1 Running title: *Solanum* trichome metabolite diversity

2

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

4

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

6 [1]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing,

7 MI 48823 USA

8 [2]Department of Plant Biology, Michigan State University, East Lansing, MI 48823 USA

9

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

11

12 Corresponding Author: Robert L. Last, Michigan State University, Department of Biochemistry

13 and Molecular Biology, 603 Wilson Road, East Lansing, MI 48824; Phone: 517-432-3278;

14 email: lastr@msu.edu

15

16 ORCID:

17 Paul D. Fiesel 0000-0002-2038-5417

18 Rachel E. Kerwin 0000-0001-7395-5973

19 A. Daniel Jones 0000-0002-7408-6690

20 Robert L. Last 0000-0001-6974-9587

21

22 P.D.F. current address: Minnesota Department of Public Health

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. Published research revealed trichome acylsugars composed of glucose and sucrose cores in species across the family. In addition, acylsugars were analyzed across a small fraction of the >1200 species in the phenotypically megadiverse *Solanum* genus, with a handful containing inositol and glycosylated inositol cores. The current study sampled several dozen species across subclades of the *Solanum* to get a more detailed view of acylsugar chemodiversity. In depth characterization of acylsugars from the Clade II species *Solanum melongena* (brinjal eggplant) led to the identification of eight unusual structures with inositol or inositol glycoside cores, and hydroxyacyl chains. Liquid chromatography-mass spectrometry analysis of 31 additional species in the *Solanum* genus revealed striking acylsugar diversity with some traits restricted to specific clades and species. Acylinositols and inositol-based acyldisaccharides were detected throughout much of the genus. In contrast, acylglucoses and acylsucroses were more restricted in distribution. Analysis of tissue-specific transcriptomes and interspecific acylsugar acetylation differences led to the 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.

Introduction

Plants are remarkable synthetic chemists, producing a multitude of structurally complex specialized metabolites that differ from the products of general, or primary, metabolism in their lineage-specific distribution and tissue or cell type-specific biosynthesis. In contrast to the negative selection against changes in primary metabolism (for example, amino acids, energy metabolism intermediates and vitamin cofactors), less constrained evolution of specialized metabolism led to accumulation of hundreds of thousands of taxonomically restricted metabolites in broad classes (Pichersky and Gang, 2000; Zähner, 1979). Specialized metabolites play many critical roles, including abiotic and biotic stress protection (Agati and Tattini, 2010; De Moraes et al., 2001; Landry et al., 1995), pollinator attraction (Kretschmar and Baumann, 1999) and mediation of beneficial and pathogenic microbe interactions (Yu et al., 2021). These diverse and bioactive small molecules have historical and modern uses in human medicine, including the chemotherapeutics vinblastine and paclitaxel, antimalarial artemisinin, and painkillers such as morphine (Anand et al., 2019; Wink, 2015). Their distinct phylogenetic distributions and relatively rapid diversification make specialized metabolic enzymes excellent models to study evolution (Firn and Jones, 2009; Fraenkel, 1959; Pichersky and Lewinsohn, 2011).

Acylsugars are specialized metabolites produced across the Solanaceae (nightshade) family, aiding in defense against herbivores, fungi, and bacteria (Goffreda et al., 1989; Leckie et al., 2016; Luu et al., 2017; Weinhold and Baldwin, 2011). In types I- and IV-glandular trichomes, BAHD-type AcylSugar AcylTransferase (ASAT) enzymes assemble acylsugars from the basic building blocks of sugars, often sucrose, and short- to medium-length acyl chains derived from acyl-coenzyme A (acyl-CoA) esters (Fan et al., 201; Lou et al., 2021; Moghe et al.,

69   2017; Schilmiller et al., 2015; Schilmiller et al., 2012). Despite their simple components,

70   acylsugars exhibit remarkable chemical diversity arising from variations in sugar core

71   composition and acyl chain length, branching pattern, position, and number (Fan et al., 2019;

72   Fiesel et al., 2022; Ghosh et al., 2014; Hurney, 2018; Lou et al., 2021; Moghe et al., 2017;

73   Schenck et al., 2022). For example, acylsucroses, composed of a sucrose disaccharide core,

74   accumulate in cultivated tomato (*Solanum lycopersicum*) trichomes, while acylsucroses and

75   acylglucoses have been observed in the trichomes of the wild tomato species, *Solanum pennellii*.

76   Acylsugar structural variation also impacts biological activity; for example, differential

77   oviposition deterrence was demonstrated from naturally derived acylsugar mixtures (Leckie et

78   al., 2016).

79   Solanaceae acylsugars have become an exemplary model to study evolution of a diverse,

80   biologically relevant trait in a plant family with genomic and phylogenetic resources. Utilization

81   of this model revealed gene duplication, neofunctionalization, co-option, and loss involved in

82   acylsugar evolution (Fan et al., 2020, 2017; Leong et al., 2019; Moghe et al., 2017). For

83   example, neofunctionalization of an invertase-like enzyme, AcylSucrose FructoFuranosidase 1

84   (ASFF1) and functional divergence of core ASAT enzymes is responsible for differences in

85   sugar core type and acyl chain positions between cultivated tomato *S. lycopersicum* and wild

86   tomato *S. pennellii* acylsugars (Fan et al., 2017; Leong et al., 2019). Identifying mechanisms of

87   acylsugar evolution across the broader *Solanum* genus requires a detailed understanding of

88   acylsugar diversity and biosynthesis, which is lacking for many species.

89   Nearly half of the Solanaceae falls into the large (>1200 species) and phenotypically

90   diverse *Solanum* genus (Gagnon et al., 2022). This genus is split into several major clades,

91   including Potato, Regmandra, Dulcamaroid-Morelloid (DulMo), Valdiviense-Archaesolanum-

92   Normania non-spiny (VANAns), and Clade II (Gagnon et al., 2022). The well-studied Potato

93   clade contains cultivated tomato and potato and their wild relatives. The Regmandra, DulMo,

94   and VANAns clades contain mostly non-spiny species, including the black nightshade, *S.*

95   *nigrum.* The large Clade II contains nearly 900 chiefly spiny species, including the cultivated

96   brinjal eggplant, *S. melongena* (Bohs, 2005; Bohs and Olmstead, 1997; Gagnon et al., 2022;

97   Levin et al., 2006; PBI Solanum Project, 2022; Särkinen et al., 2013; Stern et al., 2011; Tepe et

98   al., 2016; Weese and Bohs, 2007). To date, documentation of *Solanum* acylsugar diversity

99   largely focused on a handful of species within the Potato clade. These efforts identified at least

100 38 nuclear magnetic resonance (NMR) spectroscopy-resolved acylsugar structures and liquid

101 chromatography mass spectrometry (LC-MS) supported annotations of many more (Fan et al.,

102 2017; Ghosh et al., 2014; Lybrand et al., 2020; Schilmiller et al., 2016). Though acylsugar

103 reports outside of the Potato clade are limited, they reveal novel structural variants not observed

104 among cultivated tomato relatives. For example, acylsugars with *myo*-inositol sugar cores (i.e.,

105 acylinositols) were characterized in three species: *S. nigrum*, from the DulMo clade*,* and *S.*

106 *lanceolatum* and *S. quitoense,* from Clade II (Herrera-Salgado et al., 2005; Hurney, 2018; Leong

107 et al., 2020; Lou et al., 2021). These discoveries highlight the benefits of a comprehensive

108 description of *Solanum* acylsugars within the well-developed phylogenetic framework.

109   Here we report analysis of *Solanum* acylsugar chemical diversity in species from the

110 relatively unexplored *Solanum* clades DulMo, VANAns, and Clade II, which together comprise

111 ~1000 *Solanum* species. We first established the Clade II brinjal eggplant, *S. melongena*, as a

112 reference species. Eggplant is a major worldwide fruit crop with extensive genomic,

113 transcriptomic, and germplasm resources (Barchi et al., 2021; Li et al., 2021; Mennella et al.,

114 2010). We characterized eggplant trichome acylinositols, acylinositol glycosides, and acylsugars

115   with unusual hydroxylated acyl chains using electrospray ionization LC-quadrupole time of

116   flight-MS (ESI LC-QToF-MS) and NMR. These atypical structures likely reflect participation of

117   enzymes that are distinct from the cultivated tomato acylsucrose pathway. Moving beyond this

118   model organism framework, LC-MS phylogenetic screening of 31 additional Clade II, DulMo,

119   and VANAns species led to the identification of remarkable acylsugar structural variation. LC-

120   MS features with characteristics of acylinositols were found in 25 of the 26 acylsugar-producing

121   species, suggesting one or a small number of evolutionary origins. In contrast, acylglucoses were

122   detected in DulMo and VANAns species, but not in any tested Clade II species.

123   As a first step towards revealing the molecular basis of this extensive acylsugar diversity,

124   we leveraged interspecific acylsugar variation and RNA sequencing of multiple eggplant tissues

125   to identify an acylinositol biosynthetic enzyme, *S. melongena* Acylsugar AcylTransferase 3-Like

126   1 (SmASAT3-L1; SMEL4.1_12g015780), responsible for acetylating a triacylinositol glycoside

127   acyl acceptor. SmASAT3-L1 exhibits different acyl-CoA substrate specificity than previously

128   characterized ASAT3 homologs, demonstrating how gene duplication and functional divergence

129   created acylsugar metabolic novelty in this part of the *Solanum* clade.

130

Results and Discussion

Eggplant glandular trichomes accumulate acylsugar-like metabolites

133   We began our investigation of *Solanum* acylsugar diversity with the brinjal eggplant *S.*

134   *melongena* due to its economic importance, availability of genomic resources, and phylogenetic

135   position within the monophyletic Eastern Hemisphere Spiny clade of Clade II (formerly known

136   as the 'Old World spiny clade') (Gagnon et al., 2022). Across eight eggplant accessions (listed in

137   Table S1), we observed glandular trichomes resembling the acylsugar-producing structures

138   found in other *Solanum* species (Leong et al., 2020; Lou et al., 2021; Schilmiller et al., 2012) on

139   hypocotyls, cotyledons, and the first three true leaves of young plants (Figure 1 and Figure S1).

140   In contrast, we observed only non-glandular stellate trichomes on leaves and stems of mature

141   eggplants (Figure S1), and these are types that typically do not accumulate acylsugars or other

142   specialized metabolites (Levin, 1973; Wagner, 1991).

143   We analyzed surface metabolite extracts from young and mature eggplant tissues using

144   LC-QToF-MS coupled with collision-induced dissociation (CID) in negative and positive ion

145   mode. We used these data to annotate acylsugars based on relative mass defect, molecular adduct

146   ion masses, retention times, and ions present in CID mass spectra. Briefly, acylsugars were

147   annotated from masses of fatty acid carboxylate fragment ions, as well as fragment ions

148   corresponding to stepwise losses of acyl chains from the pseudomolecular ion to a sugar core

149   fragment ion. The acylsugar annotation methods and confidence criteria are explained in detail in

150   the Methods. This analysis revealed abundant acylsugars in extracts from young, glandular

151   trichome-producing eggplant tissues, but not from mature, non-glandular trichome producing

152   tissues.

153   We annotated 38 acylsugars in leaf surface extracts of young eggplant from eight of nine

154   accessions (Table S1), consisting of 16 acylhexoses and 22 acyldisaccharides (Table 1). LC-MS

155   based acylsugar annotations are described using a shorthand nomenclature modified from

156   previous work (Leong et al., 2020) as follows: UX:Y:Z(A,B,C,D), where U – as a single or

157   multi-letter designation – represents the sugar core, X represents the number of acyl chains, Y is

158   the sum of acyl chain carbons, Z indicates the number of unsaturated bonds in the acyl chains

159   (when present), and A-D represent the number of carbon atoms in the individual acyl chains. For

160   example, the eggplant acylhexose I3:18(4,4,10) consists of a *myo*-inositol core with three acyl

161   chains with a total of 18 carbon atoms.

162   Curiously, eggplant acyldisaccharides contain an atypical pentose-hexose core, as

163   evidenced by a fragment ion mass of *m/z* 293.09 in negative-ion mode, corresponding to a fully

164   deacylated sugar core minus a hydroxyl group and a proton. Further evidence was provided by

165   positive mode CID, which promotes glycosidic bond cleavage, yielding a fragment ion

166   corresponding to the neutral loss of an unacylated pentose ring. All acylhexoses annotated with

167   medium to high confidence formed abundant fragment ions in positive mode but few fragment

168   ions in negative mode as illustrated in Figure S2; this pattern is characteristic of acylinositols

169   found in *S. quitoense* and *S. nigrum,* but has not been observed for acylglucoses, suggesting that

170   the detected *S. melongena* acylhexoses are acylinositols (Hurney, 2018; Leong et al., 2020; Lou

171   et al., 2021).

172   The eggplant acylsugars each contained three to four acylations, all on the hexose core,

173   including one medium eight-carbon (C8) to C14 acyl chain and two to three shorter C4 or C5

174   acyl chains. The medium acyl chains (C8, C10, C12, and C14) were either iso-branched or

175   straight as revealed by GC-MS acyl chain analysis (Figure S3). Additionally, we identified

176   hydroxylated C12, C14, and C16 acyl chains not previously reported in Solanaceae acylsugars

177   (Table 1). Although we did not observe large differences between the eight eggplant accessions,

178   eggplant acylsugars differ in chain length, functional groups, and acyl chain composition from

179   the reported *Solanum* acylinositols (Herrera-Salgado et al., 2005; Hurney, 2018; Lou et al.,

180   2021).

181   Although LC-MS provided information about sugar core mass, it did not reveal the sugar

182   core structure, prompting analysis using gas chromatography-mass spectrometry (GC-MS) of

183   derivatized *S. melongena* acylsugar cores. When free sugar cores, produced by metabolite extract

184   saponification, were derivatized to form alditol acetates, GC-MS peaks corresponding to

185   derivatized *myo*-inositol and glucose were detected, supporting the presence of acylinositols

186   (Figure S4A). Because the detection of glucose might have resulted from other compounds in the

187   leaf surface metabolite extracts, the disaccharide sugar core composition was determined by

188   hydrolyzing the saponified sugar cores with formic acid to cleave the glycosidic linkage.

189   Comparison of the hydrolyzed and unhydrolyzed samples revealed a peak corresponding to

190   arabinose only in the hydrolyzed plant samples (Figure S4B). Taken together, the results of

191   saponification with and without hydrolysis, followed by derivatization, confirmed identification

192   of *myo*-inositol sugar cores and identified the pentose moiety of the hexose-pentose disaccharide

193   core as arabinose.

194

195   NMR analysis of eight eggplant acylsugars

196   While MS analysis provided valuable information about acyl chain number and length,

197   sugar core mass, as well as acyl chain number, complementary information about acyl chain

198   branch structure, sugar core stereochemistry, and acyl chain positions was obtained using NMR.

199   We purified and resolved the structures of eight abundant eggplant acylsugars from accession PI

200   555598 using a combination of 1D and 2D NMR experiments (Figure 1). All eight structures are

201   newly described, and because atomic connections were determined by NMR and MS data, the

202   proposed structures meet Metabolomics Standards Initiative level 1 criteria for metabolite

203   identification (Sumner et al., 2007).

204   A series of NMR experiments confirmed that the acylhexose and acyldisaccharide sugar

205   cores are *myo*-inositol and 4-*O*-β-arabinopyranosyl *myo*-inositol, respectively (Figure 1C,D;

206 Supplemental Data File S1). We assigned all sugar ring proton signals of each sugar core's spin

207 system with total correlation spectroscopy (TOCSY). Inferences from correlation spectroscopy

208 (COSY) data then identified the order of ring protons and revealed pyranose and cyclitol ring

209 structures for the pentose and hexose rings, respectively. Relative stereochemistry at ring

210 positions was subsequently determined through comparison of spin-spin splitting (multiplicities

211 and coupling constants) referenced to expected patterns inferred from chemical principles and

212 previously reported acylinositols (Hurney, 2018; Leong et al., 2020). The disaccharide glycosidic

213 linkage position was determined by heteronuclear multiple bond correlation (HMBC)

214 correlations to be at position 4 and position 1 of *myo*-inositol and arabinose, respectively. The

215 arabinose β anomeric configuration was inferred from the anomeric carbon $^{1}J_{CH}$ (162 Hz) as

216 revealed by a coupled-heteronuclear single quantum coherence (coupled-HSQC) experiment of

217 the free disaccharide sugar core after saponification. Our sugar core assignments identifying this

218 unusual disaccharide agreed with the previous GC-MS sugar core results, supporting the efficacy

219 of the sugar core GC-MS identification method (Figure S4) This disaccharide core differs in the

220 pentose moiety identity from a recently reported acylated 4-*O*-*ß*-xylopyranosyl *myo*-inositol in

221 the *Solanum* Clade II species *S. quitoense* (Hurney, 2018). To our knowledge, this is the first

222 report of acylated 4-*O*-*ß*-arabinosyl *myo*-inositol sugars.

223   Acyl chain positions, branching patterns, and hydroxyl positions were also resolved

224 through integration of different NMR experiments. We found that all acyl chains were confined

225 to *myo*-inositol, consistent with the LC-MS results. All eight acylsugars are decorated with two

226 short iso-branched iC4 acyl chain esters at positions 1 and 2, and one medium C8 to C14 acyl

227 chain ester at position 3 (Figure 1). Compound 2 (Figure 1C), the only tetraacylated acylsugar

228 identified in eggplant, additionally carried an acetylation at position 6 (Figure 1C). The medium-

229   length acyl chains at position 3 were resolved as straight (nC14) or terminally iso-branched (iC8,

230   iC10, iC12, iC14) based on signals characteristic of protons near the branched carbons.

231   Strikingly, we identified unusual hydroxylated straight and iso-branched 3-

232   hydroxytetradecanoate acyl chains, 3-OH-nC14 and 3-OH-iC14, in compounds 7 and 8,

233   respectively (Figure 1D). We assigned hydroxylation positions of 3-OH-nC14 and 3-OH-iC14 to

234   the third acyl carbon based on a downfield shifted signal in the ^{1}H NMR spectrum at 3.92 ppm,

235   corresponding to one hydrogen at that position. In contrast, the non-hydroxylated medium acyl

236   chains observed in compounds 1-6 (Figure 1C,D), carry two hydrogen atoms at the third acyl

237   carbon, and these have a characteristic signal near 1.50 ppm. We are not aware of previous

238   reports of 3-OH-nC14 and 3-OH-iC14 chains in Solanaceae acylsugars. While 3-OH-nC14

239   chains are observed in the acylsugar-like bacterial Lipid A glycolipids, the hydroxylation

240   position differentiates these eggplant chains from hydroxyacyl chains in castor bean (*Ricinus*

241   *communis*) seed oil, *Silene gallica* gallicasides, *Ibicella lutea* fatty acid glycosides, and

242   Convolvulaceae resin glycosides (Asai et al., 2010; Asai and Fujimoto, 2010; Bah and Pereda-

243   Miranda, 1996; Smith, 1971).

244

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

246   To gain insight into the biosynthetic origin of the hydroxylated acyl chains, we tested

247   chirality of the 3-OH-nC14 acyl chain of compound 7 as this chain has available commercial

248   standards. Two-step chiral derivatization and GC-MS analysis was performed on compound 7

249   and two commercial standards, (3*R*)-OH-nC14 and the corresponding racemate (3*R/S*)-OH-

250   nC14. Fatty acids were converted to ethyl esters to aid volatilization, followed by esterification

251   of the acyl chain hydroxyl group with the Mosher acid (*R*)-(−)-α-methoxy-α-

252   (trifluoromethyl)phenylacetate (*R*-MPTA) to distinguish stereoisomers by GC separation of the

253   diastereomeric derivatives. GC-MS analysis of the derivatized (3*R/S*)-OH-nC14 standard

254   compounds yielded two peaks at retention times of 42.21 and 42.40 minutes, while one peak at

255   42.40 minutes was obtained for (3*R*)-OH-nC14 (Figure 2). We observed expected fragment ions

256   corresponding to the MPTA fragment ion, fatty acid ethyl ester, and fatty acid which supported

257   peak assignment (Figure S5). Thus (3*S*)-OH-nC14 and (3*R*)-OH-nC14 elute at 42.21 and 42.40

258   minutes, respectively. The compound 7 acyl chain was assigned as (3*R*)-OH-nC14 as its

259   derivatized 3-OH-tetradecanoate acyl chain eluted at 42.40 minutes (Figure 2).

260   The (3*R*)-OH-nC14 chain hydroxyl position and stereochemistry implicate possible

261   metabolic pathways leading to its biosynthesis. As fatty acid hydroxylases typically act on the

262   acyl chain terminal region, they are unlikely to catalyze C3 hydroxylation of a C14 fatty acid

263   (Pinot and Beisson, 2011). Of the other well-characterized pathways, only mitochondrial and

264   plastidial fatty acid metabolism mechanisms produce (3*R*)-OH-acyl-thioester intermediates.

265   Further, previous work identified two trichome-expressed, mitochondrial enzymes, Acylsugar

266   Enoyl-CoA Hydratase 1 (AECH1) and Acylsugar Acyl-CoA Synthetase 1 (AACS1), and one

267   trichome-expressed, plastidial enzyme, a beta-ketoacyl-(acyl-carrier-protein) reductase

268   (SpKAR1), involved in medium C10 and C12 acyl chain production in cultivated tomato and its

269   wild relative, *S. pennellii* (Fan et al., 2020; Ji et al., 2023). We hypothesize that acylsugar

270   hydroxyacyl chains evolved through changes in acyl-CoA substrate availability, possibly through

271   the actions of thioesterases or acyl-CoA synthetases. Future genetic and biochemical studies

272   should reveal the mechanisms behind (3*R*)-OH acyl chain production, potentially providing new

273   approaches to engineer unique acylsugars with hydroxyacyl chains and hydroxylated fatty acids

274   for polymeric building blocks.

275

276　　<u>Enormous *Solanum* acylsugar diversity revealed by LC-MS metabolite screening</u>

277　　　　　To date, most acylsugar screening across the *Solanum* has focused on cultivated tomato

278　　and its Potato clade relatives (Fan et al., 2017; Ghosh et al., 2014; Lybrand et al., 2020;

279　　Schilmiller et al., 2016), which represent a small fraction of the species diversity in this genus

280　　(PBI Solanum Project, 2022). Though limited in number, studies profiling non-Potato clade

281　　species, including *S. nigrum* (Lou et al., 2021), *S. lanceolatum* (Herrera-Salgado et al., 2005), *S.*

282　　*quitoense* (Hurney, 2018; Leong et al., 2020), and *S. melongena* (this work) demonstrate

283　　acylsugar diversity across the *Solanum* is vast, and largely uncharacterized. To address this

284　　knowledge gap, we analyzed tissue surface extracts from 31 additional *Solanum* species,

285　　including 25 Clade II, five DulMo clade, and one VANAns clade members (listed in Table S1).

286　　While NMR is the gold standard for structural elucidation, it is prohibitively time consuming for

287　　a largescale metabolite diversity survey. LC-MS coupled with CID provides substantial

288　　structural information and is compatible with high-throughput screening. To strike a balance

289　　between data quality and quantity, we performed LC-MS-CID screening on all 31 species and

290　　annotated acylsugars as described for *S. melongena*. Twenty-four of the 31 additional analyzed

291　　species produced detectable acylsugars from visible surface glandular trichomes, which were

292　　typically found only on seedlings, a pattern observed with eggplant. Species lacking detectable

293　　acylsugars had no or few observable glandular trichomes in the tissues analyzed, though we

294　　cannot rule out the possibility that acylsugar-producing glandular trichomes may be present on

295　　unobserved and untested tissue types and developmental stages. Targeted analysis of leaf surface

296　　extracts from 32 species, including *S. melongena*, and fruit surface extracts from two species,

297　　uncovered previously unreported acylsugars and identified the distributions of unusual acylsugar

298 chemical traits in the *Solanum* genus (Figure 3; Tables 1 and S2-27), revealing substantial

299 acylsugar variation between species (Figure 4; Tables 1 and S2-27).

300

301 <u>Identification of diverse acyl chain types in *Solanum* acylsugars</u>

302 Our survey of *Solanum* species uncovered acylsugars with a surprising diversity of acyl

303 chain lengths, functional groups, and combinations of chain types and positions. We identified

304 C2 to C18 acyl chains, including unsaturated and hydroxylated forms, based on analysis of fatty

305 acid fragment carboxylate ions in negative mode and neutral losses in both negative and positive

306 mode (Table S2). Acyl chain compositions differed between acylsugars within and across

307 species: acylinositols with primarily medium C8 to C14 acyl chains were observed in three

308 species (Figure 3, Tables S14, S15, S25), while acylsugars with two short C4 to C6 acyl chains

309 and one medium C8 to C14 acyl chain were observed in 17 species, an acylation pattern also

310 observed in eggplant and wild tomato (Figure 1; Table 1) (Lybrand et al., 2020). Acylsugars with

311 unsaturated acyl chains, including C5:1, C18:1, C18:2, and C18:3, were detected only in *S.*

312 *torvum* (Table S18). To our knowledge, this is the first report of unsaturated C18 chains in

313 Solanaceae acylsugars. In contrast, hydroxylated medium-length acyl chains were surprisingly

314 common: we annotated acylsugars bearing hydroxylated acyl chains in three DulMo clade

315 species and 18 Clade II species (Figure 3). These hydroxylated acyl chains are 'chemical

316 handles' susceptible to further modifications, including glycosylations and acylations, which

317 would result in even greater acylsugar diversity.

318 Indeed, we obtained evidence for glycosylated acyl chains in three Clade II species, *S.*

319 *prinophyllum*, *S. sisymbriifolium*, *S. lasiophyllum*, and one DulMo species, *S. dulcamara* (Tables

320 S12, S13, S24 and S27). Several lines of evidence led us to conclude that these species

14

321   accumulate acylsugars with glycosyl groups attached via ether linkages to medium hydroxyacyl

322   chains esterified to the hexose core. First, the compounds fragmented abnormally under negative

323   mode CID: the glycosyl group was observed as a neutral loss from the [M-H]$^-$ ion, producing a

324   [sugar – H]$^-$ fragment ion (Figure S6). This contrasts with extensive observations that CID

325   spectra of other characterized acyldisaccharides do not exhibit negative mode disaccharide

326   glycosidic linkage cleavage (Figure S2 and Figure S7) (Ghosh et al., 2014; Hurney, 2018; Lou et

327   al., 2021). Further evidence that the glycosylation is not on the primary acylated hexose ring was

328   obtained by subjecting acylsugar extracts to saponification, which induces ester linkage breakage

329   while retaining ether linkages. Upon analyzing saponified acylsugar extracts by LC-MS, we

330   detected pentose and hexose sugars with C14 or C16 acyl chains rather than glycosylated

331   hexoses, indicating that the glycosyl group is connected to a medium hydroxyacyl chain by an

332   ether linkage that is not cleaved during alkaline saponification (Figure S8). We named these

333   compounds glycohydroxyacylhexoses and annotated them by modifying the conventional

334   acylsugar nomenclature where the molecule in Figure S6 is named H3:24(4,4,16-O-p) with 16-

335   O-p representing the pentosylated (p) C16 hydroxyacyl chain (16-O). To our knowledge, this is

336   the first report of acylhexoses with glycosylated hydroxyacyl chains in Solanaceae acylsugars.

337   Resin glycosides produced by Convolvulaceae family species contain similar hydroxyacyl chain

338   glycosidic linkages. However, in resin glycosides, hydroxyacyl chains are connected to the

339   oligosaccharide core at two points – by an ester and ether linkage, respectively – forming a

340   macrocyclic structure not observed in the *Solanum* glycohydroxyacylhexoses (Bah and Pereda-

341   Miranda, 1996; Kruse et al., 2022; Pereda-Miranda et al., 1993).

342   The unusual glycohydroxyacylhexoses raises questions regarding their biosynthetic and

343   evolutionary origins. As these compounds are found in different lineages (Eastern Hemisphere

15

344   Spiny, Dulcamaroid, and Sisymbriifolium sections) – and closely related species lack detectable

345   glycohydroxyacylhexoses – this trait appears to be highly labile (Figure 3). Because we observe

346   the cognate, non-glycosylated acylinositol in each of these species, we hypothesize that

347   glycohydroxyacylhexoses consist of a *myo*-inositol hexose core. Further analysis of acylsugars

348   from additional species may identify other types of modified acylsugar hydroxyacyl chains in

349   *Solanum* acylsugars.

350   Sixteen species accumulate acylsugars with the same molecular masses, sugar core

351   masses, and acyl chain complements as the NMR-characterized eggplant acylinositols,

352   suggesting they may be identical. To test this hypothesis, we mixed purified eggplant NMR-

353   resolved standards with metabolite extracts from each species and performed liquid co-

354   chromatography. Coelution provided one line of evidence that the eggplant-like acylsugars

355   detected in 11 Clade II species are identical to characterized eggplant acylinositols (Tables S5-

356   14, S27). In contrast, the eggplant-like acylsugars from five other species, including three DulMo

357   clade members and two Clade II members, eluted separately from characterized eggplant

358   acylinositols, indicating they are acylation positional isomers or acyl chain branching isomers

359   (Figure S9). Examples of acylinositol positional isomers were described between the DulMo

360   clade species *S. nigrum* and the Clade II species *S. quitoense* and *S. melongena*, providing

361   precedent that the non-coeluting compounds might be positional isomers of the eggplant

362   acylsugars shown in Figure 1 (Leong et al., 2020; Lou et al., 2021). Our data suggest a

363   phylogenetic pattern of acylinositol distribution in which Clade II species generally accumulate

364   eggplant-like acylinositol isomers, whereas DulMo clade species accumulate non-eggplant-like

365   acylinositol isomers (Figure 3). While the eggplant-like isomers coelute with eggplant

16

366 acylsugars, further analysis with complementary analytical methods is needed to confirm

367 whether these are structurally identical across species.

368 <u>Acylhexoses and acyldisaccharides accumulate throughout Clade II, DulMo, and VANAns</u>

369 <u>*Solanum* species</u>

370   Acylhexoses were detected in all analyzed acylsugar-producing species. They were

371 annotated as acylglucoses or acylinositols based upon differences in characteristic fragmentation

372 behavior under negative ion mode MS conditions (Leong et al., 2020; Lou et al., 2021).

373 Identification of acylglucoses was based on observation of a sugar core fragment ion of m/z

374 143.03 ($C_6H_7O_4^-$) and stepwise acyl chain losses at the lowest collision energy level (0 V)

375 (Figure S10). In contrast, acylinositol annotation was based on an absence of acyl chain fragment

376 ions at 0 V and MS^e CID negative mode functions (Figure S2). Using these rules, we identified

377 acylinositols in all acylsugar-producing species tested except for *S. americanum* (Tables 1, S2-

378 27). In contrast, acylglucoses were not detected in any Clade II species but were commonly

379 detected among the DulMo and VANAns species, found in all analyzed acylsugar-producing

380 species except *S. dulcamara* (Tables S2, S19, S20, S23, S24, S26; Figure 3; Lou et al., 2021).

381   Differing from the widespread acylhexose distribution, acyldisaccharides were detected

382 sporadically in Clade II and DulMo clade species, being observed in 15 of 21 Clade II and 2 of 4

383 DulMo acylsugar-producing analyzed species (Figure 3 and Table S2). Based on negative [M-

384 H]$^-$ fragment masses and positive mode fragmentation of the glycosidic linkage, the

385 acyldisaccharides were composed of hexose-hexose, pentose-hexose, or deoxyhexose-hexose

386 sugar cores (Table S2, Figure S2 and Figure S7). Though complete structural information cannot

387 be obtained from MS fragmentation alone, two lines of evidence lead us to hypothesize that these

388 acyldisaccharides are composed of glycosylated inositol cores. First, most of the species with

389   detectable acyldisaccharides also accumulate cognate acylinositols with the same acyl chain

390   lengths (leaf and fruit surface extracts from 9 of 15 and two of two different species,

391   respectively), suggesting a shared inositol-containing core structure (Tables 1, S3-9, S12-17,

392   S20, S22). Second, acylinositol glycosides with varying sugar core sizes and stereochemistries

393   were described in *S. lanceolatum* (Herrera-Salgado et al., 2005), *S. quitoense* (Hurney, 2018),

394   and *S. melongena* (Figure 1), suggesting that the species in this study may accumulate similar

395   compounds. The ratio of acylhexose to acyldisaccharide peak number and abundance differs

396   across the surveyed species; for example, *S. mammosum* only accumulated detectable

397   acylinositols while *S. anguivi* primarily accumulated acyldisaccharides that comprised 80% of

398   the total acylsugar peak area (Tables S21 and S4). This is reminiscent of varied acylhexose

399   accumulation observed in trichomes of evolutionarily distant *Solanum* species: *S. nigrum* only

400   contains detectable acylinositol and acylglucose monosaccharides (Lou et al., 2021), while *S.*

401   *pennellii* accessions have mixtures of acylglucoses and acylsucroses ranging from 41-95% of

402   total acylsugars as acylglucoses. (Lybrand et al., 2020; Shapiro et al., 1994).

403

404   Sticky fruits accumulate acylinositols

405   We investigated the chemical basis of sticky fruit surface substances previously reported

406   in botanical species descriptions of the Clade II species *S. acerifolium* and *S. atropurpureum*

407   (Nee, 2022a, 2022b, 1991). LC-MS analysis of fruit surface extracts revealed similar acylsugar

408   profiles between the two species, with each extract containing what appeared to be the same 128

409   acylsugars. These fruit acylsugars were distinct from the 22 and 19 trichome acylsugars

410   identified in *S. acerifolium* and *S. atropurpureum*, respectively, as evidenced by their LC elution

411   and MS characteristics (Tables S14-17; Figure 5). Seventy-seven fruit acylsugars were identified

18

412   as acylinositols with two or three medium C8 to C14 acyl chains, including one to three

413   hydroxyacyl chains. The remaining 51 fruit acylsugars are acyldisaccharides containing an

414   unusual deoxyhexose-hexose core not previously reported in Solanaceae acylsugars. In contrast,

415   *S. acerifolium* and *S. atropurpureum* trichome acylsugars, composed of a pentose-hexose

416   disaccharide core decorated with short C5 or C6 acyl chains, were not detected in fruit surface

417   extracts.

418   *S. mammosum* and *S. capsicoides,* which are closely related to *S. acerifolium* and *S.*

419   *atropurpureum*, do not produce noticeably sticky fruit surfaces, suggesting this trait evolved in

420   the common ancestor of *S. acerifolium* and *S. atropurpureum,* perhaps as recently as 2 Mya

421   (Särkinen et al., 2013). Fruit surface acylsugars were previously described within the *Physalis*

422   genus and likely represent an independent evolutionary origin from the *S. acerifolium* and *S.*

423   *atropurpureum* fruit acylsugars (Bernal et al., 2018; Cao et al., 2015; Cicchetti et al., 2018;

424   Maldonado et al., 2006). Considering that, like eggplant, *S. acerifolium* and *S. atropurpureum*

425   fruits lack trichomes (Nee, 2022a, 2022b), some other organ or cell type presumably synthesizes

426   the fruit surface acylsugars. Structures similar to *Solanum fernadesii* petiolar resin glands

427   (Sampaio et al., 2021), *Hypericum androsaemum* microscopic fruit glands (Caprioli et al., 2016),

428   and fennel and chamomile fruit secretory ducts (vittae) (Zizovic et al., 2007) may be involved in

429   *S. acerifolium* and *S. atropurpureum* fruit acylsugar secretion. Future work will be needed to

430   understand the cellular, biosynthetic, and evolutionary relationships between trichome and fruit

431   acylsugars.

432

433   Sugar core evolution across the Solanaceae

434   Our screening revealed acylinositols to be broadly distributed across Clade II, DulMo,

435   and VANAns (Figure 3, Table S1). Outside of this study, acylinositols were only reported in a

436   small number of *Solanum* species, including two Clade II members, *S. quitoense* and *S.*

437   *lanceolatum*, and one DulMo clade member, *S. nigrum*. The acylinositol distribution suggests

438   that acylinositol biosynthesis arose one or more times within the *Solanum* genus. However, our

439   ability to elucidate the likely number of acylinositol origins is impeded by the lack of resolution

440   of the phylogenetic relationships between the major *Solanum* clades, including Clade II, DulMo,

441   VANAns, Potato, and Regmandra (Gagnon et al., 2022). Further biochemical and genetic

442   analyses of acylinositol biosynthesis may reveal how this trait evolved, and whether the pathway

443   is conserved between Clade II, DulMo, and VANAns species.

444   Acylglucoses were reported to be sporadically present across the Solanaceae, including in

445   *Salpiglossis sinuata*, *Petunia*, *Nicotiana* spp., *Datura*, the DulMo clade *S. nigrum*, and Potato

446   clade *S. pennellii* (Castillo et al., 1989; Chortyk et al., 1997; Fiesel et al., 2022; Fobes et al.,

447   1985; Hurney, 2018; King and Calhoun, 1988; Lou et al., 2021; Matsuzaki et al., 1989; Schenck

448   et al., 2022; Van Dam and Hare, 1998). Recent work in *S. pennellii* and *S. nigrum* revealed that

449   acylglucoses are synthesized from acylsucroses by a neofunctionalized invertase-like enzyme,

450   AcylSucrose FructoFuranosidase 1 (ASFF1) (Leong et al., 2019; Lou et al., 2021). These results

451   show that conversion of acylsucroses to acylglucoses arose at least twice in *Solanum* because *S.*

452   *nigrum* and *S. pennellii* employ non-orthologous ASFF1 enzymes. Our current study extends

453   documented cases of *Solanum* acylglucose distribution to the VANAns clade, as well as

454   documenting it in additional DulMo species. In contrast, we found no acylglucose-accumulating

455   species among the 26 *Solanum* Clade II species analyzed in this study (Figure 3 and Table S2).

456   Elucidation of acylglucose biosynthesis outside of the *Solanum* clade will be needed to resolve

457   the number of evolutionary origins and biochemical mechanisms leading to synthesis of these

458   metabolites across the Solanaceae.

459   Varied acyldisaccharides were observed across DulMo and Clade II, and we predict they

460   are biosynthesized from acylinositols. Within the Potato clade as well as outside the *Solanum*

461   genus, acylsucroses are the dominant acyldisaccharide in the reported species and tissues (Fiesel

462   et al., 2022). In contrast, no acylsucroses were detected in any species screened in this study

463   (Figure 3 and Table S2). However, acylsucrose biosynthesis might be found in other DulMo and

464   VANAns species or accessions because acylglucoses were detected in all but one DulMo and

465   VANAns species and that acylsucroses can act as intermediates in acylglucose biosynthesis

466   (Leong et al., 2019; Lou et al., 2021). The lack of both acylglucoses and acylsucroses in the

467   analyzed Clade II species (Figure 3 and Table S2) suggests a loss of acylsucrose biosynthesis in

468   Clade II. Rather, our LC-MS data are consistent with the hypothesis that the acyldisaccharides

469   observed in this study consist of glycosylated inositol cores. Recent *in vitro* biochemistry and *in*

470   *vivo* genetic evidence in *S. quitoense* suggests that acylinositols are the precursors to acylinositol

471   disaccharides in Clade II *Solanum* species trichomes (Leong et al., 2022, 2019) and tomato roots

472   (Kerwin et al., 2024). In this scenario, acylinositols would be converted to their cognate

473   glycosides by one or more glycosyltransferases, with different sugars added either by a

474   conserved promiscuous glycosyltransferase or multiple glycosyltransferases. *S. quitoense* and *S.*

475   *melongena*, with their acylinositol glycosides bearing different glycosyl moieties, are promising

476   systems for addressing the biochemical origins of these unusual *Solanum* trichome

477   acyldisaccharides.

478

479   *S. aethiopicum* is defective in expression of a trichome acylsugar acyltransferase enzyme

480   We began analyzing the biochemical mechanisms underlying the observed *Solanum*

481   acylsugar diversity in seven of 23 members from the Eggplant clade and related Anguivi grade,

482   two small subclades within Clade II that includes brinjal eggplant and its close relatives (Figure

483   S11) (Aubriot et al., 2018). In contrast to the other six analyzed eggplant species, *S. aethiopicum*

484   (scarlet or Ethiopian eggplant) is unique in that it had unacetylated AI3:16(4,4,8) (Compound 1)

485   but did not accumulate detectable levels of the acetylated form, AI4:18(2,4,4,8) – the highest

486   abundance *S. melongena* acylsugar (Compound 2; Figure 6A, Figure S11, Tables 1, S2, S5). We

487   hypothesized that the absence of AI4:18(2,4,4,8) in *S. aethiopicum* is either due to loss-of-

488   function or altered expression of an acetylating enzyme in this species. The availability of *S.*

489   *melongena* and *S. aethiopicum* genomic sequences (Barchi et al., 2021; Song et al., 2019)

490   provided the opportunity to seek a trichome expressed acetyltransferase responsible for

491   producing AI4:18 and the mechanism behind the lack of detectable AI4:18 in *S. aethiopicum.*

492   To identify enzyme candidates, we sequenced transcriptomes of isolated trichomes,

493   trichome-depleted shaved hypocotyls, and whole roots collected from 7-day-old *S. melongena*

494   seedlings and performed differential gene expression analysis (see Methods). Of the 23,251

495   eggplant genes expressed in at least one of the three tissues, 745 were significantly enriched

496   (log2 fold-change >2, *p*-value < 0.05) in trichomes compared to hypocotyls and roots, including

497   20 BAHDs (Tables S37 and S38). As shown in Table S39, we selected for further testing seven

498   BAHDs that were abundantly expressed in trichomes (TPM > 200) and homologous to

499   characterized *Solanum* ASATs (sequences in colored text and denoted with "*" in Figure 6B)

500   (D'Auria, 2006; Fan et al., 2016; Leong et al., 2022, 2020; Lou et al., 2021; Schenck et al.,

501   2022). We expressed each of the seven candidates in *Escherichia coli* and tested for acetylation

502   of Compound 1, AI3:16(4,4,8). An SlASAT3 outparalog (Koonin, 2005) SMEL4.1_12g015780,

503   ACYLSUGAR ACYLTRANSFERASE3-LIKE1 (SmASAT3-L1) was the only enzyme to

504   exhibit detectable forward activity acetylating Compound 1, AI3:16(4,4,8), to Compound 2,

505   AI4:18(2,4,4,8) (Figure 6C). As shown in Figure S13, LC-MS analysis of the SmASAT3-L1 *in*

506   *vitro* assay product revealed that it had identical molecular masses and elution times to plant-

507   derived AI4:18(2,i4,i4,i8) (Compound 1) providing evidence that SmASAT3-L1 acetylates

508   AI3:16(i4,i4,i8). Characterization of reverse activities of the seven BAHD acyltransferases

509   (Leong et al., 2020; Lou et al., 2021), in which an acyl chain can be removed from an acylsugar

510   by incubating it with the enzyme and free Coenzyme A, confirmed that SmASAT3-L1 was the

511   only tested enzyme that removed an acetyl chain from AI4:18(2,4,4,8) (Figure 6D and Figure

512   S13). Taken together, these results indicate that trichome-expressed *SmASAT3-L1* encodes an

513   acylinositol acetyltransferase responsible for AI4:18(2,4,4,8) biosynthesis.

514   We tested the hypothesis that *S. aethiopicum* trichome extracts lack detectable acetylated

515   AI4:18(2,4,4,8) due to a defect in *SmASAT3-L1* ortholog expression. Indeed, reverse

516   transcription (RT)-PCR comparing isolated trichome and trichomeless shaved hypocotyl cDNA

517   revealed that *SaASAT3-L1* (GenBank ID: OQ547782) transcript was undetectable in *S.*

518   *aethiopicum* trichomes (Figure S14). This result implicates the defect in ASAT3-L1 enzyme

519   gene expression as responsible for lack of AI4:18 in *S. aethiopicum*. The interspecific variation

520   in *ASAT3-L1* expression is reminiscent of the ASAT4 expression differences among accessions

521   of the wild tomato *Solanum habrochaites*. In that case, a subgroup of *S. habrochaites* accessions

522   possessed a functional ASAT4 copy, but its low levels of gene expression correlated with

523   reduced accumulation of acetylated acylsugars in these 'subgroup E' accessions (Kim et al.,

524   2012).

525

23

526 <u>Natural variation potentiates novel activities and uses</u>

527 Until recently, it appeared that acylsucroses dominate acylsugar profiles of Solanaceae

528 species with a small number of documented examples of acylglucose- and acylinositol-producing

529 *Solanum* species (Fiesel et al., 2022; Fobes et al., 1985; Herrera-Salgado et al., 2005; Hurney,

530 2018; Kerwin et al., 2024; King and Calhoun, 1988; Leong et al., 2020; Liu et al., 2017; Lou et

531 al., 2021; Moghe et al., 2017; Schenck et al., 2022). Our chemotaxonomic survey of nearly three

532 dozen *Solanum* species from the sparsely sampled Clade II, DulMo, and VANAns clades

533 revealed widespread occurrence of glucose-, inositol-, and non-sucrose disaccharide-based

534 acylsugars decorated with unusual acyl chains, including medium-length (C12-C16) hydroxyacyl

535 chains, glycosylated hydroxyacyl chains, and unsaturated chains. Adding to interspecific

536 variation we observe here, the recent identification of acylsugars in cultivated tomato root

537 exudates (Korenblum et al., 2020) and root tissues (Kerwin et al., 2024) supports the value of

538 metabolite screening in additional tissue types as well as species.

539 The characteristics of ASAT3-L1, the enzyme absent from the *S. aethiopicum* trichome

540 transcriptome reveals the dynamic nature of plant specialized metabolism. First, SmASAT3-L1

541 is unique among characterized ASATs in its acyl acceptor specificity, acetylating an acylinositol

542 glycoside rather than an acylated sucrose, glucose, or *myo*-inositol. Second, acetyl-CoA donor

543 activity was previously only described for a subset of Clade III BAHDs, including ASAT4 and

544 ASAT5 (Figure 7). This theme was also observed in trichomes of two other *Solanum* plants,

545 more closely related to the species screened in this study: the Clade II *S. quitoense* triacylinositol

546 acetyl transferase SqTAIAT (Leong et al., 2020) and DulMo *S. nigrum* acylglucose acetyl

547 transferase SnAGAT1 (Lou et al., 2021). Both enzymes reside outside of the ASAT3 clade, and

548 thus are phylogenetically distinct from SmASAT3-L1. Intriguingly, the characteristics of

24

549   trichome-expressed SmASAT3-L1 and *S. quitoense* SqASAT1H, (Leong et al., 2022, 2020), as

550   well as cultivated tomato root-expressed SlASAT1-L (Kerwin et al., 2024) also support the

551   hypothesis that acylinositols are synthesized through one or more pathway(s) distinct from that

552   of acylsucroses.

553   The remarkably varied acylsugar structures reported here likely have distinct

554   bioactivities. For example, hydrophobic acylsugars with longer acyl chains and few free

555   hydroxyl groups may disrupt membranes, reminiscent of less polar triterpene saponin variants

556   (Augustin et al., 2011; Baumann et al., 2000). Variation in acyl chain linkage chemistries also

557   likely influence acylsugar mode of action. For example, ester linked short chain acylsucroses are

558   digested by *Manduca sexta* (hawkmoth) larvae and the volatile organic acids released into the

559   environment, attracting ant predators (Weinhold and Baldwin, 2011). In contrast, longer chain

560   hydroxylated and ether-linked glycohydroxyacyl chains observed in the study are likely to have

561   quite different metabolic fates, perhaps persisting in the digestive systems of herbivores, and

562   alternative metabolic fates in microbes. Recent work showed that free medium-chain 3-hydroxy

563   fatty acids elicit a strong immune response in *Arabidopsis* and treating plants with 3-OH-C10

564   fatty acid increased resistance to *Pseudomonas syringae* (Kutschera et al., 2019). If hydroxylated

565   acyl chains are released from acylsugars upon tissue damage, this could trigger a similar plant

566   immune response, thereby conferring resistance to pathogen infection. Future work is needed to

567   determine the function of hydroxylated acyl chain-containing acylsugars. Hydroxyacyl chains

568   also present promising opportunities to expand upon natural acylsugar diversity through

569   synthetic biology and/or synthetic chemistry. The hydroxyl acts as a reactive chemical handle,

570   which was exploited in a subset of *Solanum* species trichomes to create

571   glycohydroxyacylhexoses, and synthetic biology approaches could be used to add unusual

572   sugars, acyl chains, and aromatic groups through promiscuous enzymes or synthetic chemistry.

573   This handle can be exploited with synthetic chemistry using these chains as new feedstocks for

574   creating polymers and may be useful to create synthetic acylsugars for use as food-grade

575   emulsifiers. These structural changes would likely impact biological activities and may be useful

576   for developing more pest resistant plants.

577

578   **Methods**

579

580   Plant material and growth conditions

581   *Solanum* spp. seeds were obtained from the sources described in Table S1. Seeds were

582   treated with 10% (v/v) bleach for 10 min while being rocked at 24 rpm with a GyroMini nutating

583   mixer (Labnet, Edison, NJ, USA), and then rinsed 5-6 times with distilled water. Unless

584   otherwise noted, seeds were germinated on Whatman filter paper (MilliporeSigma, Burlington,

585   MA, USA) at 28°C and in the dark. Germinated seedlings were transferred to peat pots (Jiffy,

586   Zwijndrecht, Netherlands), and grown at 25°C, 16/8-h day/night light cycle, and ~70 μmol $m^{-2}s^{-1}$

587   photosynthetic photon flux density with cool white fluorescent bulbs. Mature plants were grown

588   in controlled environment growth chambers or in a greenhouse. The growth chamber conditions

589   consisted of 25°/18°C day/night temperatures, 16/8-h day/night light cycle, and ~100 μmol $m^{-2}s^{-1}$

590   photosynthetic photon flux density under LED bulbs. The greenhouse conditions consisted of a

591   16/8-h day/night light cycle achieved with supplemental sodium iodide lighting. Plants were

592   fertilized weekly with 0.5X Hoagland's solution.

593

594   Surface metabolite extractions

595   Surface metabolites were extracted as described previously (Leong et al., 2019; Lou and

596   Leong, 2019). We collected tissue primarily from seedlings as many of these *Solanum* species do

597   not produce glandular trichomes on mature tissues. The only exceptions were *S. mammosum* and

598   *S. sisymbriifolium* which produce glandular trichomes throughout their reproductive stages and

599   from which we sampled mature leaf tissue. Briefly, 0.1-1 g of leaf, stem, hypocotyl, or cotyledon

600   tissue was collected into a 1.5 mL screw-cap tube (Dot Scientific, Inc., Burton, MI, USA), 1 mL

601   extraction solvent (3:3:2 acetonitrile:isopropanol:water, 0.1% formic acid, 10 µM propyl 4-

602   hydroxybenzoate (internal standard)) was added and the sample was rocked at 24 rpm by a

603   GyroMini nutating mixer for two min. After extraction, the supernatant was transferred into a

604   glass 2 mL autosampler vial (Restek, Bellefonte, PA, USA) and sealed with a 9 mm cap

605   containing a PTFE/silicone septum (J.G. Finneran, Vineland, NJ, USA). For fruit surface

606   metabolite extractions, we modified the protocol by placing 1-2 fruit into a 15 mL polypropylene

607   conical tube (Corning Inc., Corning, NY, USA) containing 5 mL extraction solvent then

608   proceeded to the nutation step above.

609

610   Bulk eggplant acylsugar extraction and purification for NMR analysis

611   Surface metabolites were bulk extracted from approximately 1000 *S. melongena* PI

612   555598 seedlings. Seeds were treated with 10% (v/v) bleach (Clorox, Oakland, CA, USA) for 10

613   min while being gently rocked at 24 rpm with a GyroMini nutating mixer and subsequently

614   rinsed 5-6 times with distilled water. Seeds were sown in moist SUREMIX soil (Michigan

615   Grower Products, Galesburg, MI, USA) in Perma-Next plant trays, 22 x 11 x 2.5 inches

616   (Growers Supply Company, Dexter, MI, USA), covered with a humidity dome, 22 x 11 x 3

617   inches (Growers Supply Company), then immediately transferred to the growth chamber.

618   Seedlings were harvested when 2-3 true leaves were observed, approximately one week after

619   their germination, by cutting them at the base and placing them into two 2L beakers each

620   containing 1L 100% acetonitrile. Surface metabolites were extracted with gentle agitation with a

621   metal spatula for five min at room temperature. Plant material and sediment was removed by

622   vacuum filtration through a Büchner funnel lined with Whatman filter paper (MilliporeSigma),

623   solvent was removed *in vacuo* by rotary evaporation and dried residue was dissolved in 20 mL of

624   acetonitrile with 0.1% formic acid. Solvent was removed again using a vacuum centrifuge and

625   the dried residue was dissolved in 1 mL of acetonitrile:water:formic acid (80:20:0.001). The

626   semi-preparative LC method is described in detail in Table S28 and is described in brief here. *S.*

627   *melongena* acylsugars were separated with a Waters 2795 HPLC (Waters Corporation, Milford,

628   MA, USA) equipped with an Acclaim 120 C18 HPLC column (4.6 x 150 mm, 5 μm; Thermo

629   Fisher Scientific, Waltham, MA, USA). Solvent A was water with 0.1% formic acid and Solvent

630   B was acetonitrile. The reverse-phase linear gradient was as follows: 5% B at 0 min, 60% B at 2

631   min, 80% B at 40 min, 100% B at 42 min, 5% B at 42.01 min, held at 5% B until 44 min. Flow

632   rate was 1.5 mL/min, injection volume was 100 μL, and the column temperature was 40□.

633   Fractions were collected automatically at 0.25 min intervals by a 2211 Superrac fraction

634   collector (LKB Bromma, Stockholm, Sweden) and assessed for acylsugar presence and purity by

635   LC-MS analysis, as described below. Column fractions were collected in the same tubes for each

636   method run which worked to pool the fractions between each run.

637

638   LC-MS acylsugar analysis

639   Acylsugars were analyzed by LC-MS with each method described below and in Tables

640   S29-33. For each LC method, the mobile phase consisted of aqueous 10 mM ammonium

bioRxiv preprint doi: https://doi.org/10.1101/2023.06.05.542877; this version posted March 19, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

641   formate, adjusted to pH 2.8 with formic acid (Solvent A) and 100% acetonitrile (Solvent B). The

642   flow rate was maintained at 0.3 mL/min.

643   Acylsugar extracts were analyzed with a 22 min LC gradient using a Waters Acquity

644   UPLC coupled to a Waters Xevo G2-XS QToF mass spectrometer (Waters Corporation, Milford,

645   MA) equipped with electrospray ionization (ESI) operating in positive (ESI$^+$) and negative (ESI$^-$)

646   modes. Ten microliter acylsugar extracts were separated on an Acquity UPLC BEH C18 column

647   (10 cm x 2.1 mm, 130 □, 1.7 µm; Waters), kept at 40□, using a binary solvent gradient. The 22

648   min linear gradient was as follows: 5% B at 0 min, 60% B at 2 min, 100% B at 16 min, held at

649   100% B until 20 min, 5% B at 20.01 min, held at 5% B until 22 min. Acylsugar extracts were

650   analyzed with both ESI$^-$, and ESI$^+$. For ESI$^-$, the following parameters were used: capillary

651   voltage, 2 kV; sampling cone voltage, 35 V; source temperature, 100°C; desolvation temperature

652   350°C; cone gas flow, 50 L/Hr; desolvation gas flow, 600 L/Hr. For ESI$^+$, the following

653   parameters were used: capillary voltage, 3 kV; sampling cone voltage, 35 V; source temperature,

654   100°C; desolvation temperature 300°C; cone gas flow, 50 L/Hr; desolvation gas flow, 600 L/Hr.

655   Acylsugars were fragmented in either MS^E or data-dependent acquisition (DDA) MS/MS modes

656   as described previously (Lou et al., 2021; Lybrand et al., 2020). DDA survey and MS/MS

657   functions acquired over *m/z* 50 to 1500 with scan times of 0.2 s. Ions selected for MS/MS were

658   fragmented with a ramped collision energy where voltage varies based on ion mass. Collision

659   energies followed a ramp with the voltage changing linearly for ions between the low mass

660   setting (*m/z* 50, 15 to 30 V) and high mass setting (*m/z* 1500, 30 to 60 V). To increase mass

661   accuracy, lock mass correction was performed during data collection, with leucine enkephalin as

662   the reference.

29

663   Semi-preparative LC fractions were analyzed with direct infusion and a 14-min LC

664   gradient using a LC-20ADvp ternary pump (Shimadzu, Kyoto, Japan) coupled to a Waters Xevo

665   G2-XS QToF mass spectrometer (Waters Corporation). The direct infusion method quickly

666   screened fractions for acylsugar presence while the 14-min LC method tested fraction purity. Ten

667   microliters of acylsugar fractions were injected into an Ascentis Express C18 HPLC column (10

668   cm x 2.1 mm, 2.7 µm; Supelco, Bellefonte, PA, USA), kept at 40□. The 14-min gradient was as

669   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

670   min, and held at 5% until 14 min. Fractions were analyzed under ESI$^+$ with the following

671   parameters: capillary voltage, 3 kV; sampling cone voltage, 40 V; source temperature, 100°C;

672   desolvation temperature 350°C; cone gas flow, 20 L/hr; desolvation gas flow, 500 L/hr. For

673   direct infusion analysis, ions were acquired from m/z 50 to 1500 with a scan time of 0.1 s and

674   one acquisition function with 0 V collision potential. For 14-min LC analysis, ions were acquired

675   from m/z 50 to 1200 with a scan time of 0.1 s and three acquisition functions with different

676   collision potentials (0, 25, 60 V). Lock mass calibration referenced to the leucine enkephalin

677   [M+H]$^+$ ion was applied during data acquisition.

678   Enzyme assays products were analyzed with a 7-min gradient using Waters Acquity

679   UPLC coupled to a Waters Xevo G2-XS QToF mass spectrometer (Waters Corporation, Milford,

680   MA) equipped with electrospray ionization (ESI) operating in negative (ESI$^-$) mode. Reaction

681   products were separated with an Ascentis Express C18 HPLC column (10 cm x 2.1 mm, 2.7 µm;

682   Supelco), kept at 40□, using a binary solvent gradient. The 7 min linear gradient was as follows:

683   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,

684   held at 5% B until 7 min. The ESI$^-$ parameters described for acylsugar extract analysis were

685   used. Ions were acquired from m/z 50 to 1200 with a scan time of 0.1 s and three acquisition

30

686 functions with different collision potentials (0, 25, 60 V). Lock mass calibration referenced to the

687 leucine enkephalin [M-H]⁻ ion was applied during data acquisition.

688 For coelution analysis between enzymatically- and plant-produced AI3:16 and AI4:18, a

689 24-min linear gradient was used with an Acquity UPLC BEH C18 column (10 cm x 2.1 mm, 130

690 □, 1.7 µm; Waters), kept at 40□C, on the same instrument used for enzyme assay analysis. The

691 binary solvent, linear gradient was as follows: 5% B at 0 min, 60% B at 18 min, 100% B at 18.01

692 min, held at 100% B until 22 min, 5% B at 22.01 min, held at 5% B until 24 min. ESI⁻

693 parameters and MS method from the above enzyme assay analysis were used.

694 Saponified acylsugars were analyzed using a Waters Acquity UPLC coupled to a Waters

695 Xevo G2-XS QToF mass spectrometer (Waters Corporation, Milford, MA). Ten microliters of

696 saponified acylsugars were injected into either an Acquity UPLC BEH C18 column (10 cm x 2.1

697 mm, 130 □, 1.7 µm; Waters) or an Acquity UPLC BEH Amide column (10 cm x 2.1 mm, 130

698 □, 1.7 µm; Waters), kept at 40□. Samples were analyzed on the C18 column with the 22-min

699 method described above. To detect free sugars resulting from saponification, samples were

700 analyzed on the BEH amide column with a 9 min method which was as follows: 95% B at 0 min

701 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.

702 Both methods used the ESI- parameters and MS acquisition parameters described for enzyme

703 assays.

704

705 Acylsugar annotation

706 We inferred acylsugar structures utilizing negative and positive mode MS and MS/MS

707 collision-induced dissociation (CID) as previously described (Ghosh et al., 2014; Hurney, 2018;

708 Leong et al., 2019; Lou et al., 2021; Lybrand et al., 2020). We developed modified stepwise

31

709   confidence assessments based on those stepwise criteria described by Lou and coworkers in their

710   *Supplementary Text* 'Acylsugar annotation' section (Lou et al., 2021). Our criteria A-F were

711   adapted directly from Lou et al: (A) Exact mass; (B) Presence of co-eluting sugar core

712   fragments; (C) Presence of co-eluting fatty acid carboxylate fragments in high CID potential

713   (30V); (D) Presence of co-eluting fragments that correspond to neutral losses of acyl chains in

714   positive ion mode for acylinositols and negative ion mode for acyldissacharides and

715   acylglucoses; (E.) Confirmed by MS/MS; (F) Consistency across samples. Putative acylsugars

716   only meeting criterion A were annotated as low confidence, while those meeting all four criteria

717   described in A-D were annotated as medium confidence. We added two additional criteria:

718   (G) **NMR structural determination.** Acylsugars from *S. melongena* were further characterized

719   by NMR experiments in which sugar core relative stereochemistry, acyl chain positions, and acyl

720   chain branching were resolved.

721   (H) **Coelution with NMR characterized compounds.** Acylsugars with the same acyl chains

722   and sugar core size were analyzed by coelution with *S. melongena* acylsugars. If acylsugars

723   coeluted with an NMR characterized peak and shared the same acyl chains and fragmentation

724   patterns, they met this criterion and were annotated with high confidence. This criterion was

725   employed for acylsugars from species other than *S. melongena.*

726   Acylsugars meeting all criteria in A-F and either G or H were annotated with highest confidence.

727   Confidence levels for each reported acylsugar are specified in their species-specific annotation

728   table (Tables S2-27).

729

730   Acylsugar metabolomic analysis

731   To collect accurate formate adduct masses and relative abundances from putative

732 acylsugars, negative mode LC-MS raw data from metabolite extracts were analyzed with

733 Progenesis QI software v3.0 (Waters, Milford, MA, USA) using retention time alignment, peak

734 detection, adduct grouping, and deconvolution. Lock mass correction was performed during data

735 collection. We used the following analysis parameters: peak picking sensitivity, default;

736 retention time range, 0-22 min; adducts detected, M+FA-H, M+Cl, M-H, 2M-H, and 2M+FA-H.

737 Average percent peak abundance was calculated for each putative acylsugar by dividing its raw

738 abundance by the sum of the total acylsugar raw abundance and then averaging that value across

739 all samples for a species.

740

741 <u>NMR analysis of acylsugars</u>

742   Purified acylsugars were dissolved in acetonitrile-d_3 (99.96 atom % D; MilliporeSigma,

743 Burlington, MA, USA) and transferred to Kontes NMR tubes (MilliporeSigma, Burlington, MA,

744 USA). Samples were analyzed at the Michigan State University Max T. Rogers NMR Core with

745 a Varian Inova 600 MHz spectrometer (Agilent , Santa Clara, CA, USA) equipped with a

746 Nalorac 5 mm PFG switchable probe, a DirectDrive2 500 MHz spectrometer (Agilent, Santa

747 Clara, CA, USA) equipped with a OneNMR probe, Bruker Avance NEO 800 MHz spectrometer

748 (Bruker, Billerica, MA, USA) equipped with a 5 mm helium cryogenic HCN probe, or a Bruker

749 Avance NEO 600 MHz spectrometer equipped with a 5 mm nitrogen cryogenic HCN Prodigy

750 probe. All acylsugars were analyzed with a series of 1D (^{1}H, and ^{13}C) and 2D (gCOSY, TOCSY,

751 gHSQC, gHMBC, gH2BC, J-resolved) experiments. Resulting spectra were referenced to

752 acetonitrile-d_3 ($^\delta$H = 1.94 and $^\delta$C = 118.70 ppm). NMR spectra were processed and analyzed

33

753   with MestReNova 12.0 software (MestreLab, A Coruña, Spain). For full NMR metadata, see

754   Supplementary Datafile 1.

755

756   <u>Sugar core composition analysis</u>

757   *S. melongena* PI 555598 acylsugar extracts were dried down *in vacuo* with centrifugation

758   (Savant, ThermoFisher Scientific). Dried acylsugar extracts were dissolved in 1 mL of methanol.

759   One mL of 3 M ammonium hydroxide was added, and the solution was mixed vigorously. This

760   initial saponification reaction proceeded at room temperature for 48 hours. Solvent was removed

761   by vacuum centrifugation. Saponified sugar cores were reduced and derivatized with acetate

762   groups as previously described (Sassaki et al., 2008). Sugars were dissolved in 50 µL of 1 M

763   ammonium hydroxide for 15 min at room temperature. Addition of 50 µL 20 mg/mL $NaBH_4$ and

764   incubation at 100°C for 10 min converted aldoses and ketoses to polyols. Excess sodium

765   borohydride was quenched with 100 µL 1M trifluoroacetic acid. Two volumes of methanol were

766   added, and the sample was dried down by a stream of nitrogen gas. Dried down reaction products

767   were redissolved in 200 µL of methanolic 0.5 M HCl and heated at 100°C for 15 min in a closed

768   ½ dram vial, 12 x 35 mm (Kimble Chase, Vineland, NJ, USA). Solvent was removed by a stream

769   of nitrogen gas. To acetylate the sugar cores, residue was dissolved in 200 µL of pyridine:acetic

770   anhydride (1:1) and incubated at 100°C for 30 min. Acetylated polyols were then dried down *in*

771   *vacuo* in a vacuum centrifuge and redissolved in 100 µL hexane for GC-MS analysis.

772   To determine the composition of disaccharide sugar cores, saponified sugars were

773   hydrolyzed prior to derivatization. Dried down sugars were redissolved in 98 µL water in a ½

774   dram vial, 12 x 35 mm (Kimble Chase). After addition of 102 µL of 88% formic acid, the vial

775   was sealed and heated at 100□ for 15 hours in a heat block. Hydrolyzed sugars were dried down

776   *in vacuo* with a vacuum centrifuge and derivatized to alditol acetates following the method

777   described above.

778

779   Hydroxyl acyl chain stereochemistry analysis

780   We derivatized hydroxyacyl chains and extracted the derivatives as previously described

781   (Jenske and Vetter, 2007). The commercial standards (3R)-OH-tetradecanoate (Cayman

782   Chemical, Ann Arbor, MI, USA) and (3R/S)-OH-tetradecanoate (Cayman Chemical, Ann Arbor,

783   MI, USA) were first derivatized to their respective ethyl ester derivatives. 600 µg of each fatty

784   acid was dissolved in 0.5 mL of 0.5 M ethanolic KOH and incubated at 80°C for 5 min. After the

785   solution was cooled on ice, 1 mL of ethanolic BF$_3$ was added, and reactions were incubated at

786   80°C for five min. Two-phase partitioning with hexane and saturated sodium chloride yielded

787   ethyl ester derivatives in the hexane partition.

788   The 3-OH-nC14 acyl chain from I3:22(iC4,iC4,3-OH-nC14) was derivatized to an ethyl

789   ester following a previously published method (Fan et al., 2020; Ning et al., 2015). Half of the

790   purified I3:22(iC4,iC4,3-OH-nC14) was dissolved in acetonitrile with 0.1% formic acid and

791   transferred to a 1.5 mL microcentrifuge tube. The purified acylsugar was dried down *in vacuo*

792   with vacuum centrifugation. The resulting dry, purified I3:22(iC4,iC4,3-OH-nC14) was

793   dissolved in 300 µL of 21% (w/w) sodium ethoxide in ethanol (MilliporeSigma, Burlington,

794   MA, USA). The reaction proceeded at room temperature for 30 min with constant rocking at 24

795   rpm by a GyroMini nutating mixer (Labnet, Edison, NJ, USA) and occasional vortexing. 400 µL

796   hexane with 55 ng/µL tetradecane (MilliporeSigma, Burlington, MA, USA) as an internal

797   standard was then added followed by vigorous vortexing. 500 µL of aqueous saturated sodium

798   chloride was added to the hexane-ethanol mixture, and vortexed. The hexane layer was pipetted

35

799   to another tube and two more two-phase partitions with saturated sodium chloride were

800   completed. The final hexane layer containing the acyl chain ethyl esters was extracted and dried

801   down by a stream of nitrogen gas.

802   The resulting fatty acid ethyl esters from the commercial standards and purified

803   compound were derivatized with (*R*)-(-)-□-methoxy-□-trifluoromethylphenylacetyl chloride

804   ((*R*)-(-)-MTPA-Cl; MilliporeSigma, Burlington, MA, USA) following a previously published

805   procedure (Jenske and Vetter, 2007). Dried fatty acid ethyl esters were redissolved in 400 µL

806   pyridine and 15 µL of (*R*)-(-)-MPTA-Cl. The reaction proceeded at room temperature for two

807   hours, after which, 5 mL of water and tert-butyl methyl ether (TBME) were added along with

808   solid K_2CO_3 (one spatula tip). The TBME phase was collected after three successive phase

809   separations with 5 mL of water. TBME was evaporated to 1 mL and subjected to GC-MS

810   analysis.

811

812   <u>GC-MS analysis</u>

813   All GC-MS analyses employed an Agilent 5890 GC and an Agilent 5975 single

814   quadrupole MS equipped with a FactorFour VF-5ms column (30 m x 0.25 mm, 0.25 um;

815   Agilent) and 10 m EZ-Guard column (Agilent, Santa Clara, CA, USA). Helium was used as the

816   carrier gas with a constant flow rate of 1 mL/min. Electron energy was set at 70 eV. MS source

817   and quadrupole were maintained at 230°C and 150°C, respectively. Parameters specific to each

818   analysis are listed below.

819   For analysis of sugar core alditol acetate derivatives, we followed previously published

820   GC-MS parameters (Sassaki et al., 2008). Inlet temperature was maintained at 275°C. GC oven

821   temperature was held at 60°C for one min and then was ramped at a rate of 40°C/min to 180°C.

822   Oven temperature was then ramped to 240°C at a rate of 5°C/min and held for three min. Total

823   run time was 19 min. The MS detector transfer line was maintained at 280°C. Split ratio was

824   10:1 with a split flow rate of 10 mL/min. A three min solvent delay was used. MS data was

825   collected in full scan mode, *m/z* 50-600. Analysis parameters are also described in Table S34.

826   For analysis of MPTA fatty acid ethyl ester derivatives, we followed previously

827   published GC-MS parameters (Jenske and Vetter, 2007). GC oven temperature was held at 60°C

828   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

829   after which the oven temperature was ramped to 230°C at a rate of 2°C/min. After 230°C was

830   held for nine min, the oven temperature was ramped to 300°C at 10°C/min and held for 7.5 min.

831   Total run time was 55 min. A 5-min solvent delay was applied. The MS detector transfer line

832   was maintained at 280°C. Selective ion monitoring detected the ions *m/z* 189, 209, and 255 with

833   a dwell time of 50 ms. Analysis parameters are also described in Table S35.

834

835   Sample collection and transcriptome sequencing

836   We sequenced 18 transcriptomes: six biological replicates each of trichomes isolated

837   from hypocotyls (trichomes), hypocotyls stripped of their trichomes (stripped hypocotyls), and

838   whole roots (Supplemental data file 1). We sampled trichomes and stripped hypocotyls from 7-

839   day-old *S. melongena* 67/3 seedlings following methods developed for root hair cell isolation

840   (Bucher et al., 1997) with modifications. We grew lawns of eggplant seedlings in soil flats as

841   described above for bulk leaf surface metabolite extraction. At 7 days post germination, we

842   removed roots and cotyledons from seedlings, then transferred hypocotyls to liquid nitrogen in a

843   plastic 2 L Dewar flask. Frozen hypocotyls were gently stirred with a glass rod for 20 minutes to

844   physically shear trichomes from hypocotyls. After confirming under a dissecting scope that a

845   sample of three hypocotyls had been stripped bare, we filtered trichomes into a 2 L glass beaker

846   by slowly pouring the contents of the Dewar flask through a 500 μm wire mesh sieve

847   (MilliporeSigma, Burlington, MA, USA). To maximize trichome recovery, stripped hypocotyls

848   were returned to the Dewar, rinsed with liquid nitrogen, and filtered through the 500 μm sieve

849   six more times. Stripped hypocotyls were divided into six pre-weighed 50 mL conical tubes,

850   quickly weighed, then transferred to -80□C for storage. Trichomes were subsequently filtered

851   through a 150 μm sieve (MilliporeSigma, Burlington, MA, USA) into a 500 mL beaker to

852   increase sample purity then transferred to a 50 mL conical tube to allow the excess liquid

853   nitrogen to evaporate. Finally, filtered trichomes were divided into six pre-weighed 2 mL screw-

854   cap tubes (Dot Scientific, Inc., Burton, MI, USA), quickly weighed, then transferred to -80□C

855   for storage.

856   We extracted total RNA using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany)

857   following the manufacturer's directions, measured RNA concentration using a Qubit 1.0

858   instrument (Thermo Fisher Scientific, Waltham, MA, USA) with the RNA HS assay, and the

859   samples were processed by the Michigan State University Research Technology Support Facility

860   Genomics Core (East Lansing, MI, USA) for library preparation and high-throughput

861   sequencing. The core checked RNA quality using a Bioanalyzer 2100 (Agilent Technologies,

862   Santa Clara, CA, USA), constructed sequencing libraries using an Illumina Stranded mRNA

863   Prep kit (Illumina, Inc. San Diego, CA, USA), and sequenced all 18 libraries on a NovaSeq 6000

864   (Illumina, Inc. San Diego, CA) using a single S4 flow cell lane, producing 150-base pair paired-

865   end reads (Supplemental datafile 1). The raw fastq files are available to download from the

866   National Center for Biotechnology Information Sequence Read Archive database under

867   BioProject PRJNA935765.

868

869 <u>Transcriptome alignment and differential gene expression analysis</u>

870   To prepare the 150-bp paired-end sequences for alignment, we employed Trimmomatic

871 (Bolger et al., 2014) to trim adapters and low-quality bases, then filter reads shorter than 75 bp,

872 removing an average of 3.7% of sequences (range: 3.3% - 4.2%). We mapped the resulting 75–

873 150-bp (average length: 149 bp) paired-end RNAseq reads separately to the eggplant V4.1 and

874 HQ reference genomes using STAR in two-pass mode, which enhances splice junction discovery

875 and mapping sensitivity (Dobin et al., 2013; Dobin and Gingeras, 2015). Using this approach, an

876 average of 79.2% (range: 60.1 - 87.1%) and 82.2% (range: 62.0 - 90.7%) of reads mapped to a

877 unique genomic location in V4.1 and HQ, respectively (Table S36). We filtered the resulting

878 transcriptome alignments according to best practices as defined by the Genome Analysis Toolkit

879 (GATK) (DePristo et al., 2011; Van der Auwera et al., 2013). Briefly, we removed optical and

880 PCR duplicates with MarkDuplicates from the Picard toolkit

881 (http://broadinstitute.github.io/picard), parsed reads into exon segments and removed intron-

882 spanning bases using SplitNCigarReads from GATK (McKenna et al., 2010). Finally, we

883 selected unique alignments by eliminating reads with a mapping quality score below Q60 with

884 the view command from SAMtools (Li et al., 2009). To generate raw read counts, we used the

885 HTSeq command, htseq-count with the nonunique parameter set to all (Anders et al., 2015).

886   Raw read counts generated by HTSeq-count were used to perform differential gene

887 expression analysis in edgeR (Robinson et al., 2010). To restrict comparisons to expressed genes,

888 only transcripts with at least one read count-per-million (CPM) in at least one sample were

889 retained for further analysis. This filtering step removed 11,665 (33.4%) and 14,544 (39.8%)

890 annotated genes in the V4.1 and HQ eggplant genomes, respectively, yielding 23,251 and 22,024

891   expressed transcripts for differential gene expression analysis. Next, we normalized transcript

892   abundances across the 18 transcriptomes in our eggplant V4.1 and HQ alignments using the

893   default trimmed mean of M-values (TMM) method in the calcNormFactors function, then

894   performed multidimensional scaling (MDS) with the plotMDS function to compare global gene

895   expression profiles (Figure S15). This showed that our samples cluster tightly by tissue (i.e.,

896   trichome, shaved hypocotyl, or root), with no obvious differences between V4.1 and HQ

897   alignments. To test differences in gene expression across tissues, we implemented a generalized

898   linear model (GLM) using a quasi-likelihood (QL) approach: we generated an experimental

899   design matrix specifying the three tissues (i.e., trichomes, trichomeless hypocotyls, and roots)

900   with the model.matrix function, then used the glmQLFit function to fit our data to a QL-GLM

901   model. To identify genes with a log2 fold-change (FC) > 2 between tissues, we used the

902   glmTreat function, which performs threshold hypothesis testing, a rigorous statistical approach

903   that evaluates variance and magnitude to detect expression differences greater than the specified

904   value (e.g., log2 FC > 2), then applies false discovery rate (FDR) p-value corrections. Genes

905   were classified as significantly differentially expressed between two tissues if log2 FC > 2 and

906   FDR-corrected p-value < 0.05. We calculated absolute transcript abundance for all expressed

907   genes as transcripts per million (TPM) with the calculateTPM function in scateR (McCarthy et

908   al., 2017). Transcriptome expression values are in Tables S37 and S38.

909

910   Phylogenetic analyses

911   Characterized ASAT enzymes fall into clade III of the BAHD family (Fan et al., 2017,

912   2015; Leong et al., 2022, 2020; Lou et al., 2021; Moghe et al., 2017; Nadakuduti et al., 2017;

913   Schenck et al., 2022; Schilmiller et al., 2015, 2012). To identify ASAT candidates in eggplant,

914   we searched for sequences containing the Pfam Transferase domain (PF02458) (Finn et al.,

915   2010), associated with all characterized catalytically active BAHD proteins. The PF02458 HMM

916   profile was obtained from the Pfam website (http://pfam.sanger.ac.uk/) and queried against the

917   V4.1 and HQ eggplant proteomes using the hmmsearch tool from HMMER v3.2.1 (hmmer.org),

918   revealing 106 and 108 putative BAHD sequences, respectively. Using MAFFT v7.471 in E-INS-

919   i mode, we built multiple sequence alignments (MSA) of amino acid sequences from four

920   sources: 1) V4.1 or HQ eggplant PF02458 hits, 2) published reference sequences for clades I-VII

921   (Moghe et al., 2023), 3) characterized ASAT sequences from other Solanaceae species, and 4)

922   the SaASAT3 and SaASAT3-L1 candidates from *S. aethiopicum* (Katoh and Standley, 2013).

923   The E-INS-i algorithm implements local alignment with a generalized affine gap cost (Altschul,

924   1998), which aligns conserved regions (e.g., the BAHD transferase domain) and essentially

925   ignores nonconserved regions. Phylogenetic reconstruction was performed using IQ-TREE

926   v2.1.3 (Minh et al., 2020). The ModelFinder tool was implemented to identify the best maximum

927   likelihood model for estimating evolutionary relationships (Kalyaanamoorthy et al., 2017),

928   leading to selection of Jones-Taylor-Thornton (JTT)+F with seven rate categories and Q.plant+F

929   with seven rate categories for the V4.1 and HQ MSAs, respectively. Phylogenetic trees were

930   inferred by maximum likelihood using the chosen model, and branch support was obtained from

931   100,000 ultrafast bootstrap iterations (Hoang et al., 2018). The resulting BAHD phylogenies

932   were visualized using the ggtree package in R (Yu et al., 2017). To generate the clade III BAHD

933   heatmap-tree, we used the viewClade function in ggtree to subset the phylogeny and used

934   gheatmap to visualize transcript abundance (log2 TPM) for the eggplant BAHDs.

935

936   BAHD acyltransferase cloning, expression, and purification

937   Candidate ASATs were cloned into pET28b(+) (MilliporeSigma, Burlington, MA, USA).

938 Open reading frames from genes were either synthesized by Twist Biosciences (South San

939 Francisco, CA, USA) with or without codon-optimization for *E. coli* expression (Table S39) or

940 amplified from genomic DNA or cDNA with primers listed in Table S40. Q5 2X Hotstart master

941 mix (New England Biolabs, Ipswich, MA, USA) was used for cloning PCRs. The source of

942 cloned DNA and whether a gene was codon optimized is described in Table S39. Amplified

943 genes were purified by agarose gel electrophoresis and extraction with the Monarch DNA Gel

944 Extraction Kit (New England Biolabs, Ipswich, MA, USA). Both the synthesized genes and PCR

945 amplified genes were then inserted into a doubly digested BamHI/XhoI pET28b(+) through

946 Gibson assembly using the 2X Gibson Assembly Master Mix (New England Biolabs, Ipswich,

947 MA, USA) according to the manufacturer's instructions. The constructs using synthesized genes

948 were transformed into BL21(DE3) (MilliporeSigma, Burlington, MA, USA) and constructs with

949 PCR amplified genes were transformed into BL21 Rosetta(DE3) cells (MilliporeSigma,

950 Burlington, MA, USA). Constructs were verified with colony PCR and Sanger sequencing using

951 T7 terminator and promoter primers (Table S40). Sanger sequencing was completed by the

952 Michigan State University Research Technology Support Facility Genomics Core (East Lansing,

953 MI, USA).

954   Protein expression occurred as previously described (Leong et al., 2022; Lou et al.,

955 2021). Briefly, 50 mL cultures of picked transformation colonies were grown overnight at 37°C,

956 shaking at 225 rpm in Luria-Bertani (LB) media (Neogen, Lansing, MI, USA) supplemented

957 with 1% glucose (w/v). Fifteen mL of the overnight cultures were inoculated into 1 L of fresh LB

958 medium, which was incubated at 37°C shaking at 225 rpm until an OD600 of 0.5 was reached.

959 The cultures were incubated on ice for 25 min, after which, isopropylthio-β-galactoside was

960   added to a final concentration of 50-500 µM. Then, cultures were incubated at 16°C shaking at

961   180 rpm overnight for 16 hours before cells were harvested by centrifugation at 4,000 rpm for 10

962   minutes at 4°C.

963   *S. melongena* BAHDs were purified as previously described (Leong et al., 2020) with the

964   following modifications. The extraction buffer contained 10 mM imidazole, the wash buffer

965   contained 20 mM imidazole, and the elution buffer contained 500 mM imidazole. Protein eluent

966   was concentrated with 30-kD Amicon Ultra centrifugal filter units (MilliporeSigma).

967

968   Enzyme assays

969   Enzyme assays were conducted in 100 mM sodium phosphate buffer at pH 6. For

970   forward assays, acetyl-CoA (MilliporeSigma, Burlington, MA, USA) was added to a final

971   concentration of 0.1 mM, and for reverse assays free CoA (MilliporeSigma, Burlington, MA,

972   USA) was added to a final concentration of 1 mM. Purified AI3:16 and AI4:18 substrates were

973   dried down using a vacuum centrifuge and redissolved in ethanol:water:formic acid (1:1:0.001).

974   One microliter of the prepared acylsugars were used as acyl acceptors. Six microliters of enzyme

975   were added to a final volume of 60 µL. For negative controls, 6 µL of enzyme that was heat

976   inactivated at 95°C for 10 minutes was substituted in place of untreated enzyme. Assays were

977   incubated at 30°C for 30 minutes after which 120 µL of acetonitrile:isopropanol:formic acid

978   (1:1:0.001) with 1.5 µM telmisartan (MilliporeSigma, Burlington, MA, USA) stop solution was

979   added. Reactions were then spun at 17,000 x *g* for 10 minutes to remove precipitate. Supernatant

980   was placed in autosampler vials and analyzed by LC-MS.

981

982   RT-PCR of *S. aethiopicum* BAHDs

983   We employed semi-quantitative RT-PCR to test *S. aethiopicum* BAHD expression in

984   glandular trichomes with cDNA (Figure S14) and genomic DNA (gDNA) (Figure S16). Total

985   RNA was isolated with the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany), including an on-

986   column DNase digestion (Qiagen, Hilden, Germany), from shaved hypocotyls from accession *S.*

987   *aethiopicum* PI 666075 and glandular trichomes from accessions PI 666075 and Grif 14165

988   isolated as described above for *S. melongena* glandular trichomes. RNA was quantified with a

989   Nanodrop 2000c instrument (Thermo Fisher Scientific, Waltham, MA, USA). We synthesized

990   cDNA using 10 ng of RNA and SuperScript III Reverse Transcriptase (Invitrogen, Waltham,

991   MA, USA). gDNA was isolated with the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany)

992   from young leaf tissue collected from mature *S. aethiopicum* PI 666075 plants. PCR reactions

993   (25 µL) were set up with GoTaq Green Master Mix (Promega, Madison, WI, USA), 200 nM of

994   forward and reverse primers (Table S40), and 1 µL of cDNA or gDNA. PCR was performed

995   under these conditions: 2 min at 95°C followed by 22, 30, and 35 cycles of 30s at 95°C, 30s at

996   58°C, and 1 min at 72°C.

997   We identified the putative *S. aethiopicum* orthologs of the *S. melongena* candidate

998   BAHDs by querying *S. aethiopicum* annotated transcripts (Song et al., 2019) with *S. melongena*

999   candidate ASAT DNA sequences using BLASTn. SaASAT3 and SaASAT3-L1 (GenBank

1000   accession: OQ547782) were not annotated in the *S. aethiopicum* genome and were identified by

1001   querying *S. aethiopicum* scaffolds with BLASTn and then identifying open reading frames with

1002   Geneious software v9.1.8 (Dotmatics, Boston, MA, USA). The *S. aethiopicum* ef1α gene was

1003   identified by querying a putative *S. melongena* ef1α, SMEL4.1_06g005890, against the *S.*

1004   *aethiopicum* scaffolds.

1005

1006   *S. melongena* acyl chain composition analysis

1007   *S. melongena* acyl chain composition was assessed with a previously developed

1008   esterification and GC-MS methods (Fan et al., 2020; Ning et al., 2015; Schenck et al., 2022).

1009   Acyl chain identities were determined through authentic reference standards for nC8, nC10,

1010   nC12, and nC14 fatty acid ethyl esters (MilliporeSigma). The QuanLynx function of MassLynx

1011   v4.1 (Waters Corporation) integrated peaks from extracted ion chromatograms. Extracted ion

1012   chromatograms for the medium acyl chains iC8, nC8, iC10, nC10, iC12, nC12, iC14, and nC14

1013   were generated for the *m/z* of 88 with a mass window of *m/z* of 0.50. Extracted ion

1014   chromatograms for the short chains iC4, aiC5, and iC5 were generated for with the *m/z* values of

1015   71, 102, and 101, respectively, with a mass window of *m/z* of 0.50.

1016

1017   Accession numbers

1018   Raw RNA sequence data can be found in the National Center for Biotechnology

1019   Information (NCBI) Sequence Read Archive database under BioProject PRJNA935765.

1020   Sequence data for SaASAT3-L1 is available from the NCBI GenBank database under accession

1021   number OQ547782.

1022   The NMR data for compounds 1-8 in Figure 1D have been deposited in the Natural

1023   Products Magnetic Resonance Database (np-mrd.org) under accession numbers NP0332515,

1024   NP0332562, NP0332563, NP0332564, NP0332571, NP0332572, NP0332573, NP0332574, and

1025   NP0332575 as documented in Table S41.

1026   The LC-MS data used to annotate compounds in Tables 1, S3-27 have been deposited in

1027   the Metabolights database (https://www.ebi.ac.uk/metabolights) under the study ID

1028   MTBLS9459.

1029

Acknowledgements

1031   We thank Dr. Christopher T. Martine for providing Australian *Solanum* ssp. seeds, Dr.

1032   Joyce van Eck for providing *Solanum prinophyllum* seeds, Dr. Guiseppe Leonardo Rotino and

1033   the Consiglio per la ricercar in agricoltura e l'analisi dell'economia agraria for providing *S.*

1034   *melongena* 67/3 and 305E40 seeds, and USDA-GRIN for providing *Solanum* ssp. seeds. We

1035   acknowledge the Michigan State University RTSF Mass Spectrometry and Metabolomics Core

1036   Facilities for LC-MS analysis support. We also acknowledge Dr. Daniel Holmes and Dr. Li Xie

1037   at the Michigan State University Max T. Rogers NMR Facility for experimental design and data

1038   analysis support. We are grateful to current and past members of the Last lab for feedback on the

1039   research and especially Drs. Jaynee Hart and Xingxing Li for comments on the manuscript.

1040

Funding

1042   This research was funded by the National Science Foundation grant IOS-PGRP-1546617

1043   (to R. L. L. and A. D. J.), National Science Foundation grant IOS-PGRP-2218206 (to R. L. L.

1044   and R. E. K.) and National Institute of General Medical Sciences of the National Institutes of

1045   Health graduate training grant no. T32-GM110523 (to P. D. F.). A. D. J. is supported in part by

1046   USDA National Institute of Food and Agriculture Hatch project number MICL02474.

1047

Author Contributions

1049   PDF: Designed and performed research, analyzed data, wrote the paper

1050   REK: Designed and performed research, analyzed data, wrote the paper

1051   ADJ: Designed research, analyzed data, wrote the paper

1052   RLL: Designed research, reviewed data, wrote the paper

References

1054 Agati, G., Tattini, M., 2010. Multiple functional roles of flavonoids in photoprotection. New Phytol 186:
1055 786–793

1056 Altschul, S.F., 1998. Generalized affine gap costs for protein sequence alignment. Proteins Struct Funct
1057 Bioinforma 32: 88–96

1058 Anand, U., Jacobo-Herrera, N., Altemimi, A., Lakhssassi, N., 2019. A comprehensive review on
1059 medicinal plants as antimicrobial therapeutics: potential avenues of biocompatible drug discovery.
1060 Metabolites 9: 258

1061 Anders, S., Pyl, P.T., Huber, W., 2015. HTSeq—a Python framework to work with high-throughput
1062 sequencing data. Bioinformatics 31: 166–169

1063 Asai, T., Fujimoto, Y., 2010. Cyclic fatty acyl glycosides in the glandular trichome exudate of *Silene*
1064 *gallica*. Phytochemistry 71: 1410–1417

1065 Asai, T., Hara, N., Fujimoto, Y., 2010. Fatty acid derivatives and dammarane triterpenes from the
1066 glandular trichome exudates of *Ibicella lutea* and *Proboscidea louisiana*. Phytochemistry 71: 877–894

1067 Aubriot, X., Knapp, S., Syfert, M.M., Poczai, P., Buerki, S., 2018. Shedding new light on the origin and
1068 spread of the brinjal eggplant (*Solanum melongena* L.) and its wild relatives. Am J Bot 105: 1175–1187

1069 Augustin, J.M., Kuzina, V., Andersen, S.B., Bak, S., 2011. Molecular activities, biosynthesis and
1070 evolution of triterpenoid saponins. Phytochemistry 72: 435–457

1071 Bah, M., Pereda-Miranda, R., 1996. Detailed FAB-mass spectrometry and high resolution NMR
1072 investigations of tricolorins A-E, individual oligosaccharides from the resins of *Ipomoea tricolor*
1073 (Convolvulaceae). Tetrahedron 52: 13063–13080

1074 Barchi, L., Rabanus-Wallace, M.T., Prohens, J., Toppino, L., Padmarasu, S., Portis, E., Rotino, G.L.,
1075 Stein, N., Lanteri, S., Giuliano, G., 2021. Improved genome assembly and pan-genome provide key
1076 insights into eggplant domestication and breeding. Plant J 107: 579–596

1077 Baumann, E., Stoya, G., Völkner, A., Richter, W., Lemke, C., Linss, W., 2000. Hemolysis of human
1078 erythrocytes with saponin affects the membrane structure. Acta Histochem 102: 21–35

1079 Bernal, C.-A., Castellanos, L., Aragón, D.M., Martínez-Matamoros, D., Jiménez, C., Baena, Y., Ramos,
1080 F.A., 2018. Peruvioses A to F, sucrose esters from the exudate of *Physalis peruviana* fruit as α-amylase
1081 inhibitors. Carbohydr Res 461: 4–10

1082 Bohs, L., 2005. Major clades in *Solanum* based on ndhF sequence data. Monogr Syst Bot 104: 27–49

1083 Bohs, L., Olmstead, R.G., 1997. Phylogenetic relationships in *Solanum* (Solanaceae) based on ndhF
1084 sequences. Syst Bot 22: 5–17

1085 Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data.
1086 Bioinformatics 30: 2114–2120

1087 Cao, C.-M., Wu, X., Kindscher, K., Xu, L., Timmermann, B.N., 2015. Withanolides and sucrose esters
1088 from *Physalis neomexicana*. J Nat Prod 78: 2488–2493

Caprioli, G., Iannarelli, R., Cianfaglione, K., Fiorini, D., Giuliani, C., Lucarini, D., Papa, F., Sagratini, G., Vittori, S., Maggi, F., 2016. Volatile profile, nutritional value and secretory structures of the berry-like fruits of *Hypericum androsaemum* L. Food Res Int 79: 1–10

Castillo, M., Connolly, J.D., Ifeadike, P., Labbé, C., Rycroft, D.S., Woods, N., 1989. Partially acylated glucose and sucrose derivatives from *Salpiglossis sinuata* (Solanaceae). J Chem Res Synop 12: 398–399

Chortyk, O.T., Kays, S.J., Teng, Q., 1997. Characterization of insecticidal sugar esters of *Petunia*. J Agric Food Chem 45: 270–275

Cicchetti, E., Duroure, L., Le Borgne, E., Laville, R., 2018. Upregulation of skin-aging biomarkers in aged NHDF cells by a sucrose ester extract from the agroindustrial waste of *Physalis peruviana* calyces. J Nat Prod 81: 1946–1955

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

De Moraes, C.M., Mescher, M.C., Tumlinson, J.H., 2001. Caterpillar-induced nocturnal plant volatiles repel conspecific females. Nature 410: 577–580

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

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29: 15–21

Dobin, A., Gingeras, T.R., 2015. Mapping RNA-seq reads with STAR. Curr Protoc Bioinforma 51: 11.14.1-11.14.19

Fan, P., Leong, B.J., Last, R.L., 2019. Tip of the trichome: evolution of acylsugar metabolic diversity in Solanaceae. Curr Opin Plant Biol 49: 8–16

Fan, P., Miller, A.M., Liu, X., Jones, A.D., Last, R.L., 2017. Evolution of a flipped pathway creates metabolic innovation in tomato trichomes through BAHD enzyme promiscuity. Nat Commun 8: 2080

Fan, P., Miller, A.M., Schilmiller, A.L., Liu, X., Ofner, I., Jones, A.D., Zamir, D., Last, R.L., 2016. *In vitro* reconstruction and analysis of evolutionary variation of the tomato acylsucrose metabolic network. Proc Natl Acad Sci USA 113: E239–E248

Fan, P., Wang, P., Lou, Y.R., Leong, B.J., Moore, B.M., Schenck, C.A., Combs, R., Cao, P., Brandizzi, F., Shiu, S.H., Last, R.L., 2020. Evolution of a plant gene cluster in Solanaceae and emergence of metabolic diversity. eLife 9: e56717.

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

Firn, R.D., Jones, C.G., 2009. A Darwinian view of metabolism: molecular properties determine fitness. J Exp Bot 60: 719–726

Fobes, J.F., Mudd, J.B., Marsden, M.P.F., 1985. Epicuticular lipid accumulation on the leaves of *Lycopersicon pennellii* (Corr.) D'Arcy and *Lycopersicon esculentum* Mill. Plant Physiol 77: 567–570

1127   Fraenkel, G.S., 1959. The raison d'être of secondary plant substances. Science 129: 1466–1470.

1128   Gagnon, E., Hilgenhof, R., Orejuela, A., McDonnell, A., Sablok, G., Aubriot, X., Giacomin, L., Gouvêa,
1129   Y., Bragionis, T., Stehmann, J.R., Bohs, L., Dodsworth, S., Martine, C., Poczai, P., Knapp, S., Särkinen,
1130   T., 2022. Phylogenomic discordance suggests polytomies along the backbone of the large genus *Solanum*.
1131   Am J Bot 109: 580–601

1132   Ghosh, B., Westbrook, T.C., Jones, A.D., 2014. Comparative structural profiling of trichome specialized
1133   metabolites in tomato (*Solanum lycopersicum*) and *S. habrochaites*: acylsugar profiles revealed by
1134   UHPLC/MS and NMR. Metabolomics 10: 496–507

1135   Goffreda, J.C., Mutschler, M.A., Avé, D.A., Tingey, W.M., Steffens, J.C., 1989. Aphid deterrence by
1136   glucose esters in glandular trichome exudate of the wild tomato, *Lycopersicon pennellii*. J Chem Ecol 15:
1137   2135–2147

1138   Herrera-Salgado, Y., Garduño-Ramírez, M.L., Vázquez, L., Rios, M.Y., Alvarez, L., 2005. Myo-inositol-
1139   derived glycolipids with anti-inflammatory activity from *Solanum lanceolatum*. J Nat Prod 68: 1031–
1140   1036

1141   Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the
1142   ultrafast bootstrap approximation. Mol Biol Evol 35: 518–522

1143   Hurney, S.M., 2018. Strategies for profiling and discovery of acylsugar specialized metabolites (Ph.D.).
1144   Michigan State University, United States -- Michigan.

1145   Jenske, R., Vetter, W., 2007. Highly selective and sensitive gas chromatography–electron-capture
1146   negative-ion mass spectrometry method for the indirect enantioselective identification of 2- and 3-
1147   hydroxy fatty acids in food and biological samples. J Chromatogr A 1146: 225–231

1148   Ji, W., Mandal, S., Rezenom, Y.H., McKnight, T.D., 2023. Specialized metabolism by trichome-enriched
1149   Rubisco and fatty acid synthase components. Plant Physiol 191:1199-1213

1150   Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder:
1151   fast model selection for accurate phylogenetic estimates. Nat Methods 14: 587–589

1152   Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements
1153   in performance and usability. Mol Biol Evol 30: 772–780

1154   Kerwin, R.E., Hart, J.E., Fiesel, P.D., Lou, Y.-R., Fan, P., Jones, A.D., Last, R.L., 2024. Tomato root
1155   specialized metabolites evolved through gene duplication and regulatory divergence within a biosynthetic
1156   gene cluster. Sci Adv (*in press*)

1157   Kim, J., Kang, K., Gonzales-Vigil, E., Shi, F., Daniel Jones, A., Barry, C.S., Last, R.L., 2012. Striking
1158   natural diversity in glandular trichome acylsugar composition is shaped by variation at the
1159   acyltransferase2 locus in the wild tomato *Solanum habrochaites*. Plant Physiol 160: 1854–1870

1160   King, R.R., Calhoun, L.A., 1988. 6 2,3-Di-O- and 1,2,3-tri-O-acylated glucose esters from the glandular
1161   trichomes of *Datura metel*. Phytochemistry 27: 3761–3763

1162   King, R.R., Pelletier, Y., Singh, R.P., Calhoun, L.A., 1986. 3,4-Di-O-isobutyryl-6-O-caprylsucrose: the
1163   major component of a novel sucrose ester complex from the type B glandular trichomes of *Solanum*
1164   *berthaultii* Hawkes (Pl 473340). J Chem Soc Chem Commun 1078–1079

1165   King, R.R., Singh, R.P., Calhoun, L.A., 1987. Isolation and characterization of 3,3′,4,6-tetra-O-acylated
1166   sucrose esters from the type B glandular trichomes of Solanum berthaultii Hawkes (PI 265857).
1167   Carbohydr Res 166: 113–121

1168   Koonin, E.V., 2005. Orthologs, paralogs, and evolutionary genomics. Annu Rev Genet 39: 309–338

1169   Korenblum, E., Dong, Y., Szymanski, J., Panda, S., Jozwiak, A., Massalha, H., Meir, S., Rogachev, I.,
1170   Aharoni, A., 2020. Rhizosphere microbiome mediates systemic root metabolite exudation by root-to-root
1171   signaling. Proc Natl Acad Sci USA 117: 3874–3883

1172   Kretschmar, J.A., Baumann, T.W., 1999. Caffeine in citrus flowers. Phytochemistry 52: 19–23

1173   Kruse, L.H., Bennett, A.A., Mahood, E.H., Lazarus, E., Park, S.J., Schroeder, F., Moghe, G.D., 2022.
1174   Illuminating the lineage-specific diversification of resin glycoside acylsugars in the morning glory
1175   (Convolvulaceae) family using computational metabolomics. Hortic Res 9: uhab079

1176   Kutschera, A., Dawid, C., Gisch, N., Schmid, C., Raasch, L., Gerster, T., Schäffer, M., Smakowska-
1177   Luzan, E., Belkhadir, Y., Vlot, A.C., Chandler, C.E., Schellenberger, R., Schwudke, D., Ernst, R.K.,
1178   Dorey, S., Hückelhoven, R., Hofmann, T., Ranf, S., 2019. Bacterial medium-chain 3-hydroxy fatty acid
1179   metabolites trigger immunity in *Arabidopsis* plants. Science 364: 178–181

1180   Landry, L.G., Chapple, CCS., Last, R.L., 1995. *Arabidopsis* mutants lacking phenolic sunscreens exhibit
1181   enhanced ultraviolet-B injury and oxidative damage. Plant Physiol 109: 1159–1166

1182   Leckie, B.M., D'Ambrosio, D.A., Chappell, T.M., Halitschke, R., De Jong, D.M., Kessler, A., Kennedy,
1183   G.G., Mutschler, M.A., 2016. Differential and synergistic functionality of acylsugars in suppressing
1184   oviposition by insect herbivores. PLoS ONE 11: e0153345

1185   Leong, B.J., Hurney, S., Fiesel, P., Anthony, T.M., Moghe, G., Jones, A.D., Last, R.L., 2022.
1186   Identification of BAHD acyltransferases associated with acylinositol biosynthesis in *Solanum quitoense*
1187   (naranjilla). Plant Direct 6: e415

1188   Leong, B.J., Hurney, S.M., Fiesel, P.D., Moghe, G.D., Jones, A.D., Last, R.L., 2020. Specialized
1189   metabolism in a nonmodel nightshade: trichome acylinositol biosynthesis. Plant Physiol 183: 915–924

1190   Leong, B.J., Lybrand, D.B., Lou, Y.R., Fan, P., Schilmiller, A.L., Last, R.L., 2019. Evolution of
1191   metabolic novelty: a trichome-expressed invertase creates specialized metabolic diversity in wild tomato.
1192   Sci Adv 5: eabj8726

1193   Levin, D.A., 1973. The role of trichomes in plant defense. Q Rev Biol 48: 3–15

1194   Levin, R.A., Myers, N.R., Bohs, L., 2006. Phylogenetic relationships among the "spiny solanums"
1195   (*Solanum* subgenus *Leptostemonum*, Solanaceae). Am J Bot 93: 157–169

1196   Li, D., Qian, J., Li, Weiliu, Yu, N., Gan, G., Jiang, Y., Li, Wenjia, Liang, X., Chen, R., Mo, Y., Lian, J.,
1197   Niu, Y., Wang, Y., 2021. A high-quality genome assembly of the eggplant provides insights into the
1198   molecular basis of disease resistance and chlorogenic acid synthesis. Mol Ecol Resour 21: 1274–1286.

1199   Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.,
1200   1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and
1201   SAMtools. Bioinformatics 25: 2078–2079

1202 Liu, X., Enright, M., Barry, C.S., Jones, A.D., 2017. Profiling, isolation and structure elucidation of
1203 specialized acylsucrose metabolites accumulating in trichomes of *Petunia* species. Metabolomics 13: 85.

1204 Lou, Y.-R., Anthony, T.M., Fiesel, P.D., Arking, R.E., Christensen, E.M., Jones, A.D., Last, R.L., 2021.
1205 It happened again: convergent evolution of acylglucose specialized metabolism in black nightshade and
1206 wild tomato. Sci Adv 7: eabj8726

1207 Lou, Y.-R., Leong, B., 2019. Yann-Ru Lou, Bryan Leong 2019. Leaf surface acylsugar extraction and
1208 LC-MS profiling - v1.0. protocols.io https://dx.doi.org/10.17504/protocols.io.xj2fkqe

1209 Luu, V.T., Weinhold, A., Ullah, C., Dressel, S., Schoettner, M., Gase, K., Gaquerel, E., Xu, S., Baldwin,
1210 I.T., 2017. O-acyl sugars protect a wild tobacco from both native fungal pathogens and a specialist
1211 herbivore. Plant Physiol 174: 370–386

1212 Lybrand, D.B., Anthony, T.M., Jones, A.D., Last, R.L., 2020. An integrated analytical approach reveals
1213 trichome acylsugar metabolite diversity in the wild tomato *Solanum pennellii*. Metabolites 10: 401.

1214 Maldonado, E., Torres, F.R., Martínez, M., Pérez-Castorena, A.L., 2006. Sucrose esters from the fruits of
1215 *Physalis nicandroides* var. *attenuata*. J Nat Prod 69: 1511–1513

1216 Matsuzaki, T., Shinozaki, Y., Suhara, S., Ninomiya, M., Shigematsu, H., Koiwai, A., 1989. Isolation of
1217 glycolipids from the surface lipids of *Nicotiana bigelovii* and their distribution in *Nicotiana* species. Agric
1218 Biol Chem 53: 3079–3082

1219 McCarthy, D.J., Campbell, K.R., Lun, A.T.L., Wills, Q.F., 2017. Scater: pre-processing, quality control,
1220 normalization and visualization of single-cell RNA-seq data in R. Bioinformatics 33: 1179–1186

1221 McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K.,
1222 Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: A MapReduce
1223 framework for analyzing next-generation DNA sequencing data. Genome Res 20: 1297–1303

1224 Mennella, G., Rotino, G.L., Fibiani, M., D'Alessandro, A., Francese, G., Toppino, L., Cavallanti, F.,
1225 Acciarri, N., Lo Scalzo, R., 2010. Characterization of health-related compounds in eggplant (*Solanum
1226 melongena* L.) lines derived from introgression of allied species. J Agric Food Chem 58: 7597–7603

1227 Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear,
1228 R., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era.
1229 Mol Biol Evol 37: 1530–1534

1230 Moghe, G., Kruse, L.H., Petersen, M., Scossa, F., Fernie, A.R., Gaquerel, E., D'Auria, J.C., 2023. BAHD
1231 company: the ever-expanding roles of the BAHD acyltransferase gene family in plants. Annu Rev Plant
1232 Biol 74: 3.1–3.30

1233 Moghe, G.D., Leong, B.J., Hurney, S.M., Jones, A.D., Last, R.L., 2017. Evolutionary routes to
1234 biochemical innovation revealed by integrative analysis of a plant-defense related specialized metabolic
1235 pathway. eLife 6: e28468

1236 Nadakuduti, S.S., Uebler, J.B., Liu, X., Jones, A.D., Barry, C.S., 2017. Characterization of trichome-
1237 expressed BAHD acyltransferases in *Petunia axillaris* reveals distinct acylsugar assembly mechanisms
1238 within the Solanaceae. Plant Physiol 175: 36–50

1239 Nee, M., 2022a. *Solanum acerifolium* [Dunal] [WWW Document]. URL
1240 https://solanaceaesource.myspecies.info/taxonomy/term/105367/descriptions (accessed 10.3.22).

51

1241   Nee, M., 2022b. *Solanum atropurpureum* [Schrank] [WWW Document]. URL
1242   https://solanaceaesource.myspecies.info/taxonomy/term/105749/descriptions (accessed 10.3.22).

1243   Nee, M., 1991. Synopsis of *Solanum* section Acanthophora: a group of interest for glycoalkaloids, in:
1244   Solanaceae III, taxonomy, chemistry, evolution. Royal Botanic Gardens for the Linnean Society of
1245   London, Kew, Richmond, Surrey, UK, pp. 257–266

1246   Ning, J., Moghe, G.D., Leong, B., Kim, J., Ofner, I., Wang, Z., Adams, C., Jones, A.D., Zamir, D., Last,
1247   R.L., 2015. A feedback-insensitive isopropylmalate synthase affects acylsugar composition in cultivated
1248   and wild tomato. Plant Physiol 169: 1821–1835

1249   PBI Solanum Project, 2022. Solanaceae Source [WWW Document]. URL https://solanaceaesource.org
1250   (accessed 10.3.22)

1251   Pereda-Miranda, R., Mata, R., Anaya, A.L., Wickramaratne, D.B.M., Pezzuto, J.M., Kinghorn, A.D.,
1252   1993. Tricolorin A, major phytogrowth inhibitor from *Ipomoea tricolor*. J Nat Prod 56: 571–582

1253   Pichersky, E., Gang, D.R., 2000. Genetics and biochemistry of secondary metabolites in plants: an
1254   evolutionary perspective. Trends Plant Sci 5: 439–445

1255   Pichersky, E., Lewinsohn, E., 2011. Convergent evolution in plant specialized metabolism. Annu. Rev.
1256   Plant Biol 62: 549–566

1257   Pinot, F., Beisson, F., 2011. Cytochrome P450 metabolizing fatty acids in plants: characterization and
1258   physiological roles. FEBS J 278: 195–205

1259   Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential
1260   expression analysis of digital gene expression data. Bioinformatics 26: 139–140

1261   Sampaio, V. da S., Coutinho, Í.A.C., Särkinen, T., Loiola, M.I.B., 2021. Secretory and ecological
1262   function of petiolar glands in *Solanum fernandesii*: first description of resin glands in the genus *Solanum*.
1263   Aust J Bot 70: 32–41.

1264   Särkinen, T., Bohs, L., Olmstead, R.G., Knapp, S., 2013. A phylogenetic framework for evolutionary
1265   study of the nightshades (Solanaceae): a dated 1000-tip tree. BMC Evol Biol 13: 214

1266   Sassaki, G.L., Souza, L.M., Serrato, R.V., Cipriani, T.R., Gorin, P.A.J., Iacomini, M., 2008. Application
1267   of acetate derivatives for gas chromatography–mass spectrometry: novel approaches on carbohydrates,
1268   lipids and amino acids analysis. J Chromatogr A 1208: 215–222

1269   Schenck, C.A., Anthony, T.M., Jacobs, M., Jones, A.D., Last, R.L., 2022. Natural variation meets
1270   synthetic biology: promiscuous trichome-expressed acyltransferases from *Nicotiana*. Plant Physiol 190:
1271   146-164

1272   Schilmiller, A.L., Charbonneau, A.L., Last, R.L., 2012. Identification of a BAHD acetyltransferase that
1273   produces protective acyl sugars in tomato trichomes. Proc Natl Acad Sci USA 109: 16377–16382.

1274   Schilmiller, A.L., Gilgallon, K., Ghosh, B., Jones, A.D., Last, R.L., 2016. Acylsugar acylhydrolases:
1275   carboxylesterase-catalyzed hydrolysis of acylsugars in tomato trichomes. Plant Physiol 170: 1331–1344

1276   Schilmiller, A.L., Moghe, G.D., Fan, P., Ghosh, B., Ning, J., Jones, A.D., Last, R.L., 2015. Functionally
1277   divergent alleles and duplicated Loci encoding an acyltransferase contribute to acylsugar metabolite
1278   diversity in *Solanum* trichomes. Plant Cell 27: 1002–17

1279 Shapiro, J.A., Steffens, J.C., Mutschler, M.A., 1994. Acylsugars of the wild tomato *Lycopersicon*
1280 *pennellii* in relation to geographic distribution of the species. Biochem Syst Ecol 22: 545–561

1281 Smith, C.R., 1971. Occurrence of unusual fatty acids in plants. Prog Chem Fats Other Lipids 11: 137–
1282 177.

1283 Song, B., Song, Y., Fu, Y., Kizito, E.B., Kamenya, S.N., Kabod, P.N., Liu, H., Muthemba, S., Kariba, R.,
1284 Njuguna, J., Maina, S., Stomeo, F., Djikeng, A., Hendre, P.S., Chen, X., Chen, W., Li, X., Sun, W.,
1285 Wang, S., Cheng, S., Muchugi, A., Jamnadass, R., Shapiro, H.-Y., Van Deynze, A., Yang, H., Wang, J.,
1286 Xu, X., Odeny, D.A., Liu, X., 2019. Draft genome sequence of *Solanum aethiopicum* provides insights
1287 into disease resistance, drought tolerance, and the evolution of the genome. GigaScience 8: giz115.

1288 Stern, S., Agra, M. de F., Bohs, L., 2011. Molecular delimitation of clades within New World species of
1289 the "spiny solanums" (*Solanum* subg. *Leptostemonum*). TAXON 60: 1429–1441

1290 Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W.-M., Fiehn, O.,
1291 Goodacre, R., Griffin, J.L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A.N.,
1292 Lindon, J.C., Marriott, P., Nicholls, A.W., Reily, M.D., Thaden, J.J., Viant, M.R., 2007. Proposed
1293 minimum reporting standards for chemical analysis. Metabolomics 3: 211–221

1294 Tepe, E.J., Anderson, G.J., Spooner, D.M., Bohs, L., 2016. Relationships among wild relatives of the
1295 tomato, potato, and pepino. TAXON 65: 262–276

1296 Van Dam, N.M., Hare, J.D., 1998. Biological activity of *Datura wrightii* glandular trichome exudate
1297 against *Manduca sexta* larvae. J Chem Ecol 24: 1529–1549

1298 Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan,
1299 T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K.V., Altshuler, D., Gabriel, S., DePristo,
1300 M.A., 2013. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices
1301 pipeline. Curr Protoc Bioinforma 43: 11.10.1-11.10.33

1302 Wagner, G.J., 1991. Secreting glandular trichomes: more than just hairs. Plant Physiol 96: 675–679

1303 Weese, T.L., Bohs, L., 2007. A three-gene phylogeny of the genus *Solanum* (Solanaceae). Syst Bot 32:
1304 445–463.

1305 Weinhold, A., Baldwin, I.T., 2011. Trichome-derived O-acyl sugars are a first meal for caterpillars that
1306 tags them for predation. Proc Natl Acad Sci USA 108: 7855–7859

1307 Wink, M., 2015. Modes of action of herbal medicines and plant secondary metabolites. Medicines 2: 251–
1308 286.

1309 Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.-Y., 2017. ggtree: an R package for visualization and
1310 annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol 8: 28–
1311 36

1312 Yu, P., He, X., Baer, M., Beirinckx, S., Tian, T., Moya, Y.A.T., Zhang, X., Deichmann, M., Frey, F.P.,
1313 Bresgen, V., Li, C., Razavi, B.S., Schaaf, G., von Wirén, N., Su, Z., Bucher, M., Tsuda, K., Goormachtig,
1314 S., Chen, X., Hochholdinger, F., 2021. Plant flavones enrich rhizosphere Oxalobacteraceae to improve
1315 maize performance under nitrogen deprivation. Nat Plants 7: 481–499.

1316 Zähner, H., 1979. What are secondary metabolites? Folia Microbiol 24: 435–443

1317   Zizovic, I., Stamenić, M., Orlović, A., Skala, D., 2007. Supercritical carbon dioxide extraction of
1318   essential oils from plants with secretory ducts: mathematical modelling on the micro-scale. J Supercrit
1319   Fluids 39: 338–346

1320

Table 1. Summary of annotated acylsugars detected in *S. melongena* leaf surface extracts.
PH = pentose-hexose; AI = arabinose-inositol, I = inositol. RT = retention time; m/z_{acc} = theoretical monoisotopic formate adduct mass; m/z_{ex} = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Compound numbers are listed for NMR characterized compounds (Figure 1).. Acylsugars composing ≥1% of total acylsugar abundance area are in boldface text. Acylsugar abundance was calculated by the software Progenesis QI v3.0 using the sum of each compound's adduct ion intensities (see Methods for parameters). Percent acylsugar abundance was calculated by dividing average acylsugar abundance for each compound by the sum of all acylsugar abundances, as reported in the right most column. These values were averaged over 15 samples. The annotation method is described in the Acylsugar Annotation subsection of the Methods. Acylsugars are sorted by number of sugar moieties and then by elution order. Some acylsugar isomers were unable to be deconvoluted by Progensis QI and were reported with shared m/z, Δm, and percent abundance values as indicated by the merged table cells.

Name	Compound #	Annotation confidence level	RT (min)	Neutral molecule chemical formula	m/z_{acc}	m/z_{ex}	Δm (ppm)	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

1336

1337

Figure Legends

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

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 derivatization 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. The minor peak from I3:22 is either from a small amount of endogenous *S* isomer or instrument noise.

Figure 3. Distribution of acylsugar traits across the *Solanum* genus. Acylsugar phenotypes were mapped onto a previously published Supermatrix maximum likelihood phylogeny (Gagnon et al., 2022). Five rings (labeled **A-E**) surrounding the phylogeny represent different acylsugar traits that vary across *Solanum* species (summarized in Table S2). **A) AS presence:** presence of detectable acylsugars in aerial surface (i.e., trichome) extracts. **B) Hex core:** identity of hexose sugar core in detected acylhexoses. **C) Disaccharide core:** identity of disaccharide sugar core in detected acyldisaccharides. If acyldisaccharides were detected but the sugar core identity could not be resolved, it was categorized as ambiguous. **D) OH-acyl chains:** identity of unusual acyl chains bearing hydroxyl or glycosylated hydroxyl groups in detected acylsugars. **E) eggplant-like AS:** presence of acylsugars with the same acyl chain and sugar core composition that also coelute with eggplant acylsugars (and are therefore likely identical). Acylsugar annotations from this study are based upon LC-MS and NMR data. Annotations from previous studies were limited to those with NMR structural data.

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

1379 numbers.Acylsugar classes are listed for each species. Acylsugar profiles for these species are
1380 detailed in Tables S5, S8, S18, S21, and S25-27.

1381

1382 **Figure 5. *S. atropurpureum* trichomes and fruit exhibit dramatically different metabolite**
1383 **profiles.** The top LC-MS base peak intensity (BPI) chromatogram displays metabolites from a
1384 leaf surface extract, while the bottom LC-MS base peak intensity chromatogram displays
1385 metabolites from a fruit surface extract. The leaf surface extract acylsugars elute from 2.2 to 7.6
1386 min (Table S17), while fruit surface extract acylsugars elute from 4.6 to 18.4 min (Table S15).

1387

1388 **Figure 6. SmASAT3-L1 acetylates AI3:16 to produce AI4:18.** Trichome expressed ASAT
1389 homologs were identified and tested for AI3:16 acetyl transferase activity. (A) LC-MS analysis
1390 of AI3:16(4,4,8) and AI4(2,4,4,8) in *S. aethiopicum* (red) and *S. melongena* (purple) leaf surface
1391 extracts. *S. aethiopicum* does not produce detectable levels of AI4:18. (B) Shown is a clade III
1392 BAHD tree subset from a phylogeny including 106 predicted BAHDs (PF002458) in the
1393 eggplant Smel_V4.1 reference genome, published reference BAHD sequences for clades I-VII,
1394 characterized ASAT sequences from other Solanaceae species, and the SaASAT3 and
1395 SaASAT3-L1 candidates from *S. aethiopicum* (see Figure S12 for full phylogeny). The
1396 maximum likelihood tree was inferred from amino acid sequences using the Jones-Taylor-
1397 Thornton algorithm with seven rate categories in IQ-TREE v2.1.3. Values at nodes indicate
1398 bootstrap support calculated from 100,000 ultrafast bootstrap replicates. The heatmap shows
1399 absolute transcript abundance (log2 TPM) across trichomes, trichomeless hypocotyls, and roots
1400 from 7-day-old eggplant seedlings. Expression data for non-eggplant sequences are not included
1401 in the heatmap. Color gradient provides a visual marker to rank the transcript abundance from
1402 high (purple) to low (white) or absent (grey). ASAT, acylsugar acyltransferase. log2TPM, log2
1403 transformed transcripts per million. Genes denoted with "*" were selected for biochemical
1404 characterization based on expression in trichomes and sequence similarity to characterized
1405 ASAT enzymes in other species. (C) Of the seven enzymes tested *in vitro*, only SmASAT3-L1
1406 successfully acetylated AI3:16(4,4,8) to form detectable AI4:18(2,4,4,8). The extracted ion
1407 chromatograms on the left display products from reverse enzyme assays and the formate adduct
1408 of AI3:16, *m/z* 623.29. The extracted ion chromatograms on the right display products from
1409 forward enzyme assays and the formate adduct of AI4:18, *m/z* 665.30. (D) Acetylation of AI3:16
1410 *in vitro* to produce AI4:18 is abolished upon boiling of SmASAT3-L1 enzyme. Forward assay
1411 extracted ion chromatograms display the formate adduct of AI4:18, *m/z* 665.30, and reverse
1412 assay extracted ion chromatograms display the formate adduct of AI3:16, *m/z* 623.29. The
1413 reverse assay chromatograms display peaks for both AI3:16 and AI4:18 due to in source
1414 fragmentation of remaining AI4:18.

1415

1416 **Figure 7. Characterized *in vitro* acyl-CoA usage by ASATs.** The maximum likelihood
1417 phylogeny was inferred from amino acid sequences of characterized ASAT enzymes using a
1418 time-reversible algorithm specifying a plant-specific empirical substitution matrix, invariant
1419 sites, and four rate categories (Q.plant+I+G4) in IQ-TREE v2.1.3. Branch support was calculated
1420 from 100,000 ultrafast bootstrap replicates. Other than SmASAT3-L1, enzyme activities were
1421 described previously (Fan et al., 2017, 2015; Leong et al., 2022, 2020; Lou et al., 2021; Moghe

58

1422   et al., 2017; Nadakuduti et al., 2017; Schenck et al., 2022; Schilmiller et al., 2015, 2012). Acyl-
1423   CoA lengths are represented by their number of carbons. Acyl chain 7^B corresponds to a
1424   benzoyl-CoA. Reported positive activities are indicated by black squares; reported negative
1425   activities are indicated by gray squares; untested activities are indicated by white squares.
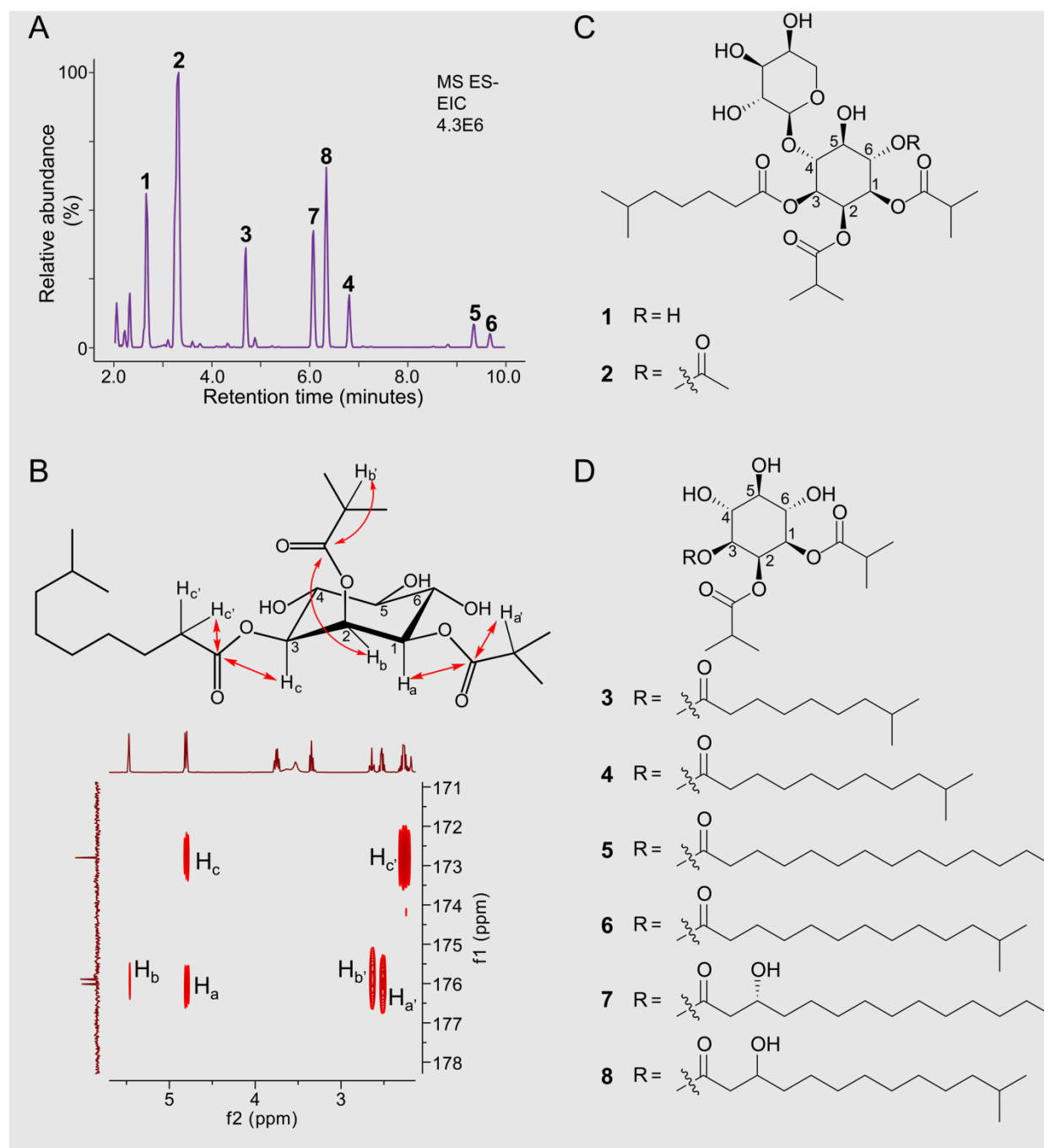
1426

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 acyldisaccharide

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

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 derivatization 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. The minor peak from I3:22 is either from a small amount of endogenous *S* isomer or instrument noise.

Figure 3. Distribution of acylsugar traits across the *Solanum* genus. Acylsugar phenotypes were mapped onto a previously published Supermatrix maximum likelihood phylogeny (Gagnon et al., 2022). Five rings (labeled **A-E**) surrounding the phylogeny represent different acylsugar traits that vary across *Solanum* species (summarized in Table S2). **A) AS presence:** presence of detectable acylsugars in aerial surface (i.e., trichome) extracts. **B) Hex core:** identity of hexose

sugar core in detected acylhexoses. **C) Disaccharide core:** identity of disaccharide sugar core in detected acyldisaccharides. If acyldisaccharides were detected but the sugar core identity could not be resolved, it was categorized as ambiguous. **D) OH-acyl chains:** identity of unusual acyl chains bearing hydroxyl or glycosylated hydroxyl groups in detected acylsugars. **E) eggplant-like AS:** presence of acylsugars with the same acyl chain and sugar core composition that also coelute with eggplant acylsugars (and are therefore likely identical). Acylsugar annotations from this study are based upon LC-MS and NMR data. Annotations from previous studies were limited to those with NMR structural data.

Figure 4. *Solanum* species produce dramatically different leaf surface metabolite profiles. The seven representative species exhibit diverse metabolite profiles as demonstrated by base peak intensity LC-MS chromatograms. Most peaks correspond to acylsugars with varied sugar cores and acyl chain types and numbers. Acylsugar classes are listed for each species. Acylsugar annotations for these species are detailed in Tables S5, S8, S18, S21, and S25-S27.
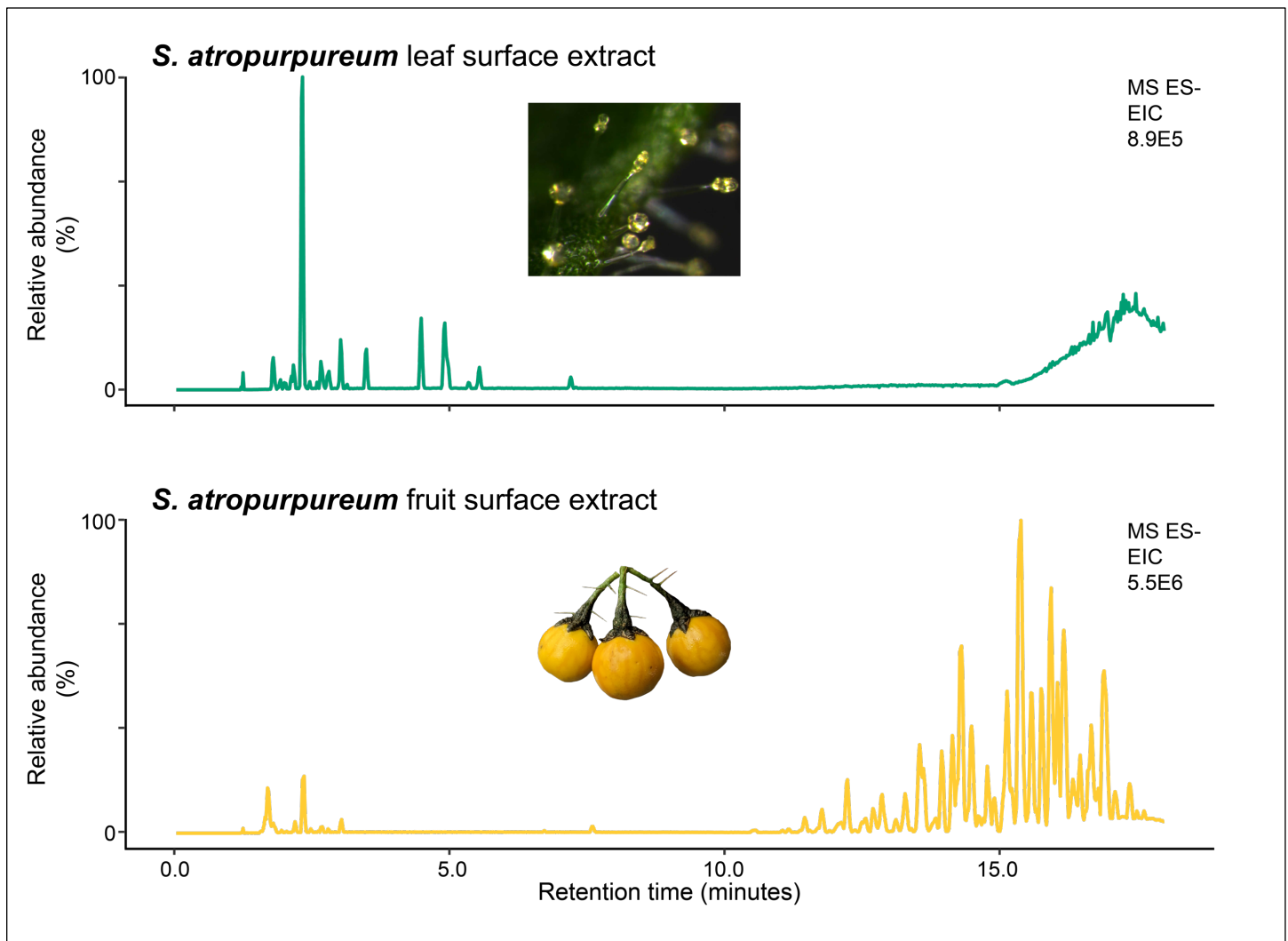
Figure 5. *S. atropurpureum* **trichomes and fruits exhibit dramatically different metabolite profiles.** The top LC-MS base peak intensity (BPI) chromatogram displays metabolites from a leaf surface extract, while the bottom LC-MS base peak intensity chromatogram displays metabolites from a fruit surface 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).
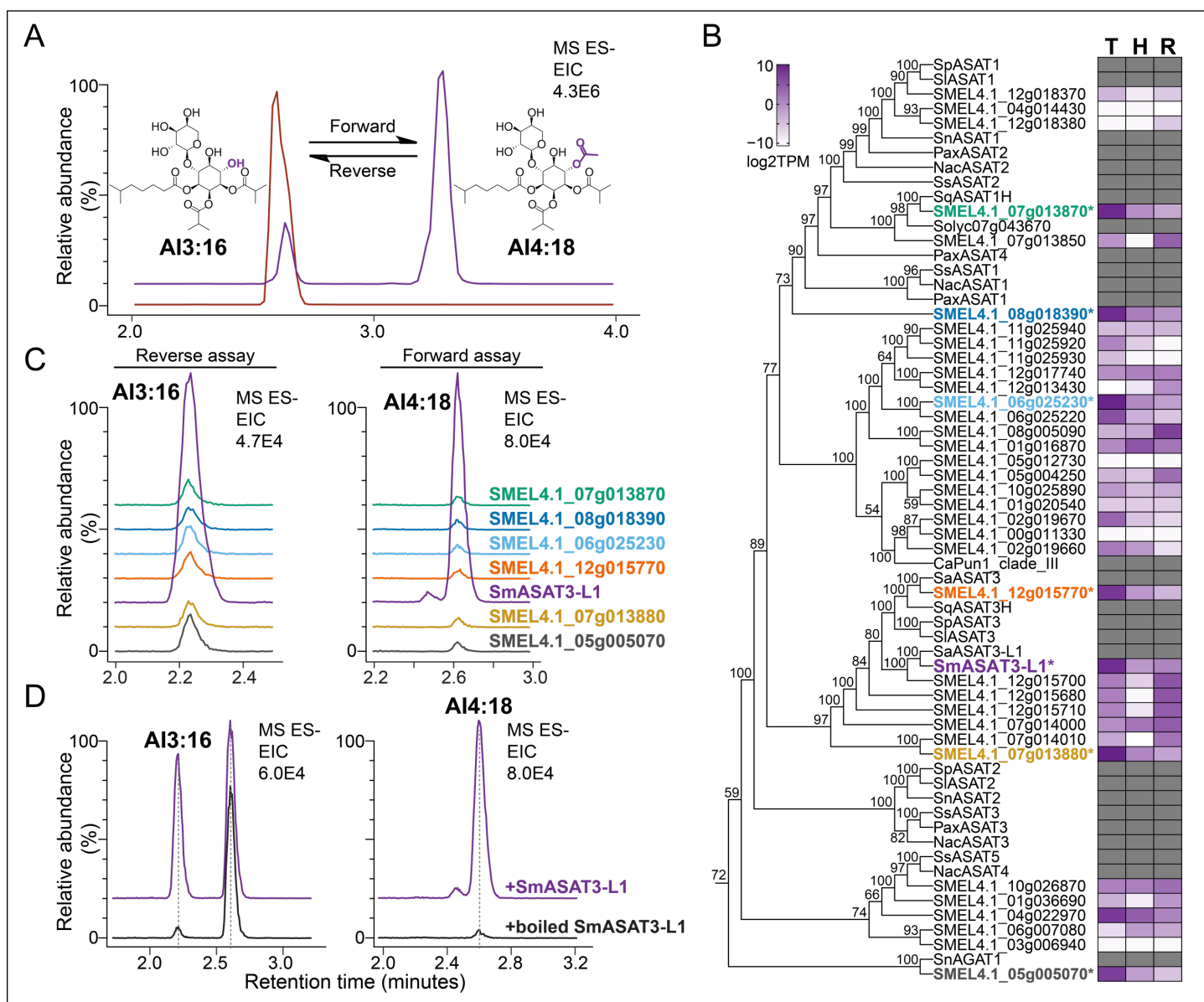
Figure 6. SmASAT3-L1 acetylates AI3:16 to produce AI4:18. Trichome expressed ASAT homologs were identified and tested for AI3:16 acetyl transferase activity. **(A)** LC-MS analysis of AI3:16(4,4,8) and AI4(2,4,4,8) in *S. aethiopicum* (red) and *S. melongena* (purple) leaf surface extracts. *S. aethiopicum* does not produce detectable levels of AI4:18. **(B)** Shown is a clade III BAHD tree subset from 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* (see Figure S12 for full phylogeny). 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. Values at nodes indicate bootstrap support calculated from 100,000 ultrafast bootstrap replicates. Heatmap shows absolute transcript abundance (log2 TPM) across trichomes, trichomeless hypocotyls, and roots from 7-

day-old eggplant seedlings. Expression data for non-eggplant sequences are not included in the heatmap. Color gradient provides a visual marker to rank the transcript abundance from high (purple) to low (white) or absent (grey). ASAT, acylsugar acyltransferase. log2TPM, log2 transformed transcripts per million. Genes denoted with "*" were selected for biochemical characterization based on expression in trichomes and sequence similarity to characterized ASAT enzymes in other species. **(C)** Of the seven enzymes tested *in vitro*, only SmASAT3-L1 successfully acetylated AI3:16(4,4,8) to form AI4:18(2,4,4,8). The extracted ion chromatograms on the left display products from reverse enzyme assays and the formate adduct of AI3:16, *m/z* 623.29. The extracted ion chromatograms on the right display products from forward enzyme assays and the formate adduct of AI4:18, *m/z* 665.30. **(D)** Acetylation of AI3:16 *in vitro* to produce AI4:18 requires functional SmASAT3-L1 enzyme. Forward assay extracted ion chromatograms display the formate adduct of AI4:18, *m/z* 665.30, and reverse assay extracted ion chromatograms display the formate adduct of AI3:16, *m/z* 623.29. The reverse assay chromatograms display peaks for both AI3:16 and AI4:18 due to in source fragmentation of unused AI4:18 substrate.
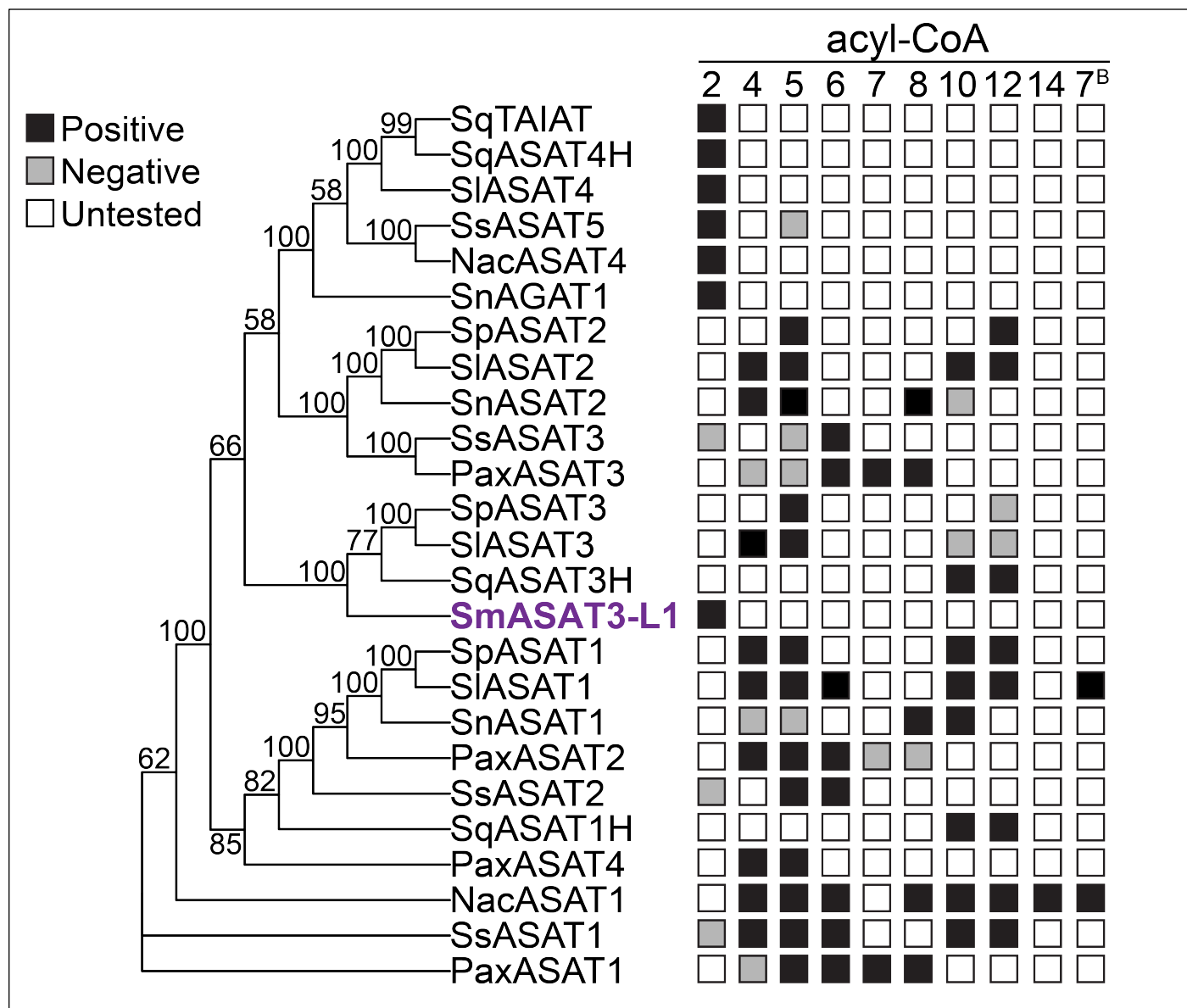
Figure 7. Characterized *in vitro* acyl-CoA usage by ASATs. The maximum likelihood phylogeny was inferred from amino acid sequences of characterized ASAT enzymes using a time-reversible algorithm specifying a plant-specific empirical substitution matrix, invariant sites, and four rate categories (Q.plant+I+G4) in IQ-TREE v2.1.3. Branch support was calculated from 100,000 ultrafast bootstrap replicates. Other than SmASAT3-L1, enzyme activities were described previously (Fan et al., 2017, 2015; Leong et al., 2022, 2020; Lou et al., 2021; Moghe et al., 2017; Nadakuduti et al., 2017; Schenck et al., 2022; Schilmiller et al., 2015, 2012). Reported positive activities are indicated by black squares; reported negative activities are indicated by gray squares; untested activities are indicated by white squares.