cholesterol
S. nigrum
GAME6/GAME8
GAME11/GAME4
GAME12/GAME15
Solasodine

D

α-solasonine
SIM *m/z* 884.50
[1] +10 UGTs
Solasodine genes [1]
[1] + UGTs – UGT93N4
[1] + UGTs – UGT73DU3
[1] + UGTs – UGT73L14
[1] + UGT73L14+ UGT73DU3+UGT93N4
α-solasonine standard
20 21 22 23
Time (min)

Pool 1
Solasodine genes [1] + all UGT candidates

Pool 2
Solasodine [1] + all UGT candidates minus one candidate at a time

Pool 3
Solasodine [1] + selected UGT candidates from pool 2

α-solamargine
SIM *m/z* 868.50
[1] +10 UGTs
Solasodine genes [1]
[1] + UGTs – UGT93M3
[1] + UGTs – UGT73DU3
[1] + UGT93M3+ UGT73DU3
α-solamargine standard
20 21 22 23
Time (min)

E

α-solasonine
UGT73L14
UGT73DU3
UGT93N4
Solasodine
UGT93M3
UGT73DU3
α-solamargine

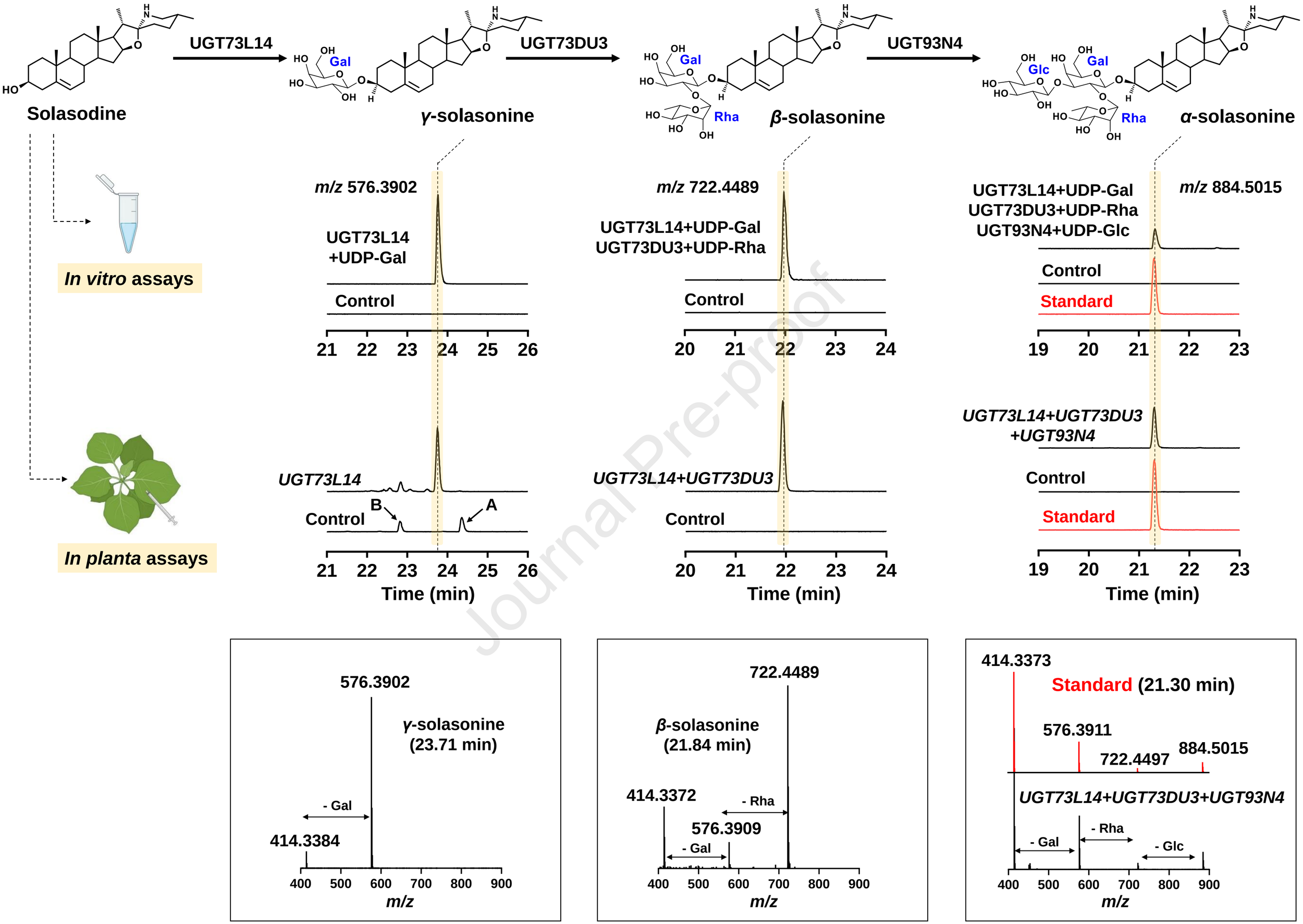
F

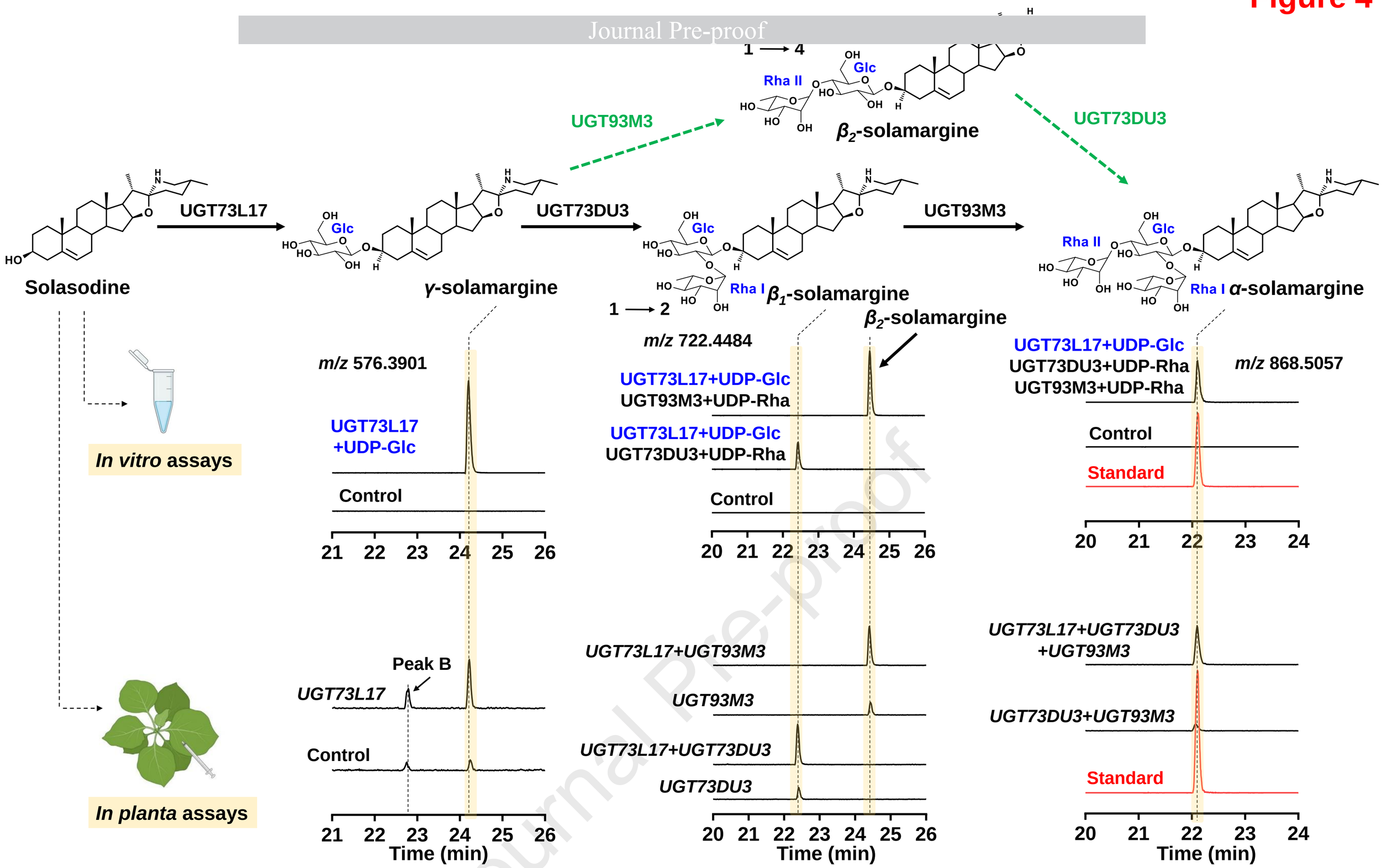
EIC
m/z 884.50
α-solasonine standard
UGT73L14+UGT73DU3+ UGT93N4 + solasodine
UGT73L14+UGT73DU3+ UGT93N4 - solasodine
Control + solasodine
19 20 21 22 23
Time (min)

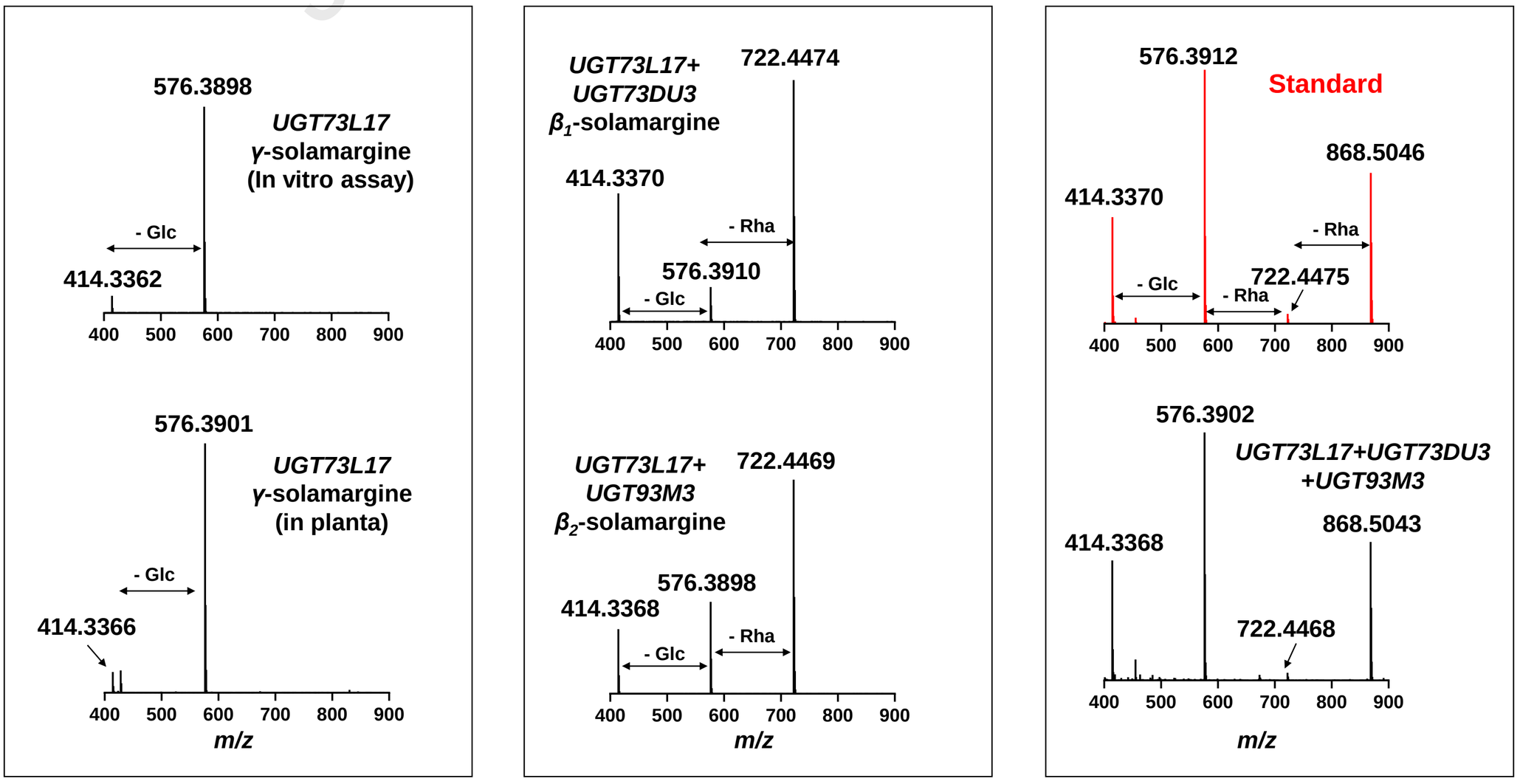
G

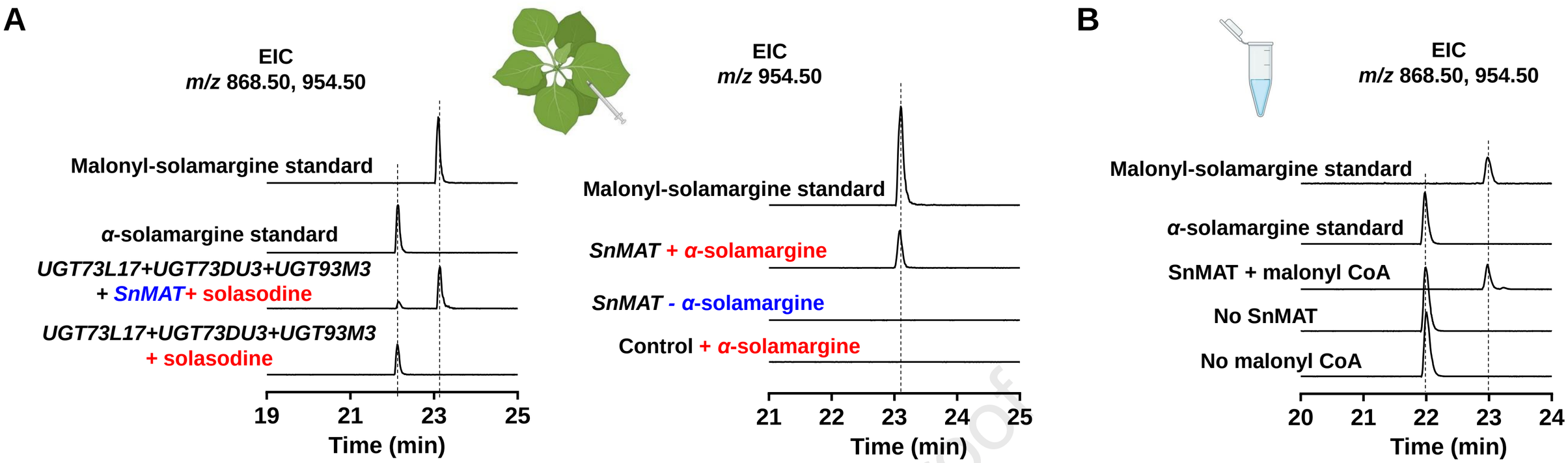
EIC
m/z 868.50
α-solamargine standard
UGT73DU3+UGT93M3 + solasodine
UGT73DU3+UGT93M3 - solasodine
Control + solasodine
20 21 22 23 24
Time (min)

Figure 3









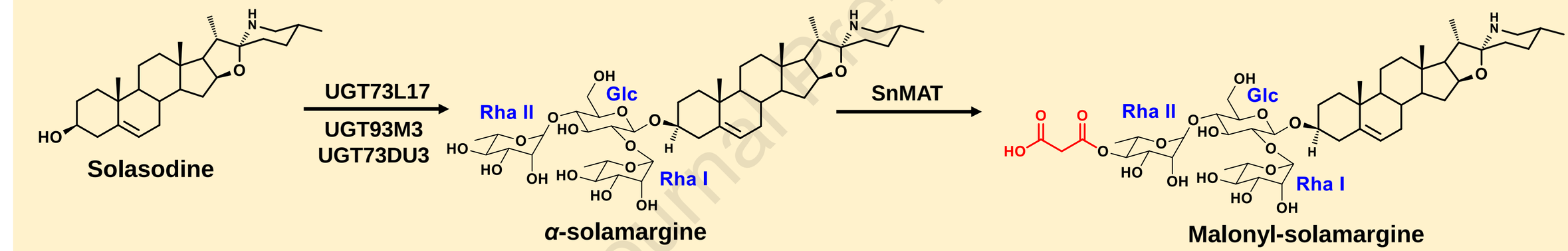
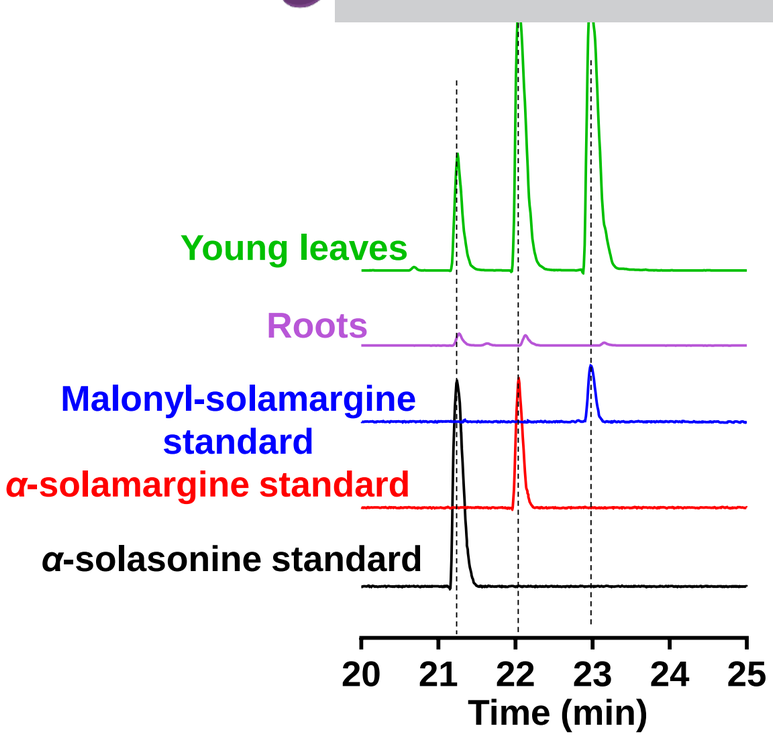


Figure 6

A



B

UGT73L15

	Eggplant	*S. nigrum*
Eggplant		82.12
S. nigrum	82.12	

UGT73DU2

	Eggplant	*S. nigrum*
Eggplant		82.36
S. nigrum	82.36	

UGT93M2

	Eggplant	*S. nigrum*
Eggplant		81.12
S. nigrum	81.12	

UGT93N3

	Eggplant	*S. nigrum*
Eggplant		81.61
S. nigrum	81.61	

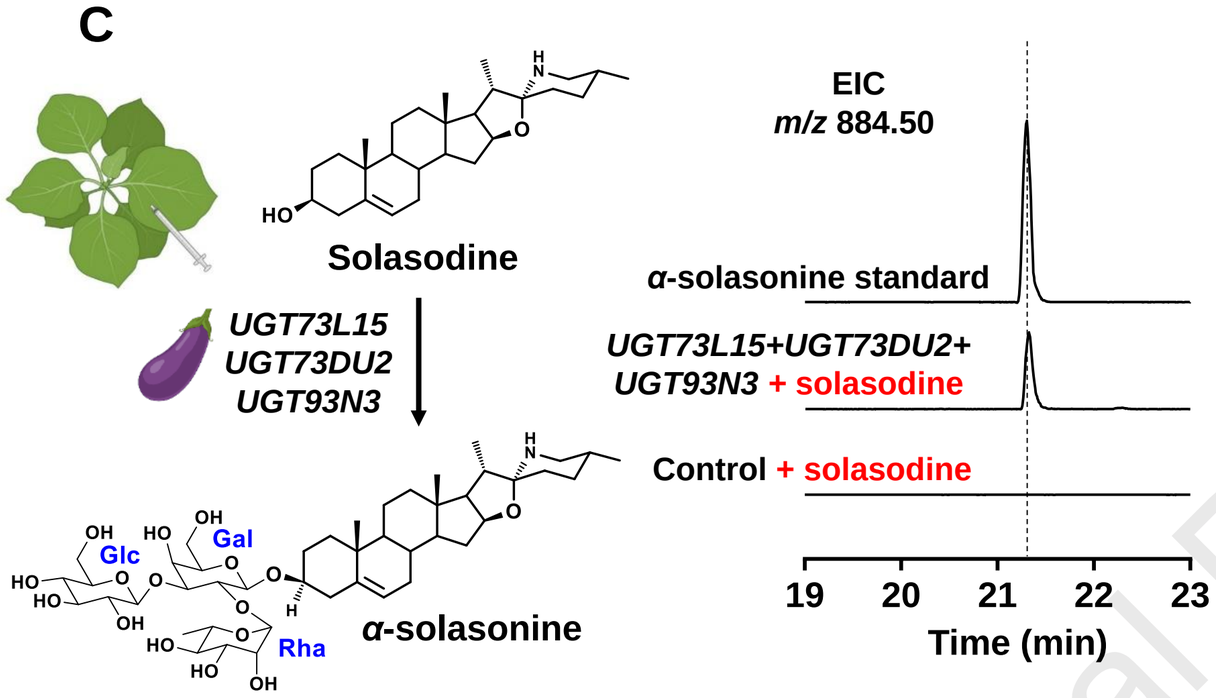
UGT73L19

	Eggplant	*S. nigrum*
Eggplant		79.96
S. nigrum	79.96	

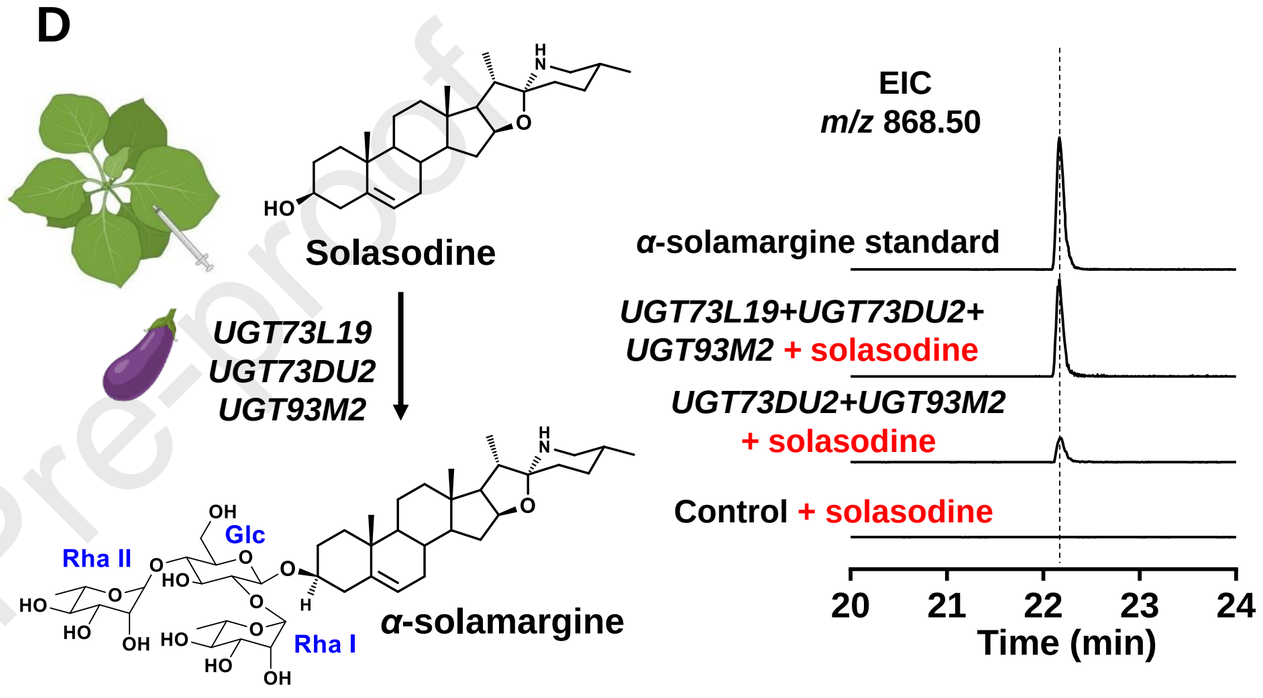
MAT

	Eggplant	*S. nigrum*
Eggplant		74.56
S. nigrum	74.56	

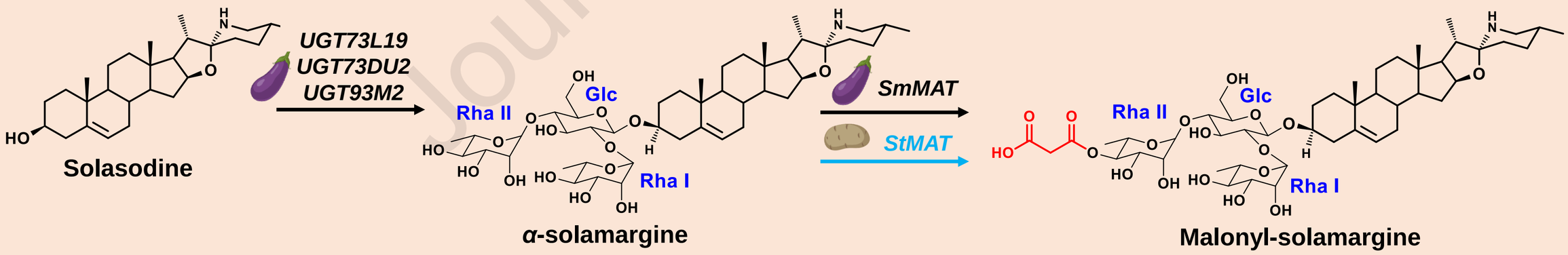
C



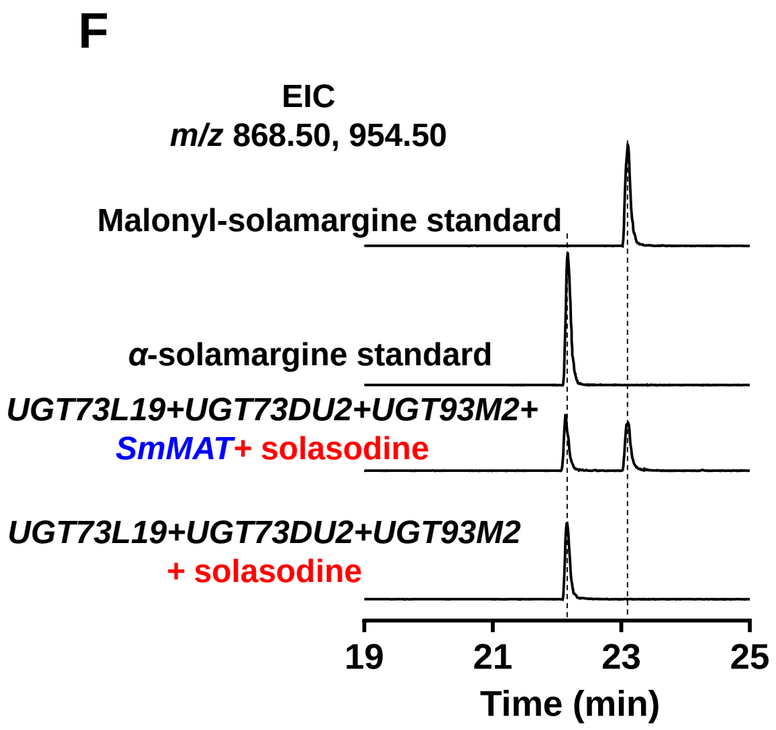
D



E



F



G

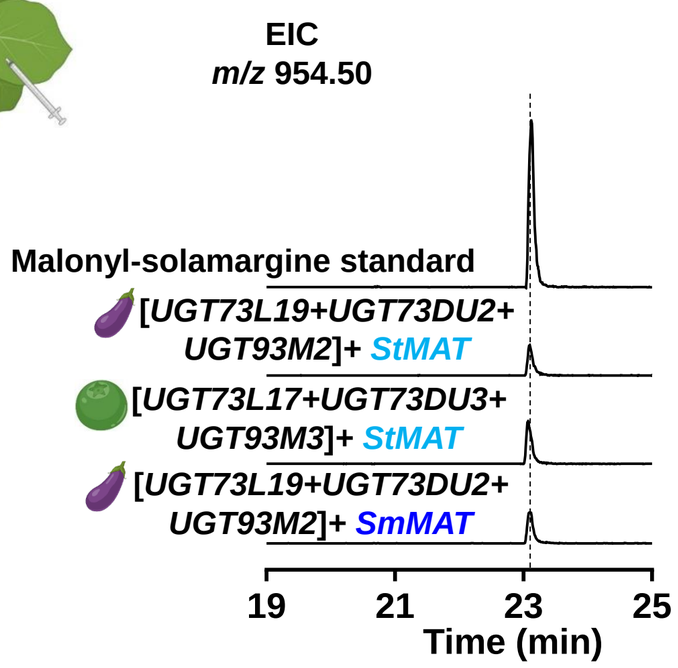


Figure 7