Running title: *Solanum* trichome metabolite diversity

Trading acyls and swapping sugars: metabolic innovations in *Solanum* trichomes

Paul D. Fiesel[1], Rachel E. Kerwin[1], A. Daniel Jones[1], and Robert L. Last[1,2]

[1]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48823 USA

[2]Department of Plant Biology, Michigan State University, East Lansing, MI 48823 USA

P.D.F and R.E.K contributed equally

Corresponding Author: Robert L. Last, Michigan State University, Department of Biochemistry and Molecular Biology, 603 Wilson Road, East Lansing, MI 48824; Phone: 517-432-3278; email: lastr@msu.edu

ORCID:

Paul D. Fiesel 0000-0002-2038-5417

Rachel E. Kerwin 0000-0001-7395-5973

A. Daniel Jones 0000-0002-7408-6690

Robert L. Last 0000-0001-6974-9587

Abstract

Solanaceae (nightshade family) species synthesize a remarkable array of clade- and tissue-specific specialized metabolites. Protective acylsugars, one such class of structurally diverse metabolites, are produced by acylsugar acyltransferases from sugars and acyl-Coenzyme A esters. We characterized trichome acylsugars of the Clade II species *Solanum melongena* (brinjal eggplant) using liquid chromatography-mass spectrometry (LC-MS), gas chromatography-MS and nuclear magnetic resonance (NMR) spectroscopy. This led to the identification of eight unusual structures with inositol or inositol glycoside cores, and hydroxyacyl chains. LC-MS analysis of 31 species in the megadiverse *Solanum* genus revealed striking acylsugar diversity with some traits restricted to specific clades and species. Acylinositols were found throughout each clade while acylglucoses were restricted to DulMo and VANAns species. Medium length hydroxyacyl chains were identified in many species. Analysis of tissue-specific transcriptomes and interspecific acylsugar acetylation differences led to the unexpected identification of the *S. melongena* AcylSugar AcylTransferase 3-Like 1 (SmASAT3-L1; SMEL4.1_12g015780) enzyme. This enzyme is distinct from previously characterized acylsugar acetyltransferases, which are in the ASAT4 clade, and appears to be a functionally divergent ASAT3. This study provides a foundation for investigating the evolution and function of diverse *Solanum* acylsugar structures and harnessing this diversity in breeding and synthetic biology.

42 **Introduction**

43 Plants are remarkable synthetic chemists, producing a multitude of structurally complex
44 specialized metabolites that differ from the products of general, or primary, metabolism in their
45 lineage-specific distribution and tissue or cell type-specific biosynthesis. In contrast to the
46 negative selection against changes in primary metabolism (for example, amino acids, energy
47 metabolism intermediates and vitamin cofactors), less constrained evolution of specialized
48 metabolism led to accumulation of hundreds of thousands of taxonomically restricted
49 metabolites in broad classes. Specialized metabolites play many critical roles such as in abiotic
50 and biotic stress adaptation (Agati and Tattini, 2010; De Moraes et al., 2001; Landry et al.,
51 1995), pollinator attraction (Kretschmar and Baumann, 1999) and mediation of interactions with
52 beneficial and pathogenic microbes (Yu et al., 2021). These diverse and bioactive small
53 molecules have historical and modern uses in human medicine, including the chemotherapeutics
54 vinblastine and paclitaxel, antimalarial artemisinin, and painkillers such as morphine. Their
55 distinct phylogenetic distributions and relatively rapid diversification make specialized metabolic
56 enzymes excellent models to study evolution.

57 Acylsugars are specialized metabolites produced across the Solanaceae (nightshade)
58 family, aiding in defense against herbivores, fungi, and bacteria (Goffreda et al., 1989; Leckie et
59 al., 2016; Luu et al., 2017; Weinhold and Baldwin, 2011). In Types I- and IV-glandular
60 trichomes, BAHD-type AcylSugar AcylTransferase (ASAT) enzymes assemble acylsugars from
61 the basic building blocks of sugars, often sucrose, and short- to medium-length acyl chains
62 derived from acyl-CoenzymeAs (acyl-CoAs) (Fan et al., 2015; Lou et al., 2021; Moghe et al.,
63 2017; Schilmiller et al., 2015; Schilmiller et al., 2012). Despite their simple components,
64 acylsugars exhibit remarkable chemical diversity arising from variations in sugar core

65 composition and acyl chain length, branching pattern, position, and number (Fan et al., 2019;
66 Fiesel et al., 2022; Ghosh et al., 2014; Hurney, 2018; Lou et al., 2021; Moghe et al., 2017;
67 Schenck et al., 2022). For example, acylsucroses, composed of a sucrose disaccharide core,
68 accumulate in cultivated tomato (*Solanum lycopersicum*) trichomes, while acylsucroses and
69 acylglucoses have been observed in the trichomes of wild tomato species, *Solanum pennellii*.
70 Acylsugar structural variation also impacts biological activity; for example, differential
71 oviposition deterrence was demonstrated from naturally derived acylsugar mixtures (Leckie et
72 al., 2016).

73 Solanaceae acylsugars have become an exemplary model to study evolution of a diverse,
74 biologically relevant trait in a plant family with genomic and phylogenetic resources. Utilization
75 of this model revealed gene duplication, neofunctionalization, co-option, and loss involved in
76 acylsugar evolution (Fan et al., 2020, 2017; Leong et al., 2019; Moghe et al., 2017). For
77 example, neofunctionalization of an invertase-like enzyme, AcylSucrose FructoFuranosidase 1
78 (ASFF1) and functional divergence of core ASAT enzymes is responsible for differences in
79 sugar core type and acyl chain positions between cultivated tomato *S. lycopersicum* and wild
80 tomato *S. pennellii* acylsugars (Fan et al., 2017; Leong et al., 2019). Identifying mechanisms of
81 acylsugar evolution across the broader *Solanum* requires a detailed understanding of acylsugar
82 diversity and biosynthesis, which is lacking for many species.

83 Nearly half of the Solanaceae falls into the large (>1200 species) and phenotypically
84 diverse *Solanum* genus. This genus is split into several major clades, including Potato,
85 Regmandra, DulMo, VANAns, and Clade II; the Potato clade contains cultivated tomato and
86 potato and their wild relatives while Clade II contains the cultivated brinjal eggplant, *Solanum*
87 *melongena*, and other 'spiny *Solanums*' (Bohs, 2004; Bohs and Olmstead, 1997; Gagnon et al.,

88 2022; Levin et al., 2006; PBI Solanum Project, 2022; Särkinen et al., 2013; Stern et al., 2011;
89 Tepe et al., 2016; Weese and Bohs, 2007). To date, documentation of Solanum acylsugar
90 diversity largely focused on a handful of species within the Potato clade, including cultivated
91 tomato and its close relatives. These efforts identified at least 38 nuclear magnetic resonance
92 (NMR) spectroscopy-resolved acylsugar structures and liquid chromatography mass
93 spectrometry (LC-MS) supported annotations of many more (Fan et al., 2017; Ghosh et al., 2014;
94 Lybrand et al., 2020; Schilmiller et al., 2016). While limited acylsugar screening outside of the
95 Potato clade was reported, novel structural variants not observed among cultivated tomato
96 relatives were identified. For example, acylsugars with *myo*-inositol sugar cores (i.e.,
97 acylinositols) were characterized in three species: *S. nigrum*, from the DulMo clade, and *S.*
98 *lanceolatum* and *S. quitoense,* from Clade II (Herrera-Salgado et al., 2005; Hurney, 2018; Leong
99 et al., 2020; Lou et al., 2021). These discoveries highlight the benefits of a comprehensive
100 description of *Solanum* acylsugars within the well-developed phylogenetic framework.
101 Here we report analysis of *Solanum* acylsugar chemical diversity in species from the
102 relatively unexplored *Solanum* clades DulMo, VANAns, and Clade II, which together comprise
103 ~1000 *Solanum* species. We first established the Clade II brinjal eggplant, *S. melongena*, as a
104 reference species. Eggplant is a major worldwide fruit crop with extensive genomic,
105 transcriptomic, and germplasm resources (Barchi et al., 2021; Li et al., 2021; Mennella et al.,
106 2010). We characterized eggplant trichome acylinositols, acylinositol glycosides, and acylsugars
107 with unusual hydroxylated acyl chains using electrospray ionization LC-quadrupole time of
108 flight-MS (ESI LC-QToF-MS) and NMR. These atypical structures likely reflect participation of
109 enzymes that are distinct from the cultivated tomato acylsucrose pathway. Moving beyond this
110 model organism framework, LC-MS phylogenetic screening of 31 Clade II, DulMo, and

111 VANAns species, including *S. melongena*, led to the identification of remarkable acylsugar
112 structural variation. LC-MS features with characteristics of acylinositols were found in 25 of the
113 26 acylsugar-producing species, suggesting one or a small number of evolutionary origins. In
114 contrast, acylglucoses were detected in DulMo and VANAns species, but not in any tested Clade
115 II species.

116 As a first step towards revealing the molecular basis of the extensive documented
117 acylsugar diversity, we utilized interspecific acylsugar differences and an eggplant tissue-
118 specific transcriptome to identify an acylinositol biosynthetic enzyme, *S. melongena* Acylsugar
119 AcylTransferase 3-Like 1 (SmASAT3-L1; SMEL4.1_12g015780), responsible for acetylating a
120 triacylinositol glycoside acyl acceptor. SmASAT3-L1 exhibits a different acyl-CoA specificity
121 than previously characterized ASAT3 homologs, highlighting how gene duplication and
122 functional divergence created acylsugar metabolic novelty in this part of the *Solanum* clade.
123

124 **Results and Discussion**

125

126 <u>Eggplant glandular trichomes accumulate acylsugar-like metabolites</u>

127 We began our investigation of *Solanum* acylsugar diversity with the brinjal eggplant *S.*
128 *melongena* due to its economic importance, genomic resources, and phylogenetic position within
129 the monophyletic Eastern Hemisphere Spiny clade of Clade II (formerly known as the 'Old
130 World spiny clade') (Gagnon et al., 2022). Using eight eggplant accessions (listed in Table S1),
131 we observed glandular trichomes on hypocotyls, cotyledons, and the first three true leaves of
132 young eggplant plants (Figure 1, S1). These resemble the acylsugar-producing structures found
133 in other *Solanum* species (Leong et al., 2020; Lou et al., 2021; Schilmiller et al., 2012). In
134 contrast, we observed only non-glandular stellate trichomes on leaves and stems of mature

135 eggplants (Figure S1), and these are types that typically do not accumulate acylsugars or other
136 specialized metabolites (Levin, 1973; Wagner, 1991).

137 We analyzed surface metabolite extracts from young and mature eggplant tissues using
138 LC-QToF-MS coupled with collision-induced dissociation (CID) in negative and positive ion
139 mode. We used these data to annotate acylsugars based on relative mass defect, molecular adduct
140 ion masses, retention times, and ions present in CID mass spectra. Briefly, acylsugars were
141 annotated from masses of fatty acid carboxylate fragment ions, as well as fragment ions
142 corresponding to stepwise losses of acyl chains from the pseudomolecular ion to a sugar core
143 fragment ion. The acylsugar annotation methods and confidence criteria are explained in detail in
144 the Methods. This analysis revealed abundant acylsugars in extracts from young, glandular
145 trichome-producing eggplant tissues, but not from mature, non-glandular trichome producing
146 tissues.

147 We annotated 38 acylsugars in leaf surface extracts of young eggplant from eight of nine
148 accessions (Table S1), consisting of 16 acylhexoses and 22 acyldisaccharides (Table 1). LC-MS
149 based acylsugar annotations are described using a shorthand nomenclature modified from
150 previous work (Leong et al., 2020) as follows: UX:Y:Z(A,B,C,D), where U – as a single or
151 multi-letter designation – represents the sugar core, X represents the number of acyl chains, Y is
152 the sum of acyl chain carbons, Z indicates the number of unsaturated bonds in the acyl chains
153 (when present), and A-D represent the number of carbon atoms in the individual acyl chains. For
154 example, the eggplant acylhexose I3:18(4,4,10) consists of a *myo*-inositol core with three acyl
155 chains with a total of 18 carbon atoms.

156 Curiously, eggplant acyldisaccharides contain an atypical pentose-hexose core, as
157 evidenced by a fragment ion mass of *m/z* 293.09 in negative-ion mode, corresponding to a fully

158 deacylated sugar core minus a hydroxyl group and a proton. Further evidence was provided by
159 positive mode CID, which promotes glycosidic bond cleavage, yielding a fragment ion
160 corresponding to the neutral loss of an unacylated pentose ring. All acylhexoses annotated with
161 medium to high confidence formed abundant fragment ions in positive mode but few fragment
162 ions in negative mode as illustrated in Figure S2; this pattern is characteristic of acylinositols
163 found in *S. quitoense* and *S. nigrum,* but has not been observed for acylglucoses, suggesting that
164 the detected *S. melongena* acylhexoses are acylinositols (Hurney, 2018; Leong et al., 2020; Lou
165 et al., 2021).

166 The eggplant acylsugars each contained three to four acylations, all on the hexose core,
167 including one medium eight-carbon (C8) to C14 acyl chain and two to three C4 or C5 acyl
168 chains. The medium acyl chains (C8, C10, C12, and C14) were either iso-branched or straight as
169 revealed by GC-MS acyl chain analysis (Figure S3). Additionally, we identified hydroxylated
170 C12, C14, and C16 acyl chains not previously reported in Solanaceae acylsugars (Table 1).
171 Although we did not observe large differences between the eight eggplant accessions, eggplant
172 acylsugars differ in chain length, functional groups, and acyl chain composition from the
173 reported *Solanum* acylinositols (Herrera-Salgado et al., 2005; Hurney, 2018; Lou et al., 2021).
174

175 **Table 1. Summary of annotated acylsugars detected in *S. melongena* leaf surface extracts.**
176 PH = pentose-hexose; AI = arabinose-inositol, I = inositol. RT = retention time; m/z_{acc} =
177 theoretical monoisotopic formate adduct mass; m/z_{ex} = experimental formate adduct mass; Δm
178 (ppm) = mass measurement error in parts per million. Compound numbers are listed for NMR
179 characterized compounds (Figure 1). Percent total peak area calculated by dividing acylsugar
180 peak area by total acylsugar peak area. Acylsugars composing ≥1% of total acylsugar peak area
181 are in bold. Acylsugar abundance was reported by Progenesis QI and averaged over 15 samples.
182 The annotation method is described in the Methods. Acylsugars are sorted by number of sugar
183 moieties and then by elution order.
184

Name	Compound #	Annotation confidence level	RT (min)	Neutral molecule chemical formula	m/z_{acc}	m/z_{ex}	Δm (ppm)	Average acylsugar abundance (%)
Acyldisaccharides								
PH2:12(4,8)		medium	1.98	C23H40O12	553.2496	553.2496	0.0	0.203
PH4:14(2,4,4,4)		medium	2.19	C25H40O14	609.2395	609.2445	8.2	0.0246
PH3:14(2,4,8)		medium	2.23	C25H42O13	595.2602	595.2621	3.2	0.0453
PH4:16(2,2,4,8)		medium	2.54	C27H44O14	637.2708	637.2720	1.9	0.0851
PH3:16(4,4,8)		**medium**	**2.61**	**C27H46O13**	**623.2915**	**623.2916**	**0.2**	**17.1**
AI3:16(4,4,8)	**1**	**high**	**2.66**	**C27H46O13**				
PH4:17(2,2,4,9)		medium	2.93	C28H46O14	651.2864	651.2924	9.2	0.114
PH3:17(4,5,8)		medium	2.96	C28H48O13	637.3071	637.3083	1.9	0.324
PH4:18(2,4,4,8)		medium	3.11	C29H48O14	665.3021	665.3033	1.8	0.901
AI4:18(2,4,4,8)	**2**	**high**	**3.27**	**C29H48O14**	**665.3021**	**665.3026**	**0.7**	**42.5**
PH3:18(4,4,10)		**medium**	**3.61**	**C29H50O13**	**651.3228**	**651.3205**	**-3.5**	**1.12**
PH3:18(4,4,10)		medium	3.74	C29H50O13	651.3228	651.3245	2.5	0.964
PH4:19(2,4,5,8)		medium	3.76	C30H50O14	679.3177	679.3146	-4.6	0.329
PH4:19(2,4,5,8)		medium	3.83	C30H50O14				
PH4:20(4,4,4,8)		medium	4.22	C31H52O14	693.3334	693.3350	2.3	0.0428
PH4:20(2,4,4,10)		medium	4.57	C31H52O14	693.3334	693.3351	2.5	0.0447
PH4:20(2,4,4,10)		medium	4.68	C31H52O14	693.3334	693.3349	2.1	0.127
PH4:20(2,4,4,10)		medium	4.85	C31H52O14	693.3334	693.3379	6.6	0.00422
PH3:20(4,4,12)		medium	5.05	C31H54O13	679.3541	679.3554	1.9	0.0389
PH3:20(4,4,12)		medium	5.23	C31H54O13	679.3541	679.3552	1.6	0.130
PH3:22(4,4,14)		medium	7.54	C33H58O13	707.3854	707.3879	3.5	0.0356
PH3:22(4,4,14)		medium	7.85	C33H58O13	707.3854	707.3892	5.3	0.0107
Acylhexoses								
I3:16(4,4,8)		**medium**	**3.27**	**C22H38O9**	**491.2492**	**491.2492**	**-0.1**	**3.99**
I3:20(4,4,12-OH)		medium	4.16	C26H46O10	563.3068	563.3071	0.5	0.0808
I3:20(4,4,12-OH)		medium	4.32	C26H46O10	563.3068	563.3068	0.0	0.500
I4:18(2,4,4,8)		medium	4.22	C24H40O10	533.2598	533.2604	1.1	0.0248
I3:18(4,4,10)	**3**	**high**	**4.70**	**C24H42O9**	**519.2805**	**519.2802**	**-0.5**	**10.6**
I3:18(4,4,10)		medium	4.88	C24H42O9	519.2805	519.2794	-2.1	0.693
I3:20(2,4,14-OH)		medium	4.68	C26H46O10	563.3068	563.3082	2.5	0.0588
I3:20(2,4,14-OH)		medium	4.90	C26H46O10	563.3068	563.3077	1.6	0.298
I3:22(4,4,14-OH)	**7**	**high**	**6.06**	**C28H50O10**	**591.3381**	**591.3385**	**0.7**	**4.96**
I3:22(4,4,14-OH)	**8**	**high**	**6.34**	**C28H50O10**	**591.3381**	**591.3380**	**-0.1**	**10.6**
I3:20(4,4,12)	**4**	**high**	**6.81**	**C26H46O9**	**547.3118**	**547.3119**	**0.1**	**2.24**
I3:20(4,4,12)		medium	7.09	C26H46O9	547.3118	547.3127	1.7	0.105
I3:24(4,4,16-OH)		medium	8.51	C30H54O10	619.3694	619.3707	2.0	0.0473
I3:24(4,4,16-OH)		medium	8.85	C30H54O10	619.3694	619.3683	-1.8	0.0355
I3:22(4,4,14)	5	high	9.34	C28H50O9	575.3431	575.3435	0.7	0.986

I3:22(4,4,14)	6	high	9.68	C28H50O9	575.3431	575.3433	0.3	0.722

185

186 Although LC-MS provided information about sugar core mass, it did not reveal the sugar
187 core structure, prompting analysis using gas chromatography-mass spectrometry (GC-MS) of
188 derivatized *S. melongena* acylsugar cores. When free sugar cores, produced by metabolite extract
189 saponification, were derivatized to form alditol acetates, GC-MS peaks corresponding to
190 derivatized *myo*-inositol and glucose were detected, supporting the presence of acylinositols
191 (Figure S4A). Because the detection of glucose might have resulted from other compounds in the
192 leaf surface metabolite extracts, the disaccharide sugar core composition was determined by
193 hydrolyzing the saponified sugar cores with formic acid to cleave the glycosidic linkage.
194 Comparison of the hydrolyzed and unhydrolyzed samples revealed a peak corresponding to
195 arabinose only in the hydrolyzed plant samples (Figure S4B). Taken together, the results of
196 saponification with and without hydrolysis, followed by derivatization, confirmed identification
197 of *myo*-inositol sugar cores and identified the pentose moiety of the hexose-pentose disaccharide
198 core as arabinose.

199

200 NMR analysis of eight eggplant acylsugars

201 While MS analysis provided valuable information about acyl chain number and length,
202 sugar core mass, as well as acyl chain number, complementary information about acyl chain
203 branch structure, sugar core stereochemistry, and acyl chain positions was obtained using NMR.
204 We purified and resolved the structures of eight abundant eggplant acylsugars from accession PI
205 555598 using a combination of 1D and 2D NMR experiments (Figure 1). All eight structures are
206 newly described, and because atomic connections were determined by NMR and MS data, the

207 proposed structures meet Metabolomics Standards Initiative level 1 criteria for metabolite
208 identification (Sumner et al., 2007).

209 A series of NMR experiments confirmed that the acylhexose and acyldisaccharide sugar
210 cores are *myo*-inositol and 4-*O*-β-arabinopyranosyl *myo*-inositol, respectively (Figure 1C,D;
211 Supplemental Data File S1). We assigned all sugar ring proton signals of each sugar core's spin
212 system with total correlation spectroscopy (TOCSY). Inferences from correlation spectroscopy
213 (COSY) data then identified the order of ring protons and revealed pyranose and cyclitol ring
214 structures for the pentose and hexose rings, respectively. Relative stereochemistry at ring
215 positions was subsequently determined through comparison of spin-spin splitting (multiplicities
216 and coupling constants) referenced to expected patterns inferred from chemical principles and
217 previously reported acylinositols (Hurney, 2018; Leong et al., 2020). The disaccharide glycosidic
218 linkage position was determined by heteronuclear multiple bond correlation (HMBC)
219 correlations to be at position 4 and position 1 of *myo*-inositol and arabinose, respectively. The
220 arabinose *β* anomeric configuration was inferred from the anomeric carbon $^{1}J_{CH}$ (162 Hz) as
221 revealed by a coupled-heteronuclear single quantum coherence (coupled-HSQC) experiment of
222 the free disaccharide sugar core after saponification. Our sugar core assignments identifying this
223 unusual disaccharide agreed with the previous GC-MS sugar core results, supporting the efficacy
224 of the sugar core GC-MS identification method (Figure S4) This disaccharide core differs in the
225 pentose moiety identity from a recently reported acylated 4-*O*-*β*-xylopyranosyl *myo*-inositol in
226 the *Solanum* Clade II species *S. quitoense* (Hurney, 2018). To our knowledge, this is the first
227 report of acylated 4-*O*-*β*-arabinosyl *myo*-inositol sugars.

228 Acyl chain positions, branching patterns, and hydroxyl positions were also resolved
229 through integration of different NMR experiments. We found that all acyl chains were confined

230 to *myo*-inositol, consistent with the LC-MS results. All eight acylsugars are decorated with two
231 short iso-branched iC4 acyl chain esters at positions 1 and 2, and one medium C8 to C14 acyl
232 chain ester at position 3 (Figure 1). Compound 2 (Figure 1C), the only tetraacylated acylsugar
233 identified in eggplant, additionally carried an acetylation at position 6 (Figure 1C). The medium-
234 length acyl chains at position 3 were resolved as straight (nC14) or terminally iso-branched (iC8,
235 iC10, iC12, iC14) based on signals characteristic of protons near the branched carbons.
236 Strikingly, we identified unusual hydroxylated straight and iso-branched 3-
237 hydroxytetradecanoate acyl chains, 3-OH-nC14 and 3-OH-iC14, in compounds 7 and 8,
238 respectively (Figure 1D). We assigned hydroxylation positions of 3-OH-nC14 and 3-OH-iC14 to
239 the third acyl carbon based on a downfield shifted signal in the ^{1}H NMR spectrum at 3.92 ppm,
240 corresponding to one hydrogen at that position. In contrast, the non-hydroxylated medium acyl
241 chains observed in compounds 1-6 (Figure 1C,D), carry two hydrogen atoms at the third acyl
242 carbon, and these have a characteristic signal near 1.50 ppm. We are not aware of previous
243 reports of 3-OH-nC14 and 3-OH-iC14 chains in Solanaceae acylsugars. While 3-OH-nC14
244 chains are observed in the acylsugar-like bacterial Lipid A glycolipids, the hydroxylation
245 position differentiates these eggplant chains from hydroxyacyl chains in castor bean (*Ricinus*
246 *communis*) seed oil, *Silene gallica* gallicasides, *Ibicella lutea* fatty acid glycosides, and
247 Convolvulaceae resin glycosides (Asai et al., 2010; Asai and Fujimoto, 2010; Bah and Pereda-
248 Miranda, 1996; Smith, 1971).

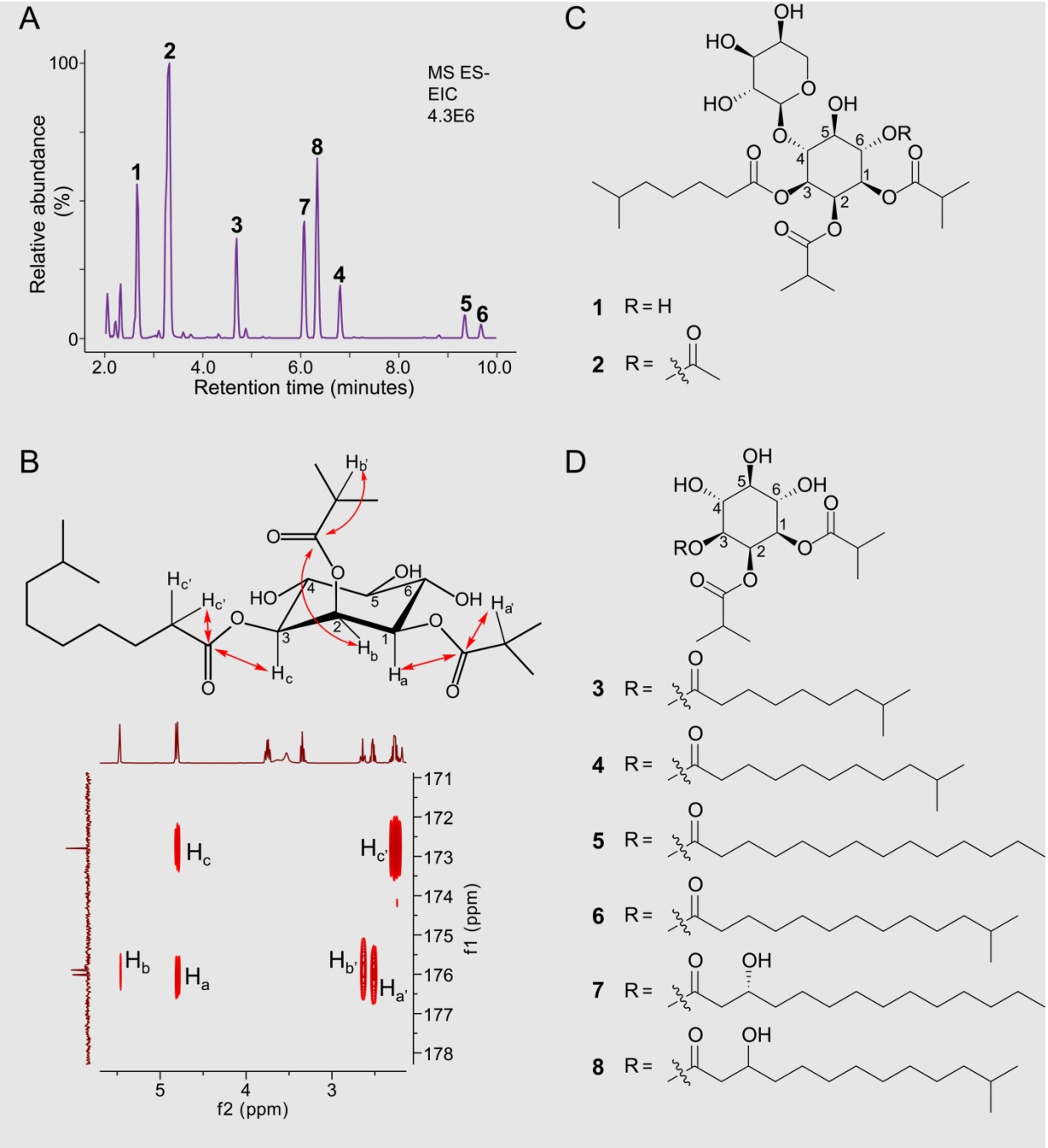


Figure 1. Profiling of *S. melongena* acylsugars. (A) LC-MS base peak intensity (BPI) chromatogram of *S. melongena* acylsugars. Peaks characterized by NMR are indicated with their compound numbers from Panels C and D. Metabolites were collected from seedling tissue with glandular trichomes similar to the ones displayed in the closeup photo of a *S. melongena* hypocotyl. (B) Annotation of HMBC experiments used to identify acylation positions of specific acyl chains. At the top, I3:18(4,4,10) is displayed in the chair conformation along with hydrogens important for determining acylation position. An HMBC NMR spectrum showing couplings between hydrogens and carbonyl carbons is shown. (C) NMR-characterized

258 acyldisaccharide structures 1 and 2, with inositol ring carbons numbered. (D) NMR-
259 characterized acylinositol monosaccharide structures 3-8.
260

261 The OH-tetradecanoate acyl chain has *R* stereochemistry

262 To gain insight into the biosynthetic origin of the hydroxylated acyl chains, we tested
263 chirality of the 3-OH-nC14 acyl chain of compound 7 as this chain has available commercial
264 standards. Two-step chiral derivatization and GC-MS analysis was performed on compound 7
265 and two commercial standards, (3*R*)-OH-nC14 and the corresponding racemate (3*R*/*S*)-OH-
266 nC14. Fatty acids were converted to ethyl esters to aid volatilization, followed by esterification
267 of the acyl chain hydroxyl group with the Mosher acid (*R*)-(−)-α-methoxy-α-
268 (trifluoromethyl)phenylacetate (*R*-MPTA) to distinguish stereoisomers by GC separation of the
269 diastereomeric derivatives. GC-MS analysis of the derivatized (3*R*/*S*)-OH-nC14 standard
270 compounds yielded two peaks at retention times of 42.21 and 42.40 minutes, while one peak at
271 42.40 minutes was obtained for (3*R*)-OH-nC14 (Figure 2). We observed expected fragment ions
272 corresponding to the MPTA fragment ion, fatty acid ethyl ester, and fatty acid which supported
273 peak assignment (Figure S5). Thus (3*S*)-OH-nC14 and (3*R*)-OH-nC14 elute at 42.21 and 42.40
274 minutes, respectively. The compound 7 acyl chain was assigned as (3*R*)-OH-nC14 as its
275 derivatized 3-OH-tetradecanoate acyl chain eluted at 42.40 minutes (Figure 2).

276 The (3*R*)-OH-nC14 chain hydroxyl position and stereochemistry implicate possible
277 metabolic pathways leading to its biosynthesis. As fatty acid hydroxylases typically act on the
278 acyl chain terminal region, they are unlikely to catalyze C3 hydroxylation of a C14 fatty acid
279 (Pinot and Beisson, 2011). Of the other well-characterized pathways, only mitochondrial and
280 plastidial fatty acid metabolism mechanisms produce (3*R*)-OH-acyl-thioester intermediates.
281 Consistent with this result, previous work identified two trichome-expressed, mitochondrial

282 enzymes, Acylsugar Enoyl-CoA Hydratase 1 (AECH1) and Acylsugar Acyl-CoA Synthetase 1
283 (AACS1), and one trichome-expressed, plastidial enzyme, a beta-ketoacyl-(acyl-carrier-protein)
284 reductase (SpKAR1), involved in medium C10 and C12 acyl chain production in cultivated
285 tomato and its wild relative, *S. pennellii* (Fan et al., 2020; Ji et al., 2022).

286 We hypothesize that acylsugar hydroxyacyl chains evolved through one of three
287 scenarios. First, changes in acyl-CoA substrate availability, possibly through the actions of
288 thioesterases or acyl-CoA synthetases, enabled a promiscuous ASAT to esterify hydroxyacyl
289 chains to acylinositols. Alternatively, substrate specificity of an existing acylinositol ASAT
290 expanded to utilize available hydroxyacyl-CoAs. Like this second hypothesis, amino acid
291 substitutions between copies of ASAT2 from different tomato species impacted the accumulation
292 of differently branched five carbon acyl chains on acylsugars (Fan et al., 2015). A third
293 possibility is that a 'new' acyltransferase, unrelated to characterized ASATs and capable of
294 utilizing hydroxyacyl-CoA substrates, was co-opted into acylinositol biosynthesis. Future genetic
295 and biochemical studies should reveal the mechanisms behind (3*R*)-OH acyl chain production,
296 potentially providing new approaches to engineer unique acylsugars with hydroxyacyl chains
297 and hydroxylated fatty acids for polymeric building blocks. Acylsugar structures can impact their
298 function, thus we hypothesize that engineered tomato plants producing these unusual
299 hydroxyacyl chains on their acylsugars may exhibit different defense capabilities than the
300 wildtype acylsucroses.

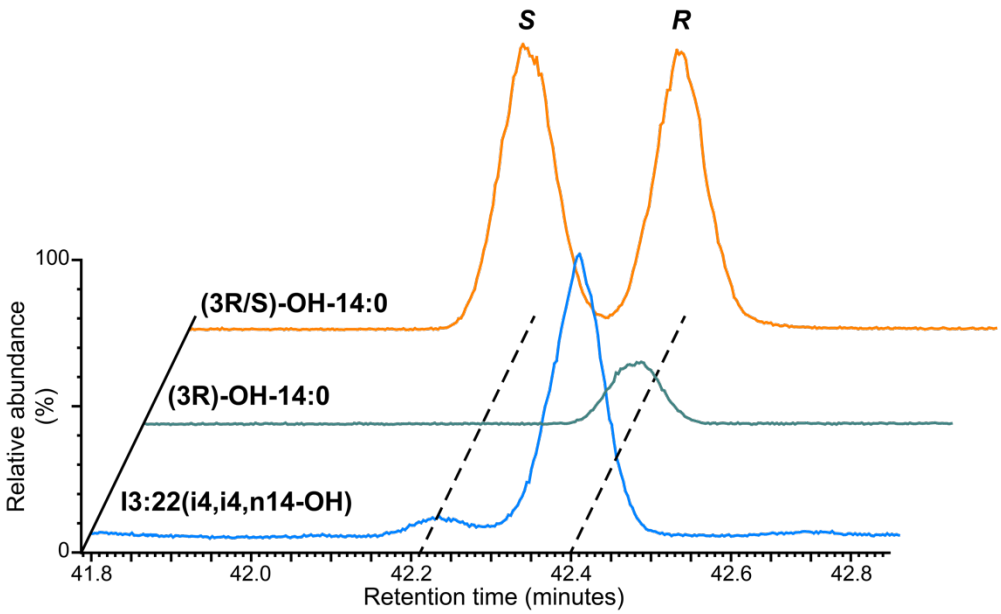
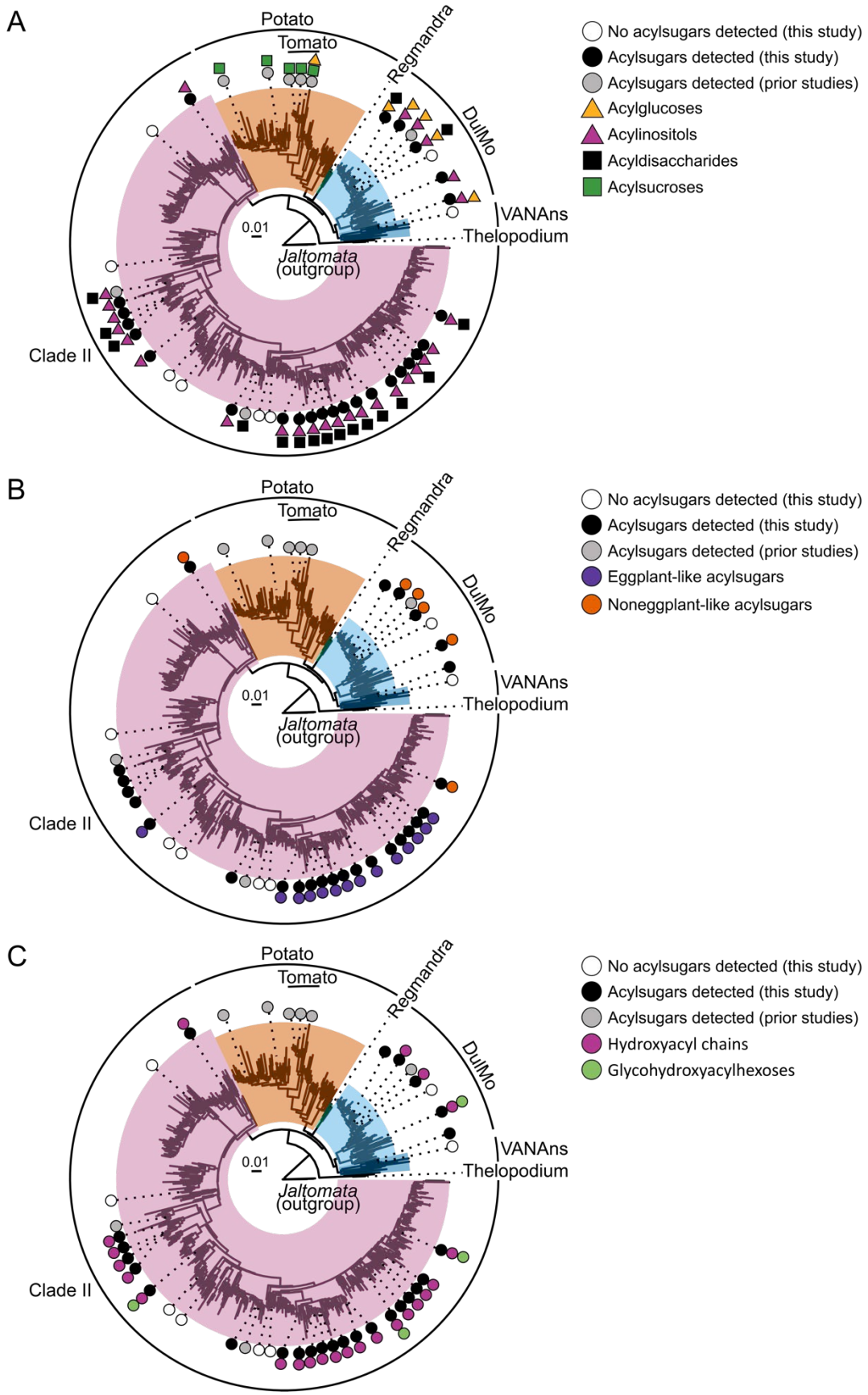


Figure 2. Evidence that the 3-OH-n14 acyl chain on Compound 7, I3:22(i4,i4,n14-OH) is in the *R* configuration. Hydroxyacyl chains analyzed by GC-selected ion monitoring-MS (GC-SIM-MS) after ethyl transesterification and derivation to their diastereomeric (*S*)-MPTA form. Fragment ions with *m/z* of 189, 209, and 255 were monitored, and relative abundance was normalized to the highest peak. Yellow and green traces represent derivatives from commercial (3*R*/*S*)- and (3*R*)-OH-14:0 standards. The main derivatized 3-OH-14:0 acyl chain from Compound 7 (Figure 1), shown with the blue trace, comigrates with the peak corresponding to the *R* enantiomer.

Enormous *Solanum* acylsugar diversity revealed by LC-MS metabolite screening

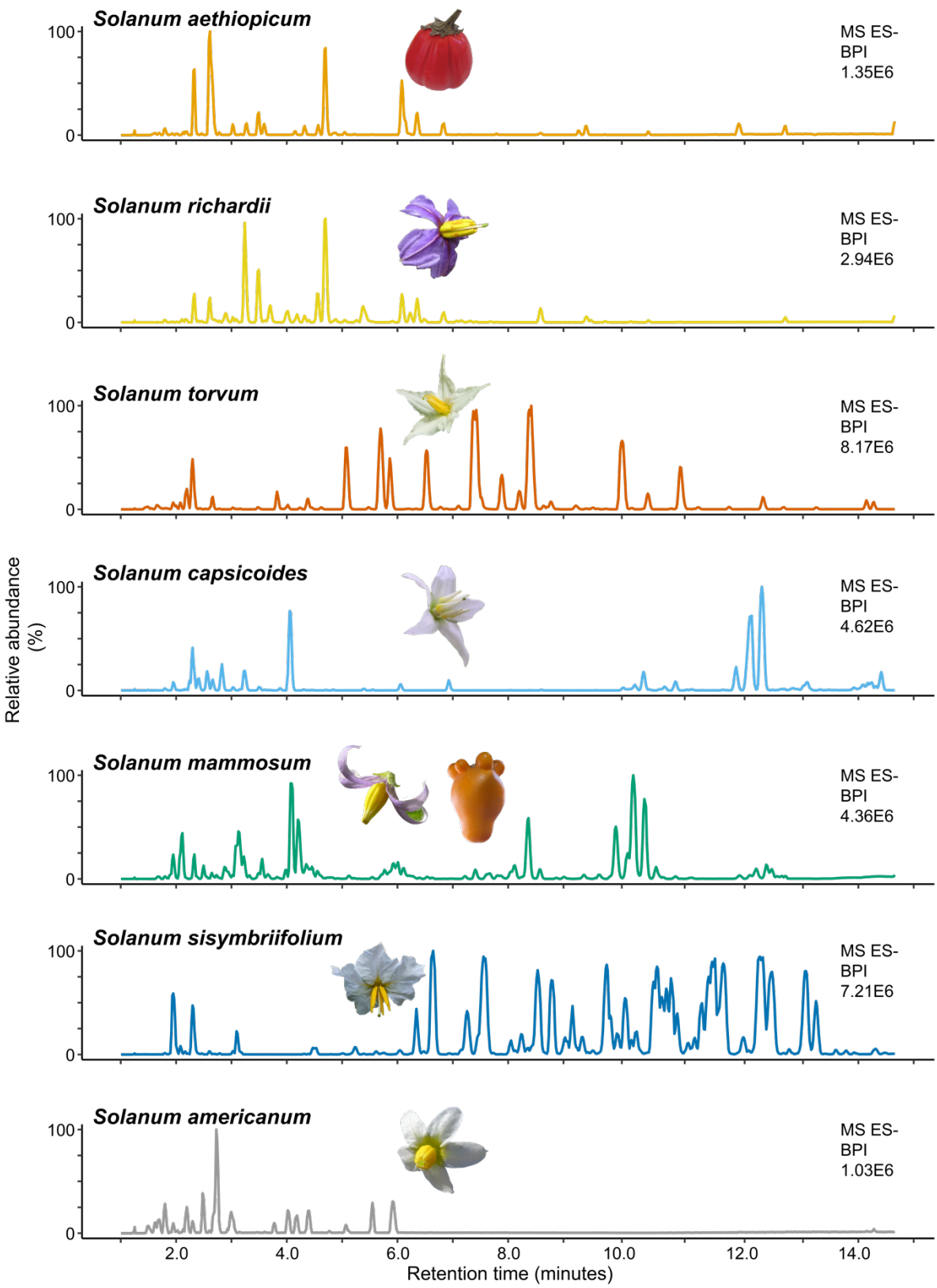
To date, most acylsugar screening across the *Solanum* has focused on cultivated tomato and its Potato clade relatives (Fan et al., 2017; Ghosh et al., 2014; Lybrand et al., 2020; Schilmiller et al., 2016), which represent a small fraction of the species diversity in this genus (PBI Solanum Project, 2022). Though limited in number, studies profiling non-Potato clade species, including *S. nigrum* (Lou et al., 2021), *S. lanceolatum* (Herrera-Salgado et al., 2005), *S. quitoense* (Hurney, 2018; Leong et al., 2020), and *S. melongena* (this work) demonstrate acylsugar diversity across the *Solanum* is vast, and largely uncharacterized. To address this knowledge gap, we analyzed tissue surface extracts from 30 additional *Solanum* species,

320 including 23 Clade II, five DulMo clade, and two VANAns clade members (listed in Table S1).
321 While NMR is the gold standard for structural elucidation, it is prohibitively time consuming for
322 a largescale metabolite diversity survey. LC-MS coupled with CID provides substantial
323 structural information and is compatible with high-throughput screening. To strike a balance
324 between data quality and quantity, we performed LC-MS-CID screening on all 30 species and
325 annotated acylsugars as described for *S. melongena*. Twenty-five of the 30 analyzed species
326 produced detectable acylsugars from visible surface glandular trichomes, while species lacking
327 acylsugars had no or few observable glandular trichomes in the tissues analyzed. Targeted
328 analysis of leaf surface extracts from 31 species, including *S. melongena*, and fruit surface
329 extracts from two species, uncovered previously unreported acylsugars, inter-species acylsugar
330 differences, and identified the distributions of unusual acylsugar chemical traits in the *Solanum*
331 genus (Figure 3; Tables 1 and S2-27). Although we did not observe intraspecific acylsugar
332 variation, acylsugars varied enormously between species (Figure 4; Tables 1 and S2-27).


A
Potato
Tomato
Regmandra
DulMo
VANAns
Thelopodium
0.01
Jaltomata
(outgroup)
Clade II
No acylsugars detected (this study)
Acylsugars detected (this study)
Acylsugars detected (prior studies)
Acylglucoses
Acylinositols
Acyldisaccharides
Acylsucroses
B
Potato
Tomato
Regmandra
DulMo
VANAns
Thelopodium
0.01
Jaltomata
(outgroup)
Clade II
No acylsugars detected (this study)
Acylsugars detected (this study)
Acylsugars detected (prior studies)
Eggplant-like acylsugars
Noneggplant-like acylsugars
C
Potato
Tomato
Regmandra
DulMo
VANAns
Thelopodium
0.01
Jaltomata
(outgroup)
Clade II
No acylsugars detected (this study)
Acylsugars detected (this study)
Acylsugars detected (prior studies)
Hydroxyacyl chains
Glycohydroxyacylhexoses

333

334 **Figure 3. Distribution of acylsugar traits within the *Solanum* genus.** Traits were mapped onto
335 a previously published Supermatrix maximum likelihood phylogeny (Gagnon et al., 2022). The
336 reported polytomy between Clade II, DulMO, VANANs, Regmandra, and Potato clades is not
337 indicated (Gagnon et al., 2022). Acylsugar trait data from this report included annotations based
338 upon LC-MS and NMR analysis (Figure 1, Tables 1 and S3-27), while data from previous
339 studies were limited to those with NMR structural data (Ghosh et al., 2014; Herrera-Salgado et
340 al., 2005; Hurney, 2018; King et al., 1986; Leong et al., 2020, 2019; Lou et al., 2021; Lybrand et
341 al., 2020). (A) Distribution of sugar core types within the *Solanum* genus. Inositol and glucose
342 cores determined by characteristic LC-MS negative mode electrospray ionization (ESI)
343 fragmentation (Figures S2 and S10). (B) Distribution of species with acylsugars that coelute with
344 eggplant acylsugars containing the same acyl chains and sugar core. (C) Distribution of
345 hydroxyacyl chains and glycosylated hydroxyacyl chains.


Solanum aethiopicum
MS ES-
BPI
1.35E6
Solanum richardii
MS ES-
BPI
2.94E6
Solanum torvum
MS ES-
BPI
8.17E6
Solanum capsicoides
MS ES-
BPI
4.62E6
Solanum mammosum
MS ES-
BPI
4.36E6
Solanum sisymbriifolium
MS ES-
BPI
7.21E6
Solanum americanum
MS ES-
BPI
1.03E6
Relative abundance
(%)
100
0
2.0
4.0
6.0
8.0
10.0
12.0
14.0
Retention time (minutes)

346

Figure 4. *Solanum* species produce dramatically different leaf surface metabolite profiles. The seven representative species display morphological differences, as demonstrated by the flower and fruit images, and they also exhibit diverse metabolite profiles as demonstrated by base peak intensity LC-MS chromatograms from their leaf surface metabolites. Most of the peaks shown consist of acylsugars with varied sugar cores and acyl chain types and numbers. Acylsugar profiles for these species are detailed in Tables S5, S8, S18, S21, and S25-27.

Identification of diverse acyl chain types in *Solanum* acylsugars

Our survey of *Solanum* species uncovered acylsugars with a surprising diversity of acyl chain lengths, functional groups, and combinations of chain types and positions. We identified C2 to C18 acyl chains, including unsaturated and hydroxylated forms, based on analysis of fatty acid fragment carboxylate ions in negative mode and neutral losses in both negative and positive mode (Table S2). Acyl chain compositions differed between acylsugars within and across species: acylinositols with primarily medium C8 to C14 acyl chains were observed in five species (Figure 3, Tables S14, S15, S23, S25), while acylsugars with two short C4 to C6 acyl chains and one medium C8 to C14 acyl chain were observed in 17 species, an acylation pattern also observed in eggplant and wild tomato (Figure 1; Table 1) (Lybrand et al., 2020). Acylsugars with unsaturated acyl chains, including C5:1, C18:1, C18:2, and C18:3, were detected only in *S. torvum* (Table S18). To the best of our knowledge, this is the first report of unsaturated C18 chains in Solanaceae acylsugars. In contrast, hydroxylated medium-length acyl chains were surprisingly common: we annotated acylsugars bearing hydroxylated acyl chains in three DulMo clade species and 25 Clade II species (Figure 3C). These hydroxylated acyl chains are 'chemical handles' susceptible to further modifications, including glycosylations and acylations, which would result in even greater acylsugar diversity.

Indeed, we obtained evidence for glycosylated acyl chains in three Clade II species, *S. prinophyllum*, *S. sisymbriifolium*, *S. lasiophyllum*, and one DulMo species, *S. dulcamara* (Tables

373 S12, S13, S24 and S27). Several lines of evidence led us to conclude that these species
374 accumulate acylsugars with glycosyl groups attached via ether linkages to medium hydroxyacyl
375 chains esterified to the hexose core. First, the compounds fragmented abnormally under negative
376 mode CID: the glycosyl group was observed as a neutral loss from the $[M-H]^-$ ion, producing a
377 $[sugar – H]^-$ fragment ion (Figure S6). This contrasts with extensive observations that CID
378 spectra of other characterized acyldisaccharides do not exhibit negative mode disaccharide
379 glycosidic linkage cleavage (Figures S2 and S7) (Ghosh et al., 2014; Hurney, 2018; Lou et al.,
380 2021). Further evidence that the glycosylation is not on the primary acylated hexose ring was
381 obtained by subjecting acylsugar extracts to saponification, which induces ester linkage breakage
382 while retaining ether linkages. Upon analyzing saponified acylsugar extracts by LC-MS, we
383 detected pentose and hexose sugars with C14 or C16 acyl chains rather than glycosylated
384 hexoses, indicating that the glycosyl group is connected to a medium hydroxyacyl chain by an
385 ether linkage that is not cleaved during alkaline saponification (Figure S8). We named these
386 compounds glycohydroxyacylhexoses and annotated them by modifying the conventional
387 acylsugar nomenclature where the molecule in Figure S6 is named H3:22(4,4,14-O-p) with 14-
388 O-p representing the pentosylated (p) C14 hydroxyacyl chain (14-O). To our knowledge, this is
389 the first report of acylhexoses with glycosylated hydroxyacyl chains in Solanaceae acylsugars.
390 Resin glycosides produced by Convolvulaceae family species contain similar hydroxyacyl chain
391 glycosidic linkages. However, in resin glycosides, hydroxyacyl chains are connected to the
392 oligosaccharide core at two points – by an ester and ether linkage, respectively – forming a
393 macrocyclic structure not observed in the *Solanum* glycohydroxyacylhexoses (Bah and Pereda-
394 Miranda, 1996; Kruse et al., 2022; Pereda-Miranda et al., 1993).

395 The unusual glycohydroxyacylhexoses raises questions regarding their biosynthetic and
396 evolutionary origins. As these compounds are found in different lineages (Eastern Hemisphere
397 Spiny, Dulcamaroid, and Sisymbriifolium sections) – and closely related species lack detectable
398 glycohydroxyacylhexoses – this trait appears to be highly labile (Figure 3C). Because we
399 observe the cognate, non-glycosylated acylinositol in each of these species, we hypothesize that
400 glycohydroxyacylhexoses consist of a *myo*-inositol hexose core. Further analysis of acylsugars
401 from additional species may identify other types of modified acylsugar hydroxyacyl chains in
402 *Solanum* acylsugars.

403 Sixteen species accumulate acylsugars with the same molecular masses, sugar core
404 masses, and acyl chain complements as the NMR-characterized eggplant acylinositols,
405 suggesting they may be identical. To test this hypothesis, we mixed purified eggplant NMR-
406 resolved standards with metabolite extracts from each species and performed liquid co-
407 chromatography. Coelution provided one line of evidence that the eggplant-like acylsugars
408 detected in 11 Clade II species are identical to characterized eggplant acylinositols (Tables S5-
409 14, S27). In contrast, the eggplant-like acylsugars from five other species, including four DulMo
410 clade members and one Clade II member, eluted separately from characterized eggplant
411 acylinositols, indicating they are acylation positional isomers or acyl chain branching isomers
412 (Figure S9). Examples of acylinositol positional isomers were described between the DulMo
413 clade species *S. nigrum* and the Clade II species *S. quitoense* and *S. melongena*, providing
414 precedent that the non-coeluting compounds might be positional isomers of the eggplant
415 acylsugars shown in Figure 1 (Leong et al., 2020; Lou et al., 2021). Our data reveal a
416 phylogenetic pattern of acylinositol distribution in which Clade II species generally accumulate
417 eggplant-like acylinositol isomers, whereas DulMo clade species accumulate noneggplant-like

418 acylinositol isomers (Figure 3B). While the eggplant-like isomers coelute with eggplant
419 acylsugars, further analysis with complementary analytical methods is needed to confirm
420 whether these are structurally identical across species.
421
422 Acylhexoses and acyldisaccharides accumulate throughout Clade II, DulMo, and VANAns
423 *Solanum* species
424 Acylhexoses were detected in all analyzed acylsugar-producing species. They were
425 annotated as acylglucoses or acylinositols based upon differences in characteristic fragmentation
426 behavior under negative ion mode MS conditions (Leong et al., 2020; Lou et al., 2021).
427 Identification of acylglucoses was based on observation of a sugar core fragment ion of *m*/*z*
428 143.03 ($C_6H_7O_4^-$) and stepwise acyl chain losses at the lowest collision energy level (0 V)
429 (Figure S11). In contrast, acylinositol annotation was based on an absence of acyl chain fragment
430 ions at 0 V and MS^e CID negative mode functions (Figure S2). Using these rules, we identified
431 acylinositols in all acylsugar-producing species tested except for *S. americanum* (Tables 1, S2-
432 27). In contrast, acylglucoses were not detected in any Clade II species but were commonly
433 detected among the DulMo and VANAns species, found in all analyzed acylsugar-producing
434 species except *S. dulcamara* (Tables S2, S19, S20, S23, S24, S26; Figure 3A; Lou et al., 2021).
435 Differing from the widespread acylhexose distribution, acyldisaccharides were detected
436 sporadically in Clade II and DulMo clade species (Figure 3A), being observed in 14 of 21 Clade
437 II and 2 of 4 DulMo acylsugar-producing analyzed species. Based on negative $[M-H]^-$ fragment
438 masses and positive mode fragmentation of the glycosidic linkage, the acyldisaccharides were
439 composed of hexose-hexose, pentose-hexose, or deoxyhexose-hexose sugar cores (Table S2,
440 Figures S2 and S7). Though complete structural information cannot be obtained from MS

441 fragmentation alone, two lines of evidence lead us to hypothesize that these acyldisaccharides are
442 composed of glycosylated inositol cores. First, most of the species with detectable
443 acyldisaccharides also accumulate cognate acylinositols with the same acyl chain lengths (leaf
444 and fruit surface extracts from 10 of 16 and two of two different species, respectively),
445 suggesting a shared inositol-containing core structure (Tables 1, S3-9, S12-17, S20, S22).
446 Second, acylinositol glycosides with varying sugar core sizes and stereochemistries were
447 described in *S. lanceolatum* (Herrera-Salgado et al., 2005), *S. quitoense* (Hurney, 2018), and *S.*
448 *melongena* (Figure 1), suggesting that the species in this study may accumulate similar
449 compounds. The ratio of acylhexose to acyldisaccharide peak number and abundance differs
450 across the surveyed species; for example, *S. mammosum* only accumulated detectable
451 acylinositols while *S. anguivi* primarily accumulated acyldisaccharides that comprised 80% of
452 the total acylsugar peak area (Tables S21 and S4). This is reminiscent of varied acylhexose
453 accumulation observed in trichomes of evolutionarily distant *Solanum* species: *S. nigrum* only
454 contains detectable acylinositol and acylglucose monosaccharides (Lou et al., 2021), while *S.*
455 *pennellii* accessions have mixtures of acylglucoses and acylsucroses ranging from 41-95% of
456 total acylsugars as acylglucoses. (Lybrand et al., 2020; Shapiro et al., 1994).
457
458

459 Sticky fruits accumulate acylinositols

460 We investigated the chemical basis of sticky fruit surface substances previously reported
461 in botanical species descriptions of the Clade II species *S. acerifolium* and *S. atropurpureum*
462 (Nee, 2022a, 2022b, 1991). LC-MS analysis of fruit surface extracts revealed similar acylsugar
463 profiles between the two species, with each extract containing what appeared to be the same 128
464 acylsugars. These fruit acylsugars were distinct from the 22 and 19 trichome acylsugars
465 identified in *S. acerifolium* and *S. atropurpureum*, respectively, as evidenced by their LC elution
466 and MS characteristics (Tables S14-17; Figure 5). Seventy-seven fruit acylsugars were identified
467 as acylinositols with two or three medium C8 to C14 acyl chains, including one to three
468 hydroxyacyl chains. The remaining 51 fruit acylsugars are acyldisaccharides containing an
469 unusual deoxyhexose-hexose core not previously reported in Solanaceae acylsugars. In contrast,
470 *S. acerifolium* and *S. atropurpureum* trichome acylsugars, composed of a pentose-hexose
471 disaccharide core decorated with short C5 or C6 acyl chains, were not detected in fruit surface
472 extracts.
473

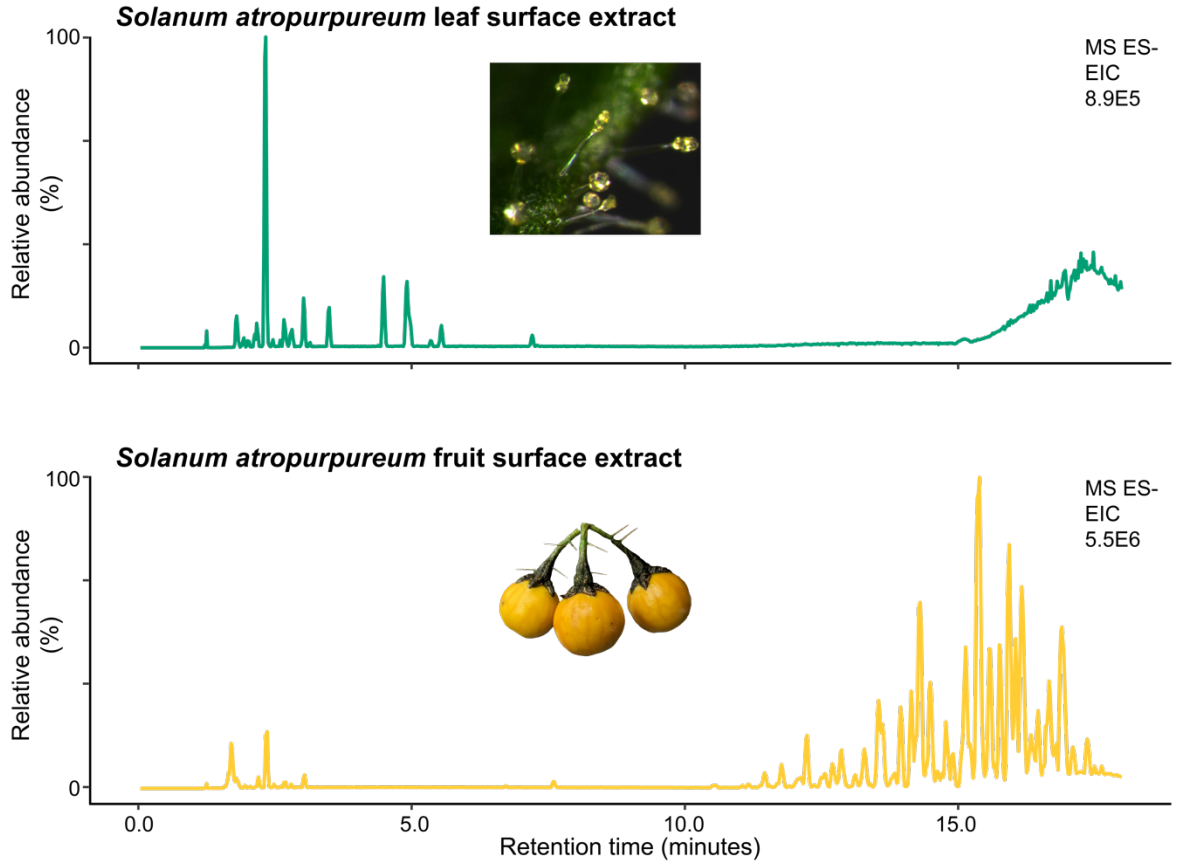


Figure 5. *S. atropurpureum* trichomes and fruit exhibit dramatically different metabolite profiles. The top LC-MS base peak intensity (BPI) chromatogram displays metabolites from a leaf surface extract, while the bottom LC-MS base peak intensity chromatogram displays metabolites from a fruit surface extract. The leaf surface extract acylsugars elute from 2.2 to 7.6 min (Table S17), while fruit surface extract acylsugars elute from 4.6 to 18.4 min (Table S15).

S. mammosum and *S. capsicoides,* which are closely related to *S. acerifolium* and *S. atropurpureum*, do not produce sticky fruit surfaces, suggesting this trait evolved in the common ancestor of *S. acerifolium* and *S. atropurpureum,* perhaps as recently as 2 Mya (Särkinen et al., 2013). Fruit surface acylsugars were described previously within the *Physalis* genus and likely represent an independent evolutionary origin from the *S. acerifolium* and *S. atropurpureum* fruit acylsugars (Bernal et al., 2018; Cao et al., 2015; Cicchetti et al., 2018; Maldonado et al., 2006). Considering that, like eggplant, *S. acerifolium* and *S. atropurpureum* fruits lack trichomes (Nee, 2022a, 2022b), some other organ or cell type presumably synthesizes the fruit surface acylsugars.

488 Structures similar to *Solanum fernadesii* petiolar resin glands (Sampaio et al., 2021), *Hypericum*
489 *androsaemum* microscopic fruit glands (Caprioli et al., 2016), and fennel and chamomile fruit
490 secretory ducts (vittae) (Zizovic et al., 2007) may be involved in *S. acerifolium* and *S.*
491 *atropurpureum* fruit acylsugar secretion. Future work will be needed to understand the cellular,
492 biosynthetic, and evolutionary relationships between trichome and fruit acylsugars.
493
494 Sugar core evolution across the Solanaceae
495 Our screening revealed acylinositols to be broadly distributed across Clade II, DulMo,
496 and VANAns (Figure 3A, Table S1). Outside of this study, acylinositols were only reported in
497 *Solanum* species, including two Clade II members, *S. quitoense* and *S. lanceolatum*, and one
498 DulMo clade member, *S. nigrum*. The acylinositol distribution suggests that acylinositol
499 biosynthesis arose one or more times within the *Solanum* genus. However, our ability to
500 elucidate the likely number of acylinositol origins is impeded by the lack of resolution of the
501 phylogenetic relationships between the major *Solanum* clades, including Clade II, DulMo,
502 VANAns, Potato, and Regmandra (Gagnon et al., 2022). Further biochemical and genetic
503 analyses of acylinositol biosynthesis may reveal how this trait evolved, and whether the pathway
504 is conserved between Clade II, DulMo, and VANAns species.
505 Acylglucoses were reported to be sporadically present across the Solanaceae, including in
506 *Salpiglossis sinuata*, *Petunia*, *Nicotiana* spp., *Datura*, the DulMo clade *S. nigrum*, and Potato
507 clade *S. pennellii* (Castillo et al., 1989; Chortyk et al., 1997; Fiesel et al., 2022; Fobes et al.,
508 1985; Hurney, 2018; King and Calhoun, 1988; Lou et al., 2021; Matsuzaki et al., 1989; Schenck
509 et al., 2022; Van Dam and Hare, 1998). Recent work in *S. pennellii* and *S. nigrum* revealed that
510 acylglucoses are synthesized from acylsucroses by a neofunctionalized invertase-like enzyme,

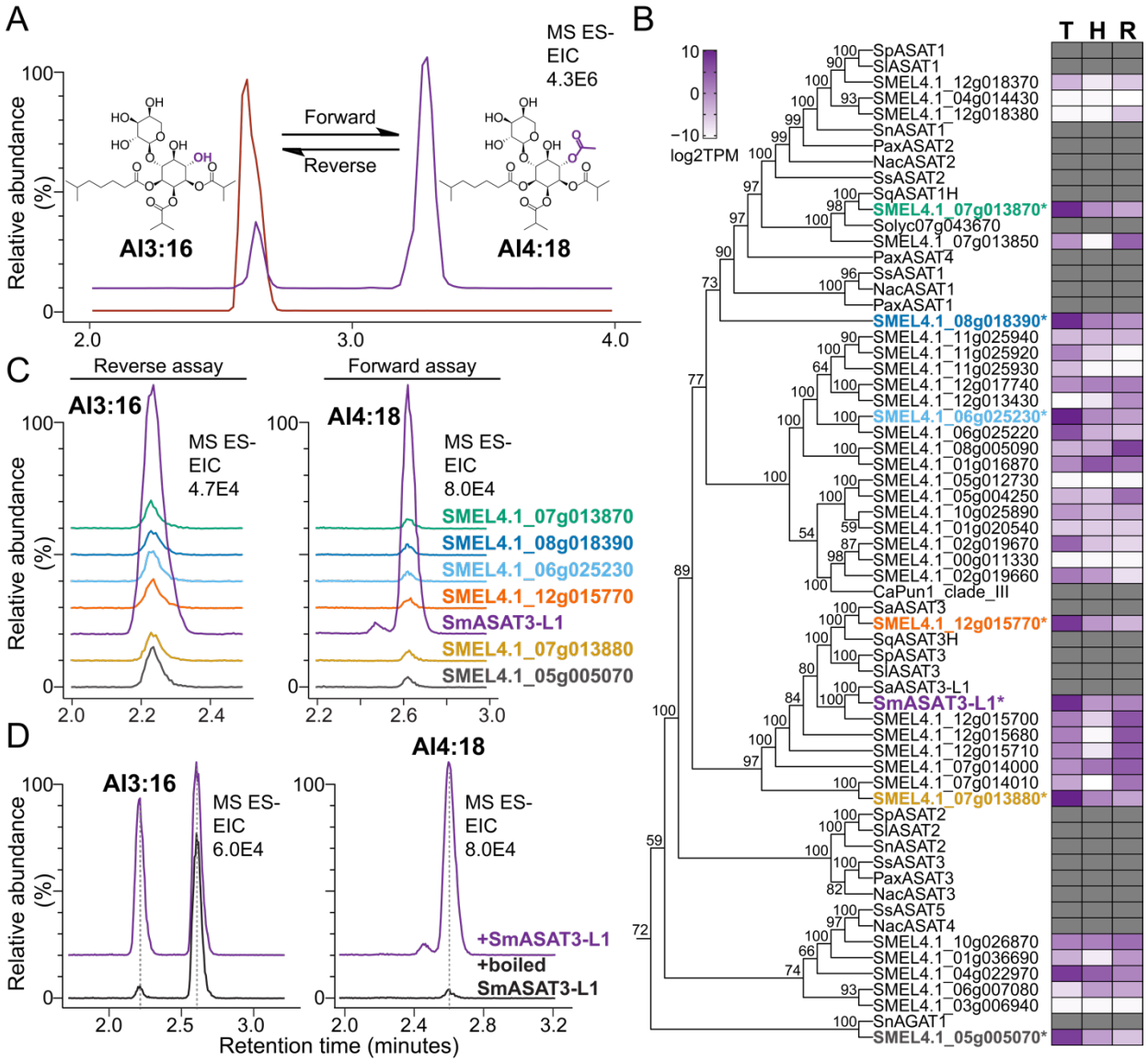
511 AcylSucrose FructoFuranosidase 1 (ASFF1) (Leong et al., 2019; Lou et al., 2021). These results
512 show that conversion of acylsucroses to acylglucoses arose at least twice in *Solanum* because *S.*
513 *nigrum* and *S. pennellii* employ non-orthologous ASFF1 enzymes. Our current study extends
514 documented cases of *Solanum* acylglucose distribution to the VANAns clade, as well as
515 documenting it in additional DulMo species. In contrast, we found no acylglucose accumulating
516 species among the 23 analyzed *Solanum* Clade II species. Elucidation of acylglucose
517 biosynthesis outside of the *Solanum* clade will be needed to resolve the number of evolutionary
518 origins and biochemical mechanisms leading to synthesis of these metabolites across the
519 Solanaceae.

520 Varied acyldisaccharides were observed across DulMo and Clade II, and we predict they
521 are biosynthesized from acylinositols. Within the Potato clade as well as outside the *Solanum*
522 genus, acylsucroses are the dominant acyldisaccharide in the reported species and tissues (Fiesel
523 et al., 2022). In contrast, no acylsucroses were detected in any species screened. However,
524 acylsucrose biosynthesis might be found in DulMo and VANAns species because acylglucoses
525 were detected in all but one DulMo and VANAns species and that acylsucroses can be
526 intermediates in acylglucose biosynthesis (Leong et al., 2019; Lou et al., 2021). The lack of both
527 acylglucoses and acylsucroses in the analyzed Clade II species suggests a loss of acylsucrose
528 biosynthesis in Clade II. Rather, our LC-MS data are consistent with the hypothesis that the
529 acyldisaccharides observed in this study consist of glycosylated inositol cores. Recent *in vitro*
530 biochemistry and *in vivo* genetic evidence in *S. quitoense* suggests that acylinositols are the
531 precursors to acylinositol disaccharides in Clade II *Solanum* species (Leong et al., 2022, 2019).
532 In this scenario, acylinositols would be converted to their cognate glycosides by one or more
533 glycosyltransferases, with different sugars added either by a conserved promiscuous

534 glycosyltransferase or multiple glycosyltransferases. *S. quitoense* and *S. melongena*, with their

535 acylinositol glycosides bearing different glycosyl moieties, are promising systems for addressing

536 the biochemical origins of these unusual *Solanum* acyldisaccharides.



537

538 **Figure 6. SmASAT3-L1 acetylates AI3:16 to produce AI4:18.** Trichome expressed ASAT
539 homologs were identified and tested for AI3:16 acetyl transferase activity. (A) LC-MS analysis

540 of AI3:16(4,4,8) and AI4(2,4,4,8) in *S. aethiopicum* (red) and *S. melongena* (purple) leaf surface
541 extracts. *S. aethiopicum* does not produce detectable levels of AI4:18. (B) Shown is a clade III
542 BAHD tree subset from a phylogeny including 106 predicted BAHDs (PF002458) in the
543 eggplant Smel_V4.1 reference genome, published reference BAHD sequences for clades I-VII,
544 characterized ASAT sequences from other selected Solanaceae species, and the SaASAT3 and
545 SaASAT3-L1 candidates from *S. aethiopicum* (see Figure S12 for full phylogeny). The
546 maximum likelihood tree was inferred from amino acid sequences using the Jones-Taylor-
547 Thornton algorithm with seven rate categories in IQ-TREE v2.1.3. Values at nodes indicate
548 bootstrap support calculated from 100,000 ultrafast bootstrap replicates. Heatmap shows absolute
549 transcript abundance (log2 TPM) across trichomes, trichomeless hypocotyls, and roots from 7-
550 day-old eggplant seedlings. Expression data for non-eggplant sequences are not included in the
551 heatmap. Color gradient provides a visual marker to rank the transcript abundance from high
552 (purple) to low (white) or absent (grey). ASAT, acylsugar acyltransferase. log2TPM, log2
553 transformed transcripts per million. Genes denoted with "*" were selected for biochemical
554 characterization based on expression in trichomes and sequence similarity to characterized
555 ASAT enzymes in other species. (C) Of the seven enzymes tested *in vitro*, only SmASAT3-L1
556 successfully acetylated AI3:16(4,4,8) to form AI4:18(2,4,4,8). The extracted ion chromatograms
557 on the left display products from reverse enzyme assays and the formate adduct of AI3:16, *m/z*
558 623.29. The extracted ion chromatograms on the right display products from forward enzyme
559 assays and the formate adduct of AI4:18, *m/z* 665.30. (D) SmASAT3-L1 acetylates AI3:16 *in*
560 *vitro* to produce AI4:18. Forward assay extracted ion chromatograms display the formate adduct
561 of AI4:18, *m/z* 665.30, and reverse assay extracted ion chromatograms display the formate
562 adduct of AI3:16, *m/z* 623.29. The reverse assay chromatograms display peaks for both AI3:16
563 and AI4:18 due to in source fragmentation of remaining AI4:18.

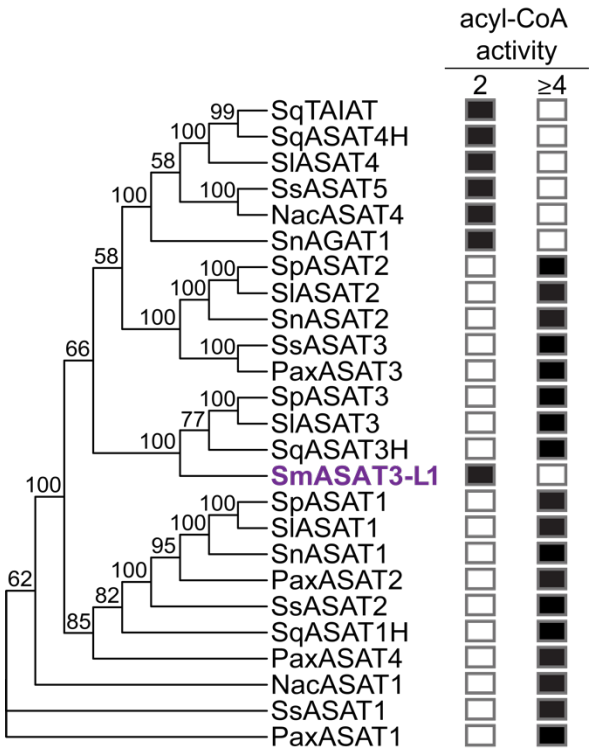
564

565 <u>*S. aethiopicum* is defective in expression of a trichome acylsugar acyltransferase enzyme</u>

566 We began analyzing the biochemical mechanisms underlying the observed *Solanum*
567 acylsugar diversity in seven of 23 members from the Eggplant clade and related Anguivi grade,
568 two small subclades within Clade II that includes brinjal eggplant and its close relatives (Figure
569 S11) (Aubriot et al., 2018). In contrast to the other six analyzed eggplant species, *S. aethiopicum*
570 (scarlet or Ethiopian eggplant) is unique in that it had unacetylated AI3:16(4,4,8) (Compound 1)
571 but did not accumulate detectable levels of the acetylated form, AI4:18(2,4,4,8) – the highest
572 abundance *S. melongena* acylsugar (Compound 2; Figures 6A, S11, Tables 1, S2, S5). We
573 hypothesized that the absence of AI4:18(2,4,4,8) in *S. aethiopicum* is either due to loss-of-
574 function or altered expression of an acetylating enzyme in this species. The availability of *S.*

575 *melongena* and *S. aethiopicum* genomic sequences (Barchi et al., 2021; Song et al., 2019)
576 provided the opportunity to seek a trichome expressed acetyltransferase responsible for
577 producing AI4:18 and the mechanism behind the lack of detectable AI4:18 in *S. aethiopicum*.
578 To identify enzyme candidates, we sequenced transcriptomes of isolated trichomes,
579 trichome-depleted shaved hypocotyls, and whole roots collected from 7-day-old *S. melongena*
580 seedlings and performed differential gene expression analysis (see Methods). Of the 23,251
581 eggplant genes expressed in at least one of the three tissues, 745 were significantly enriched
582 (log2 fold-change >2, *p*-value < 0.05) in trichomes compared to hypocotyls and roots, including
583 20 BAHDs (Tables S37 and S38). As shown in Table S39, we selected for further testing seven
584 BAHDs that were abundantly expressed in trichomes (TPM > 200) and homologous to
585 characterized *Solanum* ASATs (sequences in colored text and denoted with "*" in Figure 6B)
586 (D'Auria, 2006; Fan et al., 2015; Leong et al., 2022, 2020; Lou et al., 2021; Schenck et al.,
587 2022). We expressed each of the seven candidates in *Escherichia coli* and tested for acetylation
588 of Compound 1, AI3:16(4,4,8). An SlASAT3 outparalog (Koonin, 2005) SMEL4.1_12g015780,
589 ACYLSUGAR ACYLTRANSFERASE3-LIKE1 (SmASAT3-L1) was the only enzyme to
590 exhibit detectable forward activity acetylating Compound 1, AI3:16(4,4,8), to Compound 2,
591 AI4:18(2,4,4,8) (Figure 6C). As shown in Figure S13, LC-MS analysis of the SmASAT3-L1 *in*
592 *vitro* assay product revealed that it had identical molecular masses and elution times to plant-
593 derived AI4:18(2,i4,i4,i8) (Compound 1) providing evidence that SmASAT3-L1 acetylates
594 AI3:16(i4,i4,i8). Characterization of reverse activities of the seven BAHD acyltransferases
595 (Leong et al., 2020; Lou et al., 2021), in which an acyl chain can be removed from an acylsugar
596 by incubating it with the enzyme and free Coenzyme A, confirmed that SmASAT3-L1 was the
597 only tested enzyme that removed an acetyl chain from AI4:18(2,4,4,8) (Figures 6D and S13).

598 Taken together, these results indicate that trichome-expressed *SmASAT3-L1* encodes an
599 acylinositol acetyltransferase responsible for AI4:18(2,4,4,8) biosynthesis.
600 We tested the hypothesis that *S. aethiopicum* trichome extracts lack detectable acetylated
601 AI4:18(2,4,4,8) due to a defect in *SmASAT3-L1* ortholog expression. Indeed, reverse
602 transcription (RT)-PCR comparing isolated trichome and trichomeless shaved hypocotyl cDNA
603 revealed that *SaASAT3-L1* (GenBank ID: OQ547782) transcript was undetectable in *S.*
604 *aethiopicum* trichomes (Figure S14). This result implicates the defect in ASAT3-L1 enzyme
605 gene expression as responsible for lack of AI4:18 in *S. aethiopicum*. The interspecific variation
606 in *ASAT3-L1* expression is reminiscent of the ASAT4 expression differences among accessions
607 of the wild tomato *Solanum habrochaites*. In that case, a subgroup of *S. habrochaites* accessions
608 possessed a functional ASAT4 copy, but its low levels of gene expression correlated with
609 reduced accumulation of acetylated acylsugars in these 'subgroup E' accessions (Kim et al.,
610 2012).



611

612 **Figure 7. Characterized *in vitro* acyl-CoA usage by ASATs.** The maximum likelihood
613 phylogeny was inferred from amino acid sequences using a time-reversible algorithm specifying
614 a plant-specific empirical substitution matrix, invariant sites, and four rate categories
615 (Q.plant+I+G4) in IQ-TREE v2.1.3. Bootstrap support was calculated from 100,000 ultrafast
616 bootstrap replicates; values >=95 indicate strong support. Other than SmASAT3-L1, enzyme
617 activities were described previously (Fan et al., 2017, 2015; Leong et al., 2022, 2020; Lou et al.,
618 2021; Moghe et al., 2017; Nadakuduti et al., 2017; Schenck et al., 2022; Schilmiller et al., 2015,
619 2012). Only reported, positive activities are displayed with the black squares. Lack of activity or
620 untested activities are displayed with blank squares.

621

622 Natural variation potentiates novel activities and uses

623 Until recently, it appeared that acylsucroses dominate acylsugar profiles of Solanaceae
624 species with a small number of documented examples of acylglucose- and acylinositol-producing
625 *Solanum* species (Fiesel et al., 2022; Fobes et al., 1985; Herrera-Salgado et al., 2005; Hurney,
626 2018; King and Calhoun, 1988; Leong et al., 2020; Liu et al., 2017; Lou et al., 2021; Moghe et
627 al., 2017; Schenck et al., 2022). Our chemotaxonomic survey of nearly three dozen *Solanum*
628 species from the sparsely sampled Clade II, DulMo, and VANAns clades revealed widespread
629 occurrence of glucose-, inositol-, and non-sucrose disaccharide-based acylsugars decorated with
630 unusual acyl chains, including medium-length (C12-C16) hydroxyacyl chains, glycosylated
631 hydroxyacyl chains, and unsaturated chains. Considering that this work and previously published
632 analyses together cover <3% of *Solanum* species, this presumably is only a subset of trichome
633 acylsugar diversity within this large genus. Adding to interspecific variation we observe here, the
634 recent identification of acylsugars in cultivated tomato root exudates supports the value of
635 metabolite screening in additional tissue types as well as species (Korenblum et al., 2020).

636 With 33 acyl chains (Lou et al., 2021; Leong et al., 2020; Herrera-Salgado et al., 2005)
637 and six inositol acylation positions, there are 1.3 billion theoretically possible acylinositol
638 structures, not accounting for glycosylated acylinositols and glycohydroxyacylhexoses. This is in
639 addition to the 820 million theoretically possible acylsucrose structures (Fan et al., 2019), and

640 demonstrates that this study greatly expands our knowledge of known and potential acylsugar
641 structural diversity.
642 Knowledge of *Solanum* acylsugar diversity provides a framework for investigating the
643 molecular basis of metabolic pathway evolution. For example, the widespread occurrence of
644 acylglucoses across the DulMo clade leads us to hypothesize that the neofunctionalized invertase
645 responsible for *S. nigrum* acylglucose synthesis (Lou et al., 2021) is shared by other DulMo
646 clade species. Similarly, documented acylsugar variation between cultivated eggplant, *S.*
647 *melongena,* and its relative, *S. aethiopicum*, led to the discovery of a terminal acylsugar
648 acetyltransferase in eggplant, SmASAT3-L1. The characteristics of ASAT3-L1, the enzyme
649 absent from the *S. aethiopicum* trichome transcriptome, in turn reveals the dynamic nature of
650 plant specialized metabolism. First, SmASAT3-L1 is unique among characterized ASATs in its
651 acyl acceptor specificity, acetylating an acylinositol glycoside rather than an acylated sucrose,
652 glucose, or *myo*-inositol. Second, acetyl-CoA donor activity was previously only described for a
653 subset of Clade III BAHDs, which include ASAT4 and ASAT5 clade members (Figure 7).
654 SmASAT3-L1 breaks a pattern seen across the Solanaceae, including trichome ASATs from
655 species of early evolving lineages, *Petunia axillaris* and *Salpiglossis sinuata* (Moghe et al., 2017;
656 Nadakuduti et al., 2017), *Nicotiana attenuata* (Schenck et al., 2022), as well as cultivated and
657 wild tomatoes (Kim et al., 2012; Schilmiller et al., 2012). This theme was also observed in
658 trichomes of two other *Solanum* plants, more closely related to the species screened in this study:
659 the Clade II *S. quitoense* SqTAIAT (Leong et al., 2020) and DulMo *S. nigrum* SnAGAT1 (Lou et
660 al., 2021). Both enzymes reside outside of the ASAT3 clade, and thus are phylogenetically
661 distinct from SmASAT3-L1. SmASAT3-L1 characteristics and those of recently published *S.*
662 *quitoense* enzymes (Leong et al., 2022, 2020) also suggest that acylinositols are synthesized

663 through one or more pathway(s) distinct from that of acylsucroses. Existing genetic resources in
664 *S. nigrum*, *S. quitoense*, and *S. melongena* can be utilized to explore when and how acylinositols
665 arose and to identify the underlying mechanisms behind acylation position and chain type
666 differences between these species.

667 The remarkably varied structures of acylsugars strongly suggest that they have evolved
668 distinct bioactivities, likely in coevolutionary arms races between plants and their insect and
669 microbial pests. For example, hydrophobic acylsugars with longer acyl chains and few free
670 hydroxyl groups may disrupt membranes, reminiscent of less polar triterpene saponin variants
671 (Augustin et al., 2011; Baumann et al., 2000). Variation in acyl chain linkage chemistries also
672 likely influence acylsugar mode of action. For example, ester linked short chains acylsucroses
673 are digested by *Manduca sexta* (hawkmoth) larvae and the volatile organic acids released into the
674 environment, attracting ant predators (Weinhold and Baldwin, 2011). In contrast, longer chain
675 hydroxylated and ether-linked glycohydroxyacyl chains observed in the study are likely to have
676 quite different metabolic fates, perhaps persisting in the digestive systems of herbivores, and
677 alternative metabolic fates in microbes. Uncovering the enzymes responsible for synthesis of the
678 wide variety of natural acylsugars will enable transgenic production of isogenic variants and
679 rigorous testing of their ecological activities. Similarly, access to a wider variety of often
680 promiscuous ASAT BAHD acyltransferases (Moghe et al., 2023) creates opportunities for
681 producing novel acylsugars with a wide range of physical properties and bioactivities (Schenck
682 et al., 2022).

683 Hydroxyacyl chains also present promising opportunities to expand upon natural
684 acylsugar diversity through synthetic biology and/or synthetic chemistry. The hydroxyl acts as a
685 reactive chemical handle, which was exploited in a subset of *Solanum* species trichomes to create

686 glycohydroxyacylhexoses. This handle can also be exploited with synthetic chemistry using
687 these chains as new feedstocks for creating polymers. These chains may also be useful in
688 creating synthetic acylsugars for use as food-grade emulsifiers. We can also further modify the
689 hydroxyacyl chains to add unusual sugars, acyl chains, and aromatic groups through
690 promiscuous enzymes or synthetic chemistry. These structural changes would likely impact
691 biological activities and add another asset to developing more pest resistant plants.
692 Characterizing more natural acylsugar diversity will further enable synthetic biology approaches
693 to create completely new acylsugar structures.
694
695 **Methods**
696
697 Plant material and growth conditions
698 *Solanum* spp. seeds were obtained from the sources described in Table S1. Seeds were
699 treated with 10% (v/v) bleach for 10 min while being rocked at 24 rpm with a GyroMini nutating
700 mixer (Labnet, Edison, NJ, USA), and then rinsed 5-6 times with distilled water. Unless
701 otherwise noted, seeds were germinated on Whatman filter paper (MilliporeSigma, Burlington,
702 MA, USA) at 28°C and in the dark. Germinated seedlings were transferred to peat pots (Jiffy,
703 Zwijndrecht, Netherlands), and grown at 25°C, 16/8-h day/night light cycle, and ~70 μmol $m^{-2}s^{-2}$
704 photosynthetic photon flux density with cool white fluorescent bulbs. Mature plants were grown
705 in controlled environment growth chambers or in a greenhouse. The growth chamber conditions
706 consisted of 25°/18°C day/night temperatures, 16/8-h day/night light cycle, and ~100 μmol $m^{-2}s^{-2}$
707 photosynthetic photon flux density under LED bulbs. The greenhouse conditions consisted of a
708 16/8-h day/night light cycle achieved with supplemental sodium iodide lighting. Plants were
709 fertilized weekly with 0.5X Hoagland's solution.

710

711 Surface metabolite extractions

712 Surface metabolites were extracted as described previously (Leong et al., 2019; Lou and
713 Leong, 2019). Briefly, 0.1-1 g of leaf, stem, hypocotyl, or cotyledon tissue was collected into a
714 1.5 mL screw-cap tube (Dot Scientific, Inc., Burton, MI, USA), 1 mL extraction solvent (3:3:2
715 acetonitrile:isopropanol:water, 0.1% formic acid, 10 µM propyl 4-hydroxybenzoate (internal
716 standard)) was added and the sample was rocked at 24 rpm by a GyroMini nutating mixer for
717 two min. After extraction, the supernatant was transferred into a glass 2 mL autosampler vial
718 (Restek, Bellefonte, PA, USA) and sealed with a 9 mm cap containing a PTFE/silicone septum
719 (J.G. Finneran, Vineland, NJ, USA). For fruit surface metabolite extractions, we modified the
720 protocol by placing 1-2 fruit into a 15 mL polypropylene conical tube (Corning Inc., Corning,
721 NY, USA) containing 5 mL extraction solvent then proceeded to the nutation step above.

722

723 Bulk eggplant acylsugar extraction and purification for NMR analysis

724 Surface metabolites were bulk extracted from approximately 1000 *S. melongena* PI
725 555598 seedlings. Seeds were treated with 10% (v/v) bleach (Clorox, Oakland, CA, USA) for 10
726 min while being gently rocked at 24 rpm with a GyroMini nutating mixer and subsequently
727 rinsed 5-6 times with distilled water. Seeds were sown in moist SUREMIX soil (Michigan
728 Grower Products, Galesburg, MI, USA) in Perma-Next plant trays, 22 x 11 x 2.5 inches
729 (Growers Supply Company, Dexter, MI, USA), covered with a humidity dome, 22 x 11 x 3
730 inches (Growers Supply Company), then transferred to the growth chamber immediately.
731 Seedlings were harvested when 2-3 true leaves were observed, approximately one week, by
732 cutting them at the base and placing them into two 2L beakers each containing 1L 100%

733 acetonitrile. Surface metabolites were extracted with gentle agitation with a metal spatula for five
734 min at room temperature. Plant material and sediment was removed by vacuum filtration through
735 a Büchner funnel lined with Whatman filter paper (MilliporeSigma). solvent was removed *in*
736 *vacuo* by rotary evaporation and dried residue was dissolved in 20 mL of acetonitrile with 0.1%
737 formic acid. Solvent was removed again using a vacuum centrifuge and the dried residue was
738 dissolved in 1 mL of acetonitrile:water:formic acid (80:20:0.001). The semi-preparative LC
739 method is described in detail in Table S28 and is described in brief here. *S. melongena*
740 acylsugars were separated with a Waters 2795 HPLC (Waters Corporation, Milford, MA, USA)
741 equipped with an Acclaim 120 C18 HPLC column (4.6 x 150 mm, 5 µm; Thermo Fisher
742 Scientific, Waltham, MA, USA). Solvent A was water with 0.1% formic acid and Solvent B was
743 acetonitrile. The reverse-phase linear gradient was as follows: 5% B at 0 min, 60% B at 2 min,
744 80% B at 40 min, 100% B at 42 min, 5% B at 42.01 min, held at 5% B until 44 min. Flow rate
745 was 1.5 mL/min, injection volume was 100 µL, and the column temperature was 40°C. Fractions
746 were collected automatically at 0.25 min intervals by a 2211 Superrac fraction collector (LKB
747 Bromma, Stockholm, Sweden) and assessed for acylsugar presence and purity by LC-MS
748 analysis, as described below. Column fractions were collected in the same tubes for each method
749 run which worked to pool the fractions between each run.
750
751 <u>LC-MS acylsugar analysis</u>
752 Acylsugars were analyzed by LC-MS with each method described below and in Tables
753 S29-33. For each LC method, the mobile phase consisted of aqueous 10 mM ammonium
754 formate, adjusted to pH 2.8 with formic acid (Solvent A) and 100% acetonitrile (Solvent B). The
755 flow rate was maintained at 0.3 mL/min.

756 Acylsugar extracts were analyzed with a 22 min LC gradient using a Waters Acquity
757 UPLC coupled to a Waters Xevo G2-XS QToF mass spectrometer (Waters Corporation, Milford,
758 MA) equipped with electrospray ionization (ESI) operating in positive (ESI^+) and negative (ESI^-)
759 modes. Ten microliter acylsugar extracts were separated on an Acquity UPLC BEH C18 column
760 (10 cm x 2.1 mm, 130 Å, 1.7 µm; Waters), kept at 40°C, using a binary solvent gradient. The 22
761 min linear gradient was as follows: 5% B at 0 min, 60% B at 2 min, 100% B at 16 min, held at
762 100% B until 20 min, 5% B at 20.01 min, held at 5% B until 22 min. Acylsugar extracts were
763 analyzed with both ESI^-, and ESI^+. For ESI^-, the following parameters were used: capillary
764 voltage, 2 kV; sampling cone voltage, 35 V; source temperature, 100°C; desolvation temperature
765 350°C; cone gas flow, 50 L/Hr; desolvation gas flow, 600 L/Hr. For ESI^+, the following
766 parameters were used: capillary voltage, 3 kV; sampling cone voltage, 35 V; source temperature,
767 100°C; desolvation temperature 300°C; cone gas flow, 50 L/Hr; desolvation gas flow, 600 L/Hr.
768 Acylsugars were fragmented in either MS^E or data-dependent acquisition (DDA) MS/MS modes
769 as described previously (Lou et al., 2021; Lybrand et al., 2020). DDA survey and MS/MS
770 functions acquired over *m/z* 50 to 1500 with scan times of 0.2 s. Ions selected for MS/MS were
771 fragmented with a ramped collision energy where voltage varies based on ion mass. Collision
772 energies followed a ramp with the voltage changing linearly for ions between the low mass
773 setting (*m/z* 50, 15 to 30 V) and high mass setting (*m/z* 1500, 30 to 60 V). To increase mass
774 accuracy, lock mass correction was performed during data collection, with leucine enkephalin as
775 the reference.
776 Semi-preparative LC fractions were analyzed with direct infusion and a 14-min LC
777 gradient using a LC-20ADvp ternary pump (Shimadzu, Kyoto, Japan) coupled to a Waters Xevo
778 G2-XS QToF mass spectrometer (Waters Corporation). The direct infusion method quickly

779 screened fractions for acylsugar presence while the 14-min LC method tested fraction purity. Ten
780 microliters of acylsugar fractions were injected into an Ascentis Express C18 HPLC column (10
781 cm x 2.1 mm, 2.7 µm; Supelco, Bellefonte, PA, USA), kept at 40°C. The 14-min gradient was as
782 follows: 5% B at 0 min, 60% B at 2 min, 100% B at 10 min held until 12 min, 5% B at 12.01
783 min, and held at 5% until 14 min. Fractions were analyzed under ESI^+ with the following
784 parameters: capillary voltage, 3 kV; sampling cone voltage, 40 V; source temperature, 100°C;
785 desolvation temperature 350°C; cone gas flow, 20 L/hr; desolvation gas flow, 500 L/hr. For
786 direct infusion analysis, ions were acquired from *m/z* 50 to 1500 with a scan time of 0.1 s and
787 one acquisition function with 0 V collision potential. For 14-min LC analysis, ions were acquired
788 from *m/z* 50 to 1200 with a scan time of 0.1 s and three acquisition functions with different
789 collision potentials (0, 25, 60 V). Lock mass calibration referenced to the leucine enkephalin
790 $[M+H]^+$ ion was applied during data acquisition.

791 Enzyme assays products were analyzed with a 7-min gradient using Waters Acquity
792 UPLC coupled to a Waters Xevo G2-XS QToF mass spectrometer (Waters Corporation, Milford,
793 MA) equipped with electrospray ionization (ESI) operating in negative (ESI^-) mode. Reaction
794 products were separated with an Ascentis Express C18 HPLC column (10 cm x 2.1 mm, 2.7 µm;
795 Supelco), kept at 40°C, using a binary solvent gradient. The 7 min linear gradient was as follows:
796 5% B at 0 min, 60% B at 1 min, 100% B at 5 min, held at 100% B until 6 min, 5% B at 6.01 min,
797 held at 5% B until 7 min. The ESI^- parameters described for acylsugar extract analysis were
798 used. Ions were acquired from *m/z* 50 to 1200 with a scan time of 0.1 s and three acquisition
799 functions with different collision potentials (0, 25, 60 V). Lock mass calibration referenced to the
800 leucine enkephalin $[M+H]^-$ ion was applied during data acquisition.

801 For coelution analysis between enzymatically- and plant-produced AI3:16 and AI4:18, a
802 24-min linear gradient was used with an Acquity UPLC BEH C18 column (10 cm x 2.1 mm, 130
803 Å, 1.7 µm; Waters), kept at 40ºC, on the same instrument used for enzyme assay analysis. The
804 binary solvent, linear gradient was as follows: 5% B at 0 min, 60% B at 18 min, 100% B at 18.01
805 min, held at 100% B until 22 min, 5% B at 22.01 min, held at 5% B until 24 min. ESI^-
806 parameters and MS method from the above enzyme assay analysis were used.
807 Saponified acylsugars were analyzed using a Waters Acquity UPLC coupled to a Waters
808 Xevo G2-XS QToF mass spectrometer (Waters Corporation, Milford, MA). Ten microliters of
809 saponified acylsugars were injected into either an Acquity UPLC BEH C18 column (10 cm x 2.1
810 mm, 130 Å, 1.7 µm; Waters) or an Acquity UPLC BEH Amide column (10 cm x 2.1 mm, 130 Å,
811 1.7 µm; Waters), kept at 40°C. Samples were analyzed on the C18 column with the 22-min
812 method described above. To detect free sugars resulting from saponification, samples were
813 analyzed on the BEH amide column with a 9 min method which was as follows: 95% B at 0 min
814 held until 1 min, 60% B at 6 min, 5% B at 7 min, 95% B at 7.01 min, held at 95% B until 9 min.
815 Both methods used the ESI- parameters and MS acquisition parameters described for enzyme
816 assays.
817
818 Acylsugar annotation
819 We inferred acylsugar structures utilizing negative and positive mode MS and MS/MS
820 collision-induced dissociation (CID) as previously described (Ghosh et al., 2014; Hurney, 2018;
821 Leong et al., 2019; Lou et al., 2021; Lybrand et al., 2020). We employed modified confidence
822 criteria developed by Lou and coworkers (Lou et al., 2021). Putative acylsugars only meeting
823 criterion A were annotated as low confidence. Putative acylsugars meeting criteria A-D were

824 annotated as medium confidence. We also added two additional criteria, G and H, described
825 below. Acylsugars meeting criteria A-F and either G or H were annotated with highest
826 confidence. Confidence levels for each reported acylsugar are specified in their species-specific
827 annotation table (Tables S2-27).
828 (G) **NMR structural determination.** Acylsugars from *S. melongena* were further characterized
829 by NMR experiments in which sugar core relative stereochemistry, acyl chain positions, and acyl
830 chain branching were resolved.
831 **(H) Coelution with NMR characterized compounds.** Acylsugars with the same acyl chains
832 and sugar core size were analyzed by coelution with *S. melongena* acylsugars. If acylsugars
833 coeluted with an NMR characterized peak and shared the same acyl chains and fragmentation
834 patterns, they met this criterion and were annotated with high confidence. This criterion was
835 employed for acylsugars from species other than *S. melongena*.
836
837 Acylsugar metabolomic analysis
838 To collect accurate formate adduct masses and relative abundances from putative
839 acylsugars, negative mode LC-MS raw data from metabolite extracts were analyzed with
840 Progenesis QI software v3.0 (Waters, Milford, MA, USA) using retention time alignment, peak
841 detection, adduct grouping, and deconvolution. Lock mass correction was performed during data
842 collection. We used the following analysis parameters: peak picking sensitivity, default;
843 retention time range, 0-22 min; adducts detected, M+FA-H, M+Cl, M-H, 2M-H, and 2M+FA-H.
844 Average percent peak abundance was calculated for each putative acylsugar by dividing its raw
845 abundance by the sum of the total acylsugar raw abundance and then averaging that value across
846 all samples for a species.

847

848 NMR analysis of acylsugars

849 Purified acylsugars were dissolved in acetonitrile-d_3 (99.96 atom % D; MilliporeSigma,
850 Burlington, MA, USA) and transferred to Kontes NMR tubes (MilliporeSigma, Burlington, MA,
851 USA). Samples were analyzed at the Michigan State University Max T. Rogers NMR Core with
852 a Varian Inova 600 MHz spectrometer (Agilent , Santa Clara, CA, USA) equipped with a
853 Nalorac 5 mm PFG switchable probe, a DirectDrive2 500 MHz spectrometer (Agilent, Santa
854 Clara, CA, USA) equipped with a OneNMR probe, Bruker Avance NEO 800 MHz spectrometer
855 (Bruker, Billerica, MA, USA) equipped with a 5 mm helium cryogenic HCN probe, or a Bruker
856 Avance NEO 600 MHz spectrometer equipped with a 5 mm nitrogen cryogenic HCN Prodigy
857 probe. All acylsugars were analyzed with a series of 1D (^{1}H, and ^{13}C) and 2D (gCOSY, TOCSY,
858 gHSQC, gHMBC, gH2BC, *J*-resolved) experiments. Resulting spectra were referenced to
859 acetonitrile-d_3 ($^{\delta}$H = 1.94 and $^{\delta}$C = 118.70 ppm). NMR spectra were processed and analyzed
860 with MestReNova 12.0 software (MestreLab, A Coruña, Spain). For full NMR metadata, see
861 Supplementary Datafile 1.

862

863 Sugar core composition analysis

864 *S. melongena* PI 555598 acylsugar extracts were dried down *in vacuo* with centrifugation
865 (Savant, ThermoFisher Scientific). Dried acylsugar extracts were dissolved in 1 mL of methanol.
866 One mL of 3 M ammonium hydroxide was added, and the solution was mixed vigorously. This
867 initial saponification reaction proceeded at room temperature for 48 hours. Solvent was removed
868 by vacuum centrifugation. Saponified sugar cores were reduced and derivatized with acetate
869 groups as previously described (Sassaki et al., 2008). Sugars were dissolved in 50 µL of 1 M

870 ammonium hydroxide for 15 min at room temperature. Addition of 50 µL 20 mg/mL $NaBH_4$ and
871 incubation at 100°C for 10 min converted aldoses and ketoses to polyols. Excess sodium
872 borohydride was quenched with 100 µL 1M trifluoroacetic acid. Two volumes of methanol were
873 added, and the sample was dried down by a stream of nitrogen gas. Dried down reaction products
874 were redissolved in 200 µL of methanolic 0.5 M HCl and heated at 100°C for 15 min in a closed
875 ½ dram vial, 12 x 35 mm (Kimble Chase, Vineland, NJ, USA). Solvent was removed by a stream
876 of nitrogen gas. To acetylate the sugar cores, residue was dissolved in 200 µL of pyridine:acetic
877 anhydride (1:1) and incubated at 100°C for 30 min. Acetylated polyols were then dried down *in*
878 *vacuo* in a vacuum centrifuge and redissolved in 100 µL hexane for GC-MS analysis.
879 To determine the composition of disaccharide sugar cores, saponified sugars were
880 hydrolyzed prior to derivatization. Dried down sugars were redissolved in 98 µL water in a ½
881 dram vial, 12 x 35 mm (Kimble Chase). After addition of 102 µL of 88% formic acid, the vial
882 was sealed and heated at 100°C for 15 hours in a heat block. Hydrolyzed sugars were dried down
883 *in vacuo* with a vacuum centrifuge and derivatized to alditol acetates following the method
884 described above.
885
886 Hydroxyl acyl chain stereochemistry analysis
887 We derivatized hydroxyacyl chains and extracted the derivatives as previously described
888 (Jenske and Vetter, 2007). The commercial standards (3*R*)-OH-tetradecanoate (Cayman
889 Chemical, Ann Arbor, MI, USA) and (3*R*/*S*)-OH-tetradecanoate (Cayman Chemical, Ann Arbor,
890 MI, USA) were first derivatized to their respective ethyl ester derivatives. 600 µg of each fatty
891 acid was dissolved in 0.5 mL of 0.5 M ethanolic KOH and incubated at 80°C for 5 min. After the
892 solution was cooled on ice, 1 mL of ethanolic BF_3 was added, and reactions were incubated at

893 80°C for five min. Two-phase partitioning with hexane and saturated sodium chloride yielded
894 ethyl ester derivatives in the hexane partition.

895 The 3-OH-nC14 acyl chain from I3:22(iC4,iC4,3-OH-nC14) was derivatized to an ethyl
896 ester following a previously published method (Fan et al., 2020; Ning et al., 2015). Half of the
897 purified I3:22(iC4,iC4,3-OH-nC14) was dissolved in acetonitrile with 0.1% formic acid and
898 transferred to a 1.5 mL microcentrifuge tube. The purified acylsugar was dried down *in vacuo*
899 with vacuum centrifugation. The resulting dry, purified I3:22(iC4,iC4,3-OH-nC14) was
900 dissolved in 300 µL of 21% (w/w) sodium ethoxide in ethanol (MilliporeSigma, Burlington,
901 MA, USA). The reaction proceeded at room temperature for 30 min with constant rocking at 24
902 rpm by a GyroMini nutating mixer (Labnet, Edison, NJ, USA) and occasional vortexing. 400 µL
903 hexane with 55 ng/µL tetradacane (MilliporeSigma, Burlington, MA, USA) as an internal
904 standard was then added followed by vigorous vortexing. 500 µL of aqueous saturated sodium
905 chloride was added to the hexane-ethanol mixture, and vortexed. The hexane layer was pipetted
906 to another tube and two more two-phase partitions with saturated sodium chloride were
907 completed. The final hexane layer containing the acyl chain ethyl esters was extracted and dried
908 down by a stream of nitrogen gas.

909 The resulting fatty acid ethyl esters from the commercial standards and purified
910 compound were derivatized with (*R*)-(-)-α-methoxy-α-trifluoromethylphenylacetyl chloride ((*R*)-
911 (-)-MTPA-Cl; MilliporeSigma, Burlington, MA, USA) following a previously published
912 procedure (Jenske and Vetter, 2007). Dried fatty acid ethyl esters were redissolved in 400 µL
913 pyridine and 15 µL of (*R*)-(-)-MPTA-Cl. The reaction proceeded at room temperature for two
914 hours, after which, 5 mL of water and tert-butyl methyl ether (TBME) were added along with
915 solid K_2CO_3 (one spatula tip). The TBME phase was collected after three successive phase

916 separations with 5 mL of water. TBME was evaporated to 1 mL and subjected to GC-MS
917 analysis.
918
919 GC-MS analysis
920 All GC-MS analyses employed an Agilent 5890 GC and an Agilent 5975 single
921 quadrupole MS equipped with a FactorFour VF-5ms column (30 m x 0.25 mm, 0.25 um;
922 Agilent) and 10 m EZ-Guard column (Agilent, Santa Clara, CA, USA). Helium was used as the
923 carrier gas with a constant flow rate of 1 mL/min. Electron energy was set at 70 eV. MS source
924 and quadrupole were maintained at 230°C and 150°C, respectively. Parameters specific to each
925 analysis are listed below.
926 For analysis of sugar core alditol acetate derivatives, we followed previously published
927 GC-MS parameters (Sassaki et al., 2008). Inlet temperature was maintained at 275°C. GC oven
928 temperature was held at 60°C for one min and then was ramped at a rate of 40°C/min to 180°C.
929 Oven temperature was then ramped to 240°C at a rate of 5°C/min and held for three min. Total
930 run time was 19 min. The MS detector transfer line was maintained at 280°C. Split ratio was
931 10:1 with a split flow rate of 10 mL/min. A three min solvent delay was used. MS data was
932 collected in full scan mode, *m/z* 50-600. Analysis parameters are also described in Table S34.
933 For analysis of MPTA fatty acid ethyl ester derivatives, we followed previously
934 published GC-MS parameters (Jenske and Vetter, 2007). GC oven temperature was held at 60°C
935 for 1.5 min and then was ramped at a rate of 40°C/min to 180°C. 180°C was held for two min
936 after which the oven temperature was ramped to 230°C at a rate of 2°C/min. After 230°C was
937 held for nine min, the oven temperature was ramped to 300°C at 10°C/min and held for 7.5 min.
938 Total run time was 55 min. A 5-min solvent delay was applied. The MS detector transfer line

939 was maintained at 280°C. Selective ion monitoring detected the ions *m/z* 189, 209, and 255 with
940 a dwell time of 50 ms. Analysis parameters are also described in Table S35.
941
942 Sample collection and transcriptome sequencing
943 We sequenced 18 transcriptomes: six biological replicates each of trichomes isolated
944 from hypocotyls (trichomes), hypocotyls stripped of their trichomes (stripped hypocotyls), and
945 whole roots (Supplemental data file 1). We sampled trichomes and stripped hypocotyls from 7-
946 day-old *S. melongena* 67/3 seedlings following methods developed for root hair cell isolation
947 (Bucher et al., 1997) with modifications. We grew lawns of eggplant seedlings in soil flats as
948 described above for bulk leaf surface metabolite extraction. At 7 days post germination, we
949 removed roots and cotyledons from seedlings, then transferred hypocotyls to liquid nitrogen in a
950 plastic 2 L Dewar flask. Frozen hypocotyls were gently stirred with a glass rod for 20 minutes to
951 physically shear trichomes from hypocotyls. After confirming under a dissecting scope that a
952 sample of three hypocotyls had been stripped bare, we filtered trichomes into a 2 L glass beaker
953 by slowly pouring the contents of the Dewar flask through a 500 µm wire mesh sieve
954 (MilliporeSigma, Burlington, MA, USA). To maximize trichome recovery, stripped hypocotyls
955 were returned to the Dewar, rinsed with liquid nitrogen, and filtered through the 500 µm sieve
956 six more times. Stripped hypocotyls were divided into six pre-weighed 50 mL conical tubes,
957 quickly weighed, then transferred to -80ºC for storage. Trichomes were subsequently filtered
958 through a 150 µm sieve (MilliporeSigma, Burlington, MA, USA) into a 500 mL beaker to
959 increase sample purity then transferred to a 50 mL conical tube to allow the excess liquid
960 nitrogen to evaporate. Finally, filtered trichomes were divided into six pre-weighed 2 mL screw-

961 cap tubes (Dot Scientific, Inc., Burton, MI, USA), quickly weighed, then transferred to -80ºC for
962 storage.
963 We extracted total RNA using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany)
964 following the manufacturer's directions, measured RNA concentration using a Qubit 1.0
965 instrument (Thermo Fisher Scientific, Waltham, MA, USA) with the RNA HS assay, and the
966 samples were processed by the Michigan State University Research Technology Support Facility
967 Genomics Core (East Lansing, MI, USA) for library preparation and high-throughput
968 sequencing. The core checked RNA quality using a Bioanalyzer 2100 (Agilent Technologies,
969 Santa Clara, CA, USA), constructed sequencing libraries using an Illumina Stranded mRNA
970 Prep kit (Illumina, Inc. San Diego, CA, USA), and sequenced all 18 libraries on a NovaSeq 6000
971 (Illumina, Inc. San Diego, CA) using a single S4 flow cell lane, producing 150-base pair paired-
972 end reads (Supplemental datafile 1). The raw fastq files are available to download from the
973 National Center for Biotechnology Information Sequence Read Archive database under
974 BioProject PRJNA935765.
975
976 Transcriptome alignment and differential gene expression analysis
977 To prepare the 150-bp paired-end sequences for alignment, we employed Trimmomatic
978 (Bolger et al., 2014) to trim adapters and low-quality bases, then filter reads shorter than 75 bp,
979 removing an average of 3.7% of sequences (range: 3.3% - 4.2%). We mapped the resulting 75–
980 150-bp (average length: 149 bp) paired-end RNAseq reads separately to the eggplant V4.1 and
981 HQ reference genomes using STAR in two-pass mode, which enhances splice junction discovery
982 and mapping sensitivity (Dobin et al., 2013; Dobin and Gingeras, 2015). Using this approach, an
983 average of 79.2% (range: 60.1 - 87.1%) and 82.2% (range: 62.0 - 90.7%) of reads mapped to a

984 unique genomic location in V4.1 and HQ, respectively (Table S36). We filtered the resulting
985 transcriptome alignments according to best practices as defined by the Genome Analysis Toolkit
986 (GATK) (DePristo et al., 2011; Van der Auwera et al., 2013). Briefly, we removed optical and
987 PCR duplicates with MarkDuplicates from the Picard toolkit
988 (http://broadinstitute.github.io/picard), parsed reads into exon segments and removed intron-
989 spanning bases using SplitNCigarReads from GATK (McKenna et al., 2010). Finally, we
990 selected unique alignments by eliminating reads with a mapping quality score below Q60 with
991 the view command from SAMtools (Li et al., 2009). To generate raw read counts, we used the
992 HTSeq command, htseq-count with the nonunique parameter set to all (Anders et al., 2015).
993 Raw read counts generated by HTSeq-count were used to perform differential gene
994 expression analysis in edgeR (Robinson et al., 2010). To restrict comparisons to expressed genes,
995 only transcripts with at least one read count-per-million (CPM) in at least one sample were
996 retained for further analysis. This filtering step removed 11,665 (33.4%) and 14,544 (39.8%)
997 annotated genes in the V4.1 and HQ eggplant genomes, respectively, yielding 23,251 and 22,024
998 expressed transcripts for differential gene expression analysis. Next, we normalized transcript
999 abundances across the 18 transcriptomes in our eggplant V4.1 and HQ alignments using the
1000 default trimmed mean of M-values (TMM) method in the calcNormFactors function, then
1001 performed multidimensional scaling (MDS) with the plotMDS function to compare global gene
1002 expression profiles (Figure S15). This showed that our samples cluster tightly by tissue (i.e.,
1003 trichome, shaved hypocotyl, or root), with no obvious differences between V4.1 and HQ
1004 alignments. To test differences in gene expression across tissues, we implemented a generalized
1005 linear model (GLM) using a quasi-likelihood (QL) approach: we generated an experimental
1006 design matrix specifying the three tissues (i.e., trichomes, trichomeless hypocotyls, and roots)

1007 with the model.matrix function, then used the glmQLFit function to fit our data to a QL-GLM
1008 model. To identify genes with a log2 fold-change (FC) > 2 between tissues, we used the
1009 glmTreat function, which performs threshold hypothesis testing, a rigorous statistical approach
1010 that evaluates variance and magnitude to detect expression differences greater than the specified
1011 value (e.g., log2 FC > 2), then applies false discovery rate (FDR) *p*-value corrections. Genes
1012 were classified as significantly differentially expressed between two tissues if log2 FC > 2 and
1013 FDR-corrected *p*-value < 0.05. We calculated absolute transcript abundance for all expressed
1014 genes as transcripts per million (TPM) with the calculateTPM function in scateR (McCarthy et
1015 al., 2017). Transcriptome expression values are in Tables S37 and S38.
1016
1017 Phylogenetic analyses
1018 Characterized ASAT enzymes fall into clade III of the BAHD family (Fan et al., 2017,
1019 2015; Leong et al., 2022, 2020; Lou et al., 2021; Moghe et al., 2017; Nadakuduti et al., 2017;
1020 Schenck et al., 2022; Schilmiller et al., 2015, 2012). To identify ASAT candidates in eggplant,
1021 we searched for sequences containing the Pfam Transferase domain (PF02458) (Finn et al.,
1022 2010), associated with all characterized catalytically active BAHD proteins. The PF02458 HMM
1023 profile was obtained from the Pfam website (http://pfam.sanger.ac.uk/) and queried against the
1024 V4.1 and HQ eggplant proteomes using the hmmsearch tool from HMMER v3.2.1 (hmmer.org),
1025 revealing 106 and 108 putative BAHD sequences, respectively. Using MAFFT v7.471 in E-INS-
1026 i mode, we built multiple sequence alignments (MSA) of amino acid sequences from four
1027 sources: 1) V4.1 or HQ eggplant PF02458 hits, 2) published reference sequences for clades I-VII
1028 (Moghe et al., 2023), 3) characterized ASAT sequences from other Solanaceae species, and 4)
1029 the SaASAT3 and SaASAT3-L1 candidates from *S. aethiopicum* (Katoh and Standley, 2013).

1030 The E-INS-i algorithm implements local alignment with a generalized affine gap cost (Altschul,
1031 1998), which aligns conserved regions (e.g., the BAHD transferase domain) and essentially
1032 ignores nonconserved regions. Phylogenetic reconstruction was performed using IQ-TREE
1033 v2.1.3 (Minh et al., 2020). The ModelFinder tool was implemented to identify the best maximum
1034 likelihood model for estimating evolutionary relationships (Kalyaanamoorthy et al., 2017),
1035 leading to selection of Jones-Taylor-Thornton (JTT)+F with seven rate categories and Q.plant+F
1036 with seven rate categories for the V4.1 and HQ MSAs, respectively. Phylogenetic trees were
1037 inferred by maximum likelihood using the chosen model, and branch support was obtained from
1038 100,000 ultrafast bootstrap iterations (Hoang et al., 2018). The resulting BAHD phylogenies
1039 were visualized using the ggtree package in R (Yu et al., 2017). To generate the clade III BAHD
1040 heatmap-tree, we used the viewClade function in ggtree to subset the phylogeny and used
1041 gheatmap to visualize transcript abundance (log2 TPM) for the eggplant BAHDs.
1042
1043 <u>BAHD acyltransferase cloning, expression, and purification</u>
1044 Candidate ASATs were cloned into pET28b(+) (MilliporeSigma, Burlington, MA, USA).
1045 Open reading frames from genes were either synthesized by Twist Biosciences (South San
1046 Francisco, CA, USA) with or without codon-optimization for *E. coli* expression (Table S39) or
1047 amplified from genomic DNA or cDNA with primers listed in Table S40. Q5 2X Hotstart master
1048 mix (New England Biolabs, Ipswich, MA, USA) was used for cloning PCRs. The source of
1049 cloned DNA and whether a gene was codon optimized is described in Table S39. Amplified
1050 genes were purified by agarose gel electrophoresis and extraction with the Monarch DNA Gel
1051 Extraction Kit (New England Biolabs, Ipswich, MA, USA). Both the synthesized genes and PCR
1052 amplified genes were then inserted into a doubly digested BamHI/XhoI pET28b(+) through

1053 Gibson assembly using the 2X Gibson Assembly Master Mix (New England Biolabs, Ipswich,
1054 MA, USA) according to the manufacturer's instructions. The constructs using synthesized genes
1055 were transformed into BL21(DE3) (MilliporeSigma, Burlington, MA, USA) and constructs with
1056 PCR amplified genes were transformed into BL21 Rosetta(DE3) cells (MilliporeSigma,
1057 Burlington, MA, USA). Constructs were verified with colony PCR and Sanger sequencing using
1058 T7 terminator and promoter primers (Table S40). Sanger sequencing was completed by the
1059 Michigan State University Research Technology Support Facility Genomics Core (East Lansing,
1060 MI, USA).

1061 Protein expression occurred as previously described (Leong et al., 2022; Lou et al.,
1062 2021). Briefly, 50 mL cultures of picked transformation colonies were grown overnight at 37°C,
1063 shaking at 225 rpm in Luria-Bertani (LB) media (Neogen, Lansing, MI, USA) supplemented
1064 with 1% glucose (w/v). Fifteen mL of the overnight cultures were inoculated into 1 L of fresh LB
1065 medium, which was incubated at 37°C shaking at 225 rpm until an OD600 of 0.5 was reached.
1066 The cultures were incubated on ice for 25 min, after which, isopropylthio-β-galactoside was
1067 added to a final concentration of 50-500 μM. Then, cultures were incubated at 16°C shaking at
1068 180 rpm overnight for 16 hours before cells were harvested by centrifugation at 4,000 rpm for 10
1069 minutes at 4°C.

1070 *S. melongena* BAHDs were purified as previously described (Leong et al., 2020) with the
1071 following modifications. The extraction buffer contained 10 mM imidazole, the wash buffer
1072 contained 20 mM imidazole, and the elution buffer contained 500 mM imidazole. Protein eluent
1073 was concentrated with 30-kD Amicon Ultra centrifugal filter units (MilliporeSigma).

1074

1075 <u>Enzyme assays</u>

1076 Enzyme assays were conducted in 100 mM sodium phosphate buffer at pH 6. For
1077 forward assays, acetyl-CoA (MilliporeSigma, Burlington, MA, USA) was added to a final
1078 concentration of 0.1 mM, and for reverse assays free CoA (MilliporeSigma, Burlington, MA,
1079 USA) was added to a final concentration of 1 mM. Purified AI3:16 and AI4:18 substrates were
1080 dried down using a vacuum centrifuge and redissolved in ethanol:water:formic acid (1:1:0.001).
1081 One microliter of the prepared acylsugars were used as acyl acceptors. Six microliters of enzyme
1082 were added to a final volume of 60 µL. For negative controls, 6 µL of enzyme that was heat
1083 inactivated at 95°C for 10 minutes was substituted in place of untreated enzyme. Assays were
1084 incubated at 30°C for 30 minutes after which 120 µL of acetonitrile:isopropanol:formic acid
1085 (1:1:0.001) with 1.5 µM telmisartan (MilliporeSigma, Burlington, MA, USA) stop solution was
1086 added. Reactions were then spun at 17,000 x *g* for 10 minutes to remove precipitate. Supernatant
1087 was placed in autosampler vials and analyzed by LC-MS.
1088
1089 RT-PCR of *S. aethiopicum* BAHDs
1090 We employed semi-quantitative RT-PCR to test *S. aethiopicum* BAHD expression in
1091 glandular trichomes with cDNA (Figure S14) and genomic DNA (gDNA) (Figure S16). Total
1092 RNA was isolated with the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany), including an on-
1093 column DNase digestion (Qiagen, Hilden, Germany), from shaved hypocotyls from accession *S.*
1094 *aethiopicum* PI 666075 and glandular trichomes from accessions PI 666075 and Grif 14165
1095 isolated as described above for *S. melongena* glandular trichomes. RNA was quantified with a
1096 Nanodrop 2000c instrument (Thermo Fisher Scientific, Waltham, MA, USA). We synthesized
1097 cDNA using 10 ng of RNA and SuperScript III Reverse Transcriptase (Invitrogen, Waltham,
1098 MA, USA). gDNA was isolated with the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany)

1099 from young leaf tissue collected from mature *S. aethiopicum* PI 666075 plants. PCR reactions
1100 (25 µL) were set up with GoTaq Green Master Mix (Promega, Madison, WI, USA), 200 nM of
1101 forward and reverse primers (Table S40), and 1 µL of cDNA or gDNA. PCR was performed
1102 under these conditions: 2 min at 95°C followed by 22, 30, and 35 cycles of 30s at 95°C, 30s at
1103 58°C, and 1 min at 72°C.

1104 We identified the putative *S. aethiopicum* orthologs of the *S. melongena* candidate
1105 BAHDs by querying *S. aethiopicum* annotated transcripts (Song et al., 2019) with *S. melongena*
1106 candidate ASAT DNA sequences using BLASTn. SaASAT3 and SaASAT3-L1 (GenBank
1107 accession: OQ547782) were not annotated in the *S. aethiopicum* genome and were identified by
1108 querying *S. aethiopicum* scaffolds with BLASTn and then identifying open reading frames with
1109 Geneious software v9.1.8 (Dotmatics, Boston, MA, USA). The *S. aethiopicum* ef1α gene was
1110 identified by querying a putative *S. melongena* ef1α, SMEL4.1_06g005890, against the *S.*
1111 *aethiopicum* scaffolds.

1112

1113 <u>*S. melongena* acyl chain composition analysis</u>

1114 *S. melongena* acyl chain composition was assessed with a previously developed
1115 esterification and GC-MS methods (Fan et al., 2020; Ning et al., 2015; Schenck et al., 2022).
1116 Acyl chain identifies were determined through authentic reference standards for nC8, nC10,
1117 nC12, and nC14 fatty acid ethyl esters (MilliporeSigma). The QuanLynx function of MassLynx
1118 v4.1 (Waters Corporation) integrated peaks from extracted ion chromatograms. Extracted ion
1119 chromatograms for the medium acyl chains iC8, nC8, iC10, nC10, iC12, nC12, iC14, and nC14
1120 were generated for the *m/z* of 88 with a mass window of *m/z* of 0.50. Extracted ion

1121 chromatograms for the short chains iC4, aiC5, and iC5 were generated for with the *m/*z values of
1122 71, 102, and 101, respectively, with a mass window of *m/z* of 0.50.
1123
1124 Accession numbers
1125 Raw RNA sequence data can be found in the National Center for Biotechnology
1126 Information (NCBI) Sequence Read Archive database under BioProject PRJNA935765.
1127 Sequence data for SaASAT3-L1 is available from the NCBI GenBank database under accession
1128 number OQ547782.
1129
1130 **Acknowledgements**
1131 We thank Dr. Christopher T. Martine for providing Australian *Solanum* ssp. seeds, Dr.
1132 Joyce van Eck for providing *Solanum prinophyllum* seeds, Dr. Guiseppe Leonardo Rotino and
1133 the Consiglio per la ricercar in agricoltura e l'analisi dell'economia agraria for providing *S.*
1134 *melongena* 67/3 and 305E40 seeds, and USDA-GRIN for providing *Solanum* ssp. seeds. We
1135 acknowledge the Michigan State University RTSF Mass Spectrometry and Metabolomics Core
1136 Facilities for LC-MS analysis support. We also acknowledge Dr. Daniel Holmes and Dr. Li Xie
1137 at the Michigan State University Max T. Rogers NMR Facility for experimental design and data
1138 analysis support. We are grateful to current and past members of the Last lab for feedback on the
1139 research and especially Drs. Jaynee Hart and Xingxing Li for comments on the manuscript.
1140
1141 **Funding**
1142 This research was funded by the National Science Foundation grant IOS-PGRP-1546617
1143 (to R. L. L. and A. D. J.), National Science Foundation grant IOS-PGRP-2218206 (to R. L. L.

1144 and R. E. K.) and National Institute of General Medical Sciences of the National Institutes of
1145 Health graduate training grant no. T32-GM110523 (to P. D. F.). A. D. J. is supported in part by
1146 USDA National Institute of Food and Agriculture Hatch project number MICL02474.
1147

1148 **References**

1149 Agati, G., Tattini, M., 2010. Multiple functional roles of flavonoids in photoprotection. New
1150 Phytol. 186, 786–793.

1151 Altschul, S.F., 1998. Generalized affine gap costs for protein sequence alignment. Proteins
1152 Struct. Funct. Bioinforma. 32, 88–96. https://doi.org/10.1002/(SICI)1097-
1153 0134(19980701)32:1<88::AID-PROT10>3.0.CO;2-J

1154 Anders, S., Pyl, P.T., Huber, W., 2015. HTSeq—a Python framework to work with high-
1155 throughput sequencing data. Bioinformatics 31, 166–169.
1156 https://doi.org/10.1093/bioinformatics/btu638

1157 Asai, T., Fujimoto, Y., 2010. Cyclic fatty acyl glycosides in the glandular trichome exudate of
1158 *Silene gallica*. Phytochemistry 71, 1410–1417. https://doi.org/10.1016/j.phytochem.2010.05.008

1159 Asai, T., Hara, N., Fujimoto, Y., 2010. Fatty acid derivatives and dammarane triterpenes from
1160 the glandular trichome exudates of *Ibicella lutea* and *Proboscidea louisiana*. Phytochemistry 71,
1161 877–894. https://doi.org/10.1016/j.phytochem.2010.02.013

1162 Aubriot, X., Knapp, S., Syfert, M.M., Poczai, P., Buerki, S., 2018. Shedding new light on the
1163 origin and spread of the brinjal eggplant (*Solanum melongena* L.) and its wild relatives. Am. J.
1164 Bot. 105, 1175–1187. https://doi.org/10.1002/ajb2.1133

1165 Augustin, J.M., Kuzina, V., Andersen, S.B., Bak, S., 2011. Molecular activities, biosynthesis and
1166 evolution of triterpenoid saponins. Phytochemistry 72, 435–457.
1167 https://doi.org/10.1016/j.phytochem.2011.01.015

1168 Bah, M., Pereda-Miranda, R., 1996. Detailed FAB-mass spectrometry and high resolution NMR
1169 investigations of tricolorins A-E, individual oligosaccharides from the resins of *Ipomoea tricolor*
1170 (Convolvulaceae). Tetrahedron 52, 13063–13080. https://doi.org/10.1016/0040-4020(96)00789-
1171 2

1172 Barchi, L., Rabanus-Wallace, M.T., Prohens, J., Toppino, L., Padmarasu, S., Portis, E., Rotino,
1173 G.L., Stein, N., Lanteri, S., Giuliano, G., 2021. Improved genome assembly and pan-genome
1174 provide key insights into eggplant domestication and breeding. Plant J. 107, 579–596.
1175 https://doi.org/10.1111/tpj.15313

1176 Baumann, E., Stoya, G., Völkner, A., Richter, W., Lemke, C., Linss, W., 2000. Hemolysis of
1177 human erythrocytes with saponin affects the membrane structure. Acta Histochem. 102, 21–35.
1178 https://doi.org/10.1078/0065-1281-00534

1179 Bernal, C.-A., Castellanos, L., Aragón, D.M., Martínez-Matamoros, D., Jiménez, C., Baena, Y.,
1180 Ramos, F.A., 2018. Peruvioses A to F, sucrose esters from the exudate of *Physalis peruviana*
1181 fruit as α-amylase inhibitors. Carbohydr. Res. 461, 4–10.
1182 https://doi.org/10.1016/j.carres.2018.03.003

1183 Bohs, L., 2004. Major clades in Solanum based on ndhF sequence data. Monogr. Syst. Bot. 27–
1184 49.

1185 Bohs, L., Olmstead, R.G., 1997. Phylogenetic relationships in Solanum (Solanaceae) based on
1186 ndhF sequences. Syst. Bot. 22, 5–17. https://doi.org/10.2307/2419674

1187 Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina
1188 sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

1189 Cao, C.-M., Wu, X., Kindscher, K., Xu, L., Timmermann, B.N., 2015. Withanolides and sucrose
1190 esters from *Physalis neomexicana*. J. Nat. Prod. 78, 2488–2493.
1191 https://doi.org/10.1021/acs.jnatprod.5b00698

1192 Caprioli, G., Iannarelli, R., Cianfaglione, K., Fiorini, D., Giuliani, C., Lucarini, D., Papa, F.,
1193 Sagratini, G., Vittori, S., Maggi, F., 2016. Volatile profile, nutritional value and secretory
1194 structures of the berry-like fruits of *Hypericum androsaemum* L. Food Res. Int. 79, 1–10.
1195 https://doi.org/10.1016/j.foodres.2015.11.021

1196 Castillo, M., Connolly, J.D., Ifeadike, P., Labbé, C., Rycroft, D.S., Woods, N., 1989. Partially
1197 acylated glucose and sucrose derivatives from *Salpiglossis sinuata* (Solanaceae). J. Chem. Res.
1198 Synop. 398–399.

1199 Chortyk, O.T., Kays, S.J., Teng, Q., 1997. Characterization of insecticidal sugar esters of
1200 *Petunia*. J. Agric. Food Chem. 45, 270–275. https://doi.org/10.1021/jf960322f

1201 Cicchetti, E., Duroure, L., Le Borgne, E., Laville, R., 2018. Upregulation of skin-aging
1202 biomarkers in aged NHDF cells by a sucrose ester extract from the agroindustrial waste of
1203 *Physalis peruviana* calyces. J. Nat. Prod. 81, 1946–1955.
1204 https://doi.org/10.1021/acs.jnatprod.7b01069

1205 D'Auria, J.C., 2006. Acyltransferases in plants: a good time to be BAHD. Curr. Opin. Plant Biol.
1206 9, 331–340. https://doi.org/10.1016/j.pbi.2006.03.016

1207 De Moraes, C.M., Mescher, M.C., Tumlinson, J.H., 2001. Caterpillar-induced nocturnal plant
1208 volatiles repel conspecific females. Nature 410, 577–580. https://doi.org/10.1038/35069058

1209 DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis,
1210 A.A., del Angel, G., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky, A.M.,
1211 Sivachenko, A.Y., Cibulskis, K., Gabriel, S.B., Altshuler, D., Daly, M.J., 2011. A framework for
1212 variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43,
1213 491–498. https://doi.org/10.1038/ng.806

1214 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson,
1215 M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.
1216 https://doi.org/10.1093/bioinformatics/bts635

1217 Dobin, A., Gingeras, T.R., 2015. Mapping RNA-seq reads with STAR. Curr. Protoc. Bioinforma.
1218 51, 11.14.1-11.14.19. https://doi.org/10.1002/0471250953.bi1114s51

1219 Fan, P., Leong, B.J., Last, R.L., 2019. Tip of the trichome: evolution of acylsugar metabolic
1220 diversity in Solanaceae. Curr. Opin. Plant Biol., Physiology and metabolism 49, 8–16.
1221 https://doi.org/10.1016/j.pbi.2019.03.005

1222 Fan, P., Miller, A.M., Liu, X., Jones, A.D., Last, R.L., 2017. Evolution of a flipped pathway
1223 creates metabolic innovation in tomato trichomes through BAHD enzyme promiscuity. Nat.
1224 Commun. 8, 2080. https://doi.org/10.1038/s41467-017-02045-7

1225 Fan, P., Miller, A.M., Schilmiller, A.L., Liu, X., Ofner, I., Jones, A.D., Zamir, D., Last, R.L.,
1226 2015. In vitro reconstruction and analysis of evolutionary variation of the tomato acylsucrose
1227 metabolic network. Proc. Natl. Acad. Sci. U.S.A https://doi.org/10.1073/pnas.1517930113

1228 Fan, P., Wang, P., Lou, Y.R., Leong, B.J., Moore, B.M., Schenck, C.A., Combs, R., Cao, P.,
1229 Brandizzi, F., Shiu, S.H., Last, R.L., 2020. Evolution of a plant gene cluster in solanaceae and
1230 emergence of metabolic diversity. eLife 9, 1–26. https://doi.org/10.7554/eLife.56717

1231 Fiesel, P.D., Parks, H.M., Last, R.L., Barry, C.S., 2022. Fruity, sticky, stinky, spicy, bitter,
1232 addictive, and deadly: evolutionary signatures of metabolic complexity in the Solanaceae. Nat.
1233 Prod. Rep. 39, 1438–1464. https://doi.org/10.1039/D2NP00003B

1234 Fobes, J.F., Mudd, J.B., Marsden, M.P.F., 1985. Epicuticular lipid accumulation on the leaves of
1235 *Lycopersicon pennellii* (Corr.) D'Arcy and Lycopersicon esculentum Mill. Plant Physiol. 77,
1236 567–570. https://doi.org/10.1104/pp.77.3.567

1237 Gagnon, E., Hilgenhof, R., Orejuela, A., McDonnell, A., Sablok, G., Aubriot, X., Giacomin, L.,
1238 Gouvêa, Y., Bragionis, T., Stehmann, J.R., Bohs, L., Dodsworth, S., Martine, C., Poczai, P.,
1239 Knapp, S., Särkinen, T., 2022. Phylogenomic discordance suggests polytomies along the
1240 backbone of the large genus Solanum. Am. J. Bot. 109, 580–601.
1241 https://doi.org/10.1002/ajb2.1827

1242 Ghosh, B., Westbrook, T.C., Jones, A.D., 2014a. Comparative structural profiling of trichome
1243 specialized metabolites in tomato (*Solanum lycopersicum*) and *S. habrochaites*: acylsugar
1244 profiles revealed by UHPLC/MS and NMR. Metabolomics 10, 496–507.
1245 https://doi.org/10.1007/s11306-013-0585-y

1246 Goffreda, J.C., Mutschler, M.A., Avé, D.A., Tingey, W.M., Steffens, J.C., 1989. Aphid
1247 deterrence by glucose esters in glandular trichome exudate of the wild tomato, *Lycopersicon*
1248 *pennellii*. J. Chem. Ecol. 15, 2135–2147. https://doi.org/10.1007/BF01207444

1249 Herrera-Salgado, Y., Garduño-Ramírez, M.L., Vázquez, L., Rios, M.Y., Alvarez, L., 2005. Myo-
1250 inositol-derived glycolipids with anti-inflammatory activity from *Solanum lanceolatum*. J. Nat.
1251 Prod. 68, 1031–1036. https://doi.org/10.1021/np050054s

1252 Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2:
1253 improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.
1254 https://doi.org/10.1093/molbev/msx281

1255 Hurney, S.M., 2018. Strategies for profiling and discovery of acylsugar specialized metabolites
1256 (Ph.D.). Michigan State University, United States -- Michigan.

1257 Jenske, R., Vetter, W., 2007. Highly selective and sensitive gas chromatography–electron-
1258 capture negative-ion mass spectrometry method for the indirect enantioselective identification of
1259 2- and 3-hydroxy fatty acids in food and biological samples. J. Chromatogr. A 1146, 225–231.
1260 https://doi.org/10.1016/j.chroma.2007.01.102

1261 Ji, W., Mandal, S., Rezenom, Y.H., McKnight, T.D., 2022. Specialized metabolism by trichome-
1262 enriched Rubisco and fatty acid synthase components. Plant Physiol. kiac487.
1263 https://doi.org/10.1093/plphys/kiac487

1264 Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017.
1265 ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–
1266 589. https://doi.org/10.1038/nmeth.4285

1267 Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7:
1268 improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
1269 https://doi.org/10.1093/molbev/mst010

1270 Kim, J., Kang, K., Gonzales-Vigil, E., Shi, F., Daniel Jones, A., Barry, C.S., Last, R.L., 2012.
1271 Striking natural diversity in glandular trichome acylsugar composition is shaped by variation at
1272 the acyltransferase2 locus in the wild tomato *Solanum habrochaites*. Plant Physiol. 160, 1854–
1273 1870. https://doi.org/10.1104/pp.112.204735

1274 King, R.R., Calhoun, L.A., 1988. 6 2,3-Di-O- and 1,2,3-tri-O-acylated glucose esters from the
1275 glandular trichomes of *Datura metel*. Phytochemistry 27, 3761–3763.
1276 https://doi.org/10.1016/0031-9422(88)83013-9

1277 King, R.R., Pelletier, Y., Singh, R.P., Calhoun, L.A., 1986. 3,4-Di-O-isobutyryl-6-O-
1278 caprylsucrose: the major component of a novel sucrose ester complex from the type B glandular
1279 trichomes of *Solanum berthaultii* Hawkes (Pl 473340). J. Chem. Soc. Chem. Commun. 1078–
1280 1079. https://doi.org/10.1039/C39860001078

1281 Koonin, E.V., 2005. Orthologs, paralogs, and evolutionary genomics. Annu. Rev. Genet. 39,
1282 309–338. https://doi.org/10.1146/annurev.genet.39.073003.114725

1283 Korenblum, E., Dong, Y., Szymanski, J., Panda, S., Jozwiak, A., Massalha, H., Meir, S.,
1284 Rogachev, I., Aharoni, A., 2020. Rhizosphere microbiome mediates systemic root metabolite
1285 exudation by root-to-root signaling. Proc. Natl. Acad. Sci. U. S. A. 117, 3874–3883.
1286 https://doi.org/10.1073/pnas.1912130117

1287 Kretschmar, J.A., Baumann, T.W., 1999. Caffeine in Citrus flowers. Phytochemistry 52, 19–23.
1288 https://doi.org/10.1016/S0031-9422(99)00119-3

1289 Kruse, L.H., Bennett, A.A., Mahood, E.H., Lazarus, E., Park, S.J., Schroeder, F., Moghe, G.D.,
1290 2022. Illuminating the lineage-specific diversification of resin glycoside acylsugars in the

morning glory (Convolvulaceae) family using computational metabolomics. Hortic. Res. 9, uhab079. https://doi.org/10.1093/hr/uhab079

Landry, L.G., Chapple, CCS., Last, R.L., 1995. Arabidopsis mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. Plant Physiol. 109, 1159–1166. https://doi.org/10.1104/pp.109.4.1159

Leckie, B.M., D'Ambrosio, D.A., Chappell, T.M., Halitschke, R., De Jong, D.M., Kessler, A., Kennedy, G.G., Mutschler, M.A., 2016. Differential and synergistic functionality of acylsugars in suppressing oviposition by insect herbivores. PLoS ONE. https://doi.org/10.1371/journal.pone.0153345

Leong, B.J., Hurney, S., Fiesel, P., Anthony, T.M., Moghe, G., Jones, A.D., Last, R.L., 2022. Identification of BAHD acyltransferases associated with acylinositol biosynthesis in *Solanum quitoense* (naranjilla). Plant Direct 6, e415. https://doi.org/10.1002/pld3.415

Leong, B.J., Hurney, S.M., Fiesel, P.D., Moghe, G.D., Jones, A.D., Last, R.L., 2020. Specialized metabolism in a nonmodel nightshade: trichome acylinositol biosynthesis. Plant Physiol. 183, 915–924. https://doi.org/10.1104/pp.20.00276

Leong, B.J., Lybrand, D.B., Lou, Y.R., Fan, P., Schilmiller, A.L., Last, R.L., 2019. Evolution of metabolic novelty: a trichome-expressed invertase creates specialized metabolic diversity in wild tomato. Sci. Adv. 5, 1–14. https://doi.org/10.1126/sciadv.aaw3754

Levin, D.A., 1973. The role of trichomes in plant defense. Q. Rev. Biol. 48, 3–15.

Levin, R.A., Myers, N.R., Bohs, L., 2006. Phylogenetic relationships among the "spiny solanums" (Solanum subgenus Leptostemonum, Solanaceae). Am. J. Bot. 93, 157–169. https://doi.org/10.3732/ajb.93.1.157

Li, D., Qian, J., Li, Weiliu, Yu, N., Gan, G., Jiang, Y., Li, Wenjia, Liang, X., Chen, R., Mo, Y., Lian, J., Niu, Y., Wang, Y., 2021. A high-quality genome assembly of the eggplant provides insights into the molecular basis of disease resistance and chlorogenic acid synthesis. Mol. Ecol. Resour. 21, 1274–1286. https://doi.org/10.1111/1755-0998.13321

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liu, X., Enright, M., Barry, C.S., Jones, A.D., 2017. Profiling, isolation and structure elucidation of specialized acylsucrose metabolites accumulating in trichomes of Petunia species. Metabolomics 13, 85. https://doi.org/10.1007/s11306-017-1224-9

Lou, Y.-R., Anthony, T.M., Fiesel, P.D., Arking, R.E., Christensen, E.M., Jones, A.D., Last, R.L., 2021. It happened again: convergent evolution of acylglucose specialized metabolism in black nightshade and wild tomato. Sci. Adv. 7, eabj8726. https://doi.org/10.1126/sciadv.abj8726

Lou, Y.-R., Leong, B., 2019. Leaf surface acylsugar extraction and LC-MS profiling - v1.0.

1328 Luu, V.T., Weinhold, A., Ullah, C., Dressel, S., Schoettner, M., Gase, K., Gaquerel, E., Xu, S.,
1329 Baldwin, I.T., 2017. O-acyl sugars protect a wild tobacco from both native fungal pathogens and
1330 a specialist herbivore. Plant Physiol. 174, 370–386. https://doi.org/10.1104/pp.16.01904

1331 Lybrand, D.B., Anthony, T.M., Jones, A.D., Last, R.L., 2020. An integrated analytical approach
1332 reveals trichome acylsugar metabolite diversity in the wild tomato *Solanum pennellii*.
1333 Metabolites 10, 1–25. https://doi.org/10.3390/metabo10100401

1334 Maldonado, E., Torres, F.R., Martínez, M., Pérez-Castorena, A.L., 2006. Sucrose esters from the
1335 fruits of *Physalis nicandroides* var. attenuata. J. Nat. Prod. 69, 1511–1513.
1336 https://doi.org/10.1021/np060274l

1337 Matsuzaki, T., Shinozaki, Y., Suhara, S., Ninomiya, M., Shigematsu, H., Koiwai, A., 1989.
1338 Isolation of glycolipids from the surface lipids of *Nicotiana bigelovii* and their distribution in
1339 *Nicotiana* species. Agric. Biol. Chem. 53, 3079–3082. https://doi.org/10.1271/bbb1961.53.3079

1340 McCarthy, D.J., Campbell, K.R., Lun, A.T.L., Wills, Q.F., 2017. Scater: pre-processing, quality
1341 control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics 33,
1342 1179–1186. https://doi.org/10.1093/bioinformatics/btw777

1343 McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella,
1344 K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: A
1345 MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20,
1346 1297–1303. https://doi.org/10.1101/gr.107524.110

1347 Mennella, G., Rotino, G.L., Fibiani, M., D'Alessandro, A., Francese, G., Toppino, L., Cavallanti,
1348 F., Acciarri, N., Lo Scalzo, R., 2010. Characterization of health-related compounds in eggplant
1349 (*Solanum melongena* L.) lines derived from introgression of allied species. J. Agric. Food Chem.
1350 58, 7597–7603. https://doi.org/10.1021/jf101004z

1351 Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A.,
1352 Lanfear, R., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in
1353 the genomic era. Mol. Biol. Evol. 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015

1354 Moghe, G., Kruse, L.H., Petersen, M., Scossa, F., Fernie, A.R., Gaquerel, E., D'Auria, J.C.,
1355 2023. BAHD company: the ever-expanding roles of the BAHD acyltransferase gene family in
1356 plants. Annu. Rev. Plant Biol. 74, annurev-arplant-062922-050122.
1357 https://doi.org/10.1146/annurev-arplant-062922-050122

1358 Moghe, G.D., Leong, B.J., Hurney, S.M., Jones, A.D., Last, R.L., 2017. Evolutionary routes to
1359 biochemical innovation revealed by integrative analysis of a plant-defense related specialized
1360 metabolic pathway. eLife 6, 1–33. https://doi.org/10.7554/eLife.28468

1361 Nadakuduti, S.S., Uebler, J.B., Liu, X., Jones, A.D., Barry, C.S., 2017. Characterization of
1362 trichome-expressed BAHD acyltransferases in *Petunia axillaris* reveals distinct acylsugar
1363 assembly mechanisms within the Solanaceae. Plant Physiol. 175, 36–50.
1364 https://doi.org/10.1104/pp.17.00538

Nee, M., 2022a. *Solanum acerifolium* [Dunal] [WWW Document]. URL https://solanaceaesource.myspecies.info/taxonomy/term/105367/descriptions (accessed 10.3.22).

Nee, M., 2022b. *Solanum atropurpureum* [Schrank] [WWW Document]. URL https://solanaceaesource.myspecies.info/taxonomy/term/105749/descriptions (accessed 10.3.22).

Nee, M., 1991. Synopsis of Solanum section Acanthophora: a group of interest for glycoalkaloids, in: Solanaceae III, taxonomy, chemistry, evolution. Royal Botanic Gardens for the Linnean Society of London, Kew, Richmond, Surrey, UK, pp. 257–266.

Ning, J., Moghe, G.D., Leong, B., Kim, J., Ofner, I., Wang, Z., Adams, C., Jones, A.D., Zamir, D., Last, R.L., 2015. A feedback-insensitive isopropylmalate synthase affects acylsugar composition in cultivated and wild tomato. Plant Physiol. 169, 1821–1835. https://doi.org/10.1104/pp.15.00474

PBI Solanum Project, 2022. Solanaceae Source [WWW Document]. URL https://solanaceaesource.org (accessed 10.3.22).

Pereda-Miranda, R., Mata, R., Anaya, A.L., Wickramaratne, D.B.M., Pezzuto, J.M., Kinghorn, A.D., 1993. Tricolorin A, major phytogrowth inhibitor from *Ipomoea tricolor*. J. Nat. Prod. 56, 571–582. https://doi.org/10.1021/np50094a018

Pinot, F., Beisson, F., 2011. Cytochrome P450 metabolizing fatty acids in plants: characterization and physiological roles. FEBS J. 278, 195–205. https://doi.org/10.1111/j.1742-4658.2010.07948.x

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. https://doi.org/10.1093/bioinformatics/btp616

Sampaio, V. da S., Coutinho, Í.A.C., Särkinen, T., Loiola, M.I.B., 2021. Secretory and ecological function of petiolar glands in *Solanum fernandesii*: first description of resin glands in the genus Solanum. Aust. J. Bot. 70, 32–41. https://doi.org/10.1071/BT21001

Särkinen, T., Bohs, L., Olmstead, R.G., Knapp, S., 2013. A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. BMC Evol. Biol. 13, 214. https://doi.org/10.1186/1471-2148-13-214

Sassaki, G.L., Souza, L.M., Serrato, R.V., Cipriani, T.R., Gorin, P.A.J., Iacomini, M., 2008. Application of acetate derivatives for gas chromatography–mass spectrometry: novel approaches on carbohydrates, lipids and amino acids analysis. J. Chromatogr. A 1208, 215–222. https://doi.org/10.1016/j.chroma.2008.08.083

Schenck, C.A., Anthony, T.M., Jacobs, M., Jones, A.D., Last, R.L., 2022. Natural variation meets synthetic biology: promiscuous trichome-expressed acyltransferases from Nicotiana. Plant Physiol. kiac192. https://doi.org/10.1093/plphys/kiac192

1400 Schilmiller, A.L., Charbonneau, A.L., Last, R.L., 2012. Identification of a BAHD
1401 acetyltransferase that produces protective acyl sugars in tomato trichomes. Proc. Natl. Acad. Sci.
1402 U.S.A. 109, 16377–16382. https://doi.org/10.1073/pnas.1207906109

1403 Schilmiller, A.L., Gilgallon, K., Ghosh, B., Jones, A.D., Last, R.L., 2016. Acylsugar
1404 acylhydrolases: carboxylesterase-catalyzed hydrolysis of acylsugars in tomato trichomes. Plant
1405 Physiol. 170, 1331–1344. https://doi.org/10.1104/pp.15.01348

1406 Schilmiller, A.L., Moghe, G.D., Fan, P., Ghosh, B., Ning, J., Jones, A.D., Last, R.L., 2015.
1407 Functionally divergent alleles and duplicated Loci encoding an acyltransferase contribute to
1408 acylsugar metabolite diversity in Solanum trichomes. Plant Cell 27, 1002–17.
1409 https://doi.org/10.1105/tpc.15.00087

1410 Shapiro, J.A., Steffens, J.C., Mutschler, M.A., 1994. Acylsugars of the wild tomato *Lycopersicon*
1411 *pennellii* in relation to geographic distribution of the species. Biochem. Syst. Ecol. 22, 545–561.
1412 https://doi.org/10.1016/0305-1978(94)90067-1

1413 Smith, C.R., 1971. Occurrence of unusual fatty acids in plants. Prog. Chem. Fats Other Lipids
1414 11, 137–177. https://doi.org/10.1016/0079-6832(71)90005-X

1415 Song, B., Song, Y., Fu, Y., Kizito, E.B., Kamenya, S.N., Kabod, P.N., Liu, H., Muthemba, S.,
1416 Kariba, R., Njuguna, J., Maina, S., Stomeo, F., Djikeng, A., Hendre, P.S., Chen, X., Chen, W.,
1417 Li, X., Sun, W., Wang, S., Cheng, S., Muchugi, A., Jamnadass, R., Shapiro, H.-Y., Van Deynze,
1418 A., Yang, H., Wang, J., Xu, X., Odeny, D.A., Liu, X., 2019. Draft genome sequence of *Solanum*
1419 *aethiopicum* provides insights into disease resistance, drought tolerance, and the evolution of the
1420 genome. GigaScience 8, giz115. https://doi.org/10.1093/gigascience/giz115

1421 Stern, S., Agra, M. de F., Bohs, L., 2011. Molecular delimitation of clades within New World
1422 species of the "spiny solanums" (Solanum subg. Leptostemonum). TAXON 60, 1429–1441.
1423 https://doi.org/10.1002/tax.605018

1424 Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W.-M.,
1425 Fiehn, O., Goodacre, R., Griffin, J.L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka,
1426 J., Lane, A.N., Lindon, J.C., Marriott, P., Nicholls, A.W., Reily, M.D., Thaden, J.J., Viant, M.R.,
1427 2007. Proposed minimum reporting standards for chemical analysis. Metabolomics 3, 211–221.
1428 https://doi.org/10.1007/s11306-007-0082-2

1429 Tepe, E.J., Anderson, G.J., Spooner, D.M., Bohs, L., 2016. Relationships among wild relatives
1430 of the tomato, potato, and pepino. TAXON 65, 262–276. https://doi.org/10.12705/652.4

1431 Van Dam, N.M., Hare, J.D., 1998. Biological activity of *Datura wrightii* glandular trichome
1432 exudate against *Manduca sexta* larvae. J. Chem. Ecol. 24, 1529–1549.
1433 https://doi.org/10.1023/A:1020963817685

1434 Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine,
1435 A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K.V., Altshuler, D.,
1436 Gabriel, S., DePristo, M.A., 2013. From FastQ data to high-confidence variant calls: the genome

analysis toolkit best practices pipeline. Curr. Protoc. Bioinforma. 43, 11.10.1-11.10.33. https://doi.org/10.1002/0471250953.bi1110s43

Wagner, G.J., 1991. Secreting glandular trichomes: more than just hairs. Plant Physiol. 96, 675–679. https://doi.org/10.1104/pp.96.3.675

Weese, T.L., Bohs, L., 2007. A three-gene phylogeny of the genus Solanum (Solanaceae). Syst. Bot. 32, 445–463.

Weinhold, A., Baldwin, I.T., 2011. Trichome-derived O-acyl sugars are a first meal for caterpillars that tags them for predation. Proc. Natl. Acad. Sci. 108, 7855–7859. https://doi.org/10.1073/pnas.1101306108

Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.-Y., 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 8, 28–36. https://doi.org/10.1111/2041-210X.12628

Yu, P., He, X., Baer, M., Beirinckx, S., Tian, T., Moya, Y.A.T., Zhang, X., Deichmann, M., Frey, F.P., Bresgen, V., Li, C., Razavi, B.S., Schaaf, G., von Wirén, N., Su, Z., Bucher, M., Tsuda, K., Goormachtig, S., Chen, X., Hochholdinger, F., 2021. Plant flavones enrich rhizosphere Oxalobacteraceae to improve maize performance under nitrogen deprivation. Nat. Plants 7, 481–499. https://doi.org/10.1038/s41477-021-00897-y

Zizovic, I., Stamenić, M., Orlović, A., Skala, D., 2007. Supercritical carbon dioxide extraction of essential oils from plants with secretory ducts: mathematical modelling on the micro-scale. J. Supercrit. Fluids 39, 338–346. https://doi.org/10.1016/j.supflu.2006.03.009

Supplementary Figures for

Trading acyls and swapping sugars: metabolic innovations in *Solanum* trichomes

Paul D. Fiesel[1], Rachel E. Kerwin[1], A. Daniel Jones[1], and Robert L. Last[1,2]

[1]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48823 USA

[2]Department of Plant Biology, Michigan State University, East Lansing, MI 48823 USA

Corresponding Author: Robert L. Last, Michigan State University, Department of Biochemistry and Molecular Biology, 603 Wilson Road, East Lansing, MI 48824; Phone: 517-432-3278; email: lastr@msu.edu

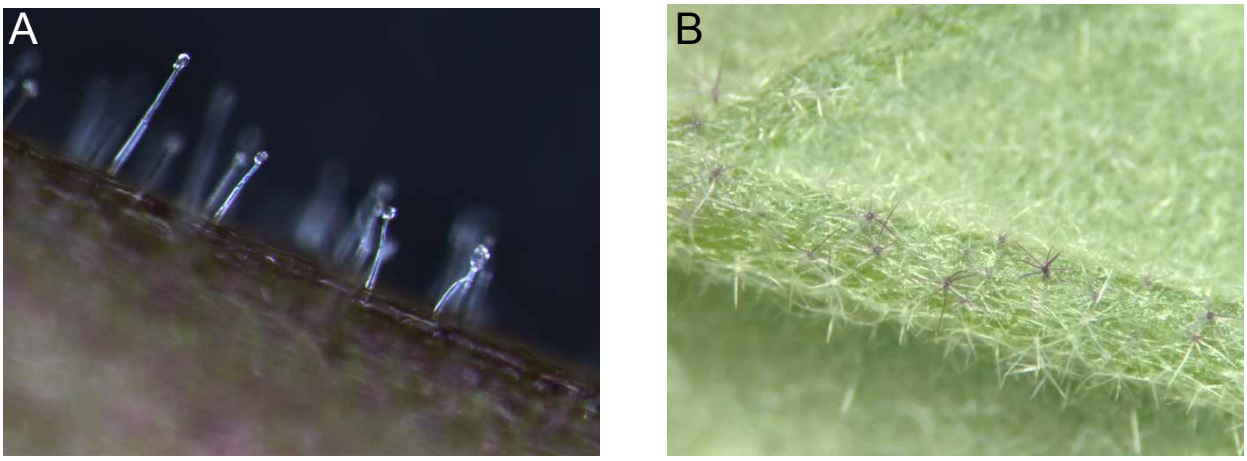


Figure S1. *S. melongena* produces single stalked glandular trichomes on seedling above ground tissue but stellate nonglandular trichomes on non-seedling above ground tissue. (A) Close up photo of a *S. melongena* hypocotyl displaying glandular trichomes with similar morphology to the acylsugar-producing *S. lycopersicum* Type I/IV trichomes (Luckwill, 1943; Schilmiller et al., 2012). (B) Close up photo of a *S. melongena* leaf from a reproductive stage plant displaying non-glandular stellate trichomes. Leaf surface metabolite extracts from this tissue do not have detectable acylsugars by LC-MS.

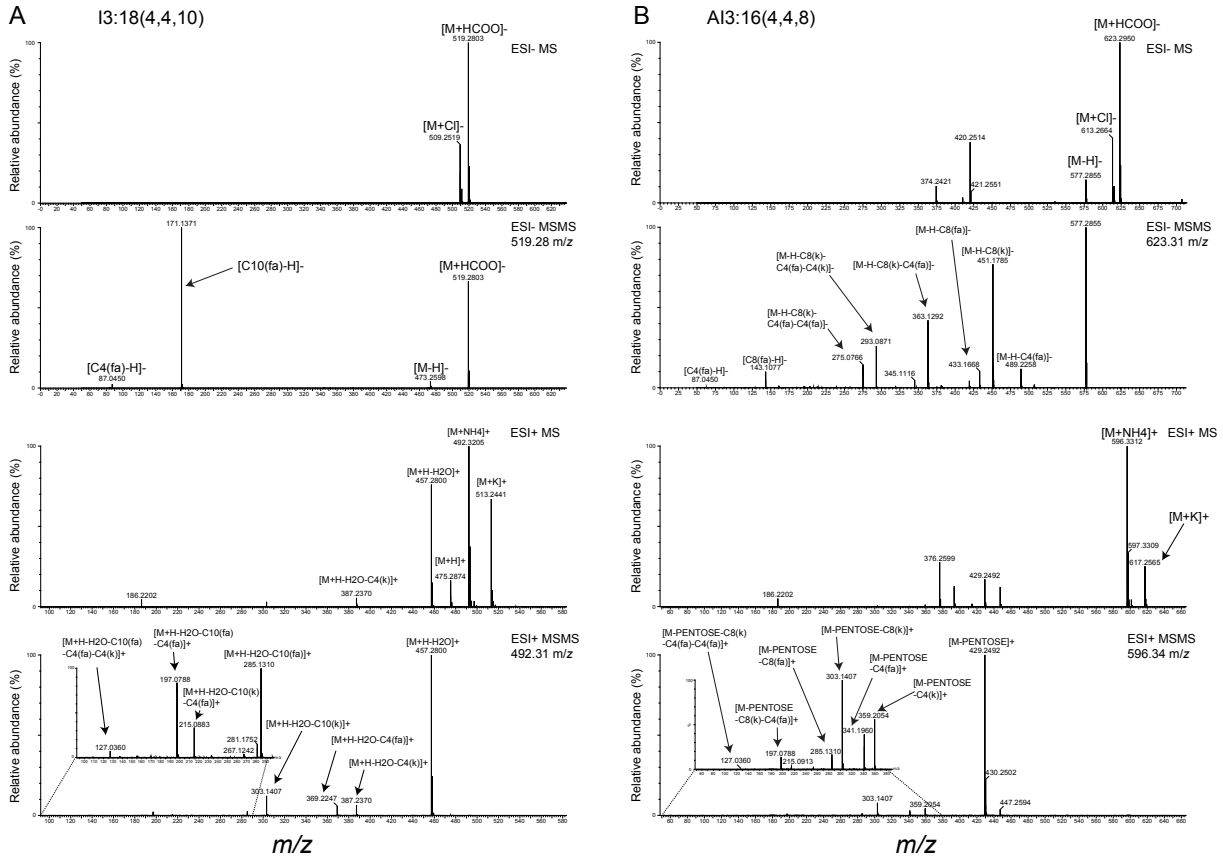


Figure S2. Annotation of I3:18(4,4,10) and AI3:16(4,4,8) from *S. melongena* 67/3 using negative and positive mode MS and CID MS/MS fragmentation. (A) I3:18(4,4,10) negative and positive mode MS and MS/MS fragmentation. (B) AI3:16(4,4,8) negative and positive MS and MS/MS fragmentation. For each compound, ESI- MS (top panels) display the formate adduct accurate mass which was used to determine chemical formulas. ESI- MS/MS (second from the top panels) exhibit acyl chain carboxylate fragment ions for both compounds and stepwise loss of acyl chains for only AI3:16. ESI+ MS/MS exhibit stepwise losses of acyl chains for both compounds. Positive mode CID of AI3:16 also produces a fragment ion (*m/z* 429.2492) corresponding to the neutral loss of the pentose moiety. This indicates that all acyl chains reside on the hexose ring. The acylsugar LC-MS annotation method is described in the Materials and Methods. fa = fatty acid; k = ketene.

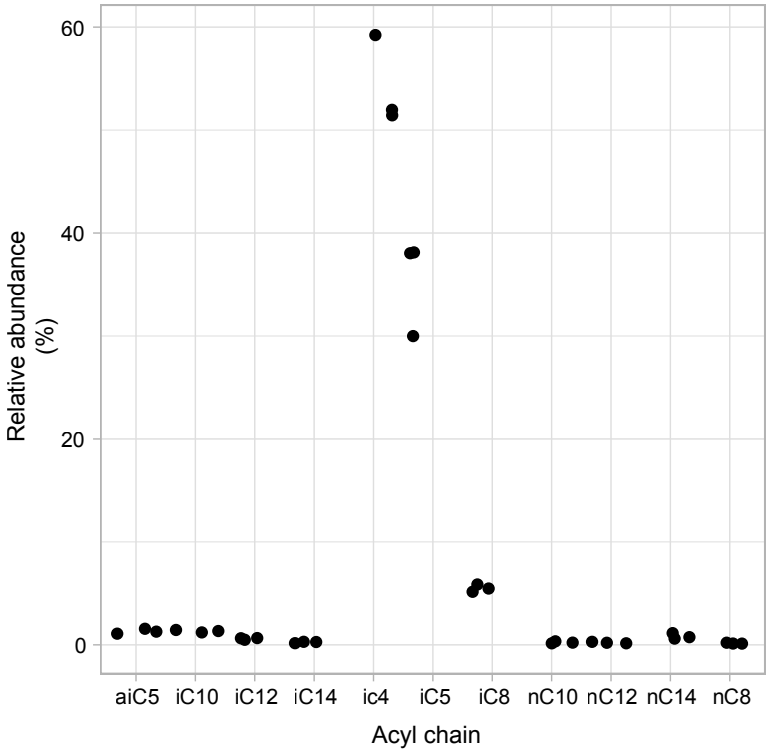


Figure S3. Acyl chain composition of *S. melongena* acylsugars. Relative abundance of acyl chains shown between three *S. melongena* PI 555598 leaf surface extracts from different seedlings. Straight acyl chains were identified with authentic reference standards and iso-branched chains were identified with NIST mass spectral library searches. Hydroxylated acyl chains were not included in this analysis and only detected acyl chains are included.

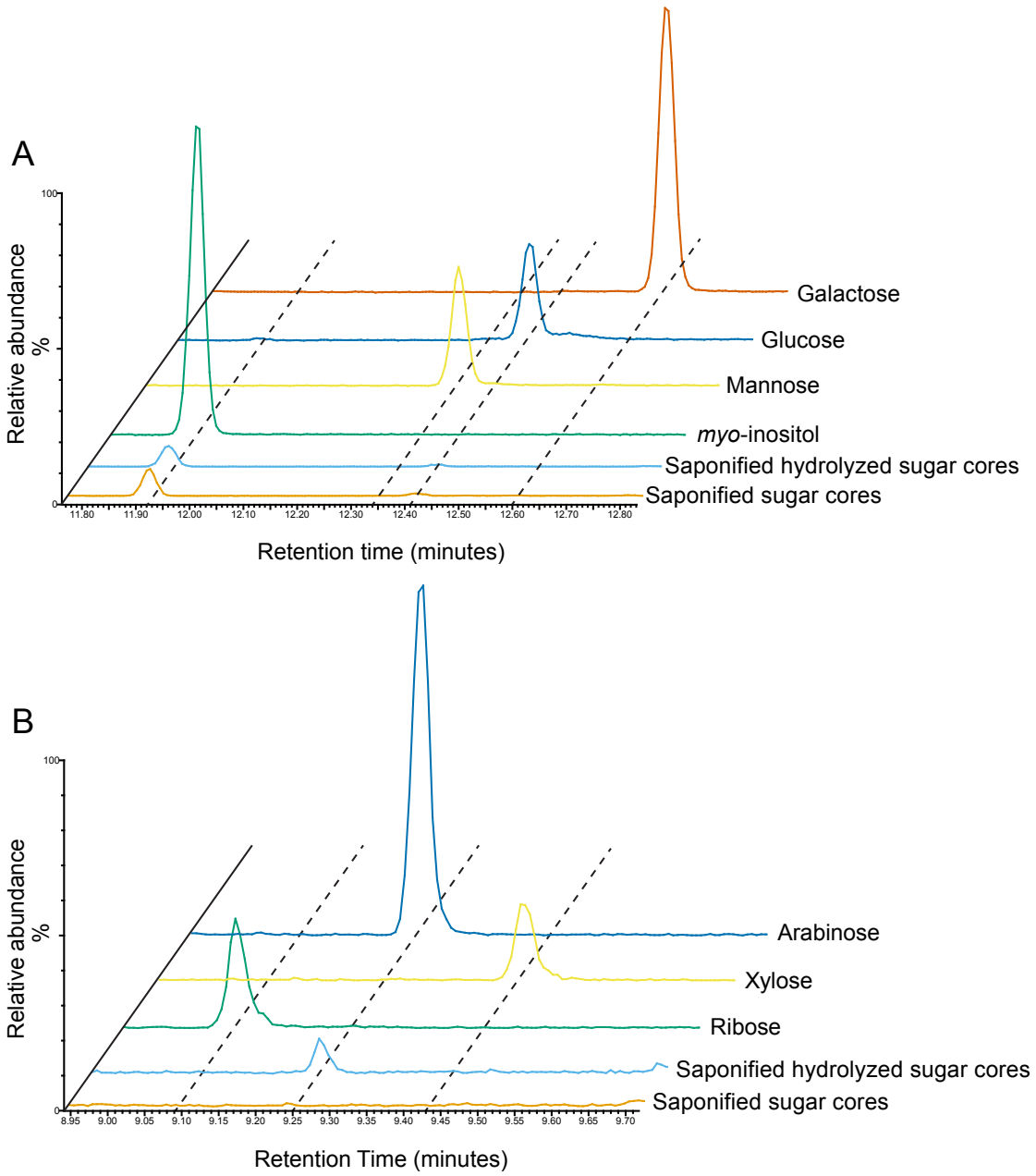


Figure S4: Identification of *S. melongena* acylsugar core composition through GC-MS analysis of alditol acetate sugar derivatives. *S. melongena* acylsugars collected from surface extracts were first saponified to remove acyl chains, and then, with or without acid hydrolysis to break the glycosidic linkages, sugar cores were derivatized to alditol acetates. (A) Alditol acetate derivatization of saponified *S. melongena* acylsugars yield a peak that comigrated with that of a *myo*-inositol standard. The traces displayed are GC-MS total ion chromatograms (TICs). (B) Alditol acetate pentose derivatives of saponified and hydrolyzed acylsugar cores comigrate with

that of a derivatized arabinose standard. The traces displayed are GC-MS TICs. Vertical scales for both panels are normalized to the largest signal within the displayed region.

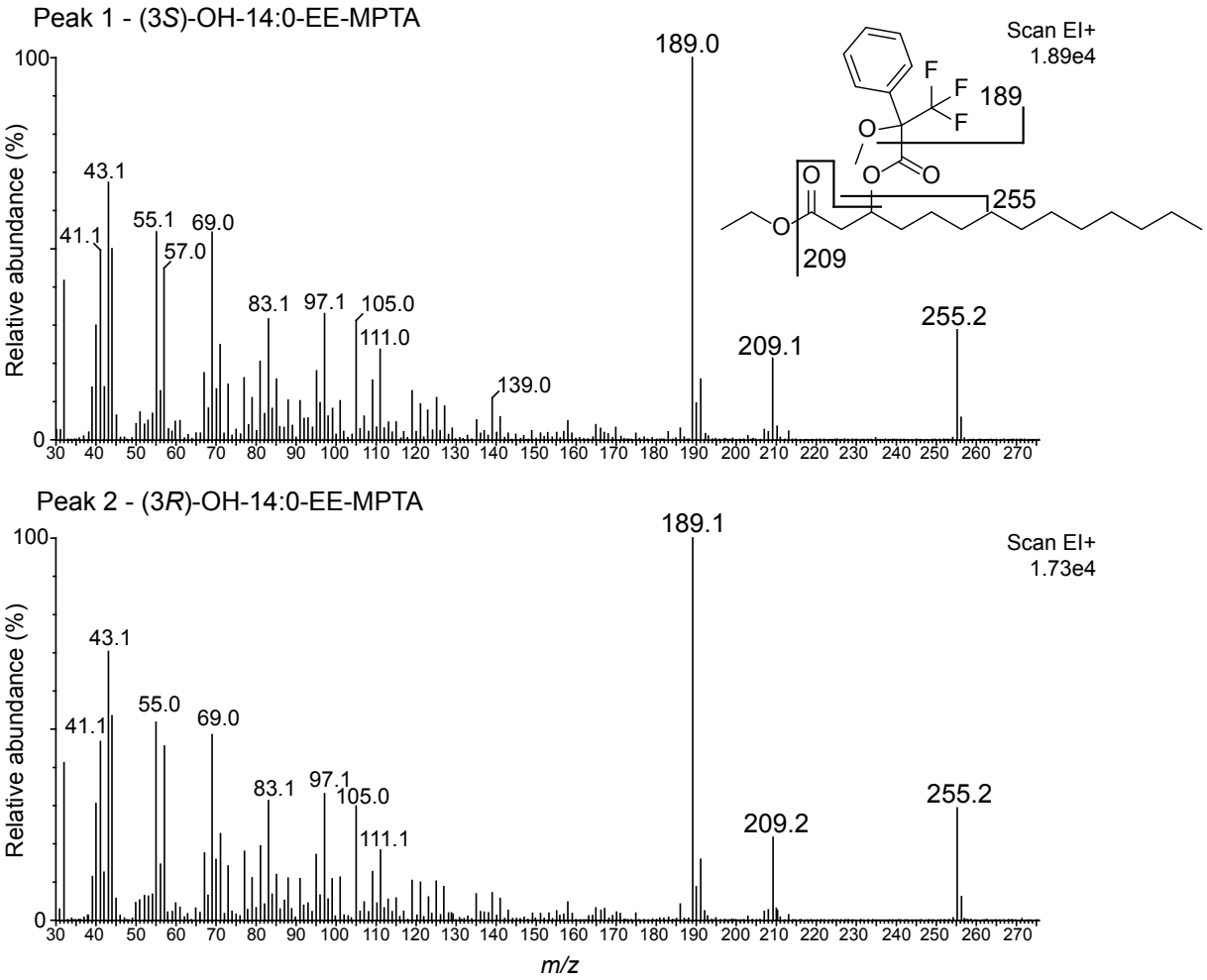


Figure S5. GC-MS full scan mass spectra of 3-OH-14:0 ethyl ester stereoisomers as their MPTA derivatives. The mass spectra contain ions at *m/z* 189, 209, and 255 corresponding to the expected fragmentation of the derivatized fatty acid. EE = ethyl ester. Peak numbers correspond to the peaks displayed in Figure 2.

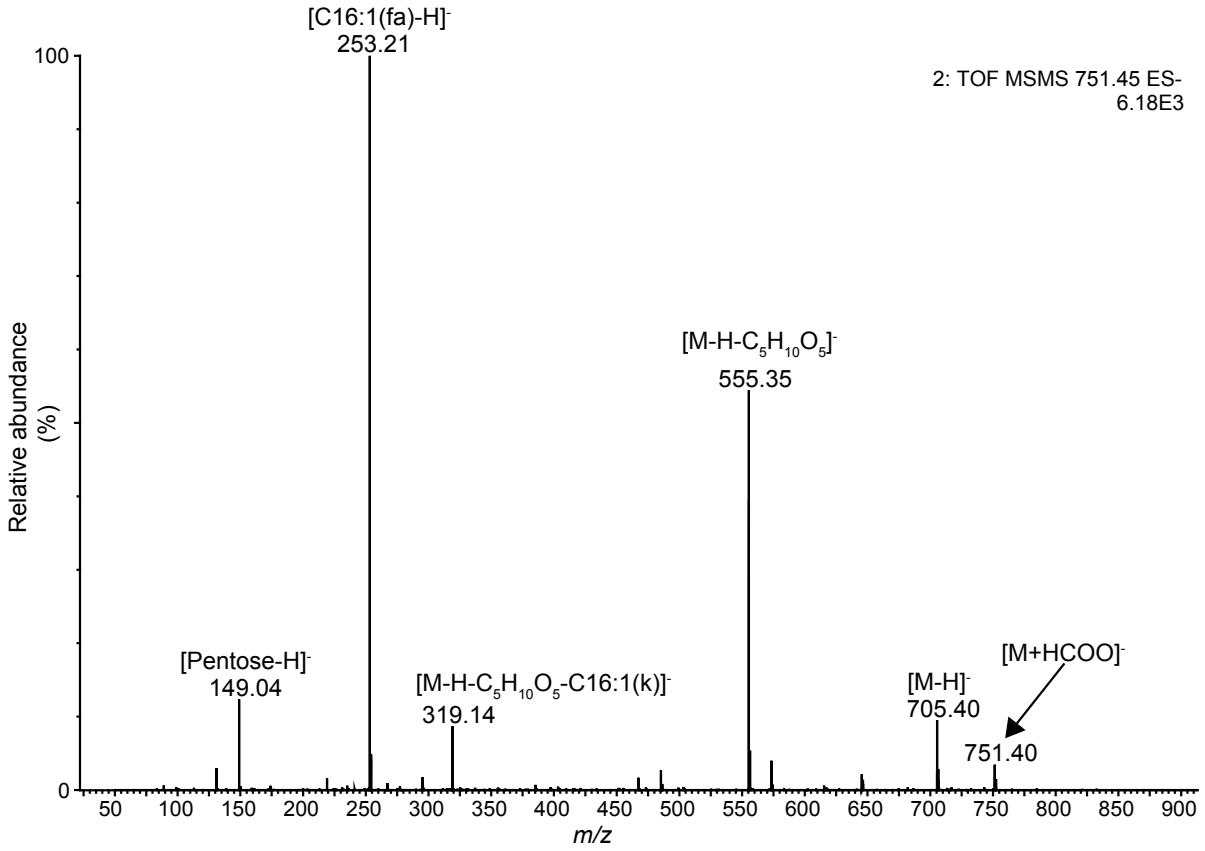


Figure S6. Negative mode MS/MS fragmentation of H3:22(4,4,16-O-p) from *S. sisymbriifolium* results in neutral loss of a pentose moiety. This neutral loss of a sugar group ($C_5H_{10}O_5$ plus formic acid) from the [M+formate]$^-$ ion *m/z* 751.4025 to form *m/z* 555.3522 is not observed with negative CID of other acyldisaccharides (Figures S2 and S7) suggesting an unusual glycosidic linkage on the acyl chain. fa = fatty acid; k = ketene; 16-O-p = pentosylated 16 carbon acyl chain.

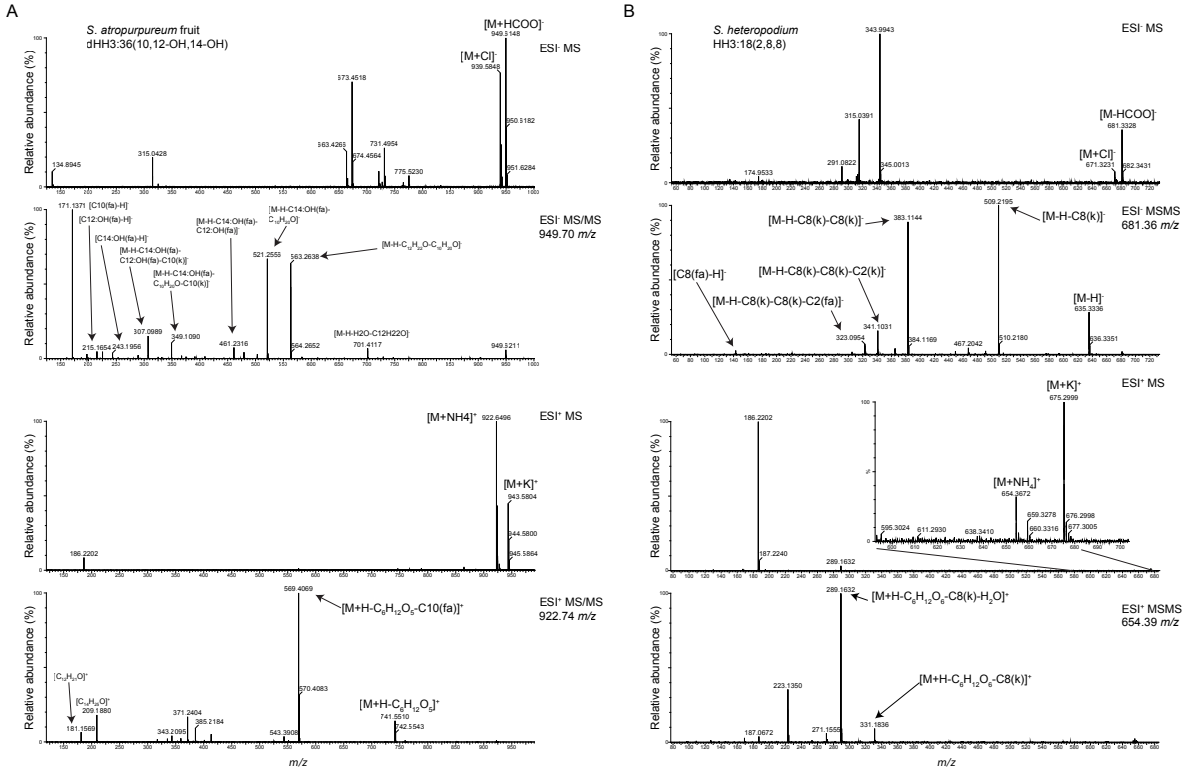


Figure S7. Negative and positive mode MS and CID MS/MS fragmentation of dHH3:36(10,12-OH,14-OH) from a *S. atropurpureum* fruit surface extract and HH3:18(2,8,8) from a *S. heteropodium* leaf surface extract. Formate and ammonium adduct ions were selected by data-dependent acquisition software and fragmented with a ramped collision energy detailed in the Methods. Under negative mode CID MS/MS, each compound produces fragment ions corresponding to stepwise acyl chain loss down to a free sugar core fragment ion. In contrast, under positive mode CID MS/MS each compound produces fragment ions corresponding to neutral loss of a sugar moiety. (A) dHH3:36(10,12-OH,14-OH) MS and MS/MS fragmentation. Positive mode CID produces a fragment ion (*m/z* 569.4069) corresponding to a neutral loss of a deoxyhexose acylated with a ten carbon acyl chain. dHH = deoxyhexose-hexose (B) HH3:18(2,8,8) MS and MS/MS fragmentation. Positive mode CID produces a fragment ion (*m/z* 331.1836) corresponding to a neutral loss of a hexose acylated with an eight carbon acyl chain. HH = hexose-hexose; fa = fatty acid; k = ketene.

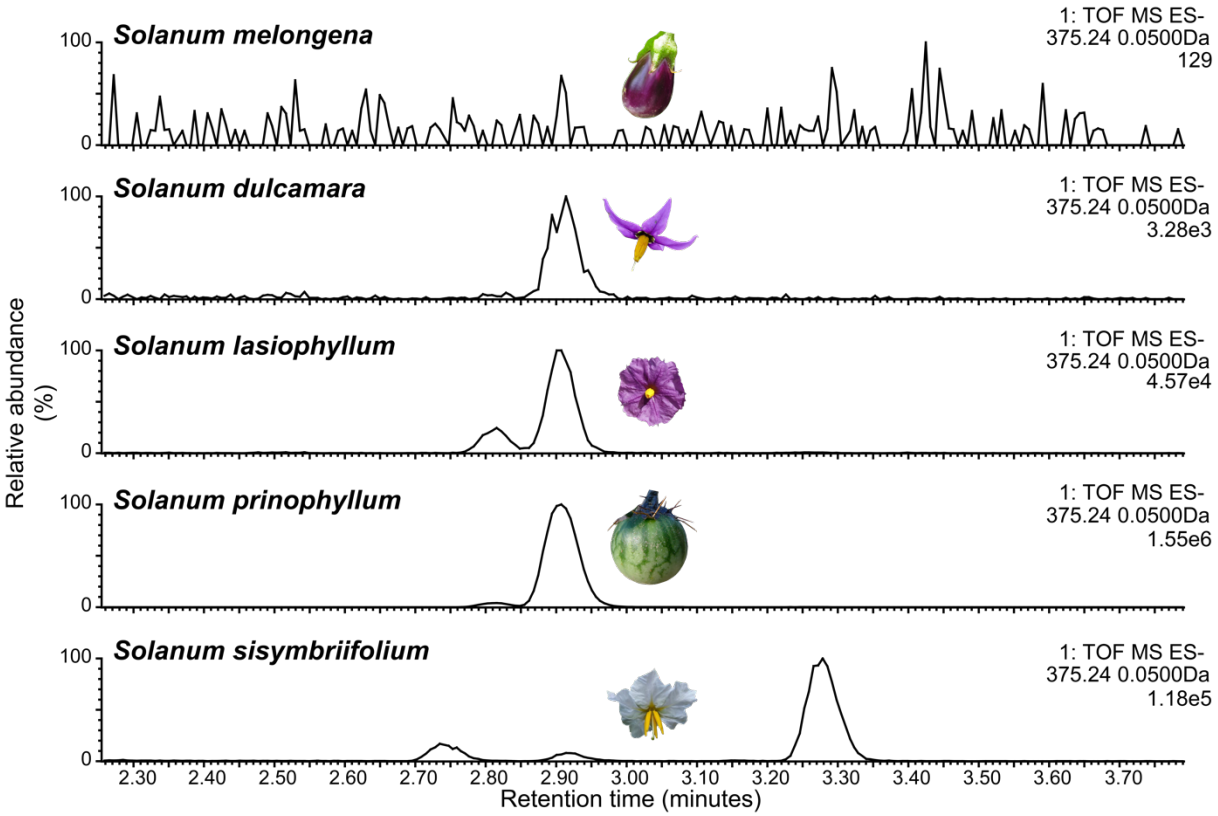


Figure S8. Acylsugar saponification produces free glycosylated fatty acids in four *Solanum* species supporting the identification of glycohydroxyacyl hexoses. All traces are extracted ion LC/MS chromatograms for pentosylated hydroxytetradecanoic acid, *m/z* 375.24±0.05 (retention time 2.90 min) from saponified leaf surface extracts. *S. melongena* does not contain detectable glycosylated fatty acids and acts as a negative control. Vertical scale is normalized to the largest signal within the displayed region (values in the upper right of each chromatogram).

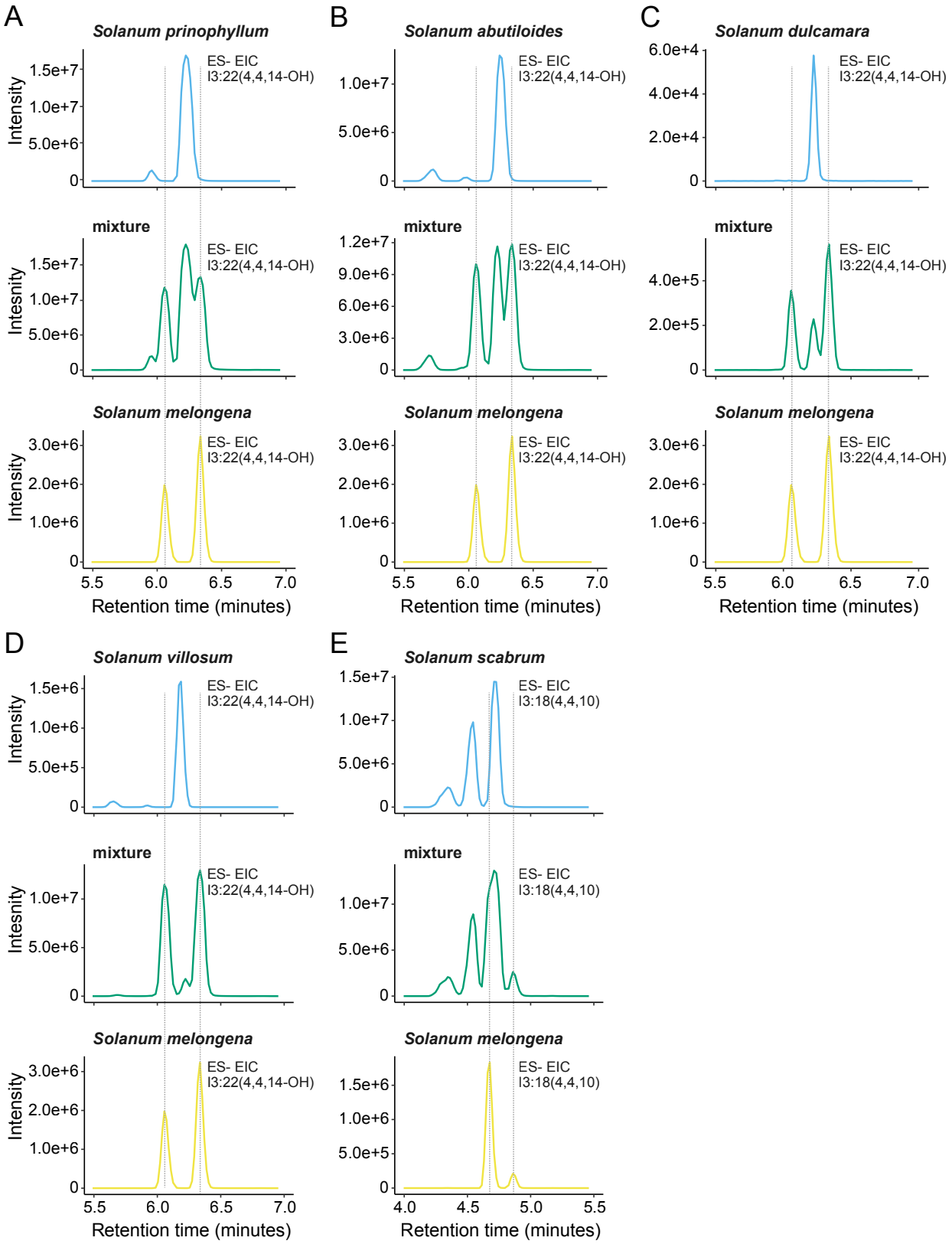

A
Solanum prinophyllum
mixture
Solanum melongena
ES- EIC
I3:22(4,4,14-OH)
Intensity
Intesnity
Retention time (minutes)
B
Solanum abutiloides
C
Solanum dulcamara
D
Solanum villosum
E
Solanum scabrum
ES- EIC
I3:18(4,4,10)

Figure S9. Coelution analysis reveals that five *Solanum* species accumulate acylsugar positional or branching isomers that differ from *S. melongena* acylsugars. Extracted ion chromatograms were generated for combined signals of *m/z* 591.34 and 519.28 corresponding to I3:22(4,4,14-OH) and I3:18(4,4,10) $[M+formate]^-$ adducts, respectively. The yellow traces (bottom row) represent *S. melongena*, the blue traces represent the five other *Solanum* species (*S. prinophylum, S. abutiloides, S. dulcamara, S. villosum*, and *S. scabrum*), and the middle row green traces represent acylsugar mixtures between *S. melongena* and the species in the trace below. The mixed samples allow for corrections of retention time drift that may occur between samples.

S. americanum
G3:15(2,5,8)

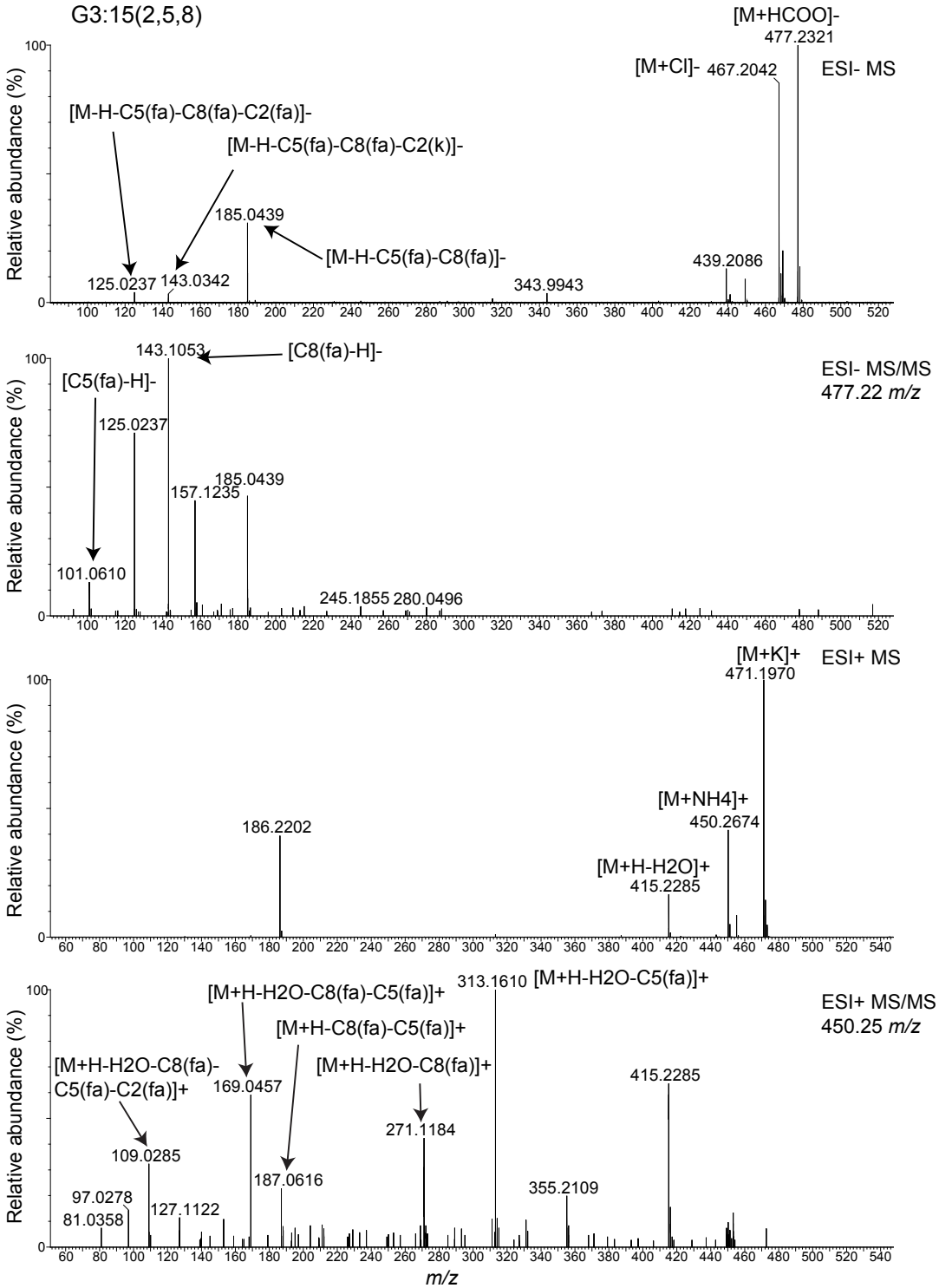


Figure S10. Negative and positive mode MS and CID MS/MS fragmentation of G3:15(2,5,8) from *S. americanum*. Acylglucoses fragment characteristically in negative mode MS and MS/MS functions producing fragment ions corresponding to stepwise acyl chain loss. This is in stark contrast to acylinositol negative mode fragmentation which does not produce major fragment ions corresponding to stepwise acyl chain loss (Figure S2). This fragmentation difference distinguishes the two sugar cores and enables their annotation by LC-MS. G3:15 formate and ammonium adduct ions were selected by data-dependent acquisition software and fragmented with a ramped collision energy detailed in the Methods. fa = fatty acid; k = ketene.

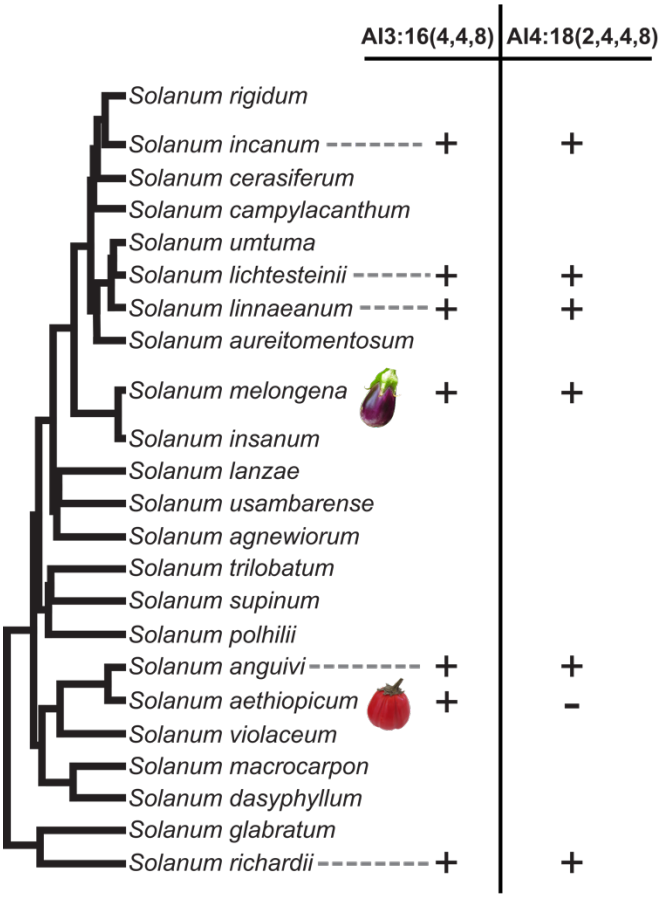


Figure S11. Phylogenetic analysis of Eggplant Clade and Anguivi Grade species reveals *S. aethiopicum* is the only analyzed acylsugar-producing species to not accumulate detectable AI4:18(2,4,4,8). AI3:16(4,4,8) and AI4:18(2,4,4,8) presence and absence in surface extracts as detected by LC-MS are plotted on a phylogeny of the Eggplant Clade and Anguivi Grade. Species that either do not accumulate detectable acylsugars (*S. macrocarpon* and *S. virginianum)* or were not analyzed did not have AI3:16 and AI4:18 presence plotted. The phylogeny was modified from a previously published version (Aubriot et al., 2018). Full acylsugar profiles are detailed Tables 1, S3-8.

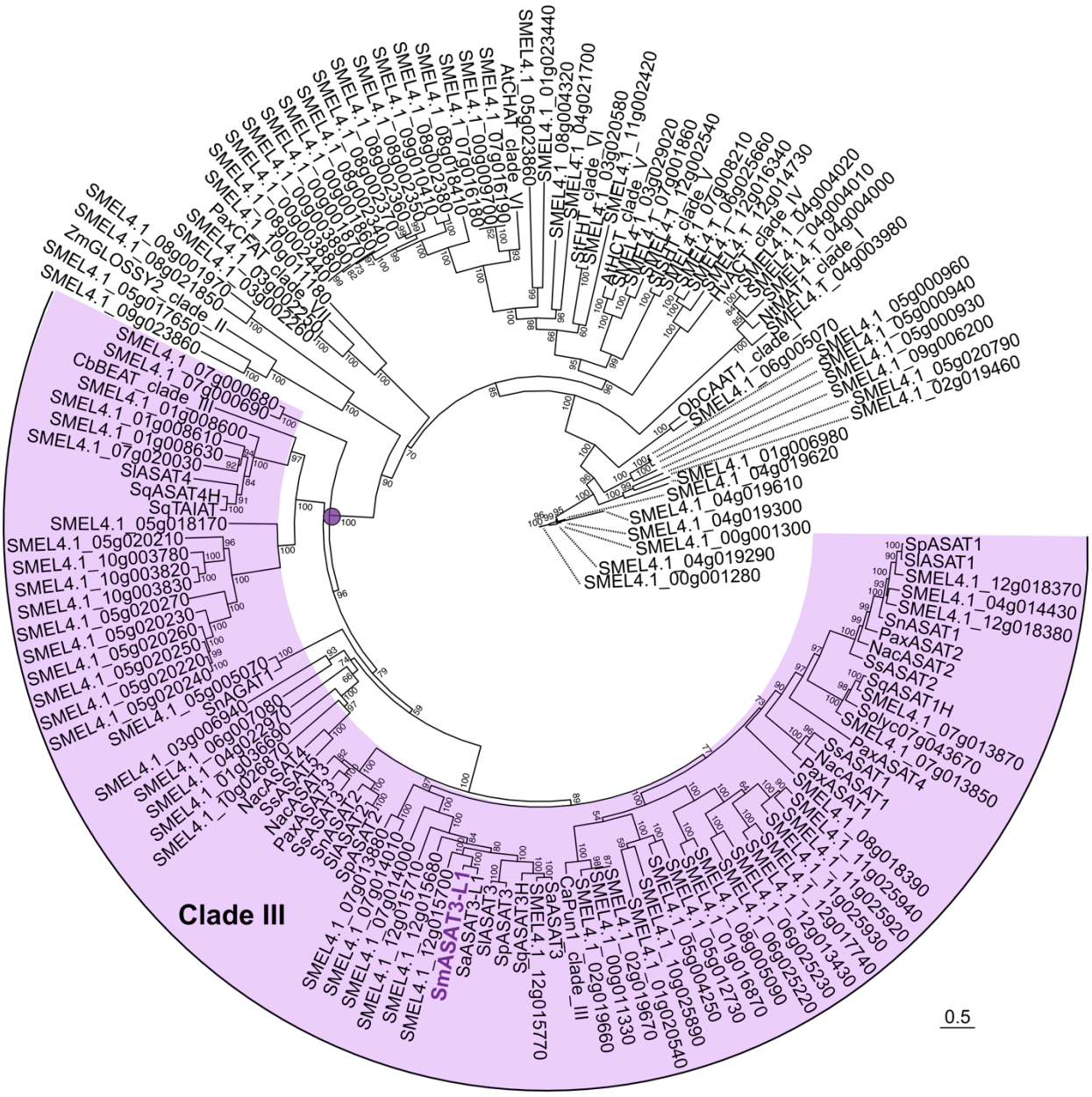


Figure S12. ***S. melongena*** **BAHD tree.** Shown is a phylogeny including 106 predicted BAHDs (PF002458) in the eggplant Smel_V4.1 reference genome, published reference BAHD sequences for clades I-VII, characterized ASAT sequences from other Solanaceae species, and the SaASAT3 and SaASAT3-L1 candidates from *S. aethiopicum*. The maximum likelihood tree was inferred from amino acid sequences using the Jones-Taylor-Thornton algorithm with seven rate categories in IQ-TREE v2.1.3. Bootstrap support calculated from 100,000 ultrafast bootstrap replicates and values >=95 indicate strong support. Clade III BAHDs are highlighted in purple,

with a purple dot marking the node. Branch lengths represent substitution rates (substitutions per site) as denoted by scale bar.

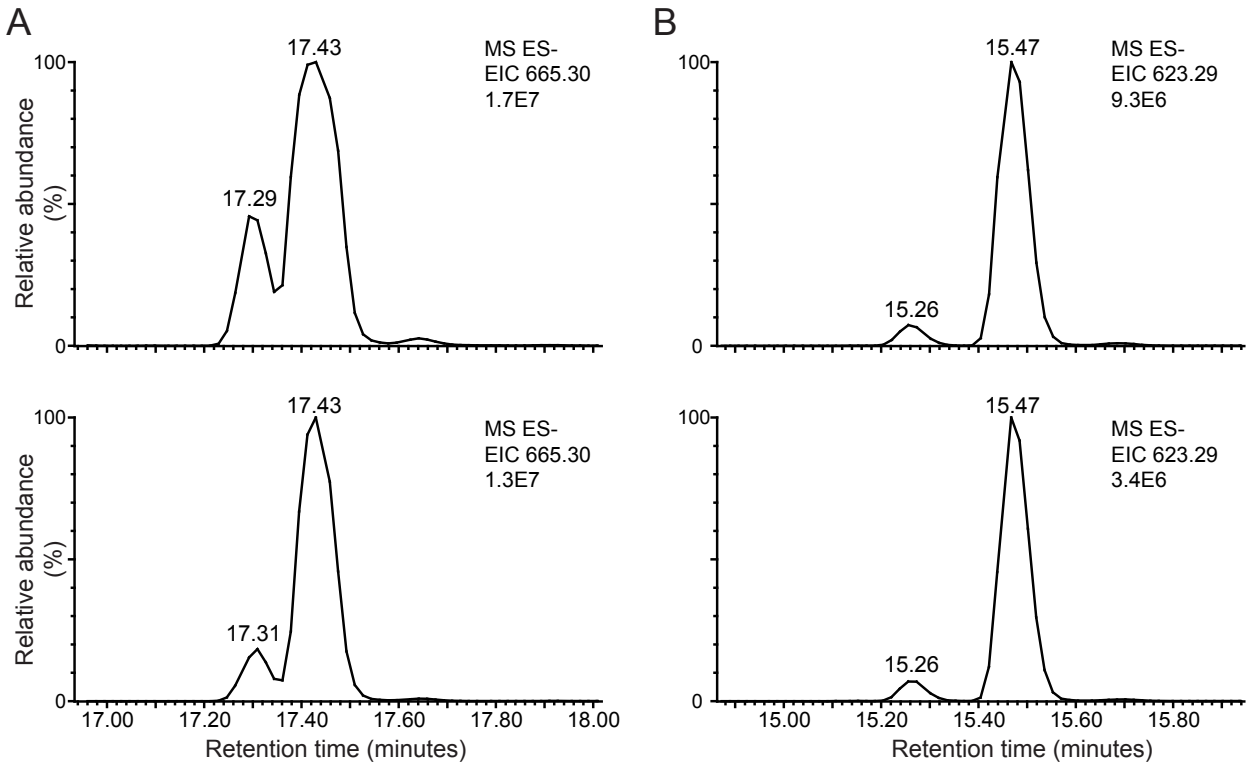


Figure S13. SmASAT3-L1 forward and reverse assay produced AI4:18 and AI3:16, respectively, coelute with plant produced AI4:18 and AI3:16. (A) Forward assay produced AI4:18 (top) coelutes with AI4:18 from a *S. melongena* 555598 leaf surface extract (bottom). The extracted ion chromatograms display the formate adduct of AI4:18, *m/z* 665.30. (B) Reverse assay produced AI3:16 (top) coelutes with AI3:16 from a *S. melongena* 555598 leaf surface extract (bottom). The extracted ion chromatograms display the formate adduct of AI3:16, *m/z* 623.20.

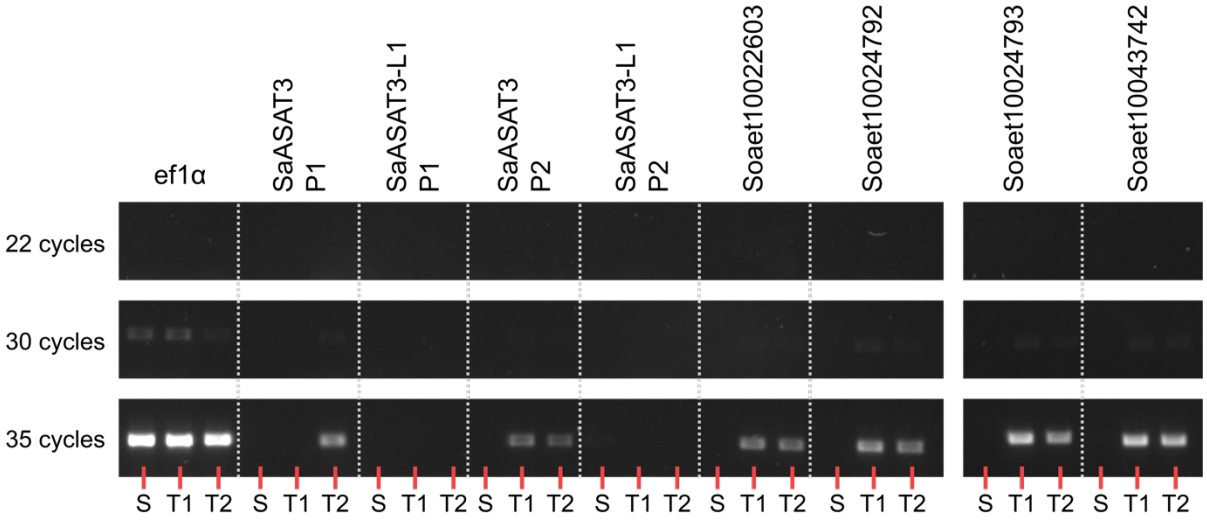


Figure S14. Semi-quantitative RT-PCR analysis of *S. aethiopicum* BAHD expression in glandular trichomes. *Elongation factor1α* (EF1α) was used as a positive control. P1 = primer pair 1; P2 = primer pair 2; S = shaved hypocotyl; T1 = PI 666075 glandular trichomes; T2 = Grif 14165 glandular trichomes.

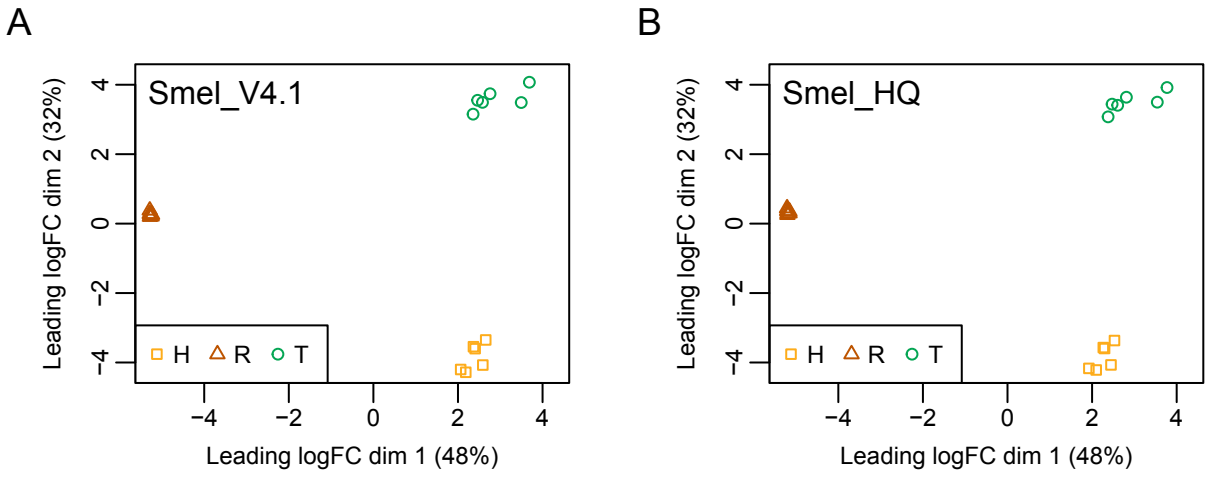


Figure S15. MDS plot of *S. melongena* RNAseq data. Genome-wide expression patterns across eggplant trichomes, trichomeless hypocotyls, and roots. Depicted is a multidimensional scaling (MDS) plot demonstrating that our 18 eggplant RNAseq samples cluster tightly by tissue identity. Distance between points illustrates expression differences between pairs of samples, calculated as leading (i.e., largest absolute) log2 fold-change (FC) in two dimensions, dim 1 and dim 2. Along the y-axis (dim 2), samples are separated into three tissue-specific groups. Along the x-axis (dim 1), samples are separated into two groups representing aerial (trichomes and trichomeless hypocotyls) and subterranean (roots) tissues, respectively. Colors and shapes represent the three tissues. T, trichomes (green circles); H, trichomeless hypocotyls (yellow squares); R, roots (brown triangles).

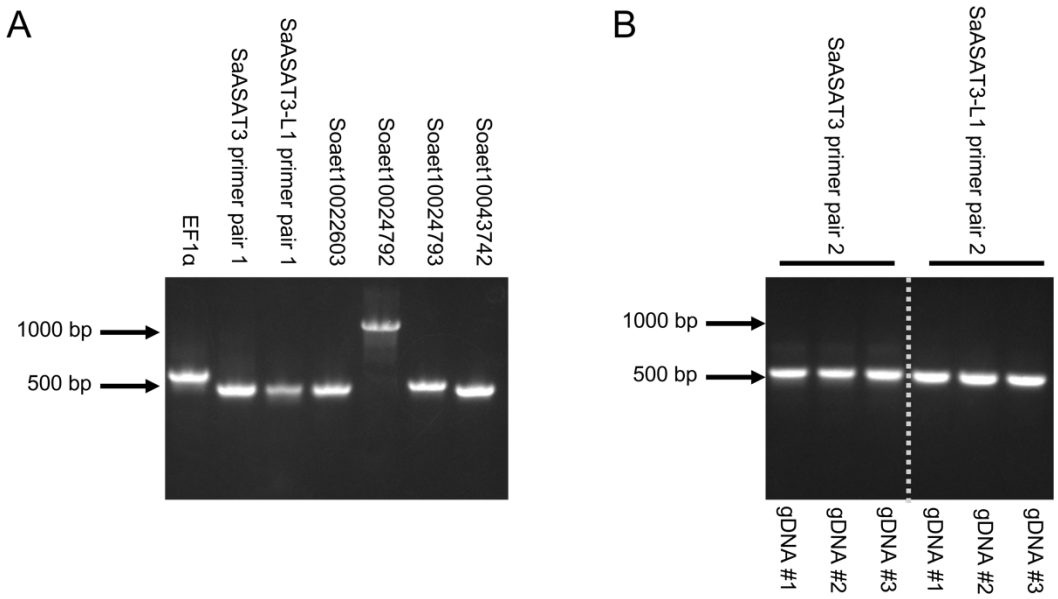


Figure S16. *S. aethiopicum* RT-PCR primers validated with gDNA controls. Soaet10024792 primers amplified a region containing an intron resulting in the larger amplicon length.