EVOLUTION OF ACYLSUGARS WITHIN THE *SOLANUM* GENUS

By

Paul D. Fiesel

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Biochemistry and Molecular Biology – Doctor of Philosophy

2023

ABSTRACT

Plants synthesize a remarkable number of lineage- and tissue-specific specialized metabolites.

These compounds exhibit diverse functions for plants, e.g., communication and defense, as well

as for humans, e.g., medicine and food. The anti-insect and anti-microbial acylsugars are one class

of specialized metabolites and accumulate in Solanaceae species. Despite being composed of the

simple building blocks of sugar cores and acyl chains, acylsugars exhibit incredible structural

diversity. This variation was previously demonstrated to impact plant pest mortality and

oviposition. These factors suggest that characterizing the acylsugar diversity within Solanaceae

species and understanding their biosynthesis can uncover how a biologically relevant trait has

evolved. While acylsucroses are the most well-characterized acylsugar type, unusual acylinositols

were characterized in three species of the large and megadiverse *Solanum* genus. In this study, the

diversity and distribution of *Solanum* genus acylinositols were characterized and their

biosynthetic pathway was investigated. I first characterized the trichome acylsugars of Clade II

species *Solanum melongena* (brinjal eggplant) using liquid chromatography-mass spectrometry

(LC-MS), gas chromatography (GC)-MS and nuclear magnetic resonance (NMR) spectroscopy,

identifying eight unusual structures with inositol cores, inositol glycoside cores, and hydroxyacyl

chains. LC-MS analysis of 31 *Solanum* DulMo clade, VANAns clade, and Clade II species

revealed striking acylsugar diversity with some traits restricted to specific clades and species.

Acylinositols were found in all three major clades while acylglucoses were restricted to the

DulMo and VANAns species characterized. Unusual disaccharide sugar cores and medium-

length hydroxyacyl chains were found to be widespread within the surveyed species. This

investigation revealed inositol sugar cores as a predominant sugar core type and prompted an

investigation into their biosynthesis. Utilizing an eggplant tissue-specific transcriptome and *in*

vitro biochemistry, an acetyltransferase ACYLSUGAR ACYLTRANSFERASE 3-LIKE 1 (SmASAT3-L1) was characterized to act upon a triacylinositol glycoside. Analysis of *S. melongena* triacylinositol biosynthesis uncovered an *in vitro* pathway producing a triacylinositol identical to a plant triacylinositol, however, production of the correct products only occurred when accompanied by nonenzymatic acyl chain rearrangement. Using this pathway knowledge and previously developed transcriptomes and gene silencing methods, I determined that two other acylinositol-producing species, *Solanum quitoense* and *Solanum nigrum*, contain an analogous acylinositol biosynthetic pathway. These results support the hypothesis that there is a conserved pathway within two major *Solanum* clades, DulMo and Clade II, which evolved in part due to gene duplications and altered substrate specificity. This study not only highlights the enormous amount of plant chemical diversity but also the usefulness of comparative biochemistry to uncover evolutionary mechanisms underlying metabolic novelty.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

CHAPTER 1:

FRUITY, STICKY, STINKY, SPICY, BITTER, ADDICTIVE, AND DEADLY:
EVOLUTIONARY SIGNATURES OF METABOLIC COMPLEXITY IN THE
SOLANACEAE

Works presented in this chapter have been published and are reproduced with permission from

the Royal Society of Chemistry:

*Fiesel, P. D., *Parks, H. M., Last, R. L., and Barry, C. S. (2022). Fruity, sticky, stinky, spicy,

bitter, addictive, and deadly: evolutionary signatures of metabolic complexity in the Solanaceae.

Nat. Prod. Rep., 2022, **39**, 1438-1464. DOI: 10.1039/D2NP00003B

*These authors contributed equally.

Abstract

Plants collectively synthesize a huge repertoire of metabolites. General metabolites, also referred to as primary metabolites, are conserved across the plant kingdom and are required for processes essential to growth and development. These include amino acids, sugars, lipids, and organic acids. In contrast, specialized metabolites, historically termed secondary metabolites, are structurally diverse, exhibit lineage-specific distribution and provide selective advantage to host species to facilitate reproduction and environmental adaptation. Due to their potent bioactivities, plant specialized metabolites attract considerable attention for use as flavorings, fragrances, pharmaceuticals, and bio-pesticides. The Solanaceae (Nightshade family) consists of approximately 2700 species and includes crops of significant economic, cultural, and scientific importance: these include potato, tomato, pepper, eggplant, tobacco, and petunia. The Solanaceae has emerged as a model family for studying the biochemical evolution of plant specialized metabolism and multiple examples exist of lineage-specific metabolites that influence the senses and physiology of commensal and harmful organisms, including humans. These include, alcohols, phenylpropanoids, and carotenoids that contribute to fruit aroma and color in tomato (fruity), glandular trichome-derived terpenoids and acylsugars that contribute to plant defense (stinky & sticky, respectively), capsaicinoids in chili-peppers that influence seed dispersal (spicy), and steroidal glycoalkaloids (bitter) from *Solanum*, nicotine (addictive) from tobacco, as well as tropane alkaloids (deadly) from Deadly Nightshade that deter herbivory. Advances in genomics and metabolomics, coupled with the adoption of comparative phylogenetic approaches, resulted in deeper knowledge of the biosynthesis and evolution of these metabolites. This review highlights recent progress in this area and outlines opportunities

for – and challenges of-developing a more comprehensive understanding of Solanaceae metabolism.

The Solanaceae: a phylogenetic framework for exploring metabolism

Metabolism is a window into micro- and macro-evolutionary processes. Plant metabolic diversity is vast and collectively plants are hypothesized to synthesize $\sim10^6$ metabolites (Afendi et al., 2012). Many of these metabolites, including sugars, amino acids, fatty acids, and organic acids – referred to as general or primary metabolites – are conserved across the plant kingdom, and essential for growth and development. However, specialized metabolites (SM), also referred to in the literature as secondary metabolites, comprise the majority of plant metabolic complexity. Specialized metabolites are chemically diverse, display taxonomically restricted distribution, and are often synthesized in individual tissues or cell types. Plants evolved the capacity to synthesize specific classes of specialized metabolites to facilitate ecological adaptations. The advent of genomics, coupled with the ability to test the function of candidate genes in host species or heterologous systems, advanced our understanding of the biosynthesis and evolution of plant specialized metabolism (Fossati et al., 2014; Lau and Sattely, 2015; Nett et al., 2020).

Although plant specialized metabolites exhibit considerable chemical complexity, they are ultimately derived from a pool of general metabolites formed through photosynthesis, glycolysis, the TCA cycle, amino acid metabolism, and the MEP-pathway (Vogt, 2010). General metabolites undergo transformations, including ligation and cyclization to generate scaffold molecules that are modified by glycosylation, acylation, methylation, prenylation, oxidation, and reduction to dramatically increase chemical complexity. In plants, the formation of these scaffold molecules and their subsequent decorations are catalyzed by large enzyme families formed by

repeated gene duplication followed by subfunctionalization, neofunctionalization, and gene loss to ultimately produce lineage-specific metabolites. The evolutionary mechanisms that create SM diversity are numerous but include co-option of general metabolism enzymes, evolution of catalytic promiscuity, enzyme compartment switching, the formation of biosynthetic gene clusters, and gene expression changes (Akiyama et al., 2021b; Itkin et al., 2013; Leong and Last, 2017; Schenck and Last, 2020; Sonawane et al., 2020). These evolutionary processes occur across different taxonomic scales, including inter-specific and intra-specific, to generate the chemical variation observed across the plant kingdom.

The Solanaceae, or nightshade family, contains approximately 2700 documented species found on six continents, which collectively have evolved morphological and metabolic adaptations for nearly every environment (Särkinen et al., 2013). A single genus – the *Solanum* – accounts for nearly half of these species (Gagnon et al., 2022). Nightshades grow in environments ranging from deserts to rainforests, with growth habits that vary from epiphytes to trees. The family includes four major food crops (potato, tomato, pepper, and eggplant), a host of minor food crops (including tomatillo, naranjilla, tamarillo, and groundcherry) as well as the several ornamental crops (including petunia, salpiglossis, schizanthus, and brugmansia) and weed species (Jimson weed and bittersweet). In addition, several Solanaceae species are grown for their narcotic or medicinal properties (tobacco, corkwood tree, deadly nightshade, henbane, and *Datura* species).

The Solanaceae family has become a model system for investigating biodiversity. The Solanaceae community concept was proposed nearly two decades ago, with the idea of using the nightshade family to connect genomics and biodiversity (Knapp et al., 2004). This concept envisioned harnessing Solanaceae natural diversity for evolutionary studies by creating the

necessary network of resources. One important tool was a detailed understanding of Solanaceae

phylogenetic relationships (https://www.solanaceaesource.org). This framework provides a basis

for evolutionary studies within the family. In parallel, the community-driven releases of the first

tomato and potato genomes created a genomic foundation. These successful projects spawned

numerous additional projects (*e.g.*, SOL-100, Varitome Project, 100 Tomato Genomes Project),

resulting in chromosome-scale genome assemblies draft genomes, pan-genomes, resequencing of

numerous wild tomato species and cultivars, and an online database for genetic resources

(Alonge et al., 2020; Barchi et al., 2021; Gao et al., 2019; Mueller et al., 2005; Sato et al., 2012;

Song et al., 2019; Xu et al., 2011). As of early 2022, genome sequences are available for more

than 30 Solanaceae species (https://plabipd.de/), and it seems likely that many more will follow

over the next few years.

These genomic tools are augmented by the availability of comprehensive germplasm

resources, particularly for the major crop species of the Solanaceae. These resources allow

genetic analysis of phenotypes of interest, facilitate genotype to phenotype comparisons and

allow exploration of natural phenotypic diversity. The pioneering work of Charles Rick – and

creation of seed stock centers (e.g., GRIN-Global and C. M. Rick Tomato Genetics Resource

Center) provide access to crop and wild relative germplasm. Notably, connecting genotype to

phenotype within tomato has been greatly accelerated by the development of the introgression

lines (ILs) and backcrossed introgression lines (BILs) of wild tomato S. pennellii within a

cultivated tomato background (Eshed and Zamir, 1995; Ofner et al., 2016). These ILs and BILs

were instrumental in discovering genes underlying multiple phenotypes, including those related

to metabolism (Fridman et al., 2004; Ofner et al., 2016; Schilmiller et al., 2012; Toal et al.,

2018). In addition, the ability to perform RNA interference (RNAi), virus-induced gene silencing

(VIGS), and CRISPR/Cas9 tools in multiple Solanaceae species allows the functional

characterization of candidate genes and a more precise connection of genotype and phenotype

(Brooks et al., 2014; Liu et al., 2002; Schijlen et al., 2007; Van Eck, 2018).

The Solanaceae has emerged as a model system for investigating the biosynthesis and

evolution of specialized metabolism (Figure 1). Members of the family have evolved to

synthesize several classes of bioactive and lineage-specific specialized metabolites, including

phenylpropanoids, acylsugars, terpenes and distinct groups of alkaloids (Figure 2). These

specialized metabolites are of interest because they influence fruit aroma and quality and are of

potential use as biopesticides and pharmaceuticals. The development of genomic resources,

coupled with the ability to survey metabolite variation across diverse germplasm, and to place

the resulting data within a phylogenetic context, enabled elucidation of the biosynthesis and

evolutionary trajectories of several major classes of Solanaceae SMs.
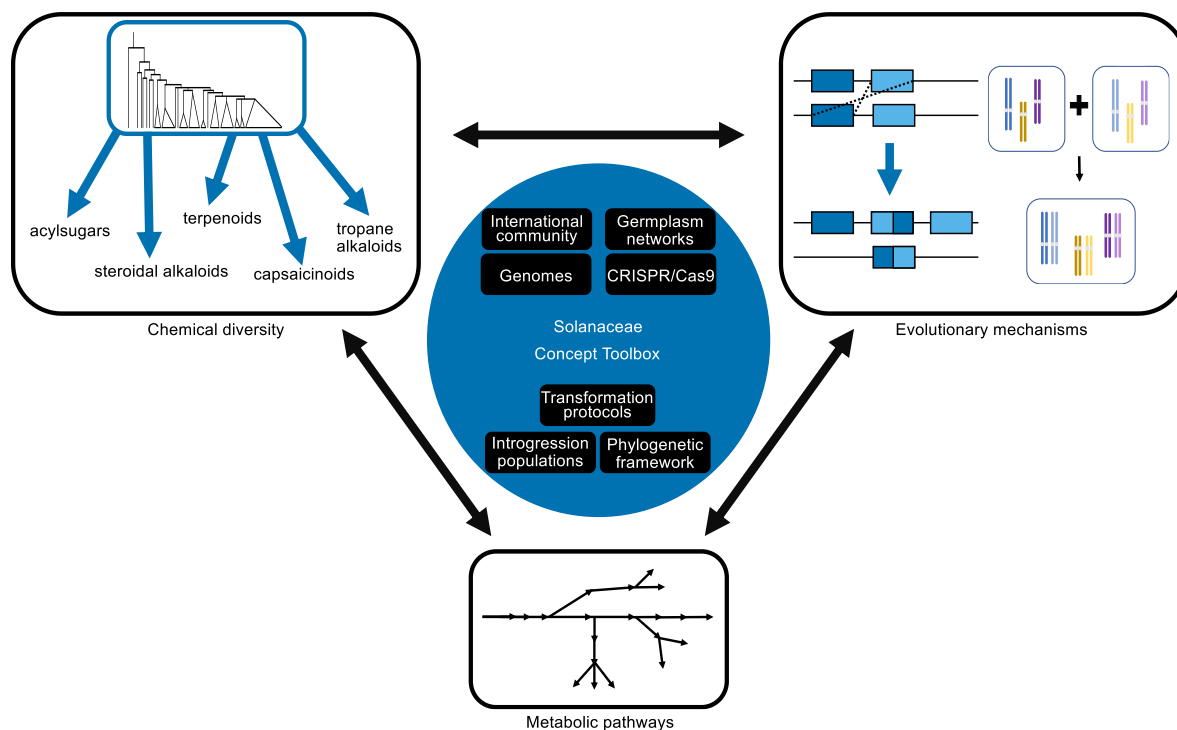
Figure 1.1. Solanaceae as a model family for specialized metabolism evolution studies. The
Solanaceae concept toolbox connects biodiversity, genetics, and evolutionary mechanisms to

6

Figure 1.1. (cont'd)
each other. Chemical diversity informs metabolic pathway discovery, which in turn reveals evolutionary mechanisms underlying chemical diversity.
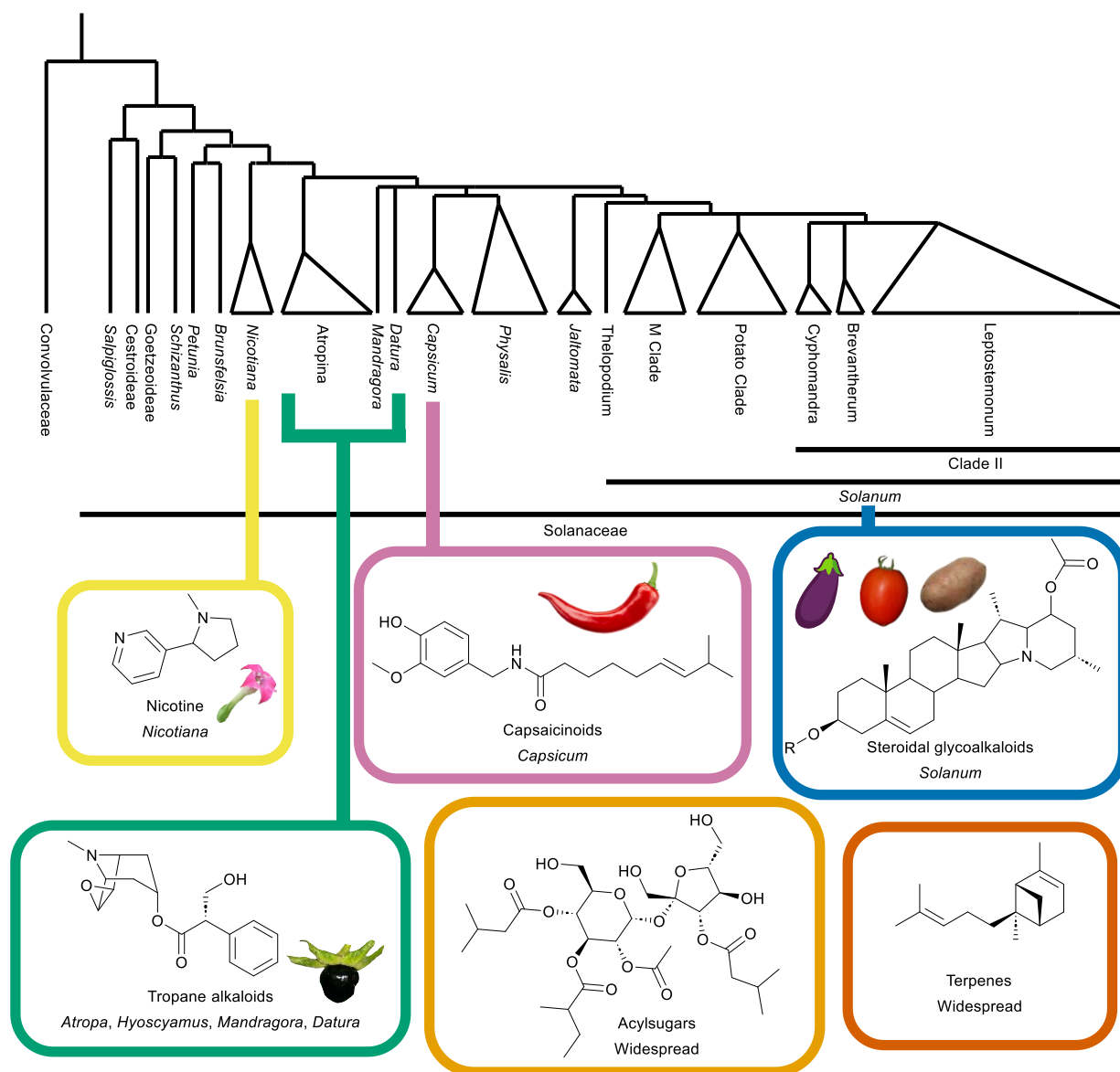
Figure 1.2. Phylogenetic distribution of major Solanaceae specialized metabolite classes. The Solanaceae family produces specialized metabolites of multiple chemical classes. A simplified phylogeny of the Solanaceae family is shown based on prior determination of phylogenetic relationships (Gagnon et al., 2022; Särkinen et al., 2013). Major metabolite classes are mapped to the corresponding clades that produce high amounts of those metabolites and/or act as model species for studying their biosynthesis and evolution. Metabolites may not be distributed solely in the noted phylogenetic group. Additional information on metabolite distribution is provided throughout the text of this article.

Fruity: GWAS-enabled discovery of aroma variation during ripening

The ripening of fleshy fruits is an agriculturally- and ecologically- important developmental process that makes fruits palatable and facilitates seed dispersal. Although fleshy fruits are highly diverse in morphology and flavor, ripening generally involves cell wall disassembly and associated softening, the conversion of starch into sugars, changes in color, and the biosynthesis of aroma volatiles. Fruit flavor and aroma is a complex species-specific quantitative trait involving the interaction between GM pathways, such as those influencing the accumulation of sugars and organic acids, as well as multiple SM pathways that yield aroma volatiles (Tieman et al., 2017). Tomato is the long-standing model crop species for investigating ripening mechanisms, including flavor and aroma biosynthesis.

Recent progress in understanding the genetic and biochemical basis of tomato flavor was facilitated by large-scale genome sequencing and resequencing projects involving hundreds of phenotypically diverse cultivated tomato accessions and wild relatives. These studies revealed insights into the nature of the tomato pan-genome and sequence variation associated with crop domestication and improvement, including gene duplication, single nucleotide polymorphisms, insertion–deletions, and large-scale structural variants (Alonge et al., 2020; Gao et al., 2019; Tieman et al., 2017; Zhu et al., 2018). The development of these resources facilitates the identification of genetic variation underlying phenotypic traits via genome-wide association studies. Notably, this approach was successfully deployed for the identification of genetic components underlying variation in tomato fruit flavor and aroma, revealing how human selection for visible traits such as fruit size, yield, and color can lead to alternative outcomes and unintentionally influence SM pathways that contribute to fruit quality.

Several hundred volatiles are detectable in ripening tomato fruits, but consumer taste panels identified 33 metabolites associated with consumer liking and 37 correlated with flavor intensity (Tieman et al., 2017). These influential aroma volatiles are derived through diversion of general metabolites, including carotenoids, phenylalanine, isoleucine/leucine, and fatty acids into diverse SM pathways. Genetic variation is evident across tomato varieties and 13 fruit aroma volatiles are significantly reduced in a collection of 48 modern cultivars when compared to 236 heirloom tomato varieties. This work shows that breeding of modern varieties for traits such as yield, shelf-life, and disease resistance has inadvertently and negatively altered SM pathways that produce aroma volatiles associated with consumer preference (Tieman et al., 2017). Subsequent GWAS analyses performed using a panel of 398 diverse tomato accessions analyzed for 27 volatiles along with glucose, fructose, malic acid, and citric acid revealed the existence of 251 association signals for 20 traits, including 15 correlated with aroma volatile production.

Among these associations are five loci that influence the production of carotenoid-derived volatiles. Two loci specifically influence the production of geranylacetone, which is formed by oxidative cleavage of the minor tomato fruit carotenoids phytoene, phytofluene, ζ-carotene, and neurosporene. A single locus specifically influences 6-methyl-5-hepten-2one (MHO) accumulation, which is derived from lycopene, the main carotenoid pigment in red-fruited tomato varieties. Two additional loci are associated with the production of both geranylacetone and MHO. Analysis of allele frequencies at these loci indicate that genetic complexity was progressively lost during breeding to the point where essentially only two allele combinations associated with accumulation of both volatiles persist in most modern cultivars. Analysis of MHO levels in genotypes with distinct allele combinations revealed that, as breeders selected for high lycopene in red-fruited varieties, they inadvertently selected favorable alleles

that increase MHO production. In contrast, the favorable alleles that promote geranylacetone accumulation are absent in modern cultivars (Tieman et al., 2017).

GWAS also revealed the identity of loci important for producing lipid and phenylalanine-derived volatiles. Ripening tomato fruit accumulate C5 and C6 volatiles derived from the breakdown of linolenic and linoleic acid, which are released from glycerolipids such as triacylglycerol. GWAS analyses of the panel of 398 tomato accessions described above identified a chromosome 9-localized SNP that is significantly associated with the fatty acid derived volatiles Z-3-hexen-1-ol and hexyl alcohol (X. Li et al., 2020). This SNP lies within a metabolic QTL region known to influence lipid content in tomato fruit (Garbowicz et al., 2018). Solyc09g091050 (Sl-LIP8) was identified as a candidate gene close to this SNP and gene expression analysis revealed that accessions possessing the reference allele from the Heinz 1706 variety had increased levels of Z-3-hexen-1-ol and hexyl alcohol together with elevated Soly09g091050 transcripts. Confirmation that Sl-LIP8 is responsible for lipid-derived volatile synthesis was achieved through CRISPR/Cas9 gene editing and in vitro biochemical assays. The knock-out mutants showed reductions in two C5 (1-pentanol and 1-penten-3-ol) and three C6 (Z-3-hexen-1-ol, E-2-hexen-1-ol, and hexyl alcohol) volatiles, while the recombinant enzyme catalyzed release of fatty acids from various glycerolipids (X. Li et al., 2020). The resultant free fatty acids undergo peroxidation at either the C9 or C13 positions in reactions catalyzed by 9-lipoxygenases and 13-lipoxygenases, respectively to yield aroma volatiles.

The phenylalanine-derived volatiles guaiacol, eugenol, and methylsalicylate contribute to the aroma of tomato fruits and are associated with smoky and medicinal-like aromas, which are often negatively correlated with consumer liking (Zanor et al., 2009). Guaiacol, eugenol, and methylsalicylate accumulate in tomato fruits as diglycosides, and cleavage of the glycoside

groups leads to release of the volatiles in "smoky" cultivars. In contrast, in "non-smoky"
varieties these metabolites exist as non-cleavable triglycosides resulting in reduced levels of
volatile release (Tikunov et al., 2013). Formation of guaiacol, eugenol, and methylsalicylate
triglycosides from their diglycoside precursors is catalyzed by the UDP-glucosyltransferase
enzyme, NON-SMOKY GLYCOSYLTRANSFERASE1 (NSGT1). The NSGT1 gene resides at
a locus on chromosome 9 that contains a second gene designated NSGT2. Both genes contain
structural changes in "smoky" cultivars that are predicted to render them non-functional although
the exact structure of the locus was unresolved (Tikunov et al., 2013).

The recent development of 14 new reference tomato genomes assembled using Oxford
Nanopore long read sequencing technology allowed the genome structure flanking the NSGT1
locus to be resolved. Five haplotypes were identified revealing evidence of intraspecific gene
duplication and loss at an SM locus that was selected during crop improvement (Alonge et al.,
2020). Haplotype I is proposed to be ancestral and contains predicted functional copies of
NSGT1 and NSGT2. All other haplotypes contain coding sequence mutations in NSGT2. In
addition, haplotypes IV and V also lack functional copies of NSGT1 and are therefore null
mutations for both NSGT1 and NSGT2. Analysis of guaiacol levels across two GWAS panels
and within an F2 population segregating for haplotype V and a functional copy of NSGT1
demonstrated that fruit guaiacol levels are reduced in individuals that contain a functional copy
of NSGT1. Together, these data illustrate the combined power of genome sequences developed
using long-read sequencing data and GWAS to investigate the evolution of loci associated with
SM phenotypes, particularly when the variation is mediated by tandem gene duplication that may
be unresolved in genome assemblies derived from short-read data. Overall, these studies

represent an example of fundamental science that provides opportunities to breed tomato varieties with favorable aroma volatile alleles.

Sticky: single-cell biochemical genetics reveals acylsugar metabolic complexity

Acylsugars are specialized metabolites produced in numerous plant families including the Solanaceae, Convolvulaceae, Geraniaceae, Martyniaceae, Rosaceae, Brassicaceae, and Caryophyllaceae (Asai et al., 2011, 2010; Asai and Fujimoto, 2011, 2010; Bah and Pereda-Miranda, 1996; Maldonado et al., 2006; Moghe et al., 2017; Ono et al., 2015; Pereda-Miranda et al., 1993; Wu et al., 2013). Many species across the Solanaceae produce acylsugars in hair-like Type I- and IV-glandular trichomes, while some species are documented to accumulate acylsugars in fruit pericarp or root exudates (Korenblum et al., 2020; Li et al., 2014; Maldonado et al., 2006; Nakashima et al., 2016). Acylsugars are composed of a sugar core, most commonly sucrose, and various fatty acids esterified to the core (Figure 1.3). Despite these simple components, variations in acylation position, chain length, chain branching pattern, and sugar core can result in hundreds of chromatographically separable acylsugars in a single species (Moghe et al., 2017). Solanaceae acylsugars are the most extensively characterized acylsugar type with more than 100 distinct NMR-resolved chemical structures (Bernal et al., 2018; Cao et al., 2015; Chortyk et al., 1997; Cicchetti et al., 2018; Ghosh et al., 2014; Hurney, 2018; Liu et al., 2017; Lou et al., 2021; Lybrand et al., 2020; Maldonado et al., 2006; C.-R. Zhang et al., 2016; C.-Y. Zhang et al., 2016). Acylsugars defend against microbes and insects; for example, deterring whitefly oviposition (Leckie et al., 2016), aphid settling (Goffreda et al., 1989), fungal growth (Luu et al., 2017), and mediating an ant-hornworm-tobacco interaction (Weinhold and Baldwin, 2011).

A

- 🟢 Sucrose sugar core
- 🔵 Glucose sugar core
- 🟡 Inositol sugar core

Convolvulaceae
Salpiglossis
Cestroideae
Goetzeoideae
Schizanthus
Petunia
Brunsfelsia
Nicotiana
Atropina
Mandragora
Datura
Capsicum
Physalis
Jaltomata
Thelopodium
M Clade
Potato Clade
Cyphomandra
Brevantherum
Leptostemonum

Clade II

Solanum

Solanaceae

B

S5:21(2,2,ai5,ai6,ai6)
Salpiglossis sinuata

S4:17(2,i5,i5,ai5)
Solanum lycopersicum

G3:19(i4,ai5,i10)
Solanum pennellii

I4:24(2,2,10,10)
Solanum quitoense

S5:25(ai5,ai5,ai5,ai7,malonyl)
Petunia axillaris

S3:19(i4,ai5,i10)
Solanum pennellii

G3:19(2,i4,i10)
Solanum nigrum

I3:17(i4,ai5,8)
Solanum nigrum

Figure 1.3. Phylogenetic distribution of acylsugar core types. (A) Simplified Solanaceae phylogeny with acylsugar core type placed on each lineage with characterized acylsugars. The phylogenetic tree is based upon previously published Solanaceae and *Solanum* trees (Gagnon et

13

Figure 1.3. (cont'd)
al., 2022; Särkinen et al., 2013). (B) Characteristic acylsugar structures produced by Solanaceae species (Chortyk et al., 1997, 1993; Ghosh et al., 2014; Hurney, 2018; King et al., 1986; Liu et al., 2017; Lou et al., 2021; Lybrand et al., 2020; Maldonado et al., 2006; Moghe et al., 2017). Acylsugar nomenclature is given for each compound where the first letter represents the sugar core (S for sucrose, G for glucose, I for inositol); the first number represents the number of acylations; the number after the colon represents the number of carbons in acyl chains; and the individual acyl chains are listed inside parentheses (ai = anteiso, i = iso).

Harnessing acylsugar genotypic diversity for tomato pathway determination

Tomato acylsugar diversity was employed to uncover the acylsugar biosynthesis pathway within cultivated tomato, *S. lycopersicum*. Analysis of *S. lycopersicum* introgression lines carrying *S. pennellii* chromosomal segments was instrumental in identifying loci required for acylsugar biosynthesis (Schilmiller et al., 2010, 2012). The identification and subsequent validation of candidate genes was facilitated by trichome-specific transcriptome, *in vitro* enzyme assays, and *in vivo* gene VIGS knockdown and CRISPR/Cas9 knockout. These approaches uncovered the core acylsugar pathway in *S. lycopersicum* glandular trichomes. A series of evolutionarily related BAHD acyltransferases, named AcylSucrose AcylTransferase 1–4 (ASAT1-4), acylate sucrose sequentially to produce tetraacylsucroses consisting of acyl chains at R_2, R_3, R_4, and $R_{3'}$ (Fan et al., 2015; Schilmiller et al., 2015, 2012) (Figure 1.4). Each enzyme selectively acylates specific sucrose hydroxyls with varying promiscuity for acyl-CoA substrates. Documenting this pathway enabled discovery of mechanisms responsible for acylsugar diversity in wild tomato relatives.
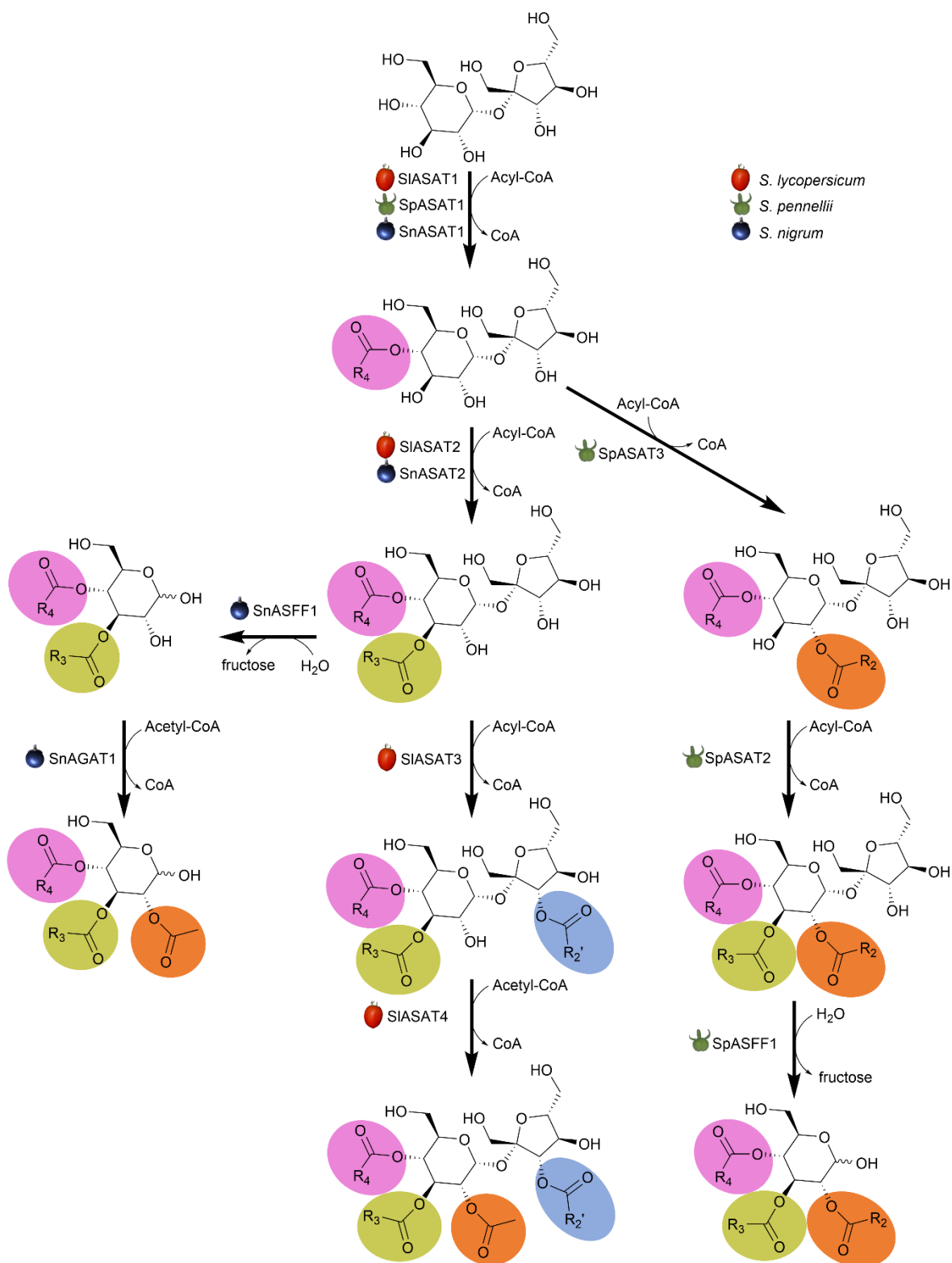
Figure 1.4. Acylsucrose and acylglucose pathway diversity in *Solanum* species. The acylsucrose and acylglucose biosynthesis pathways for *S. nigrum*, *S. lycopersicum* and *S. pennellii*. All three biosynthetic pathways begin by acylating sucrose (Fan et al., 2015; Leong et

Figure 1.4. (cont'd)

al., 2019; Lou et al., 2021; Schilmiller et al., 2015, 2012). Sequential acylations produce tetraacylsucroses, triacylsucroses, and diacylsucroses for *S. lycopersicum*, *S. pennellii*, and *S. nigrum*, respectively. *S. pennellii* triacylsucroses and *S. nigrum* diacylsucroses are cleaved by ASFF enzymes to form triacylglucoses and diacylglucoses, respectively (Leong et al., 2019; Lou et al., 2021). *S. nigrum* diacylglucose is acetylated by SnAGAT1 to form a triacylglucose (Lou et al., 2021). ASAT, acylsucrose acyltransferase; AGAT, acylglucose acyltransferase; ASFF, acylsugar fructofuranosidase; CoA, CoenzymeA.

Intra- and inter-specific differences in tomato acylsugar structures result in part from differing ASAT activities. Comparative biochemical analysis of cultivated and wild tomato ASAT sequences uncovered amino acid residues responsible for specific activity differences. For example, the comparison of ASAT2 sequences and *in vitro* enzyme activities across tomato species revealed two mutations that impact acyl-CoA specificity. Residues Val/Phe408 and Ile/Leu44 influence the ability to use the structurally similar iC5-CoA and aiC5-CoA, respectively, without altering activity with nC12-CoA (Fan et al., 2015). Comparison of *S. lycopersicum* and *S. habrochaites* ASAT3 homologs revealed a Tyr/Cys41 residue change impacting the enzyme's ability to use nC12-CoA (Schilmiller et al., 2015). Characterization of *S. habrochaites* ASAT4 in accessions collected from Ecuador to Southern Peru revealed variations in acetylation patterns that were explained either by changes in ASAT4 expression or coding sequence mutations (Kim et al., 2012; Landis et al., 2021). The comparative biochemistry approach revealed differences in enzyme acyl donor specificity, which impacted acylsugar phenotypes. This approach also determined evolutionary changes in enzyme acyl acceptor specificity.

S. pennellii LA0716 produces acylsucroses through a 'flipped pathway', resulting from changes in ASAT acyl acceptor specificity (Fan et al., 2017). While cultivated tomato produces acylsucroses with one furanose ring acylation (termed F-type acylsucroses), *S. pennellii* and some *S. habrochaites* accessions synthesize acylsucroses acylated exclusively on the pyranose

ring (Schilmiller et al., 2015). These 'P-type' acylsucroses are produced by alternate ASAT2 and ASAT3 homologs, which catalyze the third and second pathway steps, respectively. The published results suggest that *S. pennellii* ASAT2 likely evolved from an ancestral enzyme capable of acylating both mono- and diacylsucrose. Analogous sequence changes in ASAT3, potentiated by ASAT3 duplication, resulted in the neofunctionalized ASAT3 duplicate found in *S. habrochaites* and *S. pennellii*. This study revealed a remarkably small number of amino acid changes that caused a major change in pathway structure and product phenotypes in closely related species.

The flipped *S. pennellii* pathway and recruitment of an invertase-like enzyme appear to have potentiated evolution of *S. pennellii* acylglucose synthesis (Figure 1.4). *S. pennellii* acylglucoses are synthesized from P-type acylsucroses by a neofunctionalized glycoside hydrolase 32 family (GH32) beta-fructofuranosidase, SpASFF1 (Leong et al., 2019). The modified SpASFF1 substrate binding site correlates with a derived P-type acylsucrose cleavage activity, yet the neofunctionalized enzyme does not act on the F-type acylsucrose produced by *S. lycopersicum*. In addition, SpASFF1 lacks activity with sucrose, associated with changes to the canonical sucrose binding pocket. Instead, the modified SpASFF1 substrate binding site correlates with a derived P-type acylsucrose cleavage activity, yet the neofunctionalized enzyme does not act on the F-type acylsucrose produced by *S. lycopersicum*. SpASFF1 specificity for P-type acylsucroses supports the hypothesis that P-type acylsucroses are required for acylglucose production. Indeed, cultivated tomato lines engineered to contain both the flipped pathway and SpASFF1 accumulate acylglucoses. This indicates that acylglucose biosynthesis requires both a neofunctionalized invertase and the *S. pennellii* flipped pathway. Finally, CRISPR/Cas9 deletion of SpASFF1 led to accumulation of only acylsucroses – without detectable acylglucoses – in *S.*

pennellii, reinforcing that the neofunctionalized invertase is necessary for acylglucose synthesis in the wild tomato. SpASFF1 invertase is an example of co-option of general metabolic enzyme to specialized metabolism into acylsugar biosynthesis – in this case resulting in different sugar core composition.

The theme of GM enzymes recruitment to SM by gene duplication, changes in gene expression and enzyme structure and function also contribute to acyl chain type variation. For example, the duplicated and neofunctionalized isopropylmalate synthase gene, IPMS3, influences isoC5 acyl chain abundance (Ning et al., 2015). In contrast to the canonical Leu biosynthetic IPMS, IPMS3 expression is restricted to type I/IV glandular trichome tip cells, and the *S. lycopersicum* enzyme is insensitive to Leu-mediated feedback inhibition in vitro due to truncation of the C-terminal allosteric regulatory domain. Apparently, the lack of this domain frees the enzyme from Leu feedback regulation, enabling pathway diversion. IPMS3 allelic variation directly correlated with abundance of isoC5 and isoC4 acyl chains in wild *S. pennellii* accession acylsugars; accessions with majority isoC4 acyl chains were homozygous for a truncated, inactive IPMS3. In contrast, isoC5 acyl chains were abundant in accessions either heterozygous or homozygous for the unregulated IPMS3. These results reveal that acyl-CoA availability influences acylsugar acyl chain composition.

Further evidence for this hypothesis was provided by identification of natural chain diversity associated with allelic diversity of two acyl-CoA biosynthesis genes (Fan et al., 2020). These trichome-expressed genes, an enoyl-CoA hydratase (AECH1) and acyl-CoA synthetase (AACS1), reside in a gene cluster syntenic to the chromosomal region containing ASAT1. The Solanaceae family shares the syntenic region, which was likely derived from a Solanaceae-specific polyploidy event. Silencing AECH1 and AACS1 in *S. lycopersicum*, *S. pennellii*, and the

more distantly related *Solanum quitoense*, reduced or eliminated medium length (10–12 carbons) acyl chains from acylsugars. Additionally, the presence of AECH1 and AACS1 correlates with natural variation in medium acyl chains. For example, in the short chain producing genera *Petunia* and *Nicotiana*, AECH1 and AACS1 are either missing or present as pseudogenes. These genes represent another example of how evolutionary changes in metabolic machinery impacted acylsugar composition.

Genomics tools enable comparative biochemistry in non-model organisms

Application of DNA sequencing, modern analytical chemistry, and reverse genetic tools such as VIGS and genome editing enabled documentation of additional acylsugar evolutionary mechanisms in non-model species. LC-MS screening and NMR-resolved structural analysis identified Solanaceae species that produce unique acylsugars with varying cores, acylation positions, and chain types (Hurney, 2018; Leong et al., 2020; Liu et al., 2017; Lou et al., 2021; Lybrand et al., 2020; Moghe et al., 2017). For example, extant members of early-diverging lineages produce acylsucroses with acylation patterns undocumented in cultivated and wild tomatoes. Additionally, acylated glucoses are detected in some species within the *Petunia*, *Nicotiana*, *Datura*, and *Solanum* genera (Chortyk et al., 1997, 1993; King and Calhoun, 1988; Lou et al., 2021; Matsuzaki et al., 1989). Within the large *Solanum* genus, *myo*-inositol sugar cores have been documented in *S. lanceolatum*, *S. quitoense*, and *S. nigrum* (Herrera-Salgado et al., 2005; Hurney, 2018; Leong et al., 2020; Lou et al., 2021). Evolution of acylsugar biosynthesis was investigated in four non-model species: *Salpiglossis sinuata*, *Petunia axillaris*, *S. nigrum*, and *S. quitoense*. Comparison of the enzymes and pathways in each species revealed features of long-term and clade-specific acylsugar traits.

Inferring early events in acylsugar evolution

Investigations of two members of early diverging lineages, *S. sinuata* and *P. axillaris*, revealed acylsugar biosynthesis evolutionary changes occurring over tens of millions of years (Myr), well beyond the approximately 7 Myr of *Solanum* tomato clade history (Moghe et al., 2017; Nadakuduti et al., 2017; Särkinen et al., 2013). Despite similarity of acylation positions between tomato species, *S. sinuata* and *Petunia* acylsugars, a major shift occurred in the acylsugar biosynthetic pathway. The ancestral pathway found in *S. sinuata* and *P. axillaris* begins with a sucrose-acylating ancestral ASAT1, aASAT1, which is not found in tomato clade species. Another surprise is that the SlASAT1 and SlASAT2 orthologs, aASAT2 and aASAT3, respectively catalyze the second and third acylations. The first three acylations by the early evolving aASAT1-3 pathway produce triacylsucroses with the same three positions acylated as SlASAT1-3. Coinciding with this, aASAT2 and aASAT3 retained their selectivity for the R_4 and R_3 of sucrose, respectively, but shifted acyl acceptor specificity to free and monoacylsucrose, respectively. This activity shift correlates with aASAT1 loss in species with modern acylsugar biosynthesis pathways. Transcriptome and genome analyses suggest that the aASAT1 gene disappeared from the last common ancestor of the *Capsicum* and *Solanum* genera, ~15–20 MYA. Identification of these ancestral acylsugar pathways support sucrose as the ancestral acyl acceptor. From these studies of early-diverging Solanaceae species, ASAT gene loss and neofunctionalizations were implicated in a changing acylsucrose pathway, analogous to those described above in the case of the *S. pennellii* flipped acylsucrose pathway.

The ancestral and derived acylsucrose pathways provide insight into the evolutionary origins of acylsugars (Moghe et al., 2017). Lamiidae BAHD sequence homology, phylogenetics, and known whole genome duplication events all enabled inferences regarding early acylsugar

20

evolution. One hypothesis, based on sequence analysis, is that ASAT sequences derive from an alkaloid biosynthetic enzyme ancestor. Based on nonsynonymous mutation rates and historical polyploidy events, the clade containing ASAT1,2,3 appears to have arisen via an ancient whole genome duplication before the Solanaceae-Convolvulaceae split ($\sim$50–65 MYA). Subsequent duplications prior to, and following the Solanaceae polyploidization, led to evolution of the ASATs and paralogs found in the ASAT1,2,3 clade. As described above, our model of acylsugar biosynthetic pathway evolution invokes loss of aASAT1, refinement of ASAT1 and ASAT2 activities, and recruitment of ASAT3 occurred later in Solanaceae diversification.

Acylhexoses in non-model plants

Metabolite profiling revealed that, like *S. pennellii*, black nightshade (*Solanum nigrum*) also produces acylglucoses, an observation that enabled discovery of convergent and new acylsugar enzyme activities. *S. nigrum* creates di- and triacylglucoses through a similar, yet distinct, pathway when compared to *S. pennellii* acylglucose biosynthesis (Lou et al., 2021) (Fig. 4). Both pathways proceed through a series of sucrose acylations, followed by action of an acylsugar fructofuranosidase. The *S. nigrum* invertase, SnASFF1, and SpASFF1 enzymes share similarities including a modified DDTK sucrose binding pocket, loss of canonical invertase activity cleaving sucrose, and neofunctionalized activity with acylsucroses. However, each ASFF1 enzyme resides in a distinct glycoside hydrolase subfamily 32 clade and cleaves different substrates: triacylsucroses by SpASFF1 and diacylsucroses by SnASFF1.

Sn<u>A</u>cyl<u>G</u>lucose<u>A</u>cetyl<u>T</u>ransferase1, SnAGAT1, catalyzes the third *S. nigrum* acylation, marking yet another distinction between *S. nigrum* and *S. pennellii* triacylglucose biosynthesis; this is the only enzyme to acylate an acylglucose described to date. As the two characterized *Solanum*

acylglucose biosynthetic pathways include distinct invertases, it is plausible that this mechanism evolved in other acylglucose-producing genera.

In contrast to the detailed information available for acylsucrose and acylglucose biosynthesis, the pathway leading to acylinositol synthesis in the *Solanum* remains largely enigmatic. So far only one enzyme was demonstrated in acylinositol biosynthesis: the *S. quitoense* enzyme TriAcylInositolAcetylTransferase, SqTAIAT, acetylates triacylinositols to produce tetraacylinositols (Leong et al., 2020). SqTAIAT is the closest known *S. quitoense* homolog to the final enzyme in tomato acylsucrose biosynthesis, SlASAT4, indicating conservation of acetyltransferases across acylinositol and acylsucrose biosynthesis. Both enzymes acetylate triacylsugars differing in their sugar core. Similar enzymatic activity and high sequence similarity suggest a common evolutionary origin for acylinositol and acylsucrose biosynthesis. However, the initial steps of acylinositol biosynthesis remain unresolved. Further pathway elucidation in *S. quitoense* and *S. nigrum* may uncover the evolutionary innovations underlying acylinositol production.

Into the depths with acylsugars

It was recently shown that cultivated tomato accumulates acylsugars in roots and root exudates (Korenblum et al., 2020). Tomato root acylsugars structurally differ from those in trichomes, contrasting in acyl chain type, acyl chain number, and sugar core type. For example, six- and seven-carbon acyl chains and glucose sugar cores are only detected in the roots. These structural differences suggest evolutionary changes in the underlying biochemistry. One key observation is that characterized tomato trichome-expressed ASAT transcripts were not detected in root tissue, although they do express closely related homologs. These expression data suggest the hypothesis that roots produce acylsugars through an alternative pathway. In fact, expression

of two ASAT4 paralogs correlates with acylsugar abundance in roots. While the function of root acylsugars is unknown, different microbial communities systemically impacted root exudate acylsugar abundances (Korenblum et al., 2020). Investigating root acylsugar metabolism may unearth a root-specific acylsugar biosynthetic pathway among other tantalizing prospects.

Stinky: variations on a theme define terpene diversity across *Solanum*

Terpenoids are structurally diverse and are produced across all kingdoms of life, yet all are derived from the simple five-carbon isomers, dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP). These precursors are formed through either the mevalonate (MVA) or 2-C-methyl-D-erythritol-4-phosphate (MEP) pathways (Zhou and Pichersky, 2020a). Plants are unique in that they contain both the cytosolic MVA pathways and the plastid localized MEP pathway; having evolved to generate substantial flux towards DMAPP and IPP as well as create separate subcellular pools of these metabolites for different pathways (Zhou and Pichersky, 2020a). Terpenoids have diverse functions ranging from the production of photosynthetic pigments and ubiquinone in the electron transport chain to the production of several classes of plant hormones. However, most plant terpenoids are lineage-specific specialized metabolites with C10–C30 carbon skeletons that provide a fitness benefit to the host organism through signaling and defense (Zhou and Pichersky, 2020a).

Plant terpenoid diversity is created at multiple levels. Firstly, small gene families produce cis and trans-prenyltransferases that initially condense a single molecule of DMAPP and IPP to form either geranyl diphosphate (GPP) (trans isomer) or neryl diphosphate (NPP) (cis isomer). These C10 metabolites can then be extended by five carbon units, through condensation with additional units of IPP, to yield trans- or cis-farnesyl diphosphate (E,E-FPP or Z,Z-FPP, C15), geranylgeranyl or nerylneryl diphosphate (GGPP or NNPP, C20), or longer chain prenyl

diphosphates (Zhou and Pichersky, 2020a). Short-chain prenyl diphosphates (C10–C20) are

substrates for terpene synthases (TPS), which exist as moderately large gene families (up to

~100 members) and catalyze the formation of hydrocarbon terpene skeletons via rearrangements

and cyclization. TPS enzymes possess considerable catalytic potential. They frequently utilize

more than one substrate, and catalysis by a single enzyme often generates multiple products

(Karunanithi and Zerbe, 2019; Pazouki and Niinemets, 2016; Zhou and Pichersky, 2020a). These

hydrocarbon terpene skeletons are often functionalized by the addition of hydroxyl groups,

which provide targets for modifications such as epoxidation, methylation, acylation, and

glycosylation, ultimately generating the vast complexity of terpenoids observed across the plant

kingdom.

The availability of a high-quality reference genome assembly for cultivated tomato

(*Solanum lycopersicum*) facilitated what is likely the most comprehensive published catalogue of

terpene scaffold biosynthesis in plants. The data highlight considerable chemical complexity

with *in vitro* biochemical data revealing the potential to synthesize 53 known hydrocarbon

terpene scaffolds plus several unidentified products. These terpenes arise through combined

catalysis of seven cis-prenyltransferases and 10 trans-prenyltransferases that form C10, C15, and

C20 prenyl diphosphates, together with 34 functional TPS enzymes (Akhtar et al., 2013; Zhou

and Pichersky, 2020b). Consistent with the known catalytic promiscuity of TPS enzymes, many

of the tomato TPSs can utilize more than one substrate, particularly the sesquiterpene synthases

that use both *E*,*E*-FPP and *Z*,*Z*-FPP, and yield multiple products. In addition, considerable

catalytic redundancy exists. For example, eight distinct TPSs catalyze the formation of the

monoterpene β-myrcene. Individual CPT, TPT, and TPS enzymes are localized to the cytosol,

plastids, as well as mitochondria, and the corresponding genes are differentially expressed across

tomato tissues: this highlights the spatial separation of terpene synthesis modules across tomato. Metabolite profiling of 13 tomato tissues identified 29 out of 53 terpenes *in planta*, suggesting that some terpenes are either below the limit of detection in tomato grown under standard cultural conditions or are further modified to produce more structurally complex metabolites.

Genomic clustering is a key feature of terpene biosynthetic genes in plants (Boutanaev et al., 2015). These clusters generally consist of both paralogs and non-homologous genes encoding enzymes of terpene biosynthesis, creating a reservoir for the evolution of chemical novelty and facilitating the inheritance of SM modules that promote plant adaptation. Gene duplication within these clusters is often followed by pseudogenization and gene loss to create additional chemical variation. The majority of the 52 TPS loci in tomato, including 18 predicted pseudogenes, are located within gene clusters dispersed across the genome (Zhou and Pichersky, 2020b). In addition, the TPS gene clusters on chromosomes 6, 7, 8, and 12 also contain combinations of cis or trans prenyltransferases, cytochromes P450, methyltransferases, acyltransferases, and glycosyltransferases (Matsuba et al., 2013; Zhou and Pichersky, 2020b). While most of the potential terpene modifying enzymes within these clusters await functional characterization, a three-gene subcluster on chromosome 8 comprising SlTPS21-CYP71D51-SlCPT2 was demonstrated to synthesize (+)-lycosantalonol from NNPP (Zi et al., 2014).

Along with the existence of the 18 TPS pseudogenes in the tomato genome, three TPS-related gene clusters on chromosomes 6, 8, and 12 also contain inactive cytochromes P450 genes (Zhou and Pichersky, 2020b). The high prevalence of pseudogenes within these tomato terpene biosynthetic gene clusters suggests that there is potential for considerable genetic variation. For example, a gene that is pseudogenized in one accession or species may be functional in another. Thus, variation in terpene-related gene clusters may exist between distinct accessions of *S.*

lycopersicum but also more likely across the genomes of diverse Solanaceae species. The increasing availability of high-quality chromosome scale reference genomes assembled from long-read sequencing will facilitate identification of additional gene clusters and future comparative evolutionary analysis of terpene biosynthesis across the Solanaceae.

Within the *Solanum* genus, distinct evolutionary trajectories associated with trichome-derived terpene-related gene clusters are indeed apparent between cultivated tomato and wild relatives that diverged from a common ancestor approximately two-three million years ago (Särkinen et al., 2013). Notably, while limited terpene diversity exists in trichomes between cultivated tomato accessions, considerable variation is observed across distinct populations of *Solanum habrochaites* and between *S. habrochaites* and *S. lycopersicum* (Gonzales-Vigil et al., 2012). This genetic variation determines whether specific accessions preferentially synthesize monoterpenes (C10) or sesquiterpenes (C15), and results from differences at the cis-prenyltransferase 1 (CPT1) locus and associated TPS-e/f enzymes that are located within the chromosome 8 terpene gene cluster (Matsuba et al., 2013). For example, trichomes of cultivated tomato predominantly accumulate the monoterpene β-phellandrene, which is synthesized from NPP by neryl diphosphate synthase1 (NDPS1) (Schilmiller et al., 2009). While select monoterpene-producing accessions of *S. habrochaites* also contain an ortholog of NDPS1, a separate group of sesquiterpene producing accessions of *S. habrochaites* possess the C15-producing *Z,Z*-farnesyl diphosphate synthase (zFPS) at the CPT1 locus (Kang et al., 2014; Sallaud et al., 2009) (Fig. 5). Comparative sequence analysis, homology modeling, and site-directed mutagenesis revealed that the relative positioning of bulky aromatic amino acid residues within a hydrophobic cleft specifies substrate binding and prenyl-chain elongation between

CPT1 isoforms with NDPS1 and zFPS activity and that this contributes to intraspecific terpene variation in *S. habrochaites* (Kang et al., 2014).
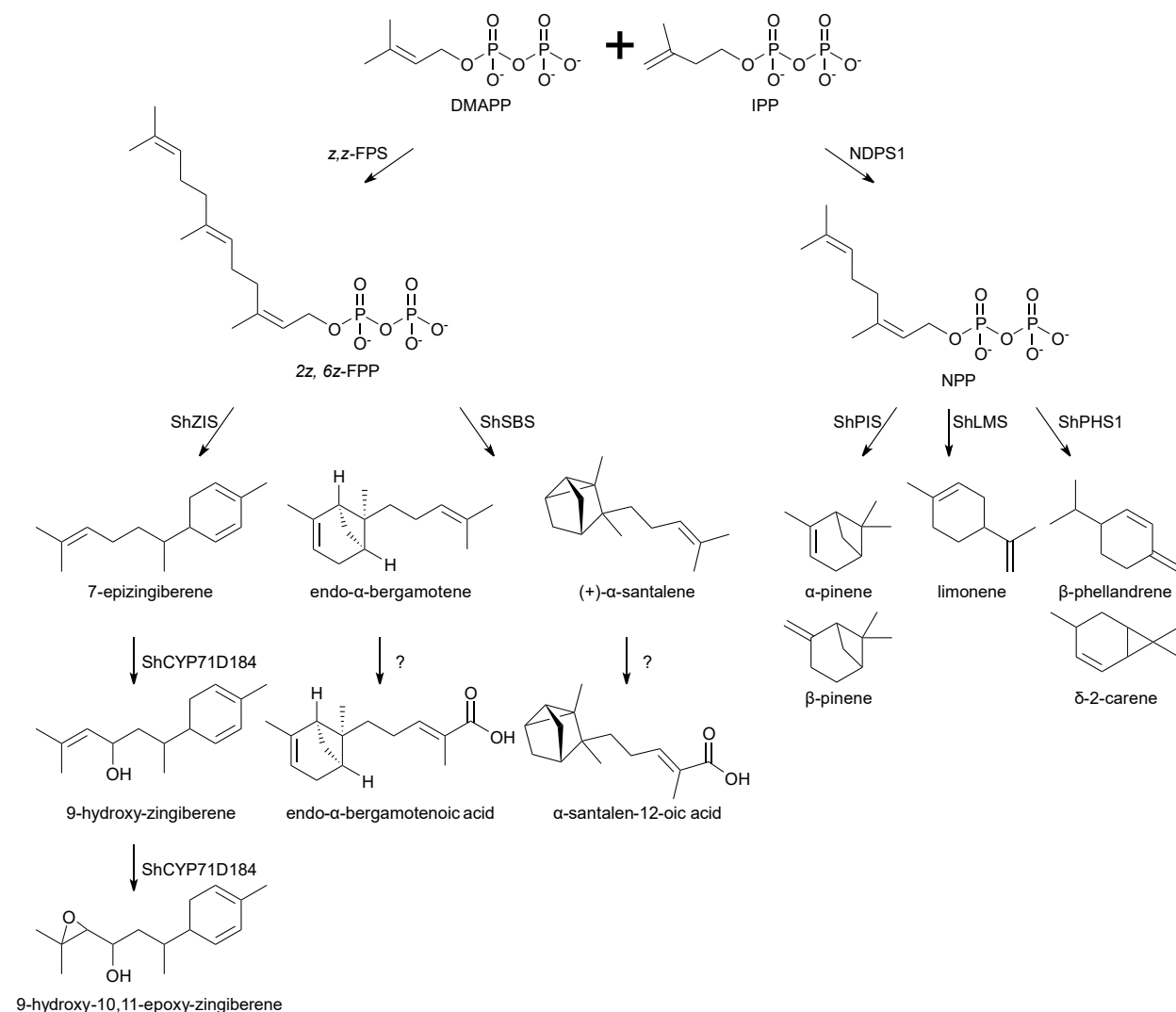
Figure 1.5. Terpenoid biosynthesis in the trichomes of *Solanum habrochaites* derived from cisoid substrates. NDPS1 catalyzes the condensation of a single molecule of DMAPP and IPP to form NPP (C10) (Schilmiller et al., 2009). In contrast, *z,z*-FPS catalyzes the formation of 2Z,6Z-FPP (C15) through sequential condensation of two molecules of IPP with a single molecule of DMAPP (Sallaud et al., 2009). In distinct NPP producing accessions of *S. habrochaites* the monoterpene synthases, ShPIS, ShLMS, and ShPHS1 catalyze the cyclization of NPP to form monoterpenes (Gonzales-Vigil et al., 2012). In a subset of 2Z,6Z-FPP forming accessions, the sesquiterpene synthase, ShSBS catalyzes the formation of endo-α-bergamotene and (+)-α-santalene (Gonzales-Vigil et al., 2012; Sallaud et al., 2009). These sesquiterpenes are converted to their corresponding acids by unknown enzymes. In a distinct subset of 2Z,6Z-FPP producing accessions, ShZIS catalyzes the formation of 7-epizingiberene, which is sequentially oxidized by ShCYP71D184 to 9-hydroxy-zingiberene and 9-hydroxy-10, 11-epoxy-zingiberene (Bleeker et al., 2012; Gonzales-Vigil et al., 2012; Zabel et al., 2021). In trichomes of cultivated

Figure 1.5. (cont'd)

tomato, *S. lycopersicum*, only orthologs of NDPS1 and ShPHS1 are present resulting in the formation of β-phellandrene and δ-2-carene (Schilmiller et al., 2009). Thus, cisoid substrate derived terpene diversity is attenuated in *S. lycopersicum* in comparison to *S. habrochaites*. Abbreviations are as follows: DMAPP, dimethylallyl diphosphate; IPP, isopentenyl diphosphate; NPP, neryl diphosphate; 2Z,6Z-FPP, 2Z,6Z-farnesyl diphosphate; ShZIS, zingiberene synthase; ShSBS, santalene and bergamotene synthase; ShPIS, pinene synthase; ShLMS, limonene synthase; ShPHS1, β-phellandrene synthase.

Together with divergent CPT1 enzymes, terpene diversity in *S. habrochaites* trichomes is also driven by natural variation in chromosome 8 cluster TPS-e/f subfamily members. *S. lycopersicum*, synthesizes a cocktail of monoterpenes in trichomes from NPP using the TPS-e/f enzyme, β-phellandrene synthase (SlPHS1/SlTPS20) (Schilmiller et al., 2009). PHS1 activity is conserved in some *S. habrochaites* accessions while others contain the TPS-e/f paralogs limonene synthase (ShLMS) and pinene synthase (ShPIS), which catalyze the formation of limonene and α-pinene from NPP, respectively (Gonzales-Vigil et al., 2012). In addition to this intraspecific variation in monoterpene biosynthesis, two additional groups of *S. habrochaites* accessions possess TPS-e/f enzymes that synthesize sesquiterpenes from Z,Z-FPP produced by zFPS: santalene and bergamotene synthase (ShSBS) catalyzes the formation of a mixture of santalene and bergamotene isomers (Gonzales-Vigil et al., 2012; Sallaud et al., 2009). In contrast, a distinct, yet closely related enzyme, zingiberene synthase (ShZIS) catalyzes the formation of 7-epizingiberene (Gonzales-Vigil et al., 2012) (Fig. 5). These sesquiterpene forming TPS-e/f enzymes are not present in *S. lycopersicum* and, to date, appear to be restricted to a subset of *S. habrochaites* accessions. Overall, together with variation at the CPT1 locus, these examples illustrate the evolutionary potential of SM associated gene clusters to create and maintain inter-specific and intra-specific chemical diversity. This relatively rapid intra-specific evolution of chemical variation in specific populations of plants may confer selective advantage against diverse biotic challenges.

The ability of trichomes of select *S. habrochaites* accessions to synthesize the sesquiterpenes santalene and bergamotene as well as 7-epizingiberene and their derivatives is known to confer increased tolerance to insect pests and pathogens when compared to trichomes that synthesize *S. lycopersicum* type monoterpenes (Bleeker et al., 2012, 2011; Coates et al., 1988; Frelichowski and Juvik, 2001). Santalene and bergamotene backbones are oxidized into sesquiterpene acids via unknown enzymes (Coates et al., 1988). In contrast, 7-epizingiberene is sequentially oxidized to a combination of 9-hydroxy-zingiberene and 9-hydroxy-10,11-epoxy-zingiberene in reactions catalyzed by the trichome-expressed cytochrome P450, ShCYP71D184 (Zabel et al., 2021) (Fig. 1.5). 9-Hydroxy-10,11-epoxy-zingiberene is particularly effective in bioactivity assays against whiteflies (*Bemisia tabaci*) and the microbial pathogens, *Phytophthora infestans* and *Botrytis cinerea*. ShCYP71D184 is encoded by the Sohab01g008670 locus and is therefore not located in the chromosome 8 TPS cluster responsible for the synthesis of the 7-epizingiberene substrate. The predicted ShCYP71D184 protein is 94% identical to its putative ortholog from *S. lycopersicum* SlCYP71D184/Solyc01g008670. The function of SlCYP71D184 is unknown but *S. lycopersicum* trichomes do not synthesize 7-epizingiberene and this enzyme is incapable of catalyzing the formation of 9-hydroxy-zingiberene and 9-hydroxy-10,11-epoxy-zingiberene. Although not completely understood, these data suggest that, like other loci that influence terpene biosynthesis in glandular trichomes of *Solanum*, genetic variation exists at the CYP71D184 locus that specifies chemical diversity.

Spicy: lineage-specific biosynthesis of capsaicinoids in pepper

Species within the *Capsicum* genus of the Solanaceae possess the capacity to synthesize a group of specialized metabolites known as capsaicinoids, including capsaicin, the principal determinant of pungency in chili peppers. These specialized metabolites are of culinary and

cultural importance but also possess applications as topical pain medications and show efficacy

as anti-inflammatories, treatments for cancer and weight-loss, and possess anti-microbial

activities (Duranova et al., 2022; Friedman et al., 2019; Spiller et al., 2008; Varghese et al.,

2017). Capsaicinoids are synthesized within the placenta that surrounds the seeds of developing

fruit and act as feeding deterrents for small mammals such as rodents, but not birds (Tewksbury

and Nabhan, 2001). This deterrence is mediated by the mammalian vanilloid receptor 1 (VR1)

ion channel that is localized to sensory nerve endings and responds to heat stimuli (Caterina et

al., 2000). The ortholog of VR1 from birds does not respond to capsaicin and as such, birds,

which are more efficient seed dispersers than small mammals, are unaffected by the pungency of

pepper fruits (Jordt and Julius, 2002).

The biosynthesis of capsaicinoids is not fully understood, particularly at the biochemical

level and this pathway is yet to be reconstructed in a heterologous system. However, capsaicin

biosynthesis is considered a derived trait within *Capsicum*, as species from the more ancient

Andean clade of the genus are non-pungent (Carrizo García et al., 2016). Within *Capsicum*

species, intra-specific variation exists resulting in loss of pungency (Carrizo García et al., 2016).

Most notably, this intra-specific variation occurs in the major crop species *Capsicum annuum*

and gives rise to both pungent and sweet pepper cultivars (Carrizo García et al., 2016). Capsaicin

is synthesized through the condensation of vanillylamine, derived from the phenylpropanoid

pathway, with 8-methyl-6-nonenoyl-CoA, produced through branched-chain amino acid

metabolism and fatty acid synthesis (Kim et al., 2014). Genetic analyses identified loci

associated with capsaicin accumulation and genes within the phenylpropanoid, branched-chain

amino acid catabolism, and fatty acid synthesis pathways are among the candidates discovered

(Han et al., 2018; Park et al., 2019; Tripodi et al., 2021). For example, loss of function alleles at

the AMT locus, which encodes an aminotransferase that catalyzes the formation of vanillylamine from vanillin, disrupts capsaicin biosynthesis (Lang et al., 2009; Tanaka et al., 2015; Weber et al., 2014). Similarly, mutation in a ketoacyl-ACP reductase (CaKR1), an enzyme involved in fatty acid biosynthesis, resulted in undetectable levels of capsaicin and 8-methyl-6-nonenoic acid, a precursor of 8-methyl-6-nonenoyl-CoA (Koeda et al., 2019). In addition, the BAHD acyltransferase capsaicin synthase, also known as Pun1, is associated with pungency in hot pepper and proposed to catalyze the condensation of vanillylamine with 8-methyl-6-nonenoyl-CoA to form capsaicin (Stewart Jr et al., 2005). A 2.5 kb deletion allele at this locus is present in non-pungent genotypes, although biochemical evidence supporting a direct role for this enzyme in capsaicin biosynthesis is lacking (Stewart Jr et al., 2005). Overall, these studies reveal genetic variation across *Capsicum* that has likely arisen due to domestication and selection.

Bitter: evolutionary signatures of glycoalkaloid biosynthesis in *Solanum*

Steroidal glycoalkaloids (SGAs) are bitter and toxic metabolites that occur in *Solanum* including the crop species tomato, potato, and eggplant. SGAs provide protection against herbivory as well as microbial pathogens and are proposed to function through the disruption of cell membranes and inhibition of cholinesterase activity (Roddick et al., 2001). In the United States, SGA levels are monitored in potato to maintain levels below an FDA-regulated threshold due to their toxicity (Dolan et al., 2010). Evolution and domestication shaped SGA diversity in *Solanum*; metabolite profiling and chemical structure elucidation reveal hundreds of SGAs that differ among members of the genus due to gene gain and loss between species (Gu et al., 2018; Iijima et al., 2013). For example, α-tomatine and esculeoside A accumulate in tomato while α-solasonine and α-solamargine are synthesized in eggplant. In contrast, domesticated potato synthesizes α-solanine and α-chaconine, while leptines, SGAs that display efficacy against

31

Colorado potato beetle (CPB), are found in wild potato species (Figure 1.6) (Akiyama et al., 2021b; Cárdenas et al., 2019; Paudel et al., 2017; Sánchez-Mata et al., 2010; Shinde et al., 2017). SGAs arise from the modification of cholesterol produced from the mevalonate pathway and are characterized by a nitrogen-containing 27-carbon core, which can undergo multiple glycosylations to form steroidal glycoalkaloids (Sonawane et al., 2016). Comparison of genomic sequences between species revealed that several biosynthetic steps of SGA formation in tomato, potato, and eggplant, encoded by GLYCOALKALOID METABOLISM (GAME) genes, are clustered within these genomes (Barchi et al., 2019; Itkin et al., 2013).
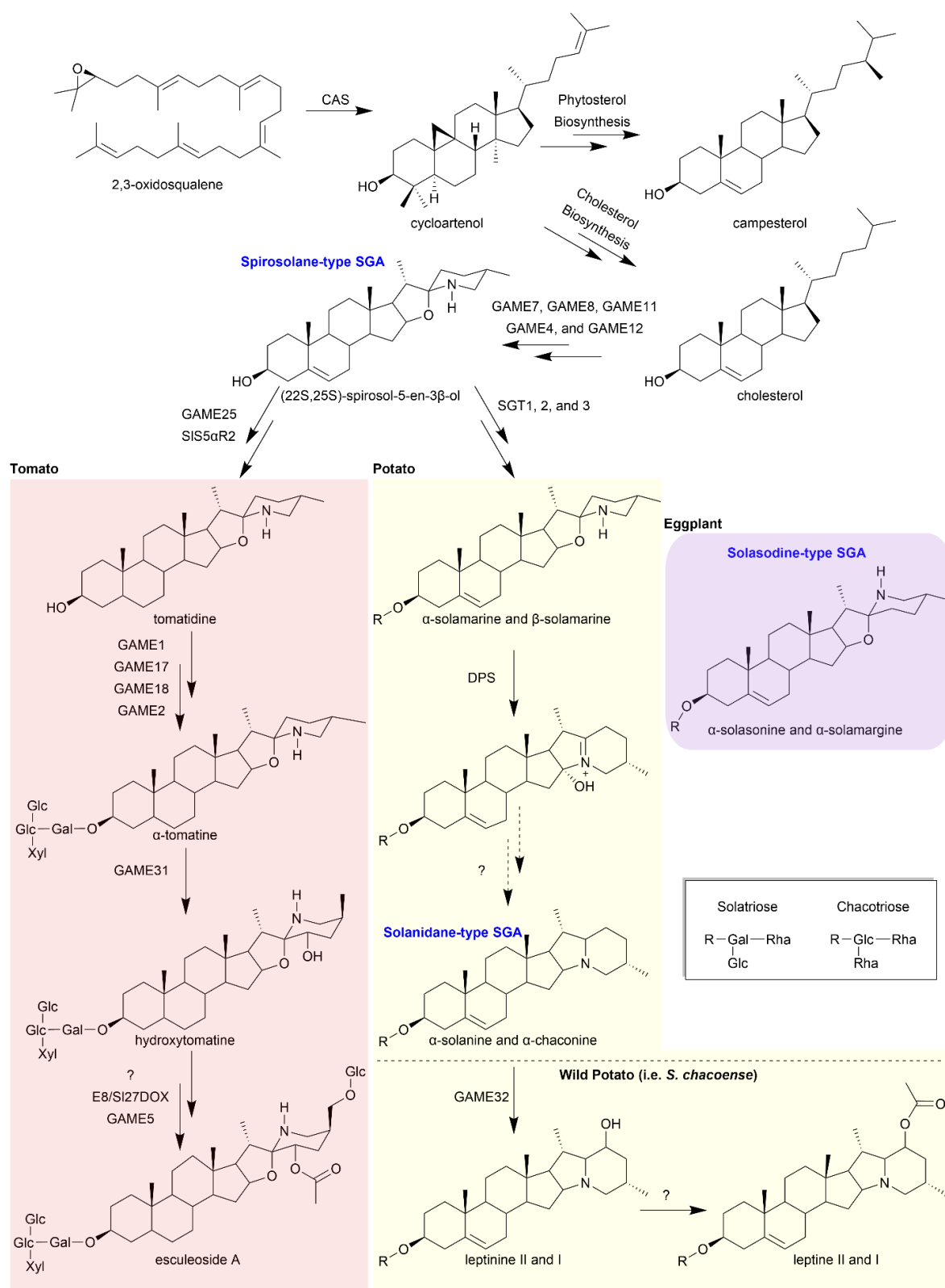
Figure 1.6. Steroidal glycoalkaloid biosynthesis in *Solanum*. CAS cyclizes 2,3-oxidosqualene from the mevalonate pathway to form cycloartenol a common metabolite in both phytosterol and cholesterol biosynthesis. Cycloartenol is converted to campesterol by a ten-

Figure 1.6. (cont'd)
step pathway and through a nine-step pathway to form cholesterol (Sonawane et al., 2016). Following the production of cholesterol, five GAME enzymes are required to produce the spirosolane-type SGA core (Itkin et al., 2013). In tomato (red shaded box), GAME25 catalyzes the first of four steps resulting in tomatidine formation via the reduction of the spirosolane-type SGA core (Akiyama et al., 2019; Sonawane et al., 2018). Subsequent sugar additions by GAME1, GAME17, GAME18, and GAME2 result in the formation of α-tomatine (Itkin et al., 2013). GAME31, E8/Sl27DOX, GAME5, and an unknown acetyltransferase catalyze the fruit ripening associated formation of esculeoside A from α-tomatine (Akiyama et al., 2021a; Cárdenas et al., 2019; Kazachkova et al., 2021; Nakayasu et al., 2020; Szymański et al., 2020). In potato (yellow shading), the addition of solatriose and chacotriose moieties by sequential sugar additions to (22*S*,25*S*)-spirosol-5-en-3β-ol results in the formation of α- and β-solamarine, respectively (Akiyama et al., 2021b). The oxidization of α- and β-solamarine by DPS represents the first step in α-solanine and α-chaconine, Solanidane-type SGA, formation (Akiyama et al., 2021b). In *S. chacoense*, α-solanine and α-chaconine are oxidized by GAME32 to form leptinines, and leptine formation requires the acetylation at the GAME32 introduced oxidation (Cárdenas et al., 2019). The solasodine-type SGAs (α-solasonine and α-solamargine) are the main SGAs in eggplant (purple shading) and contain solatriose and chacotriose moieties at the C-3 position, respectively. The biosynthetic mechanism leading to the stereochemical difference in spirosolane and solasodine cores remains uncharacterized (Akiyama et al., 2021b; Sánchez-Mata et al., 2010). Enzyme abbreviations are as follows: CAS, cycloartenol synthase; GAME, glycoalkaloid metabolism; SlS5αR2, steroid 5α-reductase 2; SGT, solanidine glycosyltransferase; DPS, dioxygenase for potato solanidane synthesis; E8/Sl27DOX, α-tomatine 27-hydroxylase; Gal, galactose; Glc, glucose; Xyl, xylose; Rha, Rhamnose.

Formation of plant SGA sterol cores requires diversion of 2,3-oxidosqualene from the mevalonate pathway into cholesterol biosynthesis, and this biosynthetic pathway appears to have evolved from the duplication and divergence of genes involved in phytosterol biosynthesis, which leads to the production of brassinosteroids, an essential class of phytohormones (Sonawane et al., 2016). Cycloartenol synthase (CAS) converts 2,3-oxidosqualene into cycloartenol, and this metabolite is the branch point between cholesterol and phytosterol biosynthesis as it serves as a substrate for both SSR2 (sterol side chain reductase 2) and SMT1 (sterol C-24 methyltransferase) to form cycloartanol or 24-methylenecycloartanol, respectively (Sonawane et al., 2016). Cholesterol biosynthesis leads to the production of the SGAs and saponins in both glycosylated and aglycone forms (Sonawane et al., 2016). Elucidation of cholesterol biosynthesis in plants revealed five enzymes shared between the cholesterol and

phytosterol pathways (Sonawane et al., 2016). Phylogenetic analysis of enzymes specific to cholesterol biosynthesis suggests that C5-SD2 (sterol C-5(6) desaturase), 7-DR2 (7-dehydrocholesterol reductase), SMO3 (C-4 sterol methyl oxidase) and SMO4 likely arose from duplication and divergence of the phytosterol pathway genes, C5-SD1, 7-DR1, SMO1 and SMO2 (Sonawane et al., 2016).

Presence-absence variation of genes involved in the conversion of dehydro-SGAs to dihydro-SGAs contributes to SGA diversity within *Solanum*. The first spirosolosane-type SGA formed, (22S, 25S)-spirosol-5-en-3β-ol, contains a Δ5,6 double bond (Akiyama et al., 2021b). In tomato, tomatidine is synthesized from a multistep process starting with the oxidation and isomerization of (22S, 25S)-spirosol-5-en-3β-ol to tomatid-4-en-3-one by GAME25, and the addition of four sugars (galactose, glucose, glucose, and xylose) to the C-3 position of tomatidine results in the production of tomatine, the major tomato SGA (Akiyama et al., 2019; Sonawane et al., 2020, 2018). Lack of a functional GAME25 is associated with the production of unsaturated SGAs, including α-solamargine, α-solasonine, and malonylsolamargine in *S. melongena* (eggplant) and expression of tomato GAME25 in eggplant results in the production of saturated SGAs (Sonawane et al., 2018). However, the mechanism underlying a lack of saturated SGA accumulation in domesticated potato is less clear. A putative GAME25 homolog is present in the genome of domesticated potato, and recombinant expression of the corresponding enzyme revealed the same activity as the tomato enzyme: 3β-hydroxyl group oxidation and isomerization of the double bond from the C-5,6 position. The potato GAME25 enzyme is active with unsaturated spirolosane- and solanidine-type SGAs although the corresponding saturated SGAs do not accumulate in domesticated potato (Sonawane et al., 2018). Overexpression of tomato GAME25 in potato hairy root cultures leads to accumulation of demissidine, a saturated

solanidine SGA found in wild potato. This suggests that the downstream enzymatic activities involved in the production of saturated SGAs exist in domesticated potato (Lee et al., 2019). However, the mechanism leading to the lack of saturated SGAs in domesticated potato remains unclear, and the *in vivo* function of the domesticated potato GAME25 and expression levels of the corresponding gene remain to be determined (Lee et al., 2019; Sonawane et al., 2018).

While the initial steps of spirolosane-type SGA formation are conserved between tomato and potato, SGA biosynthesis diverges in potato to produce solanidine-type SGAs (Akiyama et al., 2021b). Potato contains two major solanidane-type SGAs, α-solanine and α-chaconine, which differ only in the identity of the C-3 sugar additions; solanine contains galactose with rhamnose and glucose additions while chaconine contains glucose with two rhamnose additions (Akiyama et al., 2021b). The 2-oxoglutarate dependent dioxygenase, DPS (Dioxygenase for Potato Solanidane synthesis), catalyzes solanidine ring formation via C-16 hydroxylation (Akiyama et al., 2021b). While both eggplant and tomato contain DPS homologs and each recombinant enzyme is capable of C-16 hydroxylation of spirolosane-type SGAs, the expression of the corresponding genes is low or undetectable in eggplant and tomato, which likely explains the lack of solanidine-type SGAs in these species (Akiyama et al., 2021b). The DPS genes are located on chromosome 1 within a syntenic block that is conserved in *Solanum* and contains additional SM-related genes, suggesting that the DPS genes evolved prior to speciation (Akiyama et al., 2021b). While some wild potato species, such as *Solanum chacoense*, produce leptines, solanidine-type SGAs that are effective at defending against CPB, domesticated potato does not produce these SGAs. Leptine formation requires the hydroxylation of solanidine-type SGAs by GAME32 and the subsequent acetylation by an unknown enzyme. Tomato and

domesticated potato lack a functional GAME32 homolog and the corresponding leptine SGAs (Cárdenas et al., 2019).

Domestication and selection for non-bitter fruit to aid in seed dispersal influence SGA content in tomato during fruit ripening. The fruit ripening associated biosynthesis of esculeoside A from α-tomatine alleviates the bitter taste associated with SGAs (Cárdenas et al., 2019). The hydroxylation of α-tomatine at the C-23 position is the first committed step of fruit ripening associated SGA accumulation (i.e. esculeoside A), and is catalyzed by the 2-ODD enzyme, GAME31 (Cárdenas et al., 2019; Nakayasu et al., 2020). Esculeoside A formation requires an additional hydroxylation, followed by acetylation, and the glycosylation of acetoxy-hydroxytomatine by GAME5 (Akiyama et al., 2021a; Cárdenas et al., 2019; Szymański et al., 2020). The export of α-tomatine and α-tomatine derivatives out of the vacuole by a nitrate transporter 1/peptide transporter family (NPF) transporter, GORKY (meaning bitter in Russian), is essential for esculeoside A formation (Kazachkova et al., 2021). The sequestration of toxic SGAs to the vacuole likely prevents self-toxicity, and this is evidenced by the observation that tomato plants overexpressing GORKY (facilitating SGA export to the cytosol) displayed severe morphological phenotypes (Kazachkova et al., 2021). In contrast, fruit from the same overexpression lines did not display signs of self-toxicity suggesting that the conversion of toxic/bitter SGAs to esculeosides prevents self-toxicity (Kazachkova et al., 2021).

The synteny of the metabolic gene clusters involved in SGA production among *Solanum* species highlights the common origin of the trait that diverged between species through loss or gain of function of individual genes to create SGA diversity. Several of the genes involved in spirolosane-type SGA formation are found clustered on potato, eggplant, and tomato chromosomes 7 and 12 (Barchi et al., 2019; Itkin et al., 2013). Tomato possesses two extra genes

in these clusters as potato and eggplant lack homologs of GAME17 and 18, two UDP-glucosyltransferases responsible for the consecutive additions of glucose to tomatidine galactoside during α-tomatine biosynthesis in tomato (Itkin et al., 2013). Current genomic resources show that pepper (*Capsicum annuum*) does not possess the chromosome 12 cluster or putative orthologs of GAME4 and GAME12 found within the cluster, and this absence likely results in the lack of SGAs in *C. annuum* (Barchi et al., 2019). The 2-ODD genes involved in solanidine, leptine, and esculeoside SGA biosynthesis are also clustered with additional 2-ODDs of unknown function (Cárdenas et al., 2019). Changes in gene expression (i.e. low expression of DPS tomato homolog) or the presence-absence of single genes (i.e. GAME32 presence in *S. chacoense*) contribute to SGA diversity in *Solanum*.

Addictive and deadly: convergent and divergent evolution shapes nicotine and tropane alkaloid metabolism

Several Solanaceae genera, including *Datura*, *Atropa*, *Hyoscyamus*, *Mandragora*, and *Scopolia* derive medicinal and toxic qualities from the biosynthesis of tropane alkaloids. Tropane alkaloids are characterized by an eight-membered, bicyclic, nitrogen-containing core and their synthesis is reported in 10 plant families, separated by ~120 Mya of evolution (Kim et al., 2016). For example, the well-known narcotic cocaine is synthesized by *Erythroxylum coca* (Erythroxylaceae) while cochlearine is synthesized in *Cochlearia officinalis* (Brassicaceae). The Solanaceae family has emerged as a model system for studying tropane alkaloid biosynthesis, but comparative studies reveal instances of independent evolution of tropanes in distinct plant lineages (Brock et al., 2008; Jirschitzka et al., 2012).

Scopolamine and hyoscyamine are tropane aromatic esters specific to the Solanaceae, and these compounds derive their medicinal properties from anticholinergic effects, blocking activity

38

of the neurotransmitter acetylcholine. Scopolamine is used to treat a variety of illnesses including motion sickness, drooling, and for palliative care in Parkinson's disease (Clissold and Heel, 1985; Mato et al., 2010; Pérez et al., 2011). Tropane aromatic ester production requires the biosynthesis of the tropane core as well as condensation of a phenyllactic acid moiety through an ester linkage (Qiu et al., 2020). Although the biosynthesis of the tropane core intermediate and polyhydroxylated derivates, known as calystegines, occurs in many genera of the Solanaceae, including *Solanum*, the biosynthesis of tropane aromatic esters is restricted to the genera described above, suggesting that not all species in the family possess the genes required for their synthesis (Nash et al., 1993). Due to their medicinal importance, considerable effort has focused on understanding the biosynthesis of hyoscyamine and scopolamine.

Research leading to the elucidation of scopolamine biosynthesis spanned several decades, with progress driven by the available technologies of the time. Initially, approaches focused on feeding labeled forms of potential precursors to tropane producing plants and following incorporation of label into alkaloids (Kim et al., 2016). This resulted in identification of pathway precursors and intermediates, as well as the development of an overall framework of scopolamine biosynthesis. These efforts were followed by classical biochemical approaches to purify enzymes based on activity. Peptide sequencing of the resulting purified enzymes facilitated the design of oligonucleotide probes that were labeled and used to screen cDNA libraries to identify the corresponding clones. Confirmation of function was achieved through characterization of resulting recombinant enzymes expressed in *E. coli*. This led to the identification of several pathway genes, including hyoscyamine 6β-hydroxylase (H6H), tropinone reductase I/II (TRI and TRII), and putrescine *N*-methyltransferase (PMT). The development of expressed sequence tags in the mid-2000s, coupled with virus-induced gene

39

silencing (VIGS) for *in vivo* testing of function, led to the identification of littorine mutase, an enzyme that catalyzes the rearrangement of littorine into hyoscyamine aldehyde (Li et al., 2006). More recently, *Atropa belladonna* (Deadly Nightshade) emerged as a model for exploring tropane alkaloid biosynthesis following the development of a multi-tissue transcriptome assembly and the deployment of VIGS. These resources, coupled with synthetic biology, culminated in the identification of the missing steps in scopolamine formation.

The first ring of the tropane core requires the conversion of ornithine, a non-proteinogenic amino acid, into putrescine by ornithine decarboxylase (ODC). Putrescine is then N-methylated by putrescine methyltransferase (PMT) and oxidized by methylputrescine oxidase (MPO). The *N*-methyl-Δ^1-pyrrolinium cation forms through the spontaneous cyclization of *N*-methylaminobutanal, the product of MPO catalysis (Figure 1.7). PMT requires *S*-adenosyl-L-methionine (SAM) to *N*-methylate putrescine and shares high sequence similarity with spermidine synthase (SPDS), an enzyme involved in transferring the aminopropyl moiety from decarboxylated SAM (dcSAM) onto putrescine to form spermidine, a ubiquitous polyamine (Junker et al., 2013; Stenzel et al., 2006). It was hypothesized that PMT evolved from a gene duplication of SPDS and subsequent neofunctionalization, and although SPDS cannot catalyze putrescine *N*-methylation, mutation of a single SPDS amino acid, D103I, is sufficient to generate PMT activity (Junker et al., 2013). The pyrrole moiety of nicotine, a natural product produced in the *Nicotiana* genus of the Solanaceae, also requires *N*-methyl-Δ^1-pyrrolinium cation biosynthesis. The biosynthetic steps leading to *N*-methyl-Δ^1-pyrrolinium cation formation are conserved in *Nicotiana*, *Solanum*, and *Petunia* allowing the *N*-methyl-Δ^1-pyrrolinium cation to act as a core for nicotine and tropane alkaloid biosynthesis found in Solanaceae and Convolvulaceae (Kajikawa et al., 2017; Xu et al., 2017). In contrast, the genes involved in the

formation of the pyridine ring in nicotine biosynthesis are *Nicotiana*-specific indicating that

divergent evolution led to the formation of nicotine, likely through the duplication of the genes in

the nicotinamide adenine dinucleotide (NAD) cofactor biosynthetic pathway (Xu et al., 2017).
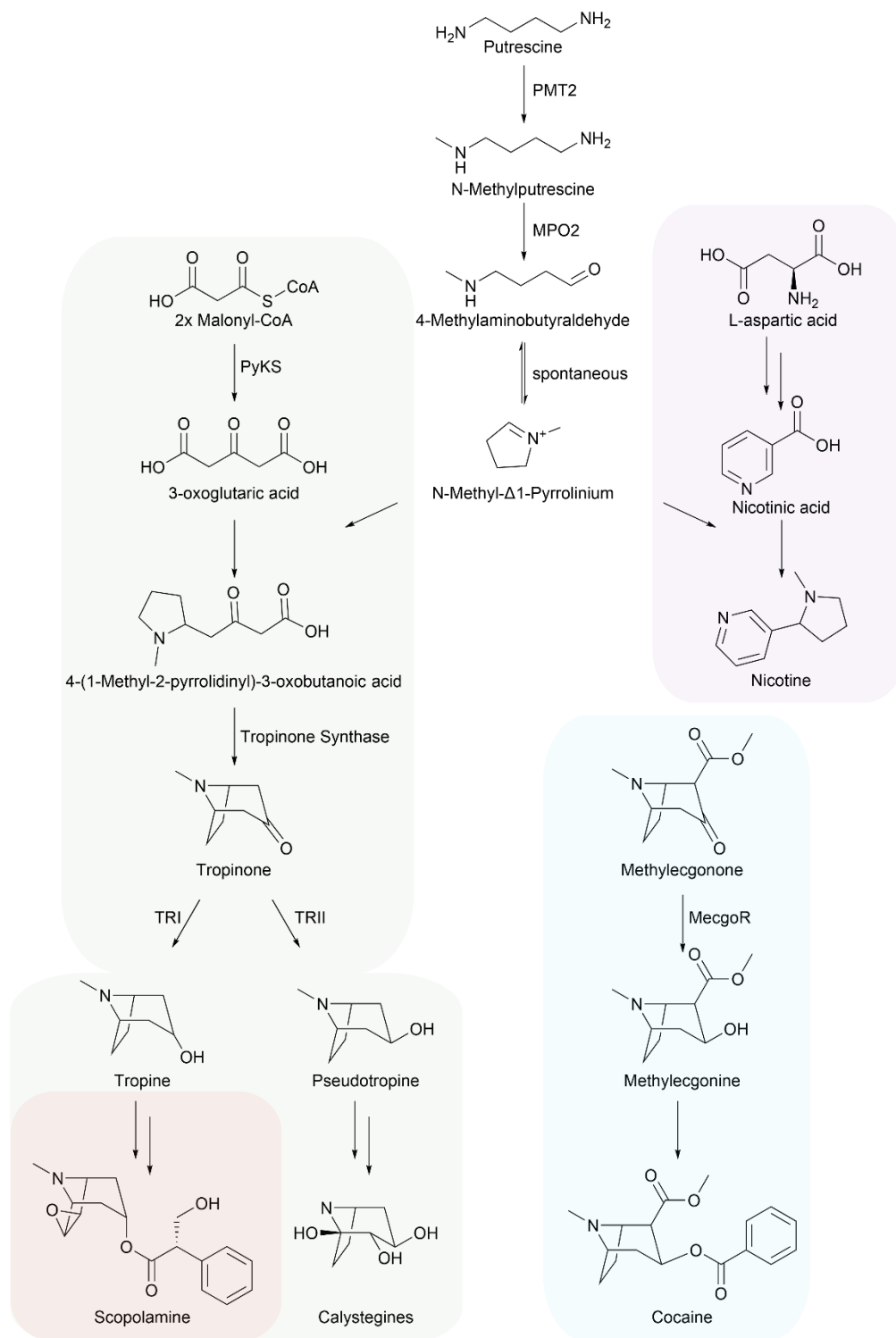
Figure 1.7. Evolutionary trajectories of tropane and nicotine formation in distinct plant lineages. Comparison of tropane and nicotine alkaloid biosynthesis reveals examples of both convergent (cocaine biosynthesis in *E. coca*) and divergent (nicotine biosynthesis) evolution (Jirschitzka et al., 2012; Xu et al., 2017). Scopolamine (orange) and nicotine (purple) represent alternative fates of the *N*-methylpyrrolinium cation in different genera of the Solanaceae. The use

Figure 1.7. (cont'd)

of an aldo-keto reductase enzyme (MecgoR) in the penultimate step of cocaine biosynthesis (blue) contrasts with catalysis by short-chain dehydrogenase/reductase (SDR) family enzymes (TRI and TRII) in scopolamine formation (green) (Jirschitzka et al., 2012). *Not shown is catalysis by a single, bifunctional SDR to produce both tropine and pseudotropine in Brassicaceae (Brock et al., 2008). Tropanol biosynthesis (green) is widely distributed across the Solanaceae compared to the biosynthesis of tropane aromatic esters such as scopolamine (orange) (Nash et al., 1993). Enzyme abbreviations are as follows: PMT2, Putrescine *N*-methyltransferase 2; MPO2, *N*-methylputrescine oxidase 2; PyKS, Polyketide Synthase; TRI, Tropinone reductase I; TRII, Tropinone Reductase II; MecgoR, Methylecgonone reductase.

Formation of the tropane core in Solanaceae species requires a second cyclization event that yields tropinone, which possesses a ketone functional group at the carbon-3 position of the core (Figure 1.7). The first step in tropinone formation is catalyzed by a type III polyketide synthase, PYKS, which uses the *N*-methyl-Δ^1-pyrrolinium cation and malonyl-Coenzyme A to form 4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoic acid (Bedewitz et al., 2018). Although PYKS can form 3-oxoglutaric acid without the *N*-methyl-Δ^1-pyrrolinium cation and these two products can react non-enzymatically, the exact mechanism of 4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoic acid formation remains unclear (Huang et al., 2019; Nett et al., 2021). Tropinone synthase (CYP82M3) converts 4-(1-methyl-2-pyrrolidinyl)-3-oxobutanoic acid to tropinone (Bedewitz et al., 2018). Although putative orthologs of PYKS and CYP82M3 are present in the genomes of several calystegine producing Solanaceae species including tomato, potato, and pepper, these genes are absent in *Nicotiana* spp.; this is consistent with the lack of detectable tropanes in these species (Bedewitz et al., 2018). In the Solanaceae, tropinone reductases I and II are members of the short-chain dehydrogenase/reductase superfamily (SDR) that catalyze the reduction of the ketone of tropinone to an alcohol to form tropine (3α-hydroxytropine) and pseudotropine (3β-hydroxytropine), respectively (Nakajima et al., 1993). TRI and TRII constitute a branch point in the tropane alkaloid biosynthetic pathway due to their stereospecificity: TRI leads to the

43

production of tropane aromatic esters, including hyoscyamine and scopolamine and TRII directs flux towards calystegine production.

Biosynthesis of the principal aromatic tropane esters in the Solanaceae, littorine, hyoscyamine, and scopolamine, requires the diversion of phenylalanine into the tropane pathway through a two-step process that yields phenyllactic acid (Bedewitz et al., 2014; Qiu et al., 2018) (Figure 1.8). Identification of the aromatic aminotransferase (AbArAT4) responsible for conversion of phenylalanine into phenylpyruvate revealed the power of transcriptomics in Solanaceae tropane alkaloid enzyme discovery (Bedewitz et al., 2014). Analogous to bacterial aromatic amino acid biosynthesis, a cytosolic aromatic aminotransferase from petunia (Ph-PPY-AT) catalyzes the formation of phenylalanine from phenylpyruvate using tyrosine as an amino donor and yielding 4-hydroxyphenylpyruvate (Yoo et al., 2013). AbArAT4 is related to Ph-PPY-AT and utilizes the same four substrates, but the *Atropa* enzyme diverts phenylalanine into the tropane pathway by virtue of a ~250-fold more active reverse reaction that yields phenylpyruvate and tyrosine. AbArAT4 is co-expressed in the roots with other tropane-related genes, and while silencing of this gene disrupts tropane alkaloid biosynthesis, it does not alter aromatic amino acid pools, further supporting its neofunctionalized and specific role in specialized metabolism (Bedewitz et al., 2014). Littorine biosynthesis requires the glycosylation of phenyllactate by a UDP-glucose dependent glycosyltransferase followed by the acylation of tropine. The serine carboxypeptidase-like (SCPL) acyltransferase (littorine synthase) acylates tropine using glycosylated phenyllactate as the acyl donor (Qiu et al., 2020).
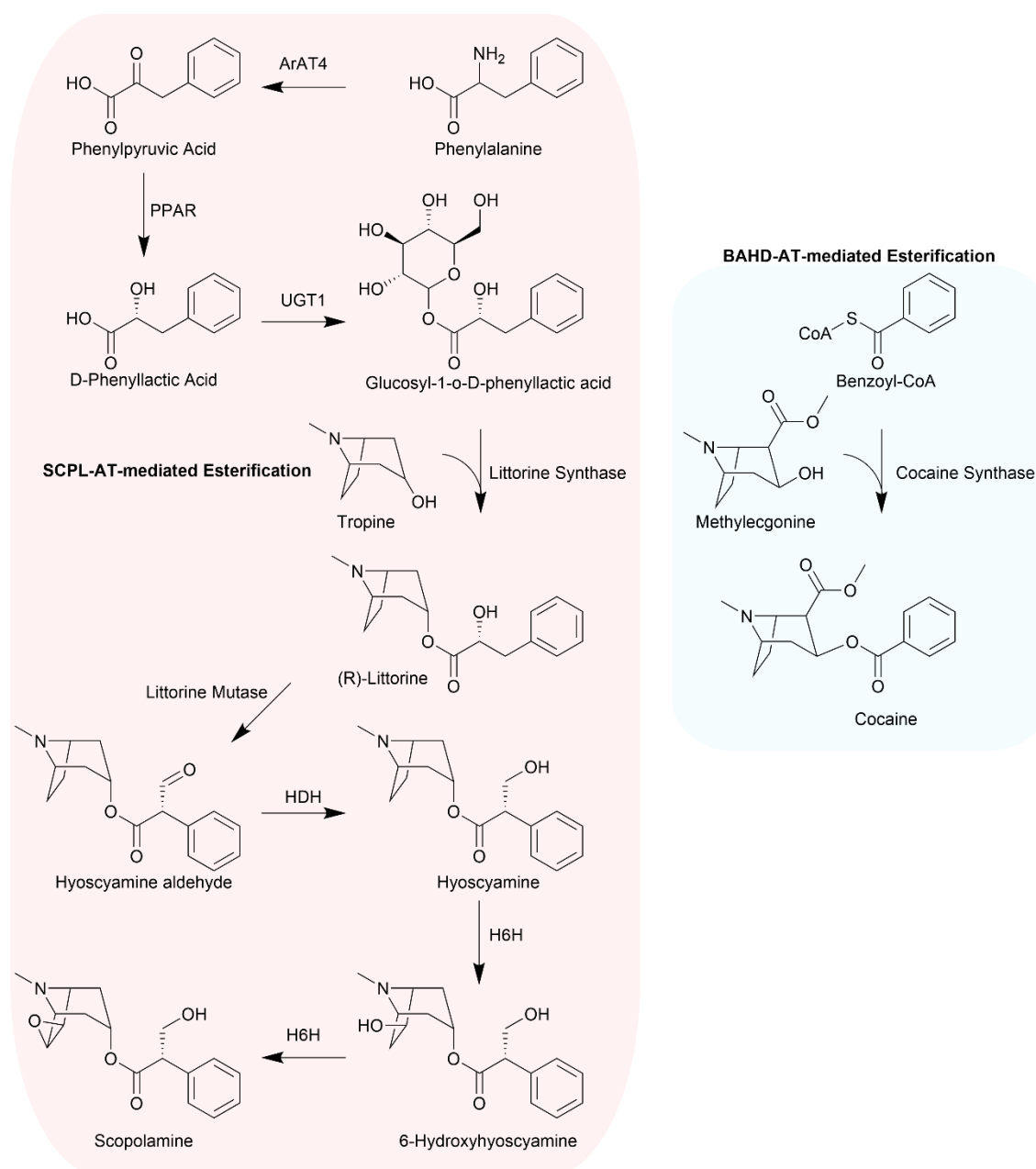
Figure 1.8. Independent evolution of tropane aromatic ester formation in Solanaceae and Erythroxylaceae. Scopolamine biosynthesis requires the biosynthesis of D-phenyllactic acid via a two-step process mediated by ArAT4 and PPAR (Bedewitz et al., 2014; Qiu et al., 2018). D-Phenyllactic acid is glycosylated by UGT1 to form a glucose ester of phenyllactic acid, which is used, along with tropine, as substrate for littorine biosynthesis by Littorine Synthase, a serine carboxypeptidase-like acyltransferase (Qiu et al., 2020). Three enzymes, Littorine Mutase, HDH, and H6H, are required for the conversion of littorine to scopolamine (Hashimoto and Yamada, 1986; Li et al., 2006; Srinivasan and Smolke, 2020). In contrast, cocaine biosynthesis utilizes a BAHD acyl-transferase and coenzyme A donor to facilitate the transfer of a benzoyl moiety on to methylecgonine, the *E. coca* tropanol, to form cocaine (Schmidt et al., 2015). Enzyme abbreviations are as follows: ArAT4, aromatic amino acid transferase 4; PPAR, phenylpyruvic acid reductase; UGT1, UDP-glycosyltransferase 1; HDH, hyoscyamine dehydrogenase; H6H,

Figure 1.8. (cont'd)
hyoscyamine-6-hydroxylase.

Synthetic biology recently was utilized both to engineer scopolamine production in yeast and facilitate the discovery of the final missing enzyme in the pathway, which had eluded discovery using *in planta* experiments. The conversion of littorine to scopolamine requires four steps catalyzed by three enzymes (Figure 1.8). Littorine mutase, a cytochrome P450, catalyzes the rearrangement of littorine to hyoscyamine aldehyde (Li et al., 2006), which is converted to hyoscyamine by hyoscyamine aldehyde dehydrogenase. Finally, hyoscyamine-6-hydroxylase catalyzes the two-step hydroxylation and epoxidation of hyoscyamine to scopolamine (Hashimoto and Yamada, 1986). The production of scopolamine in yeast was achieved through the introduction of tropane alkaloid pathway genes from several species, including *Datura stramonium*, *Datura metel*, and *Atropa belladonna* (Srinivasan and Smolke, 2020). Optimization of scopolamine production in yeast required the elimination of several native genes to reduce the flow of tropane alkaloid intermediates into side products and the introduction of a transporter from *Nicotiana tabacum* to facilitate transport of tropine into the vacuole for esterification with phenyllactic acid (Srinivasan and Smolke, 2020). Notably, the introduction of the pathway into yeast revealed the dehydrogenase responsible for the reduction of hyoscyamine aldehyde into hyoscyamine, which had not previously been identified *in planta* (Srinivasan and Smolke, 2020). For example, silencing of this gene in *A. belladonna* did not result in a decrease in downstream tropane alkaloids, likely due to promiscuous enzymatic activity of other dehydrogenases (Qiu et al., 2021). Hence, reconstruction of the pathway in a genetic host where background activities were removed facilitated the identification of the final missing step in the scopolamine pathway.

<u>Independent evolution of tropanes in distinct plant lineages</u>

Evidence for independent evolution of tropanes in distinct plant lineages is manifest at different steps throughout the pathway (Figures 1.7 and 1.8). While separate TRI and TRII enzymes reduce tropinone to tropine or pseudotropine in the Solanaceae, a single SDR enzyme catalyzes both reactions in *C. officinalis*, ultimately leading to tropine-derived cochlearine and pseudotropine-derived calystegines (Brock et al., 2008). In addition, while Solanaceae and Brassicaceae species utilize enzymes in the SDR family for the reduction of tropinone, the analogous reaction in *E. coca* cocaine biosynthesis, the reduction of methylecgonone to methylecgonine, is catalyzed by methylecgonone reductase (MecgoR) a member of the aldo-keto reductase family (Jirschitzka et al., 2012). Similarly, aromatic tropane ester biosynthesis is catalyzed by different classes of acyltransferases in the Solanaceae and Erythroxylaceae. Littorine formation is synthesized by an SCPL acyltransferase while cocaine synthase, which catalyzes the condensation of methylecgonine and benzoyl-CoA, is a member of the BAHD acyltransferase family (Schmidt et al., 2015). As additional tropane pathways in distinct plant lineages are elucidated it is likely that further examples of independent evolution will be discovered.

Challenges and unexplored frontiers in Solanaceae metabolism

There has been a rapid increase in understanding the biosynthesis and evolution of plant SM pathways during the last decade. Advances in genomics enabled gene–metabolite correlations in model and non-model species. These data – combined with development of methods to test gene function in diverse species, and transient expression in *Nicotiana benthamiana*, as well as engineering production in microbial systems – led to the elucidation of multiple plant SM pathways and identified regulators of known SM pathways (Fossati et al.,

2014; Lau and Sattely, 2015; Y. Li et al., 2020; Nett et al., 2020; Srinivasan and Smolke, 2020).

The widespread adoption of these approaches, coupled with phylogeny-guided comparative genomics and metabolomics, enabled exploration of the evolutionary trajectories of the exemplary Solanaceae SM pathways described here.

However, despite advances in understanding Solanaceae SM biosynthesis and evolution, knowledge gaps persist related to specific aspects of these well-studied pathways and opportunities exist to develop a more comprehensive understanding of these pathways and networks. As evidenced through studies of acylsugar evolution, much can be learned through adopting a broader sampling strategy to include more phylogenetically diverse species that are typically less well studied (Lou et al., 2021; Moghe et al., 2017; Nadakuduti et al., 2017). Similar, phylogenetic-guided metabolite screening approaches could be adopted to assess chemical diversity in other SM classes as the foundation for exploring metabolite evolution using comparative genomics. For example, given the tremendous chemical variation observed in trichome-derived acylsugars across the Solanaceae, and that novel acylsugars were recently identified in root and root-exudates of tomato (Korenblum et al., 2020), it will be intriguing to determine whether comparable root acylsugar diversity exists across the family and if so, to assess how this diversity evolved.

There are also several examples where the biosynthesis of exemplary SM pathways in the Solanaceae are not fully resolved. For example, the enzymes that catalyze the early steps in acylinositol biosynthesis in *Solanum* spp. are yet to be reported. Similarly, the majority of the enzymes involved in capsaicinoid biosynthesis and the final steps in nicotine biosynthesis await biochemical and functional characterization (Naves et al., 2019; Xu et al., 2017). In addition, although the biosynthesis of scopolamine is elucidated and the pathway reconstructed in yeast,

48

the steps leading to the biosynthesis of other classes of Solanaceae tropanes, including

calystegines and schizanthines, are unknown (Christen et al., 2020; Kim et al., 2016).

Comparative analyses of the evolution of SM-related gene clusters across the Solanaceae

also remains under-explored. For example, as outlined in this review, terpene and SGA-related

gene clusters exist in *Solanum* but variation across these clusters is mainly documented in a few

model species, including tomato, potato, eggplant, and closely related wild species (Barchi et al.,

2019; Itkin et al., 2013; Zhou and Pichersky, 2020b). Indeed, even for the comparatively well-

studied terpenoid-related gene clusters of tomato, many of the enzymes that reside within these

clusters, which may catalyze modifications of terpene scaffolds, remain uncharacterized.

Furthermore, the extent of conservation of terpene and other SM gene clusters across the

Solanaceae is unknown. As multiple chromosome scale genome assemblies of phylogenetically

diverse Solanaceae species are available and others will likely be generated soon, charting the

evolutionary trajectories of SM gene clusters and the metabolite variation they encode is now

possible.

Finally, it is also worth noting that the most extensively characterized Solanaceae SM

pathways are those where the identities of the major metabolites were known for decades and

their abundance is high in specific cell types or tissues, facilitating purification and structural

elucidation. It is more challenging to identify unknown metabolites and purify metabolites that

are of low abundance and technical challenges persist that impede a more comprehensive

understanding of metabolism and bridging of the gap between genotype and phenotype.

<u>Challenges in the identification and annotation of SM enzymes</u>

Advances in DNA sequencing are making development of chromosome-scale genome

assemblies more routine and recently several Solanaceae genomes were released, and the quality

of existing assemblies improved (Alonge et al., 2020; Barchi et al., 2021; Michael and VanBuren, 2020). These studies allow the gene complement of an organism to be determined. However, functional annotation of plant genomes remains incomplete, even for model species. The lack of accurate annotation is particularly problematic for large gene families encoding SM-related enzymes that catalyze common decorations of scaffold molecules, including cytochromes P450, 2-oxoglutarate dependent dioxygenases, glycosyltransferases, and acyltransferases. SM-related enzymes are often catalytically promiscuous and encoded by genes that evolved rapidly through duplication and associated subfunctionalization, neofunctionalization, and gene loss (Weng et al., 2012). Thus, annotation of SM enzymes based solely on sequence similarity, predicted orthology, or synteny is often misleading. This concept is clearly illustrated by examples identified through studying the evolution of acylsugar and terpene biosynthesis in *Solanum* glandular trichomes. These studies reveal how activity can be altered by a few amino acid differences in closely related enzymes from sister species, or diverse accessions within a species (Fan et al., 2015; Kang et al., 2014; Zabel et al., 2021). Hence, empirical determination of enzyme function remains imperative. Although characterization of enzyme activities is often technically challenging, time consuming, and limited by substrate availability, medium and high-throughput methods based on microtiter plates and microfluidics are utilized for screening natural and computationally designed enzymes and such methods could potentially be adapted for screening the activity of plant SM-related enzymes (Bunzel et al., 2018).

As documented throughout this review, co-expression is a powerful approach for predicting membership of genes in metabolic pathways, particularly when there is *a priori* knowledge about enzymes from the target pathway. Elucidation of the pathway leading to scopolamine biosynthesis, described above, is an excellent example of the use of co-expression

50

analyses to identify candidate genes co-expressed in roots. However, when results of co-expression analysis are ambiguous or multiple candidate genes are identified, as is often the case when investigating large SM-related gene families, additional filtering and refinement of gene candidates may be required prior to time-consuming functional studies. In such cases, comparative genomic analysis such as synteny or gene-cluster analysis – together with phylogenetic analysis to determine whether gene candidates exhibit lineage-specific distribution or arose through a recent duplication event – provide opportunities for refining candidate gene lists (Jacobowitz and Weng, 2020). Outside of tomato, there is a lack of publicly available transcriptome data, including data from diverse tissues, environmental perturbations, and treatments. This limits novel metabolite pathway discovery in diverse Solanaceae species and reduces the resolution of studies investigating the phylogenetic distribution and evolution of SM pathways. Furthermore, plant SM pathways are often restricted to specific cell types, and therefore the general focus on whole tissue sampling for transcriptome analysis can be limiting (Courdavault et al., 2014; Leong et al., 2019; Onoyovwe et al., 2013). The recent development of single-cell and single-nucleus transcriptome analyses holds great promise for increasing the resolution of transcriptome data and refining candidate gene lists to facilitate the identification, characterization, and cellular localization of Solanaceae SM pathways (Ryu et al., 2019; Seyfferth et al., 2021).

Machine learning is another promising approach to distinguish GM and SM-related enzymes without prior knowledge of pathway membership or gene–metabolite correlation information. Multiple features including gene expression, transcriptional network analysis, rate of evolution, and duplication mechanism allowed creation of statistical models that can distinguish GM from SM genes in *Arabidopsis*. In agreement with the established characteristics

of SM genes, machine learning models revealed that relative to GM genes, SM genes tend to be less conserved, tandemly duplicated, more narrowly expressed, and expressed at lower levels (Moore et al., 2019). The prediction models also facilitated the classification of 1220 enzyme encoding genes of unknown function as putatively SM-related. Similar machine learning strategies were deployed in tomato to predict gene association with SM or GM pathways and to determine if gene expression data can predict metabolic pathway membership (Moore et al., 2020; Wang et al., 2021). These approaches show potential to build high-quality models but are limited by the quality of the input data, including mis-annotations and the low number of functionally validated reference genes in tomato. These current limitations suggest that application of machine learning for *de novo* prediction of novel SM pathways in tomato is not yet possible at high accuracy. Furthermore, additional functional annotation, including the development of more comprehensive genome and transcriptome data, will be needed to apply machine learning approaches to predict SM pathway membership in additional members of the Solanaceae. Indeed, models predicting whether a tomato gene is associated with specialized versus general metabolism were improved when a transfer learning strategy was employed that utilized data from *Arabidopsis* models to filter tomato annotations that disagreed with *Arabidopsis* (Moore et al., 2020). This represents a promising approach to using comparative genomics data in specialized metabolic enzyme identification.

Challenges in the identification and annotation of plant metabolites.

Estimates suggest that $\sim 10^6$ metabolites are synthesized across species of the plant kingdom, collectively (Afendi et al., 2012). While we have deep knowledge of well-studied classes of plant metabolites, opportunities and challenges for improving metabolome annotation remain. Several factors make separation and annotation of metabolites challenging: for example,

their diverse chemical composition, chemical properties (polarity and hydrophilicity/hydrophobicity), and the orders of magnitude concentration range in which they occur in biological samples (Last et al., 2007; Perez de Souza et al., 2021). Improvements in analytical techniques, particularly liquid-chromatography coupled with high-resolution mass-spectrometry (LC-HRMS) based metabolite profiling, allows the detection of $>10^3$ metabolites within a single plant extract at high mass accuracy. However, a single extraction solvent and chromatographic separation method are generally selected for individual experiments, leading to unavoidable bias in the types of metabolites that are extracted and resolved and therefore an under-representation of the metabolome (Perez de Souza et al., 2021). Furthermore, most metabolites in a plant extract are uncharacterized and many are of low abundance. In such cases, annotation can be challenging. This is particularly true for specialized metabolites that are formed from diverse metabolic precursors, possess multiple chemical modifications, and frequently exist as positional or structural isomers that may be difficult to resolve. For example, even though tomato fruit ripening is one of the most extensively studied plant biological processes, a large component of this metabolome remains unannotated. In a recent study, untargeted metabolomics of tomato fruit at two different developmental stages identified >1000 semi-lipophilic metabolites but only ~170 metabolites were annotated with some degree of confidence, suggesting that the bulk of the tomato fruit metabolome remains unresolved (Szymański et al., 2020). Metabolite databases containing spectra derived from tandem mass-spectrometry of known metabolites are expanding and are useful for identifying unknown metabolites (Horai et al., 2010; Tsugawa et al., 2019; Wang et al., 2016). However, given the vast diversity of plant metabolites and their frequent lineage-specific distribution, populating and

curating such databases requires substantial research funding, effort, and community engagement.

As with spatially resolved or single cell transcriptomics, the ability to obtain spatially resolved metabolome data through mass spectrometry imaging of plant tissues represents an exciting development that will enhance understanding of metabolism. Specifically, this technology will further refine the ability to detect gene–metabolite correlations and allow the detection of metabolites that may be restricted to individual cell types and therefore fall below the limit of detection in an extract prepared from a complex tissue sample (Sumner et al., 2015). Mass spectrometry imaging has been utilized for investigating the spatial distribution of metabolites in tomato fruit, including investigating the influence of genetic perturbation on SGA accumulation (Dong et al., 2020). Similarly, the spatial separation of SGAs and acylsugars were demonstrated in tomato roots (Korenblum et al., 2020). As improved MSI technologies develop and increase in availability, they will undoubtedly be more widely adopted for exploring diverse aspects of Solanaceae metabolism.

Integration of genetic variation with metabolomics is a powerful approach to expand understanding of SM metabolic networks and bridge the gap between genotype and phenotype. As described above, both GWAS and metabolite QTL (mQTL) approaches were used to identify genomic regions and genes that influence specialized metabolism in diverse tissues of tomato. In particular, the *S. lycopersicum* x *S. pennellii* introgression line and the related backcross introgression line (BIL) populations were foundational to improving understanding of the loci that influence metabolism within the tomato clade (Alseekh et al., 2020; Cárdenas et al., 2019; Garbowicz et al., 2018; Schilmiller et al., 2010; Szymański et al., 2020). Approaches that harness natural variation are limited to species where it is possible to develop inter-specific genetic

populations or sufficient genetic variation is present within a species, to facilitate GWAS. Although not currently as extensively characterized as the genetic resources for tomato, germplasm panels and genetic populations, including introgression lines, are being developed and characterized for the three additional major food crops of the Solanaceae; potato, pepper, and eggplant (Gramazio et al., 2017; Hirsch et al., 2013; Tripodi et al., 2021). In some cases, these genetic resources are being utilized to investigate metabolic diversity via targeted and untargeted metabolomics and refinement of these efforts should facilitate linking genotype to phenotype (Levina et al., 2021; Sulli et al., 2021).

An alternative, less frequently utilized, approach to harness genetic variation to interrogate metabolism is to combine untargeted metabolite profiling with targeted disruption or over-expression of known enzymes or transcription factors (Tzin et al., 2012; Zhang et al., 2015). This approach, while more targeted than a strategy incorporating genome-wide genetic variation, can be utilized in any species where genetic manipulation is feasible and has significant potential to increase understanding of plant SM networks. For example, disruption of an SM enzyme will result in reduction of metabolites downstream of the enzyme, while the abundance of metabolites upstream of the target enzyme can increase. This approach also allows detection of alternate fates for pathway metabolites that accumulate due to gene disruption, revealing the existence of biosynthetically linked metabolites. Referred to as "silent metabolism" this component of the metabolome is likely substantial and certainly under-explored, including for engineering of novel products (Lewinsohn and Gijzen, 2009). Furthermore, as SM enzymes possess increased tendency for catalytic promiscuity, untargeted metabolite profiling of lines disrupted in an enzyme of interest may reveal the existence of previously uncharacterized catalytic activities.

While purification and structural elucidation of metabolites by NMR is a cornerstone of SM pathway discovery, it is time-consuming and typically represents a major bottleneck. This is especially problematic for metabolites that are of low abundance or co-purify with other compounds. Recent structural elucidation of acyl-hexoses from *S. nigrum* was achieved using a combination of LC-MS, GC-MS, and 2D-NMR approaches from crude and partially purified extracts without purification to homogeneity (Lou et al., 2021). Similar approaches should be adaptable to resolve the structures of other metabolites present in semi-purified plant extracts. The recent adoption of microcrystal electron diffraction (MicroED) for structural elucidation, including absolute stereochemistry, of mixtures of small organic molecules also shows great promise for structural elucidation of plant specialized metabolites (Gruene et al., 2018; Jones et al., 2018). MicroED can be used to resolve the structures of nanocrystals of ~100 nm (~10−15 g) and thus is potentially more suitable for low abundance metabolites than NMR, which typically requires hundreds of micrograms to milligram quantities of purified compound. Application of this technology to specialized metabolite discovery was recently demonstrated through a combined genome-mining, synthetic biology, and MicroED analysis that elucidated the biosynthesis and structures of several 2-pyrridone metabolites from fungi (Kim et al., 2021). Similarly, synthetic biology can be utilized to engineer production of plant SMs in heterologous systems for subsequent purification and structural elucidation. This strategy was effectively demonstrated by the synthesis of gram scale quantities of the triterpene β-amyrin by vacuum infiltration of *N. benthamiana* co-expressing a feedback insensitive variant of HMG-CoA reductase and oat β-amyrin synthase (Reed et al., 2017). Subsequent experiments combining co-expression of these enzymes with triterpene decorating cytochrome P450s from multiple species facilitated the production of novel non-natural triterpenes at sufficient scale to allow purification

and structural determination by NMR. *N. benthamiana* is widely used for transient expression of candidate genes and as demonstrated above, represents a readily scalable platform to produce metabolites for purification and subsequent structural elucidation.

Conclusions

Advances in genomics and metabolomics continue to enable greater understanding of SM pathway biosynthesis and evolution. This review focused on the catalytic steps of five well-studied SM classes that show varying degrees of lineage-specific distribution across the Solanaceae. This genetic variation, coupled with high abundance, and often restricted distribution in specific tissue or cell types, facilitated both purification and structural elucidation of these diverse metabolites as well as the identification of the enzymes responsible for their biosynthesis. For example, acylsugar and terpene biosynthesis in glandular trichomes, nicotine and tropane alkaloid biosynthesis in roots, and capsaicinoid biosynthesis in pepper fruit placenta. These studies reveal examples of both intra- and inter-specific variation as well as convergent evolution that has shaped the metabolic landscape across the Solanaceae. However, only a small fraction of the metabolome and the genes responsible for its formation are resolved. Thus, many opportunities exist to expand understanding of known pathways as well as identify novel pathways that will enable a network level understanding of metabolism across the Solanaceae and identify target molecules for agricultural and medicinal applications.

REFERENCES

Afendi, F.M., Okada, T., Yamazaki, M., Hirai-Morita, A., Nakamura, Y., Nakamura, K., Ikeda, S., Takahashi, H., Altaf-Ul-Amin, Md., Darusman, L.K., Saito, K., Kanaya, S., 2012. KNApSAcK family databases: integrated metabolite–plant species databases for multifaceted plant research. Plant Cell Physiol. 53, e1. https://doi.org/10.1093/pcp/pcr165

Akhtar, T.A., Matsuba, Y., Schauvinhold, I., Yu, G., Lees, H.A., Klein, S.E., Pichersky, E., 2013. The tomato cis–prenyltransferase gene family. Plant J. 73, 640–652. https://doi.org/10.1111/tpj.12063

Akiyama, R., Lee, H.J., Nakayasu, M., Osakabe, K., Osakabe, Y., Umemoto, N., Saito, K., Muranaka, T., Sugimoto, Y., Mizutani, M., 2019. Characterization of steroid 5α-reductase involved in α-tomatine biosynthesis in tomatoes. Plant Biotechnol. 36, 253–263. https://doi.org/10.5511/plantbiotechnology.19.1030a

Akiyama, R., Nakayasu, M., Umemoto, N., Kato, J., Kobayashi, M., Lee, H.J., Sugimoto, Y., Iijima, Y., Saito, K., Muranaka, T., Mizutani, M., 2021a. Tomato E8 encodes a C-27 hydroxylase in metabolic detoxification of α-tomatine during fruit ripening. Plant Cell Physiol. 62, 775–783. https://doi.org/10.1093/pcp/pcab080

Akiyama, R., Watanabe, B., Nakayasu, M., Lee, H.J., Kato, J., Umemoto, N., Muranaka, T., Saito, K., Sugimoto, Y., Mizutani, M., 2021b. The biosynthetic pathway of potato solanidanes diverged from that of spirosolanes due to evolution of a dioxygenase. Nat. Commun. 12, 1300. https://doi.org/10.1038/s41467-021-21546-0

Alonge, M., Wang, X., Benoit, M., Soyk, S., Pereira, L., Zhang, L., Suresh, H., Ramakrishnan, S., Maumus, F., Ciren, D., Levy, Y., Harel, T.H., Shalev-Schlosser, G., Amsellem, Z., Razifard, H., Caicedo, A.L., Tieman, D.M., Klee, H., Kirsche, M., Aganezov, S., Ranallo-Benavidez, T.R., Lemmon, Z.H., Kim, J., Robitaille, G., Kramer, M., Goodwin, S., McCombie, W.R., Hutton, S., Van Eck, J., Gillis, J., Eshed, Y., Sedlazeck, F.J., van der Knaap, E., Schatz, M.C., Lippman, Z.B., 2020. Major Impacts of Widespread Structural Variation on Gene Expression and Crop Improvement in Tomato. Cell 182, 145-161.e23. https://doi.org/10.1016/j.cell.2020.05.021

Alseekh, S., Ofner, I., Liu, Z., Osorio, S., Vallarino, J., Last, R.L., Zamir, D., Tohge, T., Fernie, A.R., 2020. Quantitative trait loci analysis of seed-specialized metabolites reveals seed-specific flavonols and differential regulation of glycoalkaloid content in tomato. Plant J. 103, 2007–2024. https://doi.org/10.1111/tpj.14879

Asai, T., Fujimoto, Y., 2011. 2-Acety-1-(3-glycosyloxyoctadecanoyl)glycerol and dammarane triterpenes in the exudates from glandular trichome-like secretory organs on the stipules and leaves of *Cerasus yedoensis*. Phytochem. Lett. 4, 38–42. https://doi.org/10.1016/j.phytol.2010.11.001

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

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

Asai, T., Sakai, T., Ohyama, K., Fujimoto, Y., 2011. *n*-Octyl α-L-rhamnopyranosyl-(1→2)-β-D-glucopyranoside derivatives from the glandular trichome exudate of *Geranium carolinianum*. Chem. Pharm. Bull. (Tokyo) 59, 747–752. https://doi.org/10.1248/cpb.59.747

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

Barchi, L., Pietrella, M., Venturini, L., Minio, A., Toppino, L., Acquadro, A., Andolfo, G., Aprea, G., Avanzato, C., Bassolino, L., Comino, C., Molin, A.D., Ferrarini, A., Maor, L.C., Portis, E., Reyes-Chin-Wo, S., Rinaldi, R., Sala, T., Scaglione, D., Sonawane, P., Tononi, P., Almekias-Siegl, E., Zago, E., Ercolano, M.R., Aharoni, A., Delledonne, M., Giuliano, G., Lanteri, S., Rotino, G.L., 2019. A chromosome-anchored eggplant genome sequence reveals key events in Solanaceae evolution. Sci. Rep. 9, 11769. https://doi.org/10.1038/s41598-019-47985-w

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

Bedewitz, M.A., Góngora-Castillo, E., Uebler, J.B., Gonzales-Vigil, E., Wiegert-Rininger, K.E., Childs, K.L., Hamilton, J.P., Vaillancourt, B., Yeo, Y.-S., Chappell, J., DellaPenna, D., Jones, A.D., Buell, C.R., Barry, C.S., 2014. A root-expressed l-phenylalanine:4-hydroxyphenylpyruvate aminotransferase is required for tropane alkaloid biosynthesis in *Atropa belladonna*. Plant Cell 26, 3745–3762. https://doi.org/10.1105/tpc.114.130534

Bedewitz, M.A., Jones, A.D., D'Auria, J.C., Barry, C.S., 2018. Tropinone synthesis via an atypical polyketide synthase and P450-mediated cyclization. Nat. Commun. 9, 5281. https://doi.org/10.1038/s41467-018-07671-3

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

Bleeker, P.M., Diergaarde, P.J., Ament, K., Schütz, S., Johne, B., Dijkink, J., Hiemstra, H., de Gelder, R., de Both, M.T.J., Sabelis, M.W., Haring, M.A., Schuurink, R.C., 2011. Tomato-produced 7-epizingiberene and R-curcumene act as repellents to whiteflies. Phytochemistry 72, 68–73. https://doi.org/10.1016/j.phytochem.2010.10.014

Bleeker, P.M., Mirabella, R., Diergaarde, P.J., VanDoorn, A., Tissier, A., Kant, M.R., Prins, M., de Vos, M., Haring, M.A., Schuurink, R.C., 2012. Improved herbivore resistance in cultivated tomato with the sesquiterpene biosynthetic pathway from a wild relative. Proc. Natl. Acad. Sci. U.S.A. 109, 20124–20129. https://doi.org/10.1073/pnas.1208756109

Boutanaev, A.M., Moses, T., Zi, J., Nelson, D.R., Mugford, S.T., Peters, R.J., Osbourn, A., 2015. Investigation of terpene diversification across multiple sequenced plant genomes. Proc. Natl. Acad. Sci. U.S.A. 112, E81–E88. https://doi.org/10.1073/pnas.1419547112

Brock, A., Brandt, W., Dräger, B., 2008. The functional divergence of short-chain dehydrogenases involved in tropinone reduction. Plant J. 54, 388–401. https://doi.org/10.1111/j.1365-313X.2008.03422.x

Brooks, C., Nekrasov, V., Lippman, Z.B., Van Eck, J., 2014. Efficient gene editing in tomato in the first generation using the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-Associated9 System. Plant Physiol. 166, 1292–1297. https://doi.org/10.1104/pp.114.247577

Bunzel, H.A., Garrabou, X., Pott, M., Hilvert, D., 2018. Speeding up enzyme discovery and engineering with ultrahigh-throughput methods. Curr. Opin. Struct. Biol. 48, 149–156. https://doi.org/10.1016/j.sbi.2017.12.010

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

Cárdenas, P.D., Sonawane, P.D., Heinig, U., Jozwiak, A., Panda, S., Abebie, B., Kazachkova, Y., Pliner, M., Unger, T., Wolf, D., Ofner, I., Vilaprinyo, E., Meir, S., Davydov, O., Gal-on, A., Burdman, S., Giri, A., Zamir, D., Scherf, T., Szymanski, J., Rogachev, I., Aharoni, A., 2019. Pathways to defense metabolites and evading fruit bitterness in genus Solanum evolved through 2-oxoglutarate-dependent dioxygenases. Nat. Commun. 10, 5169. https://doi.org/10.1038/s41467-019-13211-4

Carrizo García, C., Barfuss, M.H.J., Sehr, E.M., Barboza, G.E., Samuel, R., Moscone, E.A., Ehrendorfer, F., 2016. Phylogenetic relationships, diversification and expansion of chili peppers (Capsicum, Solanaceae). Ann. Bot. 118, 35–51. https://doi.org/10.1093/aob/mcw079

Caterina, M.J., Leffler, A., Malmberg, A.B., Martin, W.J., Trafton, J., Petersen-Zeitz, K.R., Koltzenburg, M., Basbaum, A.I., Julius, D., 2000. Impaired nociception and pain sensation in mice lacking the capsaicin receptor. Science 288, 306–313. https://doi.org/10.1126/science.288.5464.306

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

Chortyk, O.T., Severson, R.F., Cutler, H.C., Sisson, V.A., 1993. Antibiotic activities of sugar esters isolated from selected Nicotiana species. Biosci. Biotechnol. Biochem. 57, 1355–1356. https://doi.org/10.1271/bbb.57.1355

Christen, P., Cretton, S., Humam, M., Bieri, S., Muñoz, O., Joseph-Nathan, P., 2020. Chemistry and biological activity of alkaloids from the genus Schizanthus. Phytochem. Rev. 19, 615–641. https://doi.org/10.1007/s11101-018-9598-5

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. https://doi.org/10.1021/acs.jnatprod.7b01069

Clissold, S.P., Heel, R.C., 1985. Transdermal hyoscine (scopolamine). Drugs 29, 189–207. https://doi.org/10.2165/00003495-198529030-00001

Coates, R.M., Denissen, J.F., Juvik, J.A., Babka, B.A., 1988. Identification of alpha-santalenoic and endo-beta-bergamotenoic acids as moth oviposition stimulants from wild tomato leaves. J. Org. Chem. 53, 2186–2192. https://doi.org/10.1021/jo00245a012

Courdavault, V., Papon, N., Clastre, M., Giglioli-Guivarc'h, N., St-Pierre, B., Burlat, V., 2014. A look inside an alkaloid multisite plant: the Catharanthus logistics. Curr. Opin. Plant Biol., SI: Physiology and metabolism 19, 43–50. https://doi.org/10.1016/j.pbi.2014.03.010

Dolan, L.C., Matulka, R.A., Burdock, G.A., 2010. Naturally occurring food toxins. Toxins 2, 2289–2332. https://doi.org/10.3390/toxins2092289

Dong, Y., Sonawane, P., Cohen, H., Polturak, G., Feldberg, L., Avivi, S.H., Rogachev, I., Aharoni, A., 2020. High mass resolution, spatial metabolite mapping enhances the current plant gene and pathway discovery toolbox. New Phytol. 228, 1986–2002. https://doi.org/10.1111/nph.16809

Duranova, H., Valkova, V., Gabriny, L., 2022. Chili peppers (Capsicum spp.): the spice not only for cuisine purposes: an update on current knowledge. Phytochem. Rev. 21, 1379–1413. https://doi.org/10.1007/s11101-021-09789-7

Eshed, Y., Zamir, D., 1995. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141, 1147–1162. https://doi.org/10.1093/genetics/141.3.1147

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. https://doi.org/10.1038/s41467-017-02045-7

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

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, 1–26. https://doi.org/10.7554/eLife.56717

Fossati, E., Ekins, A., Narcross, L., Zhu, Y., Falgueyret, J.-P., Beaudoin, G.A.W., Facchini, P.J., Martin, V.J.J., 2014. Reconstitution of a 10-gene pathway for synthesis of the plant alkaloid dihydrosanguinarine in *Saccharomyces cerevisiae*. Nat. Commun. 5, 3283. https://doi.org/10.1038/ncomms4283

Frelichowski, J.E., Jr., Juvik, J.A., 2001. Sesquiterpene carboxylic acids from a wild tomato species affect larval feeding behavior and survival of *Helicoverpa zea* and *Spodoptera exigua* (Lepidoptera: Noctuidae). J. Econ. Entomol. 94, 1249–1259. https://doi.org/10.1603/0022-0493-94.5.1249

Fridman, E., Carrari, F., Liu, Y.-S., Fernie, A.R., Zamir, D., 2004. Zooming in on a quantitative trait for tomato yield using interspecific introgressions. Science 305, 1786–1789. https://doi.org/10.1126/science.1101666

Friedman, J.R., Richbart, S.D., Merritt, J.C., Brown, K.C., Denning, K.L., Tirona, M.T., Valentovic, M.A., Miles, S.L., Dasgupta, P., 2019. Capsaicinoids: multiple effects on angiogenesis, invasion and metastasis in human cancers. Biomed. Pharmacother. 118, 109317. https://doi.org/10.1016/j.biopha.2019.109317

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

Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D.M., Burzynski-Chang, E.A., Fish, T.L., Stromberg, K.A., Sacks, G.L., Thannhauser, T.W., Foolad, M.R., Diez, M.J., Blanca, J., Canizares, J., Xu, Y., van der Knaap, E., Huang, S., Klee, H.J., Giovannoni, J.J., Fei, Z., 2019. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. Nat. Genet. 51, 1044–1051. https://doi.org/10.1038/s41588-019-0410-2

Garbowicz, K., Liu, Z., Alseekh, S., Tieman, D., Taylor, M., Kuhalskaya, A., Ofner, I., Zamir, D., Klee, H.J., Fernie, A.R., Brotman, Y., 2018. Quantitative trait loci analysis identifies a prominent gene involved in the production of fatty acid-derived flavor volatiles in tomato. Mol. Plant 11, 1147–1165. https://doi.org/10.1016/j.molp.2018.06.003

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

Goffreda, J.C., Mutschler, M.A., Avé, D.A., Tingey, W.M., Steffens, J.C., 1989. Aphid deterrence by glucose esters in glandular trichome exudate of the wild tomato,

Lycopersicon pennellii. J. Chem. Ecol. 15, 2135–2147.
https://doi.org/10.1007/BF01207444

Gonzales-Vigil, E., Hufnagel, D.E., Kim, J., Last, R.L., Barry, C.S., 2012. Evolution of TPS20-related terpene synthases influences chemical diversity in the glandular trichomes of the wild tomato relative *Solanum habrochaites*. Plant J. 71, 921–935.
https://doi.org/10.1111/j.1365-313X.2012.05040.x

Gramazio, P., Prohens, J., Plazas, M., Mangino, G., Herraiz, F.J., Vilanova, S., 2017. Development and genetic characterization of advanced backcross materials and an introgression line population of *Solanum incanum* in a *S. melongena* background. Front. Plant Sci. 8.

Gruene, T., Wennmacher, J.T.C., Zaubitzer, C., Holstein, J.J., Heidler, J., Fecteau-Lefebvre, A., De Carlo, S., Müller, E., Goldie, K.N., Regeni, I., Li, T., Santiso-Quinones, G., Steinfeld, G., Handschin, S., van Genderen, E., van Bokhoven, J.A., Clever, G.H., Pantelic, R., 2018. Rapid structure determination of microcrystalline molecular compounds using electron diffraction. Angew. Chem. Int. Ed. 57, 16313-16317.
https://doi.org/10.1002/anie.201811318

Gu, X.-Y., Shen, X.-F., Wang, L., Wu, Z.-W., Li, F., Chen, B., Zhang, G.-L., Wang, M.-K., 2018. Bioactive steroidal alkaloids from the fruits of *Solanum nigrum*. Phytochemistry 147, 125–131. https://doi.org/10.1016/j.phytochem.2017.12.020

Han, K., Lee, H.-Y., Ro, N.-Y., Hur, O.-S., Lee, J.-H., Kwon, J.-K., Kang, B.-C., 2018. QTL mapping and GWAS reveal candidate genes controlling capsaicinoid content in Capsicum. Plant Biotechnol. J. 16, 1546–1558. https://doi.org/10.1111/pbi.12894

Hashimoto, T., Yamada, Y., 1986. Hyoscyamine 6β-hydroxylase, a 2-oxoglutarate-dependent dioxygenase, in alkaloid-producing root cultures. Plant Physiol. 81, 619–625.
https://doi.org/10.1104/pp.81.2.619

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

Hirsch, C.N., Hirsch, C.D., Felcher, K., Coombs, J., Zarka, D., Van Deynze, A., De Jong, W., Veilleux, R.E., Jansky, S., Bethke, P., Douches, D.S., Buell, C.R., 2013. Retrospective view of North American potato (*Solanum tuberosum* L.) breeding in the 20th and 21st centuries. G3: Genes Genomes Genet. 3, 1003–1013.
https://doi.org/10.1534/g3.113.005595

Horai, H., Arita, M., Kanaya, S., Nihei, Y., Ikeda, T., Suwa, K., Ojima, Y., Tanaka, Kenichi, Tanaka, S., Aoshima, K., Oda, Y., Kakazu, Y., Kusano, M., Tohge, T., Matsuda, F., Sawada, Y., Hirai, M.Y., Nakanishi, H., Ikeda, K., Akimoto, N., Maoka, T., Takahashi, H., Ara, T., Sakurai, N., Suzuki, H., Shibata, D., Neumann, S., Iida, T., Tanaka, Ken, Funatsu, K., Matsuura, F., Soga, T., Taguchi, R., Saito, K., Nishioka, T., 2010.

MassBank: a public repository for sharing mass spectral data for life sciences. J. Mass Spectrom. 45, 703–714. https://doi.org/10.1002/jms.1777

Huang, J.-P., Fang, C., Ma, X., Wang, L., Yang, J., Luo, J., Yan, Y., Zhang, Y., Huang, S.-X., 2019. Tropane alkaloids biosynthesis involves an unusual type III polyketide synthase and non-enzymatic condensation. Nat. Commun. 10, 4036. https://doi.org/10.1038/s41467-019-11987-z

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

Iijima, Y., Watanabe, B., Sasaki, R., Takenaka, M., Ono, H., Sakurai, N., Umemoto, N., Suzuki, H., Shibata, D., Aoki, K., 2013. Steroidal glycoalkaloid profiling and structures of glycoalkaloids in wild tomato fruit. Phytochemistry 95, 145–157. https://doi.org/10.1016/j.phytochem.2013.07.016

Itkin, M., Heinig, U., Tzfadia, O., Bhide, A.J., Shinde, B., Cardenas, P.D., Bocobza, S.E., Unger, T., Malitsky, S., Finkers, R., Tikunov, Y., Bovy, A., Chikate, Y., Singh, P., Rogachev, I., Beekwilder, J., Giri, A.P., Aharoni, A., 2013. Biosynthesis of antinutritional alkaloids in Solanaceous crops is mediated by clustered genes. Science 341, 175–179. https://doi.org/10.1126/science.1240230

Jacobowitz, J.R., Weng, J.-K., 2020. Exploring uncharted territories of plant specialized metabolism in the postgenomic era. Annu. Rev. Plant Biol. 71, 631–658. https://doi.org/10.1146/annurev-arplant-081519-035634

Jirschitzka, J., Schmidt, G.W., Reichelt, M., Schneider, B., Gershenzon, J., D'Auria, J.C., 2012. Plant tropane alkaloid biosynthesis evolved independently in the Solanaceae and Erythroxylaceae. Proc. Natl. Acad. Sci. U.S.A. 109, 10304–10309. https://doi.org/10.1073/pnas.1200473109

Jones, C.G., Martynowycz, M.W., Hattne, J., Fulton, T.J., Stoltz, B.M., Rodriguez, J.A., Nelson, H.M., Gonen, T., 2018. The cryoEM method microED as a powerful tool for small molecule structure determination. ACS Cent. Sci. 4, 1587–1592. https://doi.org/10.1021/acscentsci.8b00760

Jordt, S.-E., Julius, D., 2002. Molecular basis for species-specific sensitivity to "hot" chili peppers. Cell 108, 421–430. https://doi.org/10.1016/S0092-8674(02)00637-2

Junker, A., Fischer, J., Sichhart, Y., Brandt, W., Draeger, B., 2013. Evolution of the key alkaloid enzyme putrescine N-methyltransferase from spermidine synthase. Front. Plant Sci. 4.

Kajikawa, M., Sierro, N., Kawaguchi, H., Bakaher, N., Ivanov, N.V., Hashimoto, T., Shoji, T., 2017. Genomic insights into the evolution of the nicotine biosynthesis pathway in tobacco. Plant Physiol. 174, 999–1011. https://doi.org/10.1104/pp.17.00070

Kang, J.-H., Gonzales-Vigil, E., Matsuba, Y., Pichersky, E., Barry, C.S., 2014. Determination of residues responsible for substrate and product specificity of *Solanum habrochaites* short-

chain cis-prenyltransferases. Plant Physiol. 164, 80–91.
https://doi.org/10.1104/pp.113.230466

Karunanithi, P.S., Zerbe, P., 2019. Terpene synthases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity. Front. Plant Sci. 10.

Kazachkova, Y., Zemach, I., Panda, S., Bocobza, S., Vainer, A., Rogachev, I., Dong, Y., Ben-Dor, S., Veres, D., Kanstrup, C., Lambertz, S.K., Crocoll, C., Hu, Y., Shani, E., Michaeli, S., Nour-Eldin, H.H., Zamir, D., Aharoni, A., 2021. The GORKY glycoalkaloid transporter is indispensable for preventing tomato bitterness. Nat. Plants 7, 468–480. https://doi.org/10.1038/s41477-021-00865-6

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

Kim, L.J., Ohashi, M., Zhang, Z., Tan, D., Asay, M., Cascio, D., Rodriguez, J.A., Tang, Y., Nelson, H.M., 2021. Prospecting for natural products by genome mining and microcrystal electron diffraction. Nat. Chem. Biol. 17, 872–877. https://doi.org/10.1038/s41589-021-00834-2

Kim, N., Estrada, O., Chavez, B., Stewart, C., D'Auria, J.C., 2016. Tropane and granatane alkaloid biosynthesis: a systematic analysis. Molecules 21, 1510. https://doi.org/10.3390/molecules21111510

Kim, S., Park, M., Yeom, S.-I., Kim, Y.-M., Lee, J.M., Lee, H.-A., Seo, E., Choi, J., Cheong, K., Kim, K.-T., Jung, K., Lee, G.-W., Oh, S.-K., Bae, C., Kim, S.-B., Lee, H.-Y., Kim, S.-Y., Kim, M.-S., Kang, B.-C., Jo, Y.D., Yang, H.-B., Jeong, H.-J., Kang, W.-H., Kwon, J.-K., Shin, C., Lim, J.Y., Park, J.H., Huh, J.H., Kim, J.-S., Kim, B.-D., Cohen, O., Paran, I., Suh, M.C., Lee, S.B., Kim, Y.-K., Shin, Y., Noh, S.-J., Park, J., Seo, Y.S., Kwon, S.-Y., Kim, H.A., Park, J.M., Kim, H.-J., Choi, S.-B., Bosland, P.W., Reeves, G., Jo, S.-H., Lee, B.-W., Cho, H.-T., Choi, H.-S., Lee, M.-S., Yu, Y., Do Choi, Y., Park, B.-S., van Deynze, A., Ashrafi, H., Hill, T., Kim, W.T., Pai, H.-S., Ahn, H.K., Yeam, I., Giovannoni, J.J., Rose, J.K.C., Sørensen, I., Lee, S.-J., Kim, R.W., Choi, I.-Y., Choi, B.-S., Lim, J.-S., Lee, Y.-H., Choi, D., 2014. Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. Nat. Genet. 46, 270–278. https://doi.org/10.1038/ng.2877

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

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

Knapp, S., Bohs, L., Nee, M., Spooner, D.M., 2004. Solanaceae—a model for linking genomics with biodiversity. Comp. Funct. Genomics 5, 285–291. https://doi.org/10.1002/cfg.393

Koeda, S., Sato, K., Saito, H., Nagano, A.J., Yasugi, M., Kudoh, H., Tanaka, Y., 2019. Mutation in the putative ketoacyl-ACP reductase CaKR1 induces loss of pungency in Capsicum. Theor. Appl. Genet. 132, 65–80. https://doi.org/10.1007/s00122-018-3195-2

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

Landis, J.B., Miller, C.M., Broz, A.K., Bennett, A.A., Carrasquilla-Garcia, N., Cook, D.R., Last, R.L., Bedinger, P.A., Moghe, G.D., 2021. Migration through a major Andean ecogeographic disruption as a driver of genetic and phenotypic diversity in a wild tomato species. Mol. Biol. Evol. 38, 3202–3219. https://doi.org/10.1093/molbev/msab092

Lang, Y., Kisaka, H., Sugiyama, R., Nomura, K., Morita, A., Watanabe, T., Tanaka, Y., Yazawa, S., Miwa, T., 2009. Functional loss of pAMT results in biosynthesis of capsinoids, capsaicinoid analogs, in *Capsicum annuum* cv. CH-19 Sweet. Plant J. 59, 953–961. https://doi.org/10.1111/j.1365-313X.2009.03921.x

Last, R.L., Jones, A.D., Shachar-Hill, Y., 2007. Towards the plant metabolome and beyond. Nat. Rev. Mol. Cell Biol. 8, 167–174. https://doi.org/10.1038/nrm2098

Lau, W., Sattely, E.S., 2015. Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. Science 349, 1224–1228. https://doi.org/10.1126/science.aac7202

Leckie, B.M., D'Ambrosio, D.A., Chappell, T.M., Halitschke, R., Jong, D.M.D., Kessler, A., Kennedy, G.G., Mutschler, M.A., 2016. Differential and synergistic functionality of]acylsugars in suppressing oviposition by insect herbivores. PLOS ONE 11, e0153345. https://doi.org/10.1371/journal.pone.0153345

Lee, H.J., Nakayasu, M., Akiyama, R., Kobayashi, M., Miyachi, H., Sugimoto, Y., Umemoto, N., Saito, K., Muranaka, T., Mizutani, M., 2019. Identification of a 3β-hydroxysteroid dehydrogenase/ 3-ketosteroid reductase involved in α-tomatine biosynthesis in tomato. Plant Cell Physiol. 60, 1304–1315. https://doi.org/10.1093/pcp/pcz049

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

Leong, B.J., Last, R.L., 2017. Promiscuity, impersonation and accommodation: evolution of plant specialized metabolism. Curr. Opin. Struct. Biol. 47, 105–112. https://doi.org/10.1016/j.sbi.2017.07.005

Leong, B.J., Lybrand, D.B., Lou, Y.-R., Fan, P., Schilmiller, A.L., Last, R.L., 2019. Evolution of metabolic novelty: a trichome-expressed invertase creates specialized metabolic diversity in wild tomato. Sci. Adv. 5, eaaw3754. https://doi.org/10.1126/sciadv.aaw3754

Levina, A.V., Hoekenga, O., Gordin, M., Broeckling, C., De Jong, W.S., 2021. Genetic analysis of potato tuber metabolite composition: genome-wide association studies applied to a nontargeted metabolome. Crop Sci. 61, 591–603. https://doi.org/10.1002/csc2.20398

Lewinsohn, E., Gijzen, M., 2009. Phytochemical diversity: the sounds of silent metabolism. Plant Sci. 176, 161–169. https://doi.org/10.1016/j.plantsci.2008.09.018

Li, C., Wang, Z., Jones, A.D., 2014. Chemical imaging of trichome specialized metabolites using contact printing and laser desorption/ionization mass spectrometry. Anal. Bioanal. Chem. 406, 171–182. https://doi.org/10.1007/s00216-013-7444-6

Li, R., Reed, D.W., Liu, E., Nowak, J., Pelcher, L.E., Page, J.E., Covello, P.S., 2006. Functional genomic analysis of alkaloid biosynthesis in *Hyoscyamus niger* reveals a cytochrome P450 involved in littorine rearrangement. Chem. Biol. 13, 513–520. https://doi.org/10.1016/j.chembiol.2006.03.005

Li, X., Tieman, D., Liu, Z., Chen, K., Klee, H.J., 2020. Identification of a lipase gene with a role in tomato fruit short-chain fatty acid-derived flavor volatiles by genome-wide association. Plant J. 104, 631–644. https://doi.org/10.1111/tpj.14951

Li, Y., Chen, Yang, Zhou, L., You, S., Deng, H., Chen, Ya, Alseekh, S., Yuan, Y., Fu, R., Zhang, Z., Su, D., Fernie, A.R., Bouzayen, M., Ma, T., Liu, M., Zhang, Y., 2020. MicroTom metabolic network: rewiring tomato metabolic regulatory network throughout the growth cycle. Mol. Plant 13, 1203–1218. https://doi.org/10.1016/j.molp.2020.06.005

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

Liu, Y., Schiff, M., Dinesh-Kumar, S.P., 2002. Virus-induced gene silencing in tomato. Plant J. 31, 777–786. https://doi.org/10.1046/j.1365-313X.2002.01394.x

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

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

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

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

Mato, A., Limeres, J., Tomás, I., Muñoz, M., Abuín, C., Feijoo, J.F., Diz, P., 2010. Management of drooling in disabled patients with scopolamine patches. Br. J. Clin. Pharmacol. 69, 684–688. https://doi.org/10.1111/j.1365-2125.2010.03659.x

Matsuba, Y., Nguyen, T.T.H., Wiegert, K., Falara, V., Gonzales-Vigil, E., Leong, B., Schäfer, P., Kudrna, D., Wing, R.A., Bolger, A.M., Usadel, B., Tissier, A., Fernie, A.R., Barry, C.S., Pichersky, E., 2013. Evolution of a complex locus for terpene biosynthesis in Solanum. Plant Cell 25, 2022–2036. https://doi.org/10.1105/tpc.113.111013

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

Michael, T.P., VanBuren, R., 2020. Building near-complete plant genomes. Curr. Opin. Plant Biol. 54, 26–33. https://doi.org/10.1016/j.pbi.2019.12.009

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

Moore, B.M., Wang, P., Fan, P., Lee, A., Leong, B., Lou, Y.-R., Schenck, C.A., Sugimoto, K., Last, R., Lehti-Shiu, M.D., Barry, C.S., Shiu, S.-H., 2020. Within- and cross-species predictions of plant specialized metabolism genes using transfer learning. Silico Plants 2, diaa005. https://doi.org/10.1093/insilicoplants/diaa005

Moore, B.M., Wang, P., Fan, P., Leong, B., Schenck, C.A., Lloyd, J.P., Lehti-Shiu, M.D., Last, R.L., Pichersky, E., Shiu, S.-H., 2019. Robust predictions of specialized metabolism genes through machine learning. Proc. Natl. Acad. Sci. U.S.A. 116, 2344–2353. https://doi.org/10.1073/pnas.1817074116

Mueller, L.A., Tanksley, S.D., Giovannoni, J.J., van Eck, J., Stack, S., Choi, D., Kim, B.D., Chen, M., Cheng, Z., Li, C., Ling, H., Xue, Y., Seymour, G., Bishop, G., Bryan, G., Sharma, R., Khurana, J., Tyagi, A., Chattopadhyay, D., Singh, N.K., Stiekema, W., Lindhout, P., Jesse, T., Lankhorst, R.K., Bouzayen, M., Shibata, D., Tabata, S., Granell, A., Botella, M.A., Giuliano, G., Frusciante, L., Causse, M., Zamir, D., 2005. The omato Sequencing Project, the first cornerstone of the International Solanaceae Project (SOL). Comp. Funct. Genomics 6, 153–158. https://doi.org/10.1002/cfg.468

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

Nakajima, K., Hashimoto, T., Yamada, Y., 1993. Two tropinone reductases with different stereospecificities are short-chain dehydrogenases evolved from a common ancestor. Proc. Natl. Acad. Sci. U.S.A. 90, 9591–9595. https://doi.org/10.1073/pnas.90.20.9591

Nakashima, T., Wada, H., Morita, S., Erra-Balsells, R., Hiraoka, K., Nonami, H., 2016. Single-cell metabolite profiling of stalk and glandular cells of intact trichomes with internal electrode capillary pressure probe electrospray ionization mass spectrometry. Anal. Chem. 88, 3049–3057. https://doi.org/10.1021/acs.analchem.5b03366

Nakayasu, M., Akiyama, R., Kobayashi, M., Lee, H.J., Kawasaki, T., Watanabe, B., Urakawa, S., Kato, J., Sugimoto, Y., Iijima, Y., Saito, K., Muranaka, T., Umemoto, N., Mizutani, M., 2020. Identification of α-tomatine 23-hydroxylase involved in the detoxification of a bitter glycoalkaloid. Plant Cell Physiol. 61, 21–28. https://doi.org/10.1093/pcp/pcz224

Nash, R.J., Rothschild, M., Porter, E.A., Watson, A.A., Waigh, R.D., Waterman, P.G., 1993. Calystegines in Solanum and Datura species and the death's-head hawk-moth (*Acherontia atropus*). Phytochemistry, The International Journal of Plant Biochemistry 34, 1281–1283. https://doi.org/10.1016/0031-9422(91)80016-T

Naves, E.R., Silva, L. de Á., Sulpice, R., Araújo, W.L., Nunes-Nesi, A., Peres, L.E.P., Zsögön, A., 2019. Capsaicinoids: pungency beyond Capsicum. Trends Plant Sci. 24, 109–120. https://doi.org/10.1016/j.tplants.2018.11.001

Nett, R.S., Dho, Y., Low, Y.-Y., Sattely, E.S., 2021. A metabolic regulon reveals early and late acting enzymes in neuroactive Lycopodium alkaloid biosynthesis. Proc. Natl. Acad. Sci. U.S.A. 118, e2102949118. https://doi.org/10.1073/pnas.2102949118

Nett, R.S., Lau, W., Sattely, E.S., 2020. Discovery and engineering of colchicine alkaloid biosynthesis. Nature 584, 148–153. https://doi.org/10.1038/s41586-020-2546-8

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

Ofner, I., Lashbrooke, J., Pleban, T., Aharoni, A., Zamir, D., 2016. *Solanum pennellii* backcross inbred lines (BILs) link small genomic bins with tomato traits. Plant J. 87, 151–160. https://doi.org/10.1111/tpj.13194

Ono, M., Takigawa, A., Muto, H., Kabata, K., Okawa, M., Kinjo, J., Yokomizo, K., Yoshimitsu, H., Nohara, T., 2015. Antiviral activity of four new resin glycosides Calysolins XIV–XVII from *Calystegia soldanella* against herpes simplex virus. Chem. Pharm. Bull. (Tokyo) 63, 641–648. https://doi.org/10.1248/cpb.c15-00307

69

Onoyovwe, A., Hagel, J.M., Chen, X., Khan, M.F., Schriemer, D.C., Facchini, P.J., 2013. Morphine biosynthesis in opium poppy involves two cell types: sieve elements and laticifers. Plant Cell 25, 4110–4122. https://doi.org/10.1105/tpc.113.115113

Park, M., Lee, J.-H., Han, K., Jang, S., Han, J., Lim, J.-H., Jung, J.-W., Kang, B.-C., 2019. A major QTL and candidate genes for capsaicinoid biosynthesis in the pericarp of *Capsicum chinense* revealed using QTL-seq and RNA-seq. Theor. Appl. Genet. 132, 515–529. https://doi.org/10.1007/s00122-018-3238-8

Paudel, J.R., Davidson, C., Song, J., Maxim, I., Aharoni, A., Tai, H.H., 2017. Pathogen and pest responses are altered due to RNAi-mediated knockdown of GLYCOALKALOID METABOLISM 4 in *Solanum tuberosum*. Mol. Plant-Microbe Interactions 30, 876–885. https://doi.org/10.1094/MPMI-02-17-0033-R

Pazouki, L., Niinemets, Ü., 2016. Multi-substrate terpene synthases: their occurrence and physiological significance. Front. Plant Sci. 7.

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

Perez de Souza, L., Alseekh, S., Scossa, F., Fernie, A.R., 2021. Ultra-high-performance liquid chromatography high-resolution mass spectrometry variants for metabolomics research. Nat. Methods 18, 733–746. https://doi.org/10.1038/s41592-021-01116-4

Pérez, L.M., Farriols, C., Puente, V., Planas, J., Ruiz, I., 2011. The use of subcutaneous scopolamine as a palliative treatment in Parkinson's disease. Palliat. Med. 25, 92–93. https://doi.org/10.1177/0269216310381662

Qiu, F., Yan, Y., Zeng, J., Huang, J.-P., Zeng, L., Zhong, W., Lan, X., Chen, M., Huang, S.-X., Liao, Z., 2021. Biochemical and metabolic insights into hyoscyamine dehydrogenase. ACS Catal. 11, 2912–2924. https://doi.org/10.1021/acscatal.0c04667

Qiu, F., Yang, C., Yuan, L., Xiang, D., Lan, X., Chen, M., Liao, Z., 2018. A phenylpyruvic acid reductase is required for biosynthesis of tropane alkaloids. Org. Lett. 20, 7807–7810. https://doi.org/10.1021/acs.orglett.8b03236

Qiu, F., Zeng, J., Wang, J., Huang, J.-P., Zhou, W., Yang, C., Lan, X., Chen, M., Huang, S.-X., Kai, G., Liao, Z., 2020. Functional genomics analysis reveals two novel genes required for littorine biosynthesis. New Phytol. 225, 1906–1914. https://doi.org/10.1111/nph.16317

Reed, J., Stephenson, M.J., Miettinen, K., Brouwer, B., Leveau, A., Brett, P., Goss, R.J.M., Goossens, A., O'Connell, M.A., Osbourn, A., 2017. A translational synthetic biology platform for rapid access to gram-scale quantities of novel drug-like molecules. Metab. Eng. 42, 185–193. https://doi.org/10.1016/j.ymben.2017.06.012

Roddick, J.G., Weissenberg, M., Leonard, A.L., 2001. Membrane disruption and enzyme inhibition by naturally-occurring and modified chacotriose-containing Solanum steroidal glycoalkaloids. Phytochemistry 56, 603–610. https://doi.org/10.1016/S0031-9422(00)00420-9

Ryu, K.H., Huang, L., Kang, H.M., Schiefelbein, J., 2019. Single-cell RNA sequencing resolves molecular relationships among individual plant cells. Plant Physiol. 179, 1444–1456. https://doi.org/10.1104/pp.18.01482

Sallaud, C., Rontein, D., Onillon, S., Jabès, F., Duffé, P., Giacalone, C., Thoraval, S., Escoffier, C., Herbette, G., Leonhardt, N., Causse, M., Tissier, A., 2009. A novel pathway for sesquiterpene biosynthesis from Z,Z-farnesyl pyrophosphate in the wild tomato *Solanum habrochaites*. Plant Cell 21, 301–317. https://doi.org/10.1105/tpc.107.057885

Sánchez-Mata, M.-C., Yokoyama, W.E., Hong, Y.-J., Prohens, J., 2010. α-Solasonine and α-solamargine contents of gboma (*Solanum macrocarpon* L.) and scarlet (*Solanum aethiopicum* L.) eggplants. J. Agric. Food Chem. 58, 5502–5508. https://doi.org/10.1021/jf100709g

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

The Tomato Genome Consortium, 2012. The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485, 635–641. https://doi.org/10.1038/nature11119

Schenck, C.A., Last, R.L., 2020. Location, location! cellular relocalization primes specialized metabolic diversification. FEBS J. 287, 1359–1368. https://doi.org/10.1111/febs.15097

Schijlen, E.G.W.M., de Vos, C.H.R., Martens, S., Jonker, H.H., Rosin, F.M., Molthoff, J.W., Tikunov, Y.M., Angenent, G.C., van Tunen, A.J., Bovy, A.G., 2007. RNA interference silencing of chalcone synthase, the first step in the flavonoid biosynthesis pathway, leads to parthenocarpic tomato fruits. Plant Physiol. 144, 1520–1530. https://doi.org/10.1104/pp.107.100305

Schilmiller, A., Shi, F., Kim, J., Charbonneau, A.L., Holmes, D., Daniel Jones, A., Last, R.L., 2010. Mass spectrometry screening reveals widespread diversity in trichome specialized metabolites of tomato chromosomal substitution lines. Plant J. 62, 391–403. https://doi.org/10.1111/j.1365-313X.2010.04154.x

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

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

Schilmiller, A.L., Schauvinhold, I., Larson, M., Xu, R., Charbonneau, A.L., Schmidt, A., Wilkerson, C., Last, R.L., Pichersky, E., 2009. Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. Proc. Natl. Acad. Sci. U.S.A. 106, 10865–10870. https://doi.org/10.1073/pnas.0904113106

Schmidt, G.W., Jirschitzka, J., Porta, T., Reichelt, M., Luck, K., Torre, J.C.P., Dolke, F., Varesio, E., Hopfgartner, G., Gershenzon, J., D'Auria, J.C., 2015. The last step in cocaine biosynthesis is catalyzed by a BAHD acyltransferase. Plant Physiol. 167, 89–101. https://doi.org/10.1104/pp.114.248187

Seyfferth, C., Renema, J., Wendrich, J.R., Eekhout, T., Seurinck, R., Vandamme, N., Blob, B., Saeys, Y., Helariutta, Y., Birnbaum, K.D., De Rybel, B., 2021. Advances and opportunities in single-cell transcriptomics for plant research. Annu. Rev. Plant Biol. 72, 847–866. https://doi.org/10.1146/annurev-arplant-081720-010120

Shinde, B.A., Dholakia, B.B., Hussain, K., Panda, S., Meir, S., Rogachev, I., Aharoni, A., Giri, A.P., Kamble, A.C., 2017. Dynamic metabolic reprogramming of steroidal glycol-alkaloid and phenylpropanoid biosynthesis may impart early blight resistance in wild tomato (*Solanum arcanum* Peralta). Plant Mol. Biol. 95, 411–423. https://doi.org/10.1007/s11103-017-0660-2

Sonawane, P.D., Heinig, U., Panda, S., Gilboa, N.S., Yona, M., Kumar, S.P., Alkan, N., Unger, T., Bocobza, S., Pliner, M., Malitsky, S., Tkachev, M., Meir, S., Rogachev, I., Aharoni, A., 2018. Short-chain dehydrogenase/reductase governs steroidal specialized metabolites structural diversity and toxicity in the genus Solanum. Proc. Natl. Acad. Sci. U.S.A. 115, E5419–E5428. https://doi.org/10.1073/pnas.1804835115

Sonawane, P.D., Jozwiak, A., Panda, S., Aharoni, A., 2020. 'Hijacking' core metabolism: a new panache for the evolution of steroidal glycoalkaloids structural diversity. Curr. Opin. Plant Biol. 55, 118–128. https://doi.org/10.1016/j.pbi.2020.03.008

Sonawane, P.D., Pollier, J., Panda, S., Szymanski, J., Massalha, H., Yona, M., Unger, T., Malitsky, S., Arendt, P., Pauwels, L., Almekias-Siegl, E., Rogachev, I., Meir, S., Cárdenas, P.D., Masri, A., Petrikov, M., Schaller, H., Schaffer, A.A., Kamble, A., Giri, A.P., Goossens, A., Aharoni, A., 2016. Plant cholesterol biosynthetic pathway overlaps with phytosterol metabolism. Nat. Plants 3, 1–13. https://doi.org/10.1038/nplants.2016.205

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

Spiller, F., Alves, M.K., Vieira, S.M., Carvalho, T.A., Leite, C.E., Lunardelli, A., Poloni, J.A., Cunha, F.Q., de Oliveira, J.R., 2008. Anti-inflammatory effects of red pepper (*Capsicum baccatum*) on carrageenan- and antigen-induced inflammation. J. Pharm. Pharmacol. 60, 473–478. https://doi.org/10.1211/jpp.60.4.0010

Srinivasan, P., Smolke, C.D., 2020. Biosynthesis of medicinal tropane alkaloids in yeast. Nature 585, 614–619. https://doi.org/10.1038/s41586-020-2650-9

Stenzel, O., Teuber, M., Dräger, B., 2006. Putrescine N-methyltransferase in *Solanum tuberosum* L., a calystegine-forming plant. Planta 223, 200–212. https://doi.org/10.1007/s00425-005-0077-z

Stewart Jr, C., Kang, B.-C., Liu, K., Mazourek, M., Moore, S.L., Yoo, E.Y., Kim, B.-D., Paran, I., Jahn, M.M., 2005. The Pun1 gene for pungency in pepper encodes a putative acyltransferase. Plant J. 42, 675–688. https://doi.org/10.1111/j.1365-313X.2005.02410.x

Sulli, M., Barchi, L., Toppino, L., Diretto, G., Sala, T., Lanteri, S., Rotino, G.L., Giuliano, G., 2021. An eggplant recombinant inbred population allows the discovery of metabolic QTLs controlling fruit nutritional quality. Front. Plant Sci. 12.

Szymański, J., Bocobza, S., Panda, S., Sonawane, P., Cárdenas, P.D., Lashbrooke, J., Kamble, A., Shahaf, N., Meir, S., Bovy, A., Beekwilder, J., Tikunov, Y., Romero de la Fuente, I., Zamir, D., Rogachev, I., Aharoni, A., 2020. Analysis of wild tomato introgression lines elucidates the genetic basis of transcriptome and metabolome variation underlying fruit traits and pathogen response. Nat. Genet. 52, 1111–1121. https://doi.org/10.1038/s41588-020-0690-6

Tanaka, Y., Sonoyama, T., Muraga, Y., Koeda, S., Goto, T., Yoshida, Y., Yasuba, K., 2015. Multiple loss-of-function putative aminotransferase alleles contribute to low pungency and capsinoid biosynthesis in *Capsicum chinense*. Mol. Breed. 35, 142. https://doi.org/10.1007/s11032-015-0339-9

Tewksbury, J.J., Nabhan, G.P., 2001. Directed deterrence by capsaicin in chillies. Nature 412, 403–404. https://doi.org/10.1038/35086653

Tieman, D., Zhu, G., Resende, M.F.R., Lin, T., Nguyen, C., Bies, D., Rambla, J.L., Beltran, K.S.O., Taylor, M., Zhang, B., Ikeda, H., Liu, Z., Fisher, J., Zemach, I., Monforte, A., Zamir, D., Granell, A., Kirst, M., Huang, S., Klee, H., 2017. A chemical genetic roadmap to improved tomato flavor. Science 355, 391–394. https://doi.org/10.1126/science.aal1556

Tikunov, Y.M., Molthoff, J., de Vos, R.C.H., Beekwilder, J., van Houwelingen, A., van der Hooft, J.J.J., Nijenhuis-de Vries, M., Labrie, C.W., Verkerke, W., van de Geest, H., Viquez Zamora, M., Presa, S., Rambla, J.L., Granell, A., Hall, R.D., Bovy, A.G., 2013. NON-SMOKY GLYCOSYLTRANSFERASE1 prevents the release of smoky aroma from tomato fruit. Plant Cell 25, 3067–3078. https://doi.org/10.1105/tpc.113.114231

Toal, T.W., Ron, M., Gibson, D., Kajala, K., Splitt, B., Johnson, L.S., Miller, N.D., Slovak, R., Gaudinier, A., Patel, R., de Lucas, M., Provart, N.J., Spalding, E.P., Busch, W., Kliebenstein, D.J., Brady, S.M., 2018. Regulation of root angle and gravitropism. G3: Genes Genomes Genet. 8, 3841–3855. https://doi.org/10.1534/g3.118.200540

Tripodi, P., Rabanus-Wallace, M.T., Barchi, L., Kale, S., Esposito, S., Acquadro, A., Schafleitner, R., van Zonneveld, M., Prohens, J., Diez, M.J., Börner, A., Salinier, J., Caromel, B., Bovy, A., Boyaci, F., Pasev, G., Brandt, R., Himmelbach, A., Portis, E., Finkers, R., Lanteri, S., Paran, I., Lefebvre, V., Giuliano, G., Stein, N., 2021. Global range expansion history of pepper (Capsicum spp.) revealed by over 10,000 genebank accessions. Proc. Natl. Acad. Sci. U.S.A. 118, e2104315118. https://doi.org/10.1073/pnas.2104315118

Tsugawa, H., Nakabayashi, R., Mori, T., Yamada, Y., Takahashi, M., Rai, A., Sugiyama, R., Yamamoto, H., Nakaya, T., Yamazaki, M., Kooke, R., Bac-Molenaar, J.A., Oztolan-Erol, N., Keurentjes, J.J.B., Arita, M., Saito, K., 2019. A cheminformatics approach to characterize metabolomes in stable-isotope-labeled organisms. Nat. Methods 16, 295–298. https://doi.org/10.1038/s41592-019-0358-2

Tzin, V., Malitsky, S., Zvi, M.M.B., Bedair, M., Sumner, L., Aharoni, A., Galili, G., 2012. Expression of a bacterial feedback-insensitive 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase of the shikimate pathway in Arabidopsis elucidates potential metabolic bottlenecks between primary and secondary metabolism. New Phytol. 194, 430–439. https://doi.org/10.1111/j.1469-8137.2012.04052.x

Van Eck, J., 2018. Genome editing and plant transformation of solanaceous food crops. Curr. Opin. Biotechnol., Food biotechnology • Plant biotechnology 49, 35–41. https://doi.org/10.1016/j.copbio.2017.07.012

Varghese, S., Kubatka, P., Rodrigo, L., Gazdikova, K., Caprnda, M., Fedotova, J., Zulli, A., Kruzliak, P., Büsselberg, D., 2017. Chili pepper as a body weight-loss food. Int. J. Food Sci. Nutr. 68, 392–401. https://doi.org/10.1080/09637486.2016.1258044

Vogt, T., 2010. Phenylpropanoid biosynthesis. Mol. Plant 3, 2–20. https://doi.org/10.1093/mp/ssp106

Wang, M., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., Peng, Y., Nguyen, D.D., Watrous, J., Kapono, C.A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A.V., Meehan, M.J., Liu, W.-T., Crüsemann, M., Boudreau, P.D., Esquenazi, E., Sandoval-Calderón, M., Kersten, R.D., Pace, L.A., Quinn, R.A., Duncan, K.R., Hsu, C.-C., Floros, D.J., Gavilan, R.G., Kleigrewe, K., Northen, T., Dutton, R.J., Parrot, D., Carlson, E.E., Aigle, B., Michelsen, C.F., Jelsbak, L., Sohlenkamp, C., Pevzner, P., Edlund, A., McLean, J., Piel, J., Murphy, B.T., Gerwick, L., Liaw, C.-C., Yang, Y.-L., Humpf, H.-U., Maansson, M., Keyzers, R.A., Sims, A.C., Johnson, A.R., Sidebottom, A.M., Sedio, B.E., Klitgaard, A., Larson, C.B., Boya P, C.A., Torres-Mendoza, D., Gonzalez, D.J., Silva, D.B., Marques, L.M., Demarque, D.P., Pociute, E., O'Neill, E.C., Briand, E., Helfrich, E.J.N., Granatosky, E.A., Glukhov, E., Ryffel, F., Houson, H., Mohimani, H., Kharbush,

J.J., Zeng, Y., Vorholt, J.A., Kurita, K.L., Charusanti, P., McPhail, K.L., Nielsen, K.F., Vuong, L., Elfeki, M., Traxler, M.F., Engene, N., Koyama, N., Vining, O.B., Baric, R., Silva, R.R., Mascuch, S.J., Tomasi, S., Jenkins, S., Macherla, V., Hoffman, T., Agarwal, V., Williams, P.G., Dai, J., Neupane, R., Gurr, J., Rodríguez, A.M.C., Lamsa, A., Zhang, C., Dorrestein, K., Duggan, B.M., Almaliti, J., Allard, P.-M., Phapale, P., Nothias, L.-F., Alexandrov, T., Litaudon, M., Wolfender, J.-L., Kyle, J.E., Metz, T.O., Peryea, T., Nguyen, D.-T., VanLeer, D., Shinn, P., Jadhav, A., Müller, R., Waters, K.M., Shi, W., Liu, X., Zhang, L., Knight, R., Jensen, P.R., Palsson, B.Ø., Pogliano, K., Linington, R.G., Gutiérrez, M., Lopes, N.P., Gerwick, W.H., Moore, B.S., Dorrestein, P.C., Bandeira, N., 2016. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat. Biotechnol. 34, 828–837. https://doi.org/10.1038/nbt.3597

Wang, P., Moore, B.M., Uygun, S., Lehti-Shiu, M.D., Barry, C.S., Shiu, S.-H., 2021. Optimising the use of gene expression data to predict plant metabolic pathway memberships. New Phytol. 231, 475–489. https://doi.org/10.1111/nph.17355

Weber, N., Ismail, A., Gorwa-Grauslund, M., Carlquist, M., 2014. Biocatalytic potential of vanillin aminotransferase from *Capsicum chinense*. BMC Biotechnol. 14, 25. https://doi.org/10.1186/1472-6750-14-25

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

Weng, J.-K., Philippe, R.N., Noel, J.P., 2012. The rise of chemodiversity in plants. Science 336, 1667–1670. https://doi.org/10.1126/science.1217411

W. Sumner, L., Lei, Z., J. Nikolau, B., Saito, K., 2015. Modern plant metabolomics: advanced natural product gene discoveries, improved technologies, and future prospects. Nat. Prod. Rep. 32, 212–229. https://doi.org/10.1039/C4NP00072B

Wu, Q., Cho, J.-G., Lee, D.-S., Lee, D.-Y., Song, N.-Y., Kim, Y.-C., Lee, K.-T., Chung, H.-G., Choi, M.-S., Jeong, T.-S., Ahn, E.-M., Kim, G.-S., Baek, N.-I., 2013. Carbohydrate derivatives from the roots of *Brassica rapa* ssp. *campestris* and their effects on ROS production and glutamate-induced cell death in HT-22 cells. Carbohydr. Res. 372, 9–14. https://doi.org/10.1016/j.carres.2012.09.015

Xu, S., Brockmöller, T., Navarro-Quezada, A., Kuhl, H., Gase, K., Ling, Z., Zhou, W., Kreitzer, C., Stanke, M., Tang, H., Lyons, E., Pandey, P., Pandey, S.P., Timmermann, B., Gaquerel, E., Baldwin, I.T., 2017. Wild tobacco genomes reveal the evolution of nicotine biosynthesis. Proc. Natl. Acad. Sci. U.S.A. 114, 6133–6138. https://doi.org/10.1073/pnas.1700073114

The Potato Genome Sequencing Consortium, 2011. Genome sequence and analysis of the tuber crop potato. Nature 475, 189–195. https://doi.org/10.1038/nature10158

Yoo, H., Widhalm, J.R., Qian, Y., Maeda, H., Cooper, B.R., Jannasch, A.S., Gonda, I., Lewinsohn, E., Rhodes, D., Dudareva, N., 2013. An alternative pathway contributes to phenylalanine biosynthesis in plants via a cytosolic tyrosine:phenylpyruvate aminotransferase. Nat. Commun. 4, 2833. https://doi.org/10.1038/ncomms3833

Zabel, S., Brandt, W., Porzel, A., Athmer, B., Bennewitz, S., Schäfer, P., Kortbeek, R., Bleeker, P., Tissier, A., 2021. A single cytochrome P450 oxidase from *Solanum habrochaites* sequentially oxidizes 7-epi-zingiberene to derivatives toxic to whiteflies and various microorganisms. Plant J. 105, 1309–1325. https://doi.org/10.1111/tpj.15113

Zanor, M.I., Rambla, J.-L., Chaïb, J., Steppa, A., Medina, A., Granell, A., Fernie, A.R., Causse, M., 2009. Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents. J. Exp. Bot. 60, 2139–2154. https://doi.org/10.1093/jxb/erp086

Zhang, C.-R., Khan, W., Bakht, J., Nair, M.G., 2016. New antiinflammatory sucrose esters in the natural sticky coating of tomatillo (*Physalis philadelphica*), an important culinary fruit. Food Chem. 196, 726–732. https://doi.org/10.1016/j.foodchem.2015.10.007

Zhang, C.-Y., Luo, J.-G., Liu, R.-H., Lin, R., Yang, M.-H., Kong, L.-Y., 2016. 1H NMR spectroscopy-guided isolation of new sucrose esters from *Physalis alkekengi* var. *franchetii* and their antibacterial activity. Fitoterapia 114, 138–143. https://doi.org/10.1016/j.fitote.2016.09.007

Zhang, Y., Butelli, E., Alseekh, S., Tohge, T., Rallapalli, G., Luo, J., Kawar, P.G., Hill, L., Santino, A., Fernie, A.R., Martin, C., 2015. Multi-level engineering facilitates the production of phenylpropanoid compounds in tomato. Nat. Commun. 6, 8635. https://doi.org/10.1038/ncomms9635

Zhou, F., Pichersky, E., 2020a. More is better: the diversity of terpene metabolism in plants. Curr. Opin. Plant Biol., Physiology and Metabolism 55, 1–10. https://doi.org/10.1016/j.pbi.2020.01.005

Zhou, F., Pichersky, E., 2020b. The complete functional characterisation of the terpene synthase family in tomato. New Phytol. 226, 1341–1360. https://doi.org/10.1111/nph.16431

Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., Lin, T., Qin, M., Peng, M., Yang, C., Cao, X., Han, X., Wang, X., Knaap, E. van der, Zhang, Z., Cui, X., Klee, H., Fernie, A.R., Luo, J., Huang, S., 2018. Rewiring of the fruit metabolome in tomato breeding. Cell 172, 249-261.e12. https://doi.org/10.1016/j.cell.2017.12.019

Zi, J., Matsuba, Y., Hong, Y.J., Jackson, A.J., Tantillo, D.J., Pichersky, E., Peters, R.J., 2014. Biosynthesis of lycosantalonol, a cis-prenyl derived diterpenoid. J. Am. Chem. Soc. 136, 16951–16953. https://doi.org/10.1021/ja508477e

CHAPTER 2:

TRADING ACYLS AND SWAPPING SUGARS: METABOLIC INNOVATIONS IN

SOLANUM TRICHOMES

Works presented in this chapter have been published:

Fiesel, P. D., Kerwin, R.A., Jones, A.D., Last, R.L., 2023. Trading acyls and swapping sugars: metabolic innovations in *Solanum* trichomes. bioRxiv. https://doi.org/10.1101/2023.06.05.542877

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 from sugars and acyl-Coenzyme A esters by acylsugar acyltransferases in glandular trichomes. We characterized trichome acylsugars of the Clade II species *Solanum melongena* (brinjal eggplant) using liquid chromatography-mass spectrometry (LC-MS), gas chromatography-MS and nuclear magnetic resonance (NMR) spectroscopy. This led to the identification of eight unusual structures with inositol cores, inositol glycoside cores, and hydroxyacyl chains. LC-MS analysis of 31 species in the megadiverse *Solanum* genus revealed striking acylsugar diversity with some traits restricted to specific clades and species. Acylinositols were found throughout each clade while acylglucoses were restricted to DulMo and VANAns species. Medium length hydroxyacyl chains were found in many species. Analysis of tissue-specific transcriptomes and interspecific acylsugar acetylation differences led to characterization 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 is a functionally divergent ASAT3. This study provides a foundation for investigating the evolution 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. Specialized metabolites play many critical roles such as in abiotic and biotic stress adaptation (Agati and Tattini, 2010; De Moraes et al., 2001; Landry et al., 1995), pollinator attraction (Kretschmar and Baumann, 1999) and mediation of interactions with beneficial and pathogenic microbes (Yu et al., 2021). These diverse and bioactive small molecules have historical and modern uses in human medicine, including the anticancer alkaloids vinblastine and paclitaxel, antimalarial artemisinin, and painkillers such as morphine.

   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 Type 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-CoAs (Fan et al., 2015; Lou et al., 2021; Moghe et al., 2017; Schilmiller et al., 2012; Schilmiller et al., 2015). Despite their simple components, acylsugars exhibit remarkable chemical diversity arising from variations in sugar core composition and acyl chain length, branching pattern, position, and number (Fan et al., 2019; Fiesel et al., 2022; Ghosh et al., 2014a; Hurney, 2018; Lou et al., 2021; Moghe et al., 2017; Schenck et al., 2022). For example,

79

acylsucroses, composed of a sucrose disaccharide core, accumulate in cultivated tomato (*Solanum lycopersicum*) trichomes, while acylsucroses and acylglucoses have been observed in the trichomes of wild tomato species, *Solanum pennellii*. Acylsugar structural variation also impacts biological activity; for example, differential oviposition deterrence was demonstrated from naturally derived acylsugar mixtures (Leckie et al., 2016).

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

Specialized metabolism diversification is often driven by gene duplication and subsequent sequence divergence. In fact, specialized metabolism genes have higher duplication rates than general metabolism genes (Moore et al., 2019). Duplicates often arise through whole genome duplications and localized tandem duplications and can exhibit lower rates of selection leading to genes and gene products with new functions, localizations, and regulation (Panchy et al., 2016). For example, the tandem duplication of N-methyltransferases and cytochrome P450s led to evolution of the alkaloid caffeine and the benzoxazinoid DIMBOA, respectively (Dutartre et al., 2012; Denoeud et al., 2014).

Nearly half of the Solanaceae falls into the large (>1200 species) and phenotypically diverse *Solanum* genus. This genus is split into several major clades, including Potato, Regmandra, DulMo, VANAns, and Clade II; the Potato clade contains cultivated tomato and potato and their wild relatives while Clade II contains the cultivated brinjal eggplant, *Solanum melongena*, and other 'spiny *Solanums*' (Bohs, 2004; Bohs and Olmstead, 1997; Gagnon et al., 2022; Levin et al., 2006; PBI Solanum Project, 2022; Särkinen et al., 2013; Stern et al., 2011; Tepe et al., 2016; Weese and Bohs, 2007) (Gagnon et al., 2022). To date, documentation of acylsugar diversity largely focused on a handful of species within the Potato clade, including cultivated tomato and its close relatives. These efforts identified at least 38 nuclear magnetic resonance (NMR) spectroscopy-resolved acylsugar structures and liquid chromatography mass spectrometry (LC-MS) supported annotations of many more (Fan et al., 2017; Ghosh et al., 2014a; Lybrand et al., 2020; Schilmiller et al., 2016). While limited acylsugar screening outside of the Potato clade was reported, novel structural variants not observed among cultivated tomato relatives were identified. For example, acylsugars with *myo*-inositol sugar cores (i.e., acylinositols) were characterized in three species: *S. nigrum*, from the DulMo clade, and *S. lanceolatum* and *S. quitoense,* from Clade II (Herrera-Salgado et al., 2005; Hurney, 2018; Leong et al., 2020; Lou et al., 2021). These discoveries highlight the benefits of a more comprehensive description of *Solanum* acylsugars within the well-developed phylogenetic framework.

Here we report analysis of *Solanum* acylsugar chemical diversity in species from the relatively unexplored *Solanum* clades DulMo, VANAns, and Clade II, which together comprise ~1000 *Solanum* species. We first established the Clade II brinjal eggplant, *S. melongena*, as a reference species. Eggplant is a major worldwide fruit crop with extensive genomic, transcriptomic, and germplasm resources (Barchi et al., 2021; Li et al., 2021; Mennella et al.,

2010). We characterized eggplant trichome acylinositols, acylinositol glycosides, and acylsugars with unusual hydroxylated acyl chains using electrospray ionization LC-quadrupole time of flight-MS (ESI LC-QToF-MS) and NMR. These atypical structures likely reflect altered biochemistry from the cultivated tomato acylsucrose pathway. Moving out from this model organism framework, LC-MS phylogenetic screening of 31 Clade II, DulMo, and VANAns species, including *S. melongena*, led to the identification of remarkable acylsugar structural variation. LC-MS features with characteristics of acylinositols were found in 25 of the 26 acylsugar-producing species, suggesting one or a small number of evolutionary origins. In contrast, acylglucoses were detected in DulMo and VANAns species, but not in any tested Clade II species.

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

Results and Discussion

<u>Eggplant glandular trichomes accumulate acylsugar-like compounds</u>

We began our investigation of *Solanum* acylsugar diversity with the brinjal eggplant *S. melongena* due to its economic importance, genomic resources, and phylogenetic position within the monophyletic Eastern Hemisphere Spiny clade of Clade II (formerly known as the 'Old World spiny clade') (Gagnon et al., 2022). Using eight eggplant accessions (Table S2.37), we observed glandular trichomes on hypocotyls, cotyledons, and the first three true leaves of young

eggplants, which resemble the acylsugar-producing structures found in other *Solanum* species (Figure 2.2) (Leong et al., 2020; Lou et al., 2021; Schilmiller et al., 2012). In contrast, we observed only non-glandular stellate trichomes on leaves and stems of mature eggplants (Figure S2.12), which are unlikely to accumulate and synthesize acylsugars or other specialized metabolites (Levin, 1973; Wagner, 1991). We analyzed surface metabolite extracts from young and mature eggplant tissues using LC-QToF-MS coupled with collision-induced dissociation (CID) in negative and positive ion mode, and annotated acylsugars based on relative mass defect, molecular adduct ion masses, retention times, and ions present in CID mass spectra. Briefly, acylsugars were annotated from masses of fatty acid carboxylate fragment ions, as well as fragment ions corresponding to stepwise losses of acyl chains from the pseudomolecular ion to a sugar core fragment ion. The acylsugar annotation methods and confidence criteria are explained in detail in the Methods. This analysis revealed abundant acylsugars in extracts from young, glandular trichome-producing eggplant tissues, but not from mature, non-glandular trichome producing tissues.

We annotated 38 acylsugars in young eggplant extracts from eight accessions, including 16 acylhexoses and 22 acyldisaccharides (Table 2.1). LC-MS based acylsugar annotations are described using a modified shorthand nomenclature (Leong et al., 2020) as follows: UX:Y:Z(A,B,C,D), where U – as a single or multi-letter designation – represents the sugar core, X represents the number of acyl chains, Y represents the sum of acyl chain carbons, Z represents the number of unsaturated bonds in the acyl chains (when present), and A-D represent the number of carbon atoms in the individual acyl chains. For example, the eggplant acylhexose I3:18(4,4,10) consists of a *myo*-inositol core with three acyl chains with a total of 18 carbon atoms. Curiously, eggplant acyldisaccharides contain an atypical pentose-hexose core, as

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

deacylated sugar core minus a proton. Further evidence was provided by positive mode CID,

which promotes glycosidic bond cleavage, yielding a fragment ion corresponding to the neutral

loss of an unacylated pentose ring. All acylhexoses annotated with medium to high confidence

formed abundant fragment ions in positive mode but few fragment ions in negative mode as

illustrated in Figure S2.1; this pattern is characteristic of acylinositols found in *S. quitoense* and

S. nigrum, but has not been observed for acylglucoses, suggesting that all detected *S. melongena*

acylhexoses are acylinositols (Hurney, 2018; Leong et al., 2020; Lou et al., 2021). The eggplant

acylsugars contained three to four acylations, all on the hexose core, including one medium

eight-carbon (C8) to C14 acyl chain and two to three short C4 or C5 acyl chains. The medium

acyl chains (C8, C10, C12, and C14) were either iso-branched or straight as revealed by GC-MS

acyl chain analysis (Figure S2.85). Additionally, we identified hydroxylated C12, C14, and C16

acyl chains not previously reported in Solanaceae acylsugars. Although we did not observe large

differences between the eight eggplant accessions, eggplant acylsugars differ in chain length,

functional groups, and acyl chain composition from the reported *Solanum* acylinositols (Herrera-

Salgado et al., 2005; Hurney, 2018; Lou et al., 2021).

Table 2.1. 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 number is listed for NMR
characterized compounds (Figure 2.2). Percent total peak area calculated by dividing acylsugar
peak area by total acylsugar peak area. Acylsugars composing ≥1% of total acylsugar peak area
are bolded. Percent peak area calculated from 15 *S. melongena* leaf surface extracts. Annotation
method is described in Materials and Methods. Acylsugars are sorted by number of sugar
moieties and then by elution order.

Name	Compound #	Annotation confidence level	RT (min)	Neutral molecule chemical formula	m/z_{acc}	m/z_{ex}	Δm (ppm)	Average percent peak area (%)

Table 2.1. (cont'd)

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

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

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

85

derivatized *S. melongena* acylsugar cores. When free sugar cores, produced by metabolite extract saponification, were derivatized to form alditol acetates, GC-MS peaks corresponding to derivatized *myo*-inositol and glucose were detected, supporting the presence of acylinositols (Figure 2.1A). The detection of glucose might have resulted from other compounds in the leaf surface metabolite extracts. The disaccharide sugar core composition was determined by hydrolyzing the saponified sugar cores with formic acid to cleave the glycosidic linkage. Comparison of the hydrolyzed and unhydrolyzed samples revealed a peak corresponding to arabinose only in the hydrolyzed plant samples (Figure 2.1B). Taken together, the results of saponification with and without hydrolysis, followed by derivatization, confirmed identification of *myo*-inositol sugar cores, and identified the pentose moiety of the hexose-pentose disaccharide core as arabinose.
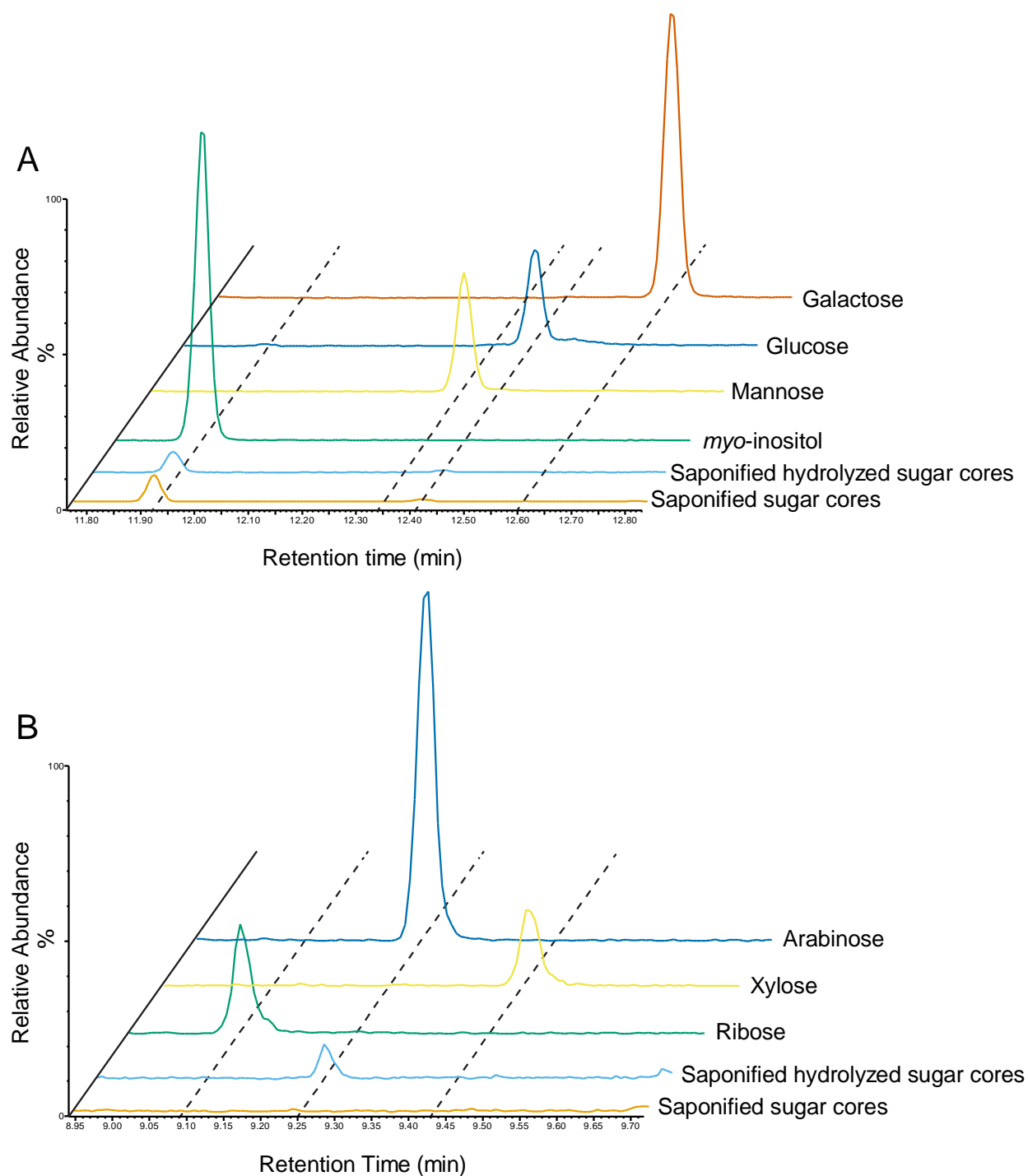
Figure 2.1. Identification of *S. melongena* acylsugar core composition through GC-MS analysis of alditol acetate sugar derivatives. *S. melongena* acylsugars collected from surface extracts were first saponified to remove acyl chains, and then, with or without acid hydrolysis to break the glycosidic linkages, sugar cores were derivatized to alditol acetates. (A) Alditol acetate derivatization of saponified *S. melongena* acylsugars yield a peak that comigrated with that of a *myo*-inositol standard. The traces displayed are GC-MS total ion chromatograms (TICs) normalized to the highest ion count in the selected traces. (B) Alditol acetate pentose derivatives

Figure 2.1. (cont'd)

of saponified and hydrolyzed acylsugar cores comigrate with that of a derivatized arabinose standard. The traces displayed are GC-MS TICs.

<u>NMR analysis of eight eggplant acylsugars</u>

While MS analysis provided valuable information about acyl chain number and length, sugar core mass, as well as acyl chain number, complementary information about acyl chain branch structure, sugar core stereochemistry, and acyl chain positions was obtained using NMR. We purified and resolved the structures of eight abundant eggplant acylsugars from accession PI 555598 using a combination of 1D and 2D NMR experiments (Figure 2.2). All eight structures are newly described, and because atomic connections were determined by NMR and MS data, the proposed structures meet Metabolomics Standards Initiative level 1 criteria for metabolite identification (Sumner et al., 2007).

A series of NMR experiments confirmed that the acylhexose and acyldisaccharide sugar cores are *myo*-inositol and 4-*O*-β-arabinopyranosyl *myo*-inositol, respectively (Figure 2.2C,D). We assigned all sugar ring proton signals of each sugar core's spin system with total correlation spectroscopy (TOCSY). Inferences from correlation spectroscopy (COSY) data then identified the order of ring protons and identified pyranose and cyclitol ring structures for the pentose and hexose rings, respectively. Relative stereochemistry at ring positions was subsequently determined through comparison of spin-spin splitting (multiplicities and coupling constants) referenced to expected patterns inferred from chemical principles and previously reported acylinositols (Hurney, 2018; Leong et al., 2020). The disaccharide glycosidic linkage position was determined by heteronuclear multiple bond correlation (HMBC) correlations to be at position 4 and position 1 of *myo*-inositol and arabinose, respectively. The arabinose *β* anomeric configuration was inferred from the anomeric carbon $^1J_{CH}$ (162 Hz) as revealed by a coupled-heteronuclear single quantum coherence (coupled-HSQC) experiment of the free disaccharide

sugar core after saponification. Our sugar core assignments identifying this unusual disaccharide agreed with the previous GC-MS sugar core results, supporting the efficacy of the sugar core GC-MS identification method. This disaccharide differs only in the pentose moiety identity from a recently reported acylated 4-*O*-*ß*-xylopyranosyl *myo*-inositol in the *Solanum* Clade II species *S. quitoense* (Hurney, 2018). To the best of our knowledge, this is the first report of acylated 4-*O*-*ß*-arabinosyl *myo*-inositol sugars.

Acyl chain positions, branching patterns, and hydroxyl positions were also resolved through integration of different NMR experiments. We found that all acyl chains were confined to *myo*-inositol, consistent with the LC-MS results. All eight acylsugars are decorated with two short iso-branched iC4 acyl chain esters at positions 1 and 2, and one medium C8 to C14 acyl chain ester at position 3 (Figure 2.2). Compound 2 (Figure 2.2C), the only tetraacylated acylsugar identified in eggplant, additionally carried an acetylation at position 6 (Figure 2.2C). The medium-length acyl chains at position 3 were resolved as straight (nC14) or terminally iso-branched (iC8, iC10, iC12, iC14) based on signals characteristic of protons near the branched carbons. Strikingly, we identified peculiar hydroxylated straight and iso-branched 3-hydroxytetradecanoate acyl chains, 3-OH-nC14 and 3-OH-iC14, in compounds 7 and 8, respectively (Figure 2.2D). We assigned hydroxylation positions of 3-OH-nC14 and 3-OH-iC14 to the third acyl carbon based on a downfield shifted signal in the ^{1}H NMR spectrum at 3.92 ppm, corresponding to one hydrogen at that position. In contrast, the non-hydroxylated medium acyl chains observed in compounds 1-6 (Figure 2.2C,D), carry two hydrogen atoms at the third acyl carbon, and these have a characteristic signal near 1.50 ppm. We believe this is the first report of 3-OH-nC14 and 3-OH-iC14 chains in Solanaceae acylsugars. While 3-OH-nC14 chains are observed in the acylsugar-like bacterial Lipid A glycolipids, the hydroxylation position

differentiates these eggplant chains from hydroxyacyl chains in castor bean (*Ricinus communis*) seed oil, *Silene gallica* gallicasides, *Ibicella lutea* fatty acid glycosides, and Convolvulaceae resin glycosides (Asai et al., 2010; Asai and Fujimoto, 2010; Bah and Pereda-Miranda, 1996; Smith, 1971).
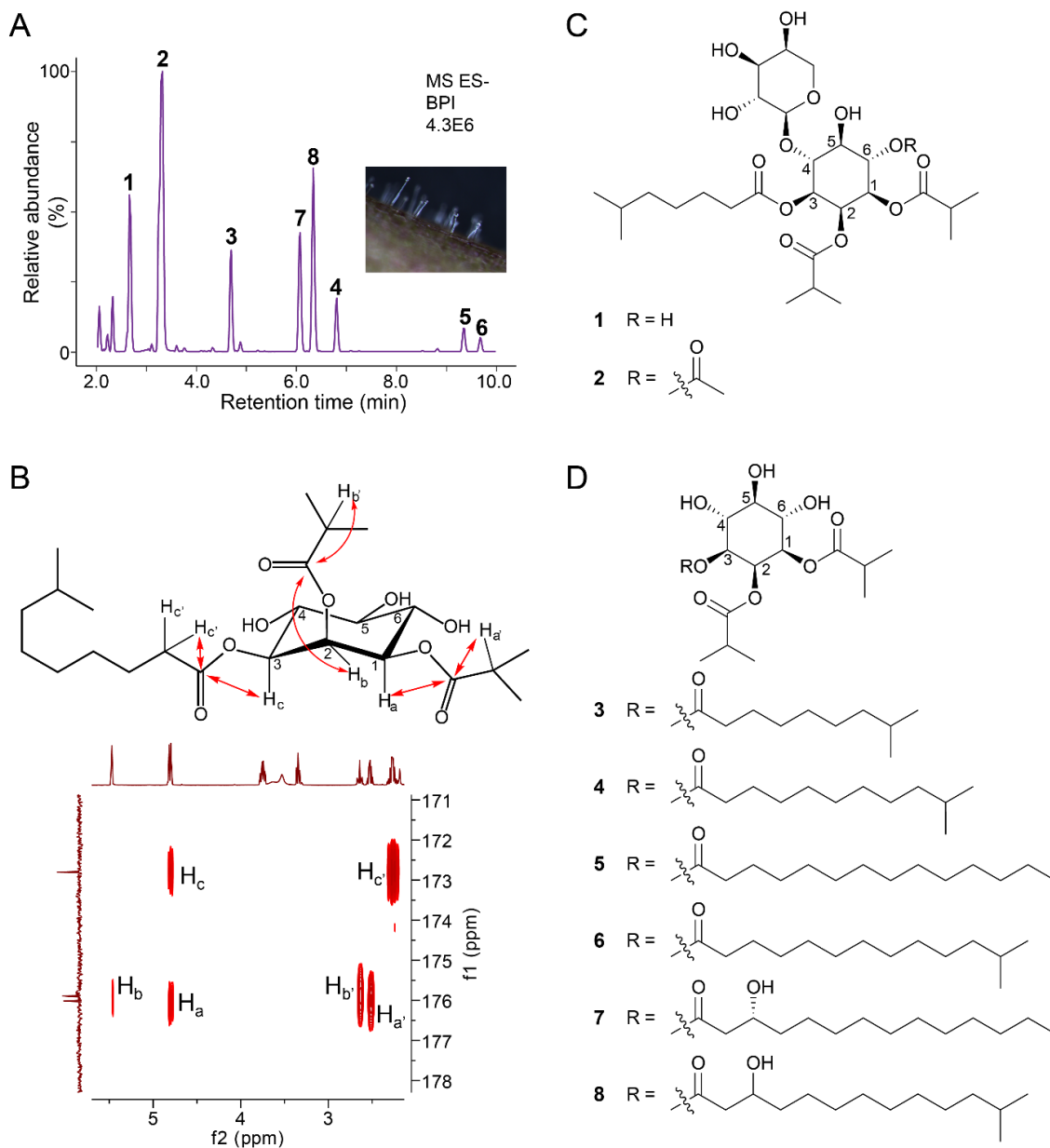
Figure 2.2. Profiling of *S. melongena* acylsugars. (A) LC-MS base peak intensity (BPI) chromatogram of *S. melongena* acylsugars. Peaks characterized by NMR are indicated with their

90

Figure 2.2. (cont'd)
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. HMBC NMR spectrum showing couplings between hydrogens and carbonyl carbons are shown. (C) NMR-characterized acyldisaccharide structures 1 and 2, with inositol ring carbons numbered. (D) NMR-characterized acylinositol monosaccharide structures 3-8.

The OH-tetradecanoate acyl chain has *R* stereochemistry

To gain insight into the biosynthetic origin of hydroxylated acyl chains, we tested chirality of the 3-OH-nC14 acyl chain of compound 7 as this chain has available commercial standards. Two-step chiral derivatization and GC-MS analysis was performed on compound 7 and two commercial standards, (3*R*)-OH-nC14 and the corresponding racemate (3*R/S*)-OH-nC14. Fatty acids were converted to ethyl esters to aid volatilization, followed by esterification of the acyl chain hydroxyl group with the Mosher acid (*R*)-(−)-α-methoxy-α-(trifluoromethyl)phenylacetate (*R*-MPTA) to distinguish stereoisomers by GC separation of the diastereomeric derivatives. GC-MS analysis of the derivatized (3*R/S*)-OH-nC14 standard compounds yielded two peaks at retention times of 42.21 and 42.40 minutes, while one peak at 42.40 minutes was obtained for (3*R*)-OH-nC14. We observed expected fragment ions corresponding to the MPTA fragment ion, fatty acid ethyl ester, and fatty acid which supported peak assignment (Figure S2.5). Thus (3*S*)-OH-nC14 and (3*R*)-OH-nC14 elute at 42.21 and 42.40 minutes, respectively. The compound 7 acyl chain was assigned as (3*R*)-OH-nC14 as its derivatized 3-OH-tetradecanoate acyl chain eluted at 42.40 minutes.

The (3*R*)-OH-nC14 chain hydroxyl position and stereochemistry have implications for the metabolic pathway leading to its biosynthesis. As fatty acid hydroxylases typically act on the acyl chain terminal region, they are unlikely to catalyze C3 hydroxylation of a C14 fatty acid (Pinot and Beisson, 2011). Of the other well-characterized pathways, only mitochondrial and

plastidial fatty acid metabolism mechanisms produce (3*R*)-OH-acyl-thioester intermediates. The (3*R*)-OH-nC14 chain hydroxyl position and stereochemistry have implications for the metabolic pathway leading to its biosynthesis. As fatty acid hydroxylases typically act on the acyl chain terminal region, they are unlikely to catalyze C3 hydroxylation of a C14 fatty acid (Pinot and Beisson, 2011). Of the other well-characterized pathways, only mitochondrial and plastidial fatty acid metabolism mechanisms produce (3*R*)-OH-acyl-thioester intermediates. Consistent with this result, previous work identified two trichome-expressed, mitochondrial enzymes, Acylsugar Enoyl-CoA Hydratase 1 (AECH1) and Acylsugar Acyl-CoA Synthetase 1 (AACS1), and one trichome-expressed, plastidial enzyme, a beta-ketoacyl-(acyl-carrier-protein) reductase (SpKAR1), involved in medium C10 and C12 acyl chain production in cultivated tomato and its wild relative, *S. pennellii* (Fan et al., 2020; Ji et al., 2022).

We hypothesize that acylsugar hydroxyacyl chains evolved through one of three scenarios. First, changes in acyl-CoA substrate availability, possibly through the actions of thioesterases or acyl-CoA synthetases, enabled a promiscuous ASAT to esterify hydroxyacyl chains to acylinositols. Alternatively, substrate specificity of an existing acylinositol ASAT expanded to utilize available hydroxyacyl-CoAs. Like this second hypothesis, amino acid substitutions between copies of ASAT2 from different tomato species impacted the accumulation of differently branched five carbon acyl chains on acylsugars (Fan et al., 2015). A third possibility is that a 'new' acyltransferase, unrelated to characterized ASATs and capable of utilizing hydroxyacyl-CoA substrates, was co-opted into acylinositol biosynthesis. Future genetic and biochemical studies should reveal the mechanisms behind (3*R*)-OH acyl chain production, potentially providing new approaches to engineer unique acylsugars with hydroxyacyl chains and hydroxylated fatty acids for polymeric building blocks. Acylsugar structures impact their

92

function, thus we hypothesize that engineered tomato producing these unusual hydroxyacyl chains on their acylsugars may exhibit different defense capabilities than the wildtype acylsucroses.
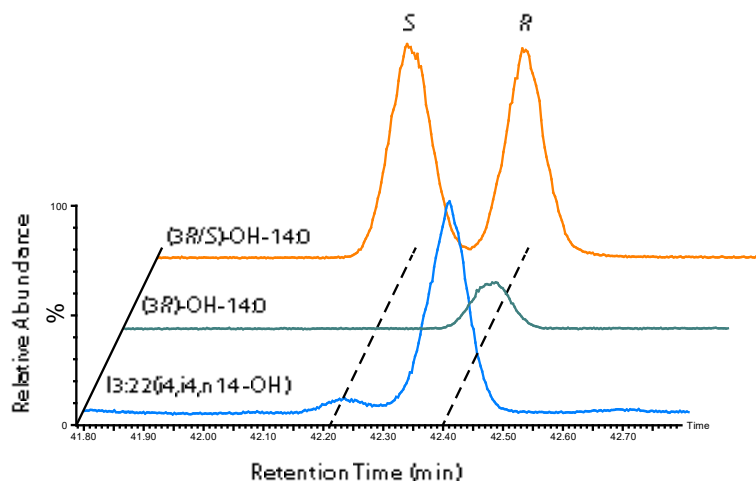
Figure 2.3. The 3-OH-n14 acyl chain on Compound 7, I3:22(i4,i4,n14-OH), is in the *R* configuration as revealed by GC-MS using a derivatization method that separates hydroxyacyl chain enantiomers. Hydroxyacyl chains analyzed by GC-SIM- MS after ethyl transesterification and derivation to their diastereomeric (*S*)-MPTA form. Fragment ions with *m/z* of 189, 209, and 255 were monitored, and relative abundance was normalized to the highest peak. Yellow and green traces represent derivatives from commercial (3*R*/*S*)- and (3*R*)-OH-14:0 standards. The main derivatized 3-OH-14:0 acyl chain from Compound 7 (Figure 2.2), shown with the blue trace, comigrates with the peak corresponding to the *R* enantiomer.

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

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

including 23 Clade II, five DulMo clade, and two VANAns clade members (listed in Table S2.37). While NMR is the gold standard for structural elucidation, it is prohibitively time-consuming for a largescale metabolite diversity survey. LC-MS coupled with CID provides substantial structural information and is compatible with high-throughput screening. To strike a balance between data quality and quantity, we performed LC-MS-CID screening on all 30 species and annotated acylsugars as described for *S. melongena*. Twenty-five of the 30 analyzed species produced detectable acylsugars from visible surface glandular trichomes, and species lacking acylsugars had no or few observable glandular trichomes in the tissues analyzed. Targeted analysis of leaf surface extracts from 31 species, including *S. melongena*, and fruit surface extracts from two species, uncovered previously unreported acylsugars, inter-species acylsugar differences, and identified the distributions of unusual acylsugar chemical traits in the *Solanum* genus (Figure 2.5; Tables 2.1,S2.1-26). Although we did not observe intraspecific acylsugar variation, acylsugars varied enormously between species (Figure 2.4; Tables 2.1,S2.1-26).
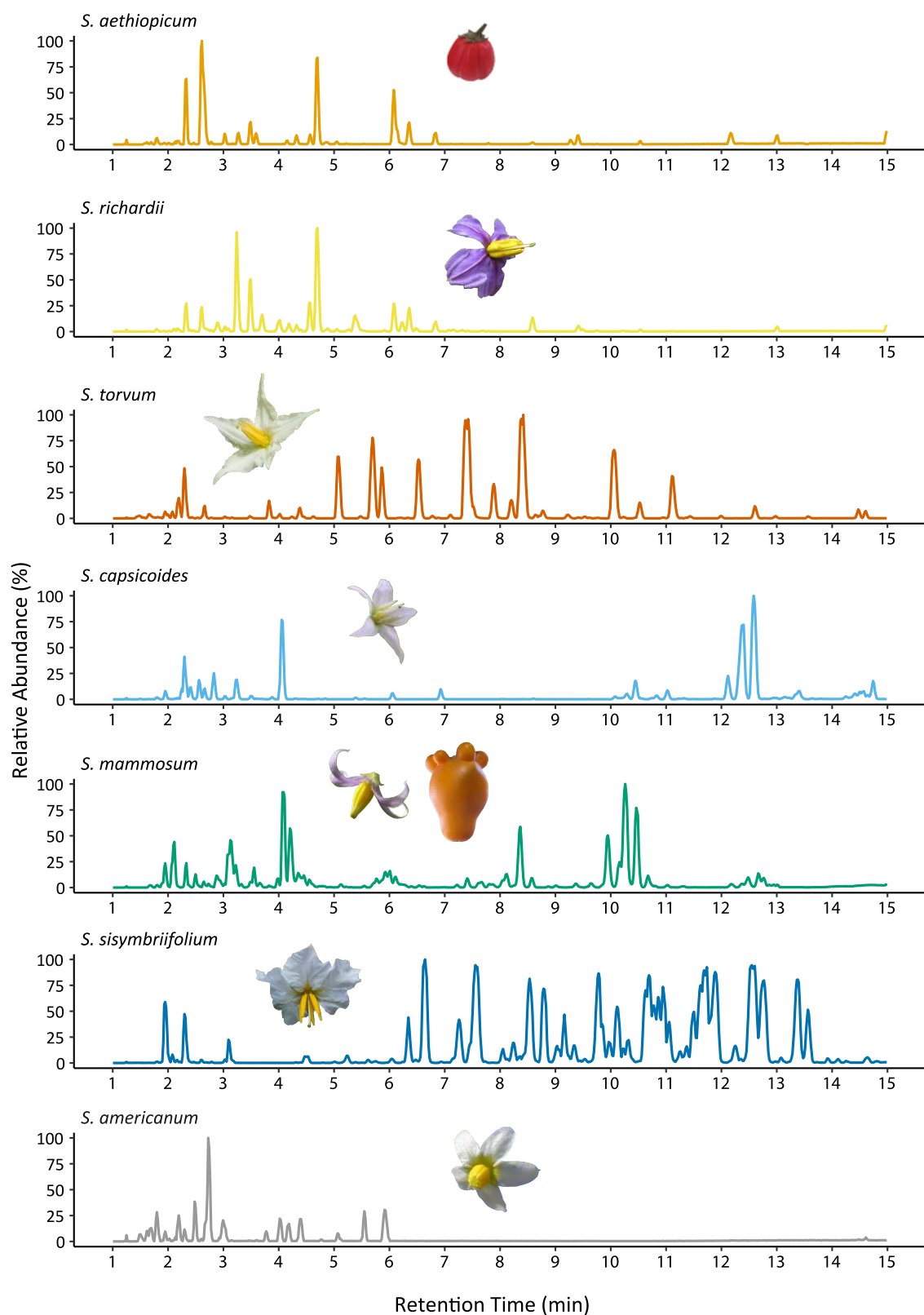
Figure 2.4. *Solanum* **species produce dramatically different leaf surface metabolite profiles.**
The seven species display morphological differences, as demonstrated by the flower and fruit

Figure 2.4. (cont'd)
images, and they also exhibit diverse metabolite profiles as demonstrated by base peak intensity LC-MS chromatograms from their leaf surface metabolites. Most of the peaks shown consist of acylsugars with varied sugar cores and acyl chain types and numbers. Acylsugar profiles for these species are detailed in Tables S2.4, S2.7, S2.17, S2.20, and S2.24-26.
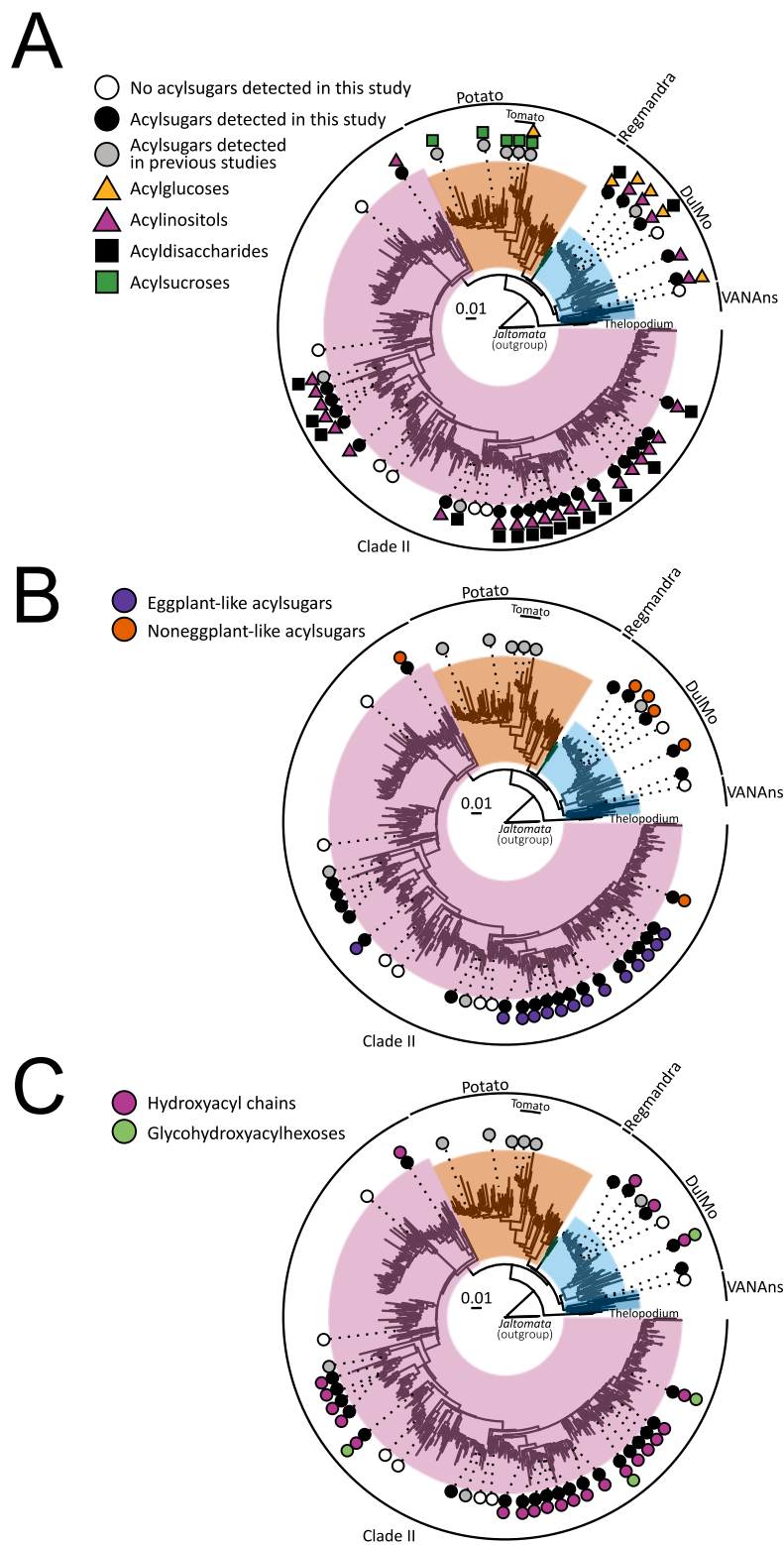
Figure 2.5. Distribution of acylsugar traits within the *Solanum* genus. Traits were mapped onto a previously published Supermatrix maximum likelihood phylogeny (Gagnon et al., 2022). The reported polytomy between Clade II, DulMO, VANANs, Regmandra, and Potato clades is

97

Figure 2.5. (cont'd)

not indicated (Gagnon et al., 2022). Acylsugar trait data from this report included annotations based upon LC-MS and NMR analysis (Figure 2.2, Tables 2.1, S2.2-26), while data from previous studies were limited to those with NMR structural data (Ghosh et al., 2014b; Herrera-Salgado et al., 2005; Hurney, 2018; King et al., 1986; Leong et al., 2020, 2019; Lou et al., 2021; Lybrand et al., 2020). (A) Distribution of sugar core types within the *Solanum* genus. Inositol and glucose cores determined by characteristic LC-MS negative mode electrospray ionization (ESI) fragmentation (Figures S2.1 and S2.7). (B) Distribution of species with acylsugars that coelute with eggplant acylsugars containing the same acyl chains and sugar core. (C) Distribution of hydroxyacyl chains and glycosylated hydroxyacyl chains.

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

Our survey of *Solanum* species uncovered acylsugars with a surprising diversity of acyl chain lengths, functional groups, and combinations of chain types and positions. We identified C2 to C18 acyl chains, including unsaturated and hydroxylated forms, based on analysis of fatty acid fragment carboxylate ions in negative mode and neutral losses in both negative and positive mode (Table S2.1). Acyl chain compositions differed between acylsugars within and between species: acylinositols with only medium C8 to C14 acyl chains were observed in 5 species (Figure 2.5, Tables S2.13-15, S2.20, S2.22, S2.24), while acylsugars with two short C4 to C6 acyl chains and one medium C8 to C14 acyl chain were observed in 17 species, an acylation pattern also observed in eggplant and wild tomato (Figures 2.2; Table 2.1) (Lybrand et al., 2020). Acylsugars with unsaturated acyl chains, including C5:1, C18:1, C18:2, and C18:3, were only detected in *S. torvum* (Table S2.17). To the best of our knowledge, this is the first report of unsaturated C18 chains in Solanaceae acylsugars. In contrast, hydroxylated medium-length acyl chains were surprisingly common: we detected acylsugars bearing hydroxylated acyl chains in three DulMo clade species and 25 Clade II species (Figure 2.5C). The additional hydroxyl group present on these hydroxylated acyl chains is susceptible to further modifications, including glycosylations and acylations, which would result in even greater acylsugar diversity.

Indeed, we obtained evidence for glycosylated acyl chains in three Clade II species, *S. prinophyllum*, *S. sisymbriifolium*, *S. lasiophyllum*, and one DulMo species, *S. dulcamara*. Several lines of evidence lead us to conclude that these species accumulate acylsugars with glycosyl groups attached via ether linkages to medium hydroxyacyl chains esterified to the hexose core. First, the compounds fragmented abnormally under negative mode CID: the glycosyl group was observed as a neutral loss from the [M-H]$^-$ ion, producing a [sugar – H]$^-$ fragment ion (Figure S2.2). This contrasts with extensive observations that CID spectra of other characterized acyldisaccharides do not exhibit negative mode disaccharide glycosidic linkage cleavage (Figures S2.1 and S2.8) (Ghosh et al., 2014; Hurney, 2018; Lou et al., 2021). Further evidence that the glycosylation is not on the primary acylated hexose ring was obtained by subjecting acylsugar extracts to saponification, which induces ester linkage breakage while retaining ether linkages. Upon analyzing saponified acylsugar extracts by LC-MS, we detected pentose and hexose sugars with C14 or C16 acyl chains rather than glycosylated hexoses, indicating that the glycosyl group is connected to a medium hydroxyacyl chain by an ether linkage that is not cleaved during alkaline saponification (Figure S2.3). We named these compounds glycohydroxyacylhexoses and annotated them by modifying the conventional acylsugar nomenclature where the molecule in Figure S2.2 is named H3:22(4,4,14-O-p) with 14-O-p representing the pentosylated (p) C14 hydroxyacyl chain (14-O). To our knowledge, this is the first report of acylhexoses with glycosylated hydroxyacyl chains in Solanaceae acylsugars. Resin glycosides produced by Convolvulaceae family species contain similar hydroxyacyl chain glycosidic linkages. However, in resin glycosides, hydroxyacyl chains are connected to the oligosaccharide core at two points – by an ester and ether linkage, respectively – forming a

macrocyclic structure not observed in the *Solanum* glycohydroxyacylhexoses (Bah and Pereda-Miranda, 1996; Kruse et al., 2022; Pereda-Miranda et al., 1993).

The unusual glycohydroxyacylhexoses raises questions regarding their biosynthetic and evolutionary origins. As these compounds are found in different lineages (Eastern Hemisphere Spiny, Dulcamaroid, and Sisymbriifolium sections) – and closely related species lack detectable glycohydroxyacylhexoses – this trait appears to be highly labile (Figure 4C). Because we observe the cognate, non-glycosylated acylinositol in each of these species, we hypothesize that glycohydroxyacylhexoses consist of a *myo*-inositol hexose core. Further analysis of acylsugars from additional species may identify other types of modified acylsugar hydroxyacyl chains in *Solanum* acylsugars.

Sixteen species accumulate acylsugars with the same molecular masses, sugar core masses, and acyl chain complements as the NMR-characterized eggplant acylinositols, suggesting they may be identical. To test this hypothesis, we mixed purified eggplant NMR-resolved standards with metabolite extracts from each species and performed liquid co-chromatography. Coelution provided one line of evidence that the eggplant-like acylsugars detected in 11 Clade II species are identical to characterized eggplant acylinositols. In contrast, the eggplant-like acylsugars from five other species, including four DulMo clade members and one Clade II member, eluted separately from characterized eggplant acylinositols, indicating they are acylation positional isomers or acyl chain branching isomers (Figure S2.4). Examples of acylinositol positional isomers were described between the DulMo clade species *S. nigrum* and the Clade II species *S. quitoense* and *S. melongena*, providing precedent that the non-coeluting compounds might be positional isomers of the eggplant acylsugars (Figure 2.2) (Leong et al., 2020; Lou et al., 2021). Our data reveal a phylogenetic pattern of acylinositol distribution, in

which Clade II species generally accumulate eggplant-like acylinositol isomers whereas DulMo clade species accumulate noneggplant-like acylinositol isomers (Figure 2.5B). While the eggplant-like isomers coelute with eggplant acylsugars further analysis with complementary analytical methods is needed to confirm whether these are structurally identical across species.

Acylhexoses and acyldisaccharides accumulate throughout Clade II, DulMo, and VANAns
Solanum species

Acylhexoses were detected in all analyzed acylsugar-producing species which we annotated as acylglucoses or acylinositols based on characteristic fragmentation behavior under negative ion mode MS conditions (Leong et al., 2020; Lou et al., 2021). Identification of acylglucoses was based on observation of a sugar core fragment ion of m/z 143.03 ($C_6H_7O_4-$) and stepwise acyl chain losses at the lowest collision energy level (0 V) (Figure S2.7). In contrast, acylinositol annotation was based on an absence of acyl chain fragment ions at 0 V and MS^e CID negative mode functions (Figure S2.1). Using these rules, we identified acylinositols in all acylsugar-producing species tested except for *S. americanum* (Tables 2.1, S2.1-26). In contrast, acylglucoses were not observed in any Clade II species tested, but were common among the DulMo and VANAns clades, present in all species with characterized acylsugars except *S. dulcamara* (Tables S2.1, S2.18, S2.19, S2.22, S2.23, S2.2520; Figure 2.5A; Lou et al., 2021).

In contrast to the acylhexose distribution, acyldisaccharides were detected in two DulMo species and 14 Clade II species. Based on negative [M-H]- fragment masses and positive mode fragmentation of the glycosidic linkage, the acyldisaccharides were composed of hexose-hexose, pentose-hexose, or deoxyhexose-hexose sugar cores (Table S2.1, Figures S2.1 and S2.8). Though complete structural information cannot be obtained from MS fragmentation alone, we hypothesize that these acyldisaccharides are composed of glycosylated inositol cores based on

two lines of evidence. First, the majority of species with detectable acyldisaccharides also accumulate cognate acylinositols with the same acyl chain lengths (leaf and fruit surface extracts from 10 of 15 and two of two different species, respectively), suggesting a shared inositol-containing core structure (Tables S2.1, S2.2-8, S2.11-16, S2.19, S2.21, S2.25). Second, acylinositol glycosides with varying sugar core sizes and stereochemistries were described in *S. lanceolatum* (Herrera-Salgado et al., 2005), *S. quitoense* (Hurney, 2018), and *S. melongena* (Figure 2.2), suggesting that the species in this study may accumulate similar compounds. The ratio of acylhexose to acyldisaccharide peak number and abundance differs across the surveyed species; for example, *S. mammosum* only accumulated detectable acylinositols while *S. anguivi* primarily accumulated acyldisaccharides that comprised 80% of the total acylsugar peak area (Tables S2.20 and S2.3). This is reminiscent of varied acylhexose accumulation observed in trichomes of evolutionarily distant *Solanum* species: *S. nigrum* only contains detectable acylinositol monosaccharides (Lou et al., 2021), while *S. pennellii* accessions have mixtures of acylglucoses and acylsucroses ranging from 41-95% of total acylsugars as acylglucoses. (Lybrand et al., 2020; Shapiro et al., 1994).

Sticky fruits accumulate acylinositols

We investigated the chemical basis of sticky fruit surface substances previously reported in botanical species descriptions of two Clade II species, *S. acerifolium* and *S. atropurpureum* (Nee, 2022a, 2022b, 1991). LC-MS analysis of fruit surface extracts revealed similar acylsugar profiles between the two species, with each extract containing what appeared to be the same 128 acylsugars. These fruit acylsugars were distinct from the 22 and 19 trichome acylsugars identified in *S. acerifolium* and *S. atropurpureum*, respectively, as evidenced by their LC elution and MS characteristics (Tables S2.13-16; Figure 2.6). Seventy-seven fruit acylsugars were

identified as acylinositols with two or three medium C8 to C14 acyl chains, including one to three hydroxyacyl chains. The remaining 51 fruit acylsugars are acyldisaccharides containing an unusual deoxyhexose-hexose core not previously reported in Solanaceae acylsugars. In contrast, *S. acerifolium* and *S. atropurpureum* trichome acylsugars, composed of a pentose-hexose disaccharide core decorated with short C5 or C6 acyl chains, were not detected in fruit surface extracts.
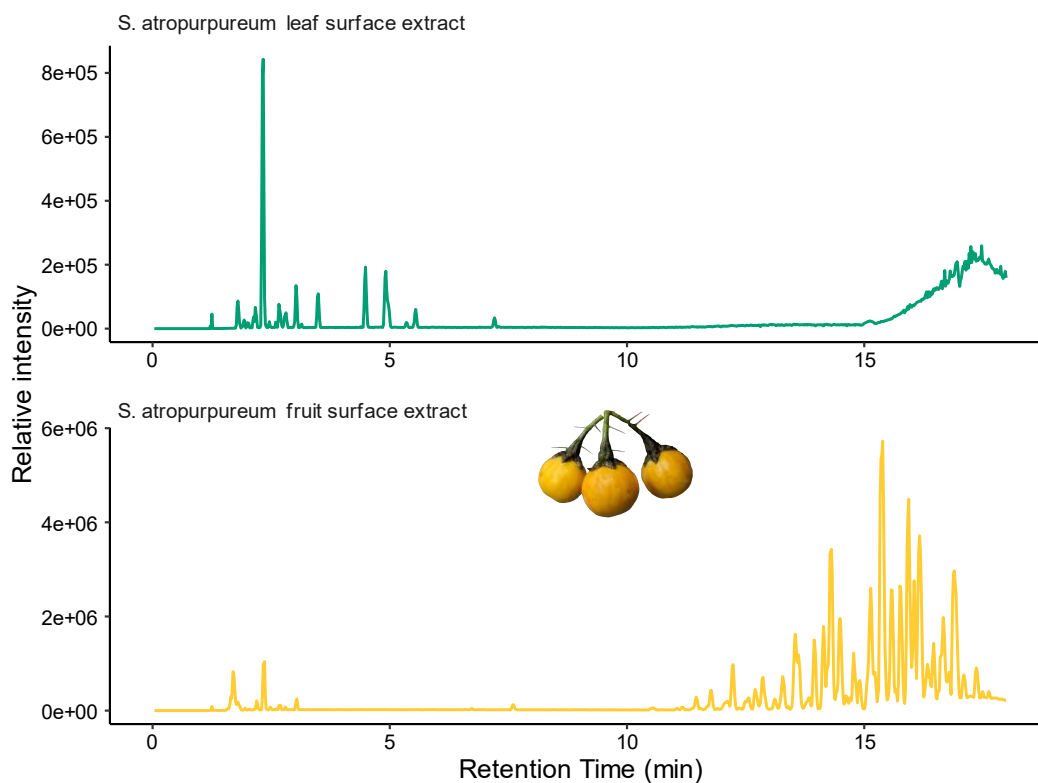
Figure 2.6. *S. atropurpureum* trichomes and fruit exhibit dramatically different metabolite profiles. The top LC-MS base peak intensity (BPI) chromatogram displays metabolites from a leaf surface extract, while the bottom LC-MS base peak intensity chromatogram displays metabolites from a fruit surface extract. The leaf surface extract acylsugars elute from 2.2 to 7.6 min (Table S2.16), while fruit surface extract acylsugars elute from 4.6 to 18.4 min (Table S2.14).

S. mammosum and *S. capsicoides,* which are closely related to *S. acerifolium* and *S. atropurpureum*, do not produce sticky fruit surfaces, suggesting this trait evolved in the common ancestor of *S. acerifolium* and *S. atropurpureum,* perhaps as recently as 2 Mya (Särkinen et al.,

2013). Fruit surface acylsugars were described previously within the *Physalis* genus and likely represent an independent evolutionary origin from the *S. acerifolium* and *S. atropurpureum* fruit acylsugars (Bernal et al., 2018; Cao et al., 2015; Cicchetti et al., 2018; Maldonado et al., 2006). Considering that, like eggplant, *S. acerifolium* and *S. atropurpureum* fruits are glabrous (Nee, 2022a, 2022b), some other organ synthesizes the fruit surface acylsugars. Structures similar to *Solanum fernadesii* petiolar resin glands (Sampaio et al., 2021), *Hypericum androsaemum* microscopic fruit glands (Caprioli et al., 2016), and fennel and chamomile fruit secretory ducts (vittae) (Zizovic et al., 2007) may be involved in *S. acerifolium* and *S. atropurpureum* fruit acylsugar secretion. Future work will be needed to understand the cellular, biosynthetic and evolutionary relationships between trichome and fruit acylsugars.

Sugar core evolution across the Solanaceae

Our screening revealed acylinositols are broadly distributed across Clade II, DulMo, and VANAns (Figure 2.5A, Table S2.1). Outside of this study, acylinositols were only reported in *Solanum* species, including two Clade II members, *S. quitoense* and *S. lanceolatum*, and one DulMo clade member, *S. nigrum*. The acylinositol distribution suggests that acylinositol biosynthesis arose one or more times within the *Solanum* genus. However, our ability to elucidate the likely number of acylinositol origins is impeded by the lack of resolution of the phylogenetic relationships between the major *Solanum* clades, including Clade II, DulMo, VANAns, Potato, and Regmandra (Gagnon et al., 2022). Further biochemical and genetic analyses of acylinositol biosynthesis may reveal how this trait evolved and whether the pathway is conserved between Clade II, DulMo, and VANAns species.

Acylglucoses were reported to be sporadically present across the Solanaceae, including in *Salpiglossis sinuata*, *Petunia*, *Nicotiana* spp., and *Datura*, as well as *S. nigrum* and *S. pennellii*,

104

members of the DulMo and Potato clades, respectively (Castillo et al., 1989; Chortyk et al.,

1997; Fiesel et al., 2022; Fobes et al., 1985; Hurney, 2018; King and Calhoun, 1988; Lou et al.,

2021; Matsuzaki et al., 1989; Schenck et al., 2022; Van Dam and Hare, 1998). We observed

acylglucoses in members of the DulMo and VANAns clades, but not in Clade II. This spotty

distribution suggests that acylglucose biosynthesis 1) arose once within the Solanaceae then

underwent repeated losses or 2) arose independently several times in different Solanaceae clades.

Recent work in *S. pennellii* and *S. nigrum* revealed that acylglucoses are synthesized from

acylsucroses by a neofunctionalized invertase-like enzyme, AcylSucrose FructoFuranosidase 1

(ASFF1) (Leong et al., 2019; Lou et al., 2021). Interestingly, *S. nigrum* and *S. pennellii* employ

non-orthologous ASFF1 enzymes, providing support for multiple acylglucose biosynthesis

origins. Elucidation of acylglucose biosynthesis outside of the *Solanum* clade will be needed to

resolve the evolutionary history of these metabolites.

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

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

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

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

acylsucrose biosynthesis likely persists in DulMo and VANAns species considering that

acylglucoses were detected in all but one DulMo and VANAns species and that acylsucroses can

be intermediates in acylglucose biosynthesis (Leong et al., 2019; Lou et al., 2021). The lack of

both acylglucoses and acylsucroses in the analyzed Clade II species suggests a loss of

acylsucrose biosynthesis in Clade II. Rather, our LC-MS data are consistent with the hypothesis

that the acyldisaccharides observed in this study consist of glycosylated inositol cores. Recent *in*

vitro biochemistry and *in vivo* genetic evidence in *S. quitoense* suggests that acylinositols are the

precursors to acylinositol disaccharides in Clade II *Solanum* species (Leong et al., 2022, 2019).

In this scenario, acylinositols would be converted to their cognate glycosides by one or more

glycosyltransferases, with different sugars added either by a conserved promiscuous

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

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

the biochemical origins of these unusual *Solanum* acyldisaccharides.
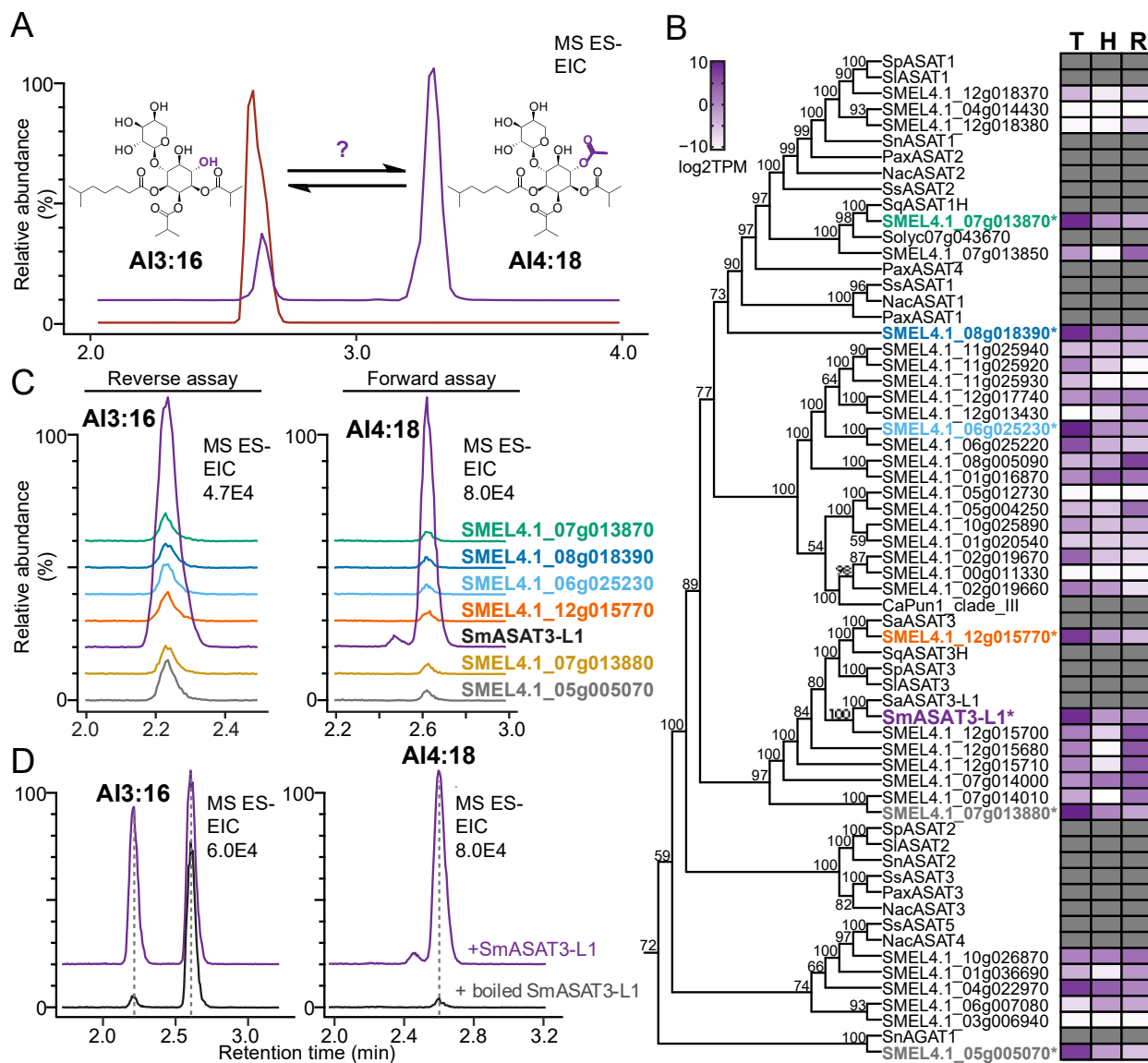
Figure 2.7. SmASAT3-L1 acetylates AI3:16 to produce AI4:18. (A) LC-MS analysis of AI3:16(4,4,8) and AI4(2,4,4,8) in *S. aethiopicum* (red) and *S. melongena* (purple) leaf surface extracts. *S. aethiopicum* does not produce detectable levels of AI4:18. (B) *S. melongena* Clade III

Figure 2.7. (cont'd)

BAHDs. 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 S2.11 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 not included in 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. (C) Of the seven tested enzymes, only SmASAT3-L1 acetylates AI3:16(4,4,8) to form AI4:18(2,4,4,8). The extracted ion chromatograms on the left display products from forward enzyme assays and the formate adduct of AI4:18, *m/z* 665.30. The extracted ion chromatograms on the right display products from reverse enzyme assays and the formate adduct of AI3:16, *m/z* 623.29. (D) SmASAT3-L1 acetylates AI3:16 *in vitro* to produce AI4:18. 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 remaining AI4:18.

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

We began analyzing the biochemical mechanisms underlying the observed *Solanum* acylsugar diversity in seven of 23 members from the Eggplant clade and related Anguivi grade, two small subclades within Clade II that includes brinjal eggplant and its close relatives. In contrast to the remaining six analyzed species, *S. aethiopicum* (scarlet or Ethiopian eggplant) does accumulate detectable AI3:16(i4,i4,i8) (Compound 1) but does not accumulate the acetylated form, AI4:18(2,i4,i4,i8) – the highest abundance *S. melongena* acylsugar (Compound 2; Figure S2.6, Tables S2.1, S2.3, S2.4). We hypothesized that the absence of AI4:18(2,i4,i4,i8) in *S. aethiopicum* is due either to mutations causing loss-of-function or loss of expression of an acetylating enzyme in this species. The availability of *S. melongena* and *S. aethiopicum* genomic sequences (Barchi et al., 2021; Song et al., 2019) provided the opportunity to seek a trichome expressed acetyltransferase responsible for producing AI4:18 and the mechanism behind the lack of detectable AI4:18 in *S. aethiopicum*.

To identify suitable enzyme candidates, we sequenced transcriptomes of isolated

trichomes, trichome-depleted shaved hypocotyls, and whole roots collected from 7-day-old

eggplant seedlings and performed differential gene expression analysis. Of the 23,251 eggplant

genes expressed in at least one of the three tissues, 745 were significantly enriched (log2 fold-

change >2, p-value < 0.05) in trichomes compared to hypocotyls and roots, including 20 BAHDs

(Supplemental datafile 1). We selected for further testing seven BAHDs that were abundantly

expressed in trichomes (TPM > 200) and homologous to characterized *Solanum* ASATs (Figure

2.7B) (D'Auria, 2006; Fan et al., 2015; Leong et al., 2022, 2020; Lou et al., 2021). We expressed

each of the seven candidates in *Escherichia coli* and tested for acetylation of Compound 1,

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

ACYLTRANSFERASE3-LIKE1 (SmASAT3-L1) was the only enzyme to exhibit forward

activity converting Compound 1, AI3:16(4,4,8), to Compound 2, AI4:18(2,4,4,8) (Figure 2.7C).

As shown in Figure S2.14, LC-MS analysis of the SmASAT3-L1 *in vitro* assay product revealed

that it had identical molecular masses and elution times to plant-derived AI4:18(2,i4,i4,i8)

(Compound 1) providing evidence that SmASAT3-L1 acetylates AI3:16(i4,i4,i8).

Characterization of reverse activities of the seven BAHD acyltransferases (Leong et al., 2020;

Lou et al., 2021), in which an acyl chain can be removed from an acylsugar by incubating it with

the enzyme and free Coenzyme A, confirmed that SmASAT3-L1 was the only candidate enzyme

that removed an acetyl chain from AI4:18(2,4,4,8) (Figures 2.7D and S2.14). Taken together,

these results indicate that trichome-expressed *SmASAT3-L1* encodes an acylinositol

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

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

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

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

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

aethiopicum trichomes (Figures S2.10). This result implicates the defect in ASAT3-L1 enzyme

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

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

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

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

reduced accumulation of acetylated acylsugars in these accessions (Kim et al., 2012).
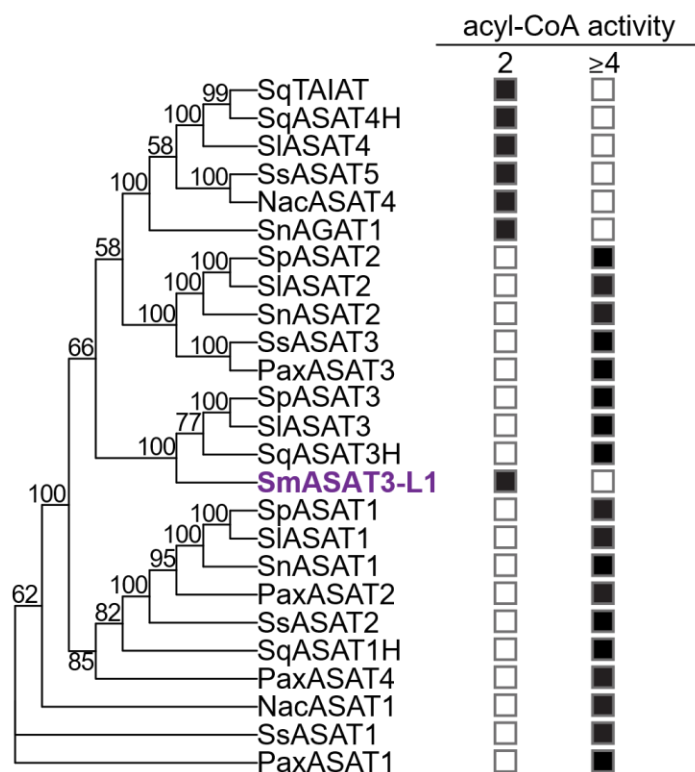
Figure 2.8. Characterized *in vitro* acyl-CoA usage by ASATs. The maximum likelihood phylogeny was inferred from amino acid sequences using a time-reversible algorithm specifying a plant-specific empirical substitution matrix, invariant sites, and four rate categories (Q.plant+I+G4) in IQ-TREE v2.1.3. Bootstrap support was calculated from 100,000 ultrafast bootstrap replicates; values >=95 indicate strong support. 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; Schilmiller et al., 2012). Only positive activities are displayed, and the blank squares may not indicate a lack of activity.

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

Knowledge of *Solanum* acylsugar diversity provides a framework for investigating the molecular basis of metabolic pathway evolution. For example, the widespread occurrence of acylglucoses across the DulMo clade leads us to hypothesize that the neofunctionalized invertase responsible for *S. nigrum* acylglucose synthesis (Lou et al., 2021) is shared by other DulMo clade species. Similarly, documented acylsugar variation between cultivated eggplant, *S. melongena,* and its relative, *S. aethiopicum*, led to the discovery of a terminal acylsugar acetyltransferase in eggplant, SmASAT3-L1. The characteristics of ASAT3-L1, the enzyme absent from the *S. aethiopicum* trichome transcriptome, in turn reveals the dynamic nature of plant specialized metabolism. First, SmASAT3-L1 is unique among characterized ASATs in its

110

acyl acceptor specificity, acetylating an acylinositol glycoside rather than an acylated sucrose, glucose, or *myo*-inositol. Second, acetyl-CoA donor activity was previously only described for a subset of Clade III BAHDs, which include ASAT4 and ASAT5 clade members (Figure 2.8). This breaks a pattern seen across the Solanaceae, including trichome ASATs from species of early evolving lineages, *Petunia axillaris* and *Salpiglossis sinuata* (Moghe et al., 2017; Nadakuduti et al., 2017), *Nicotiana attenuata* (Schenck et al., 2022), as well as cultivated and wild tomatoes (Kim et al., 2012; Schilmiller et al., 2012). This theme was also observed in trichomes of two other *Solanum* plants, more closely related to the species screened in this study: the Clade II *S. quitoense* SqTAIAT (Leong et al., 2020) and DulMo *S. nigrum* SnAGAT1 (Lou et al., 2021). Both enzymes reside outside of the ASAT3 clade, and thus are phylogenetically distinct from SmASAT3-L1. SmASAT3-L1 characteristics and those of recently published *S. quitoense* enzymes (Leong et al., 2022, 2020) also suggest that acylinositols are synthesized through one or more pathway(s) distinct from that of acylsucroses. Existing genetic resources in *S. nigrum*, *S. quitoense*, and *S. melongena* can be utilized to explore when and how acylinositols arose and to identify the underlying mechanisms behind acylation position and chain type differences between these species.

The remarkably varied structures of acylsugars, and their associated physical properties, strongly suggest that they have evolved distinct bioactivities, likely in coevolutionary arms races between plants and their insect and microbial pests. For example, hydrophobic acylsugars with longer acyl chains and few free hydroxyl groups may disrupt membranes, reminiscent of less polar triterpene saponin variants (Augustin et al., 2011; Baumann et al., 2000). Variation in acyl chain linkage chemistries also likely influence acylsugar mode of action. For example, ester linked short chains acylsucroses are digested by *Manduca sexta* (hawkmoth) larvae and the

volatile organic acids released into the environment, attracting ant predators (Weinhold and Baldwin, 2011). In contrast, longer chain hydroxylated and ether-linked glycohydroxyacyl chains observed in the study are likely to have quite different metabolic fates, perhaps persisting in the digestive systems of herbivores and alternative metabolism in microbes. Uncovering the enzymes responsible for synthesis of the wide variety of natural acylsugars will enable transgenic production of isogenic variants and rigorous testing of their ecological activities. Similarly, access to a wider variety of often promiscuous ASAT BAHD acyltransferases (Moghe et al., 2023) creates opportunities for producing novel acylsugars with a wide range of physical properties and bioactivities (Schenck et al., 2022).

Hydroxyacyl chains also present promising opportunities to expand upon natural acylsugar diversity through synthetic biology and/or synthetic chemistry. The hydroxyl acts as a reactive chemical handle which was exploited in nature to create glycohydroxyacylhexoses. We can further modify the hydroxyacyl chains to add unusual sugars, acyl chains, and aromatic groups through promiscuous enzymes or synthetic chemistry. These structural changes would likely impact biological activities and add another asset to developing more pest resistant plants. Characterizing more natural acylsugar diversity will further enable synthetic biology approaches to create completely new acylsugar structures.

Methods

Plant material and growth conditions

Solanum spp. seeds were obtained from the sources described in Table S2.37. Seeds were treated with 10% (v/v) bleach for 10 min while being rocked at 24 rpm with a GyroMini nutating mixer (Labnet, Edison, NJ, USA), and then rinsed 5-6 times with distilled water. Unless otherwise noted, seeds were germinated on Whatman filter paper (MilliporeSigma, Burlington, MA, USA)

at 28°C and in the dark. Germinated seedlings were transferred to peat pots (Jiffy, Zwijndrecht, Netherlands), and grown at 25°C, 16/8-h day/night light cycle, and ~70 µmol m^{-2}s^{-2} photosynthetic photon flux density with cool white fluorescent bulbs. Mature plants were grown in controlled environment growth chambers or in a greenhouse. The growth chamber conditions consisted of 25°/18°C day/night temperatures, 16/8-h day/night light cycle, and ~100 µmol m^{-2}s^{-2} photosynthetic photon flux density under LED bulbs. The greenhouse conditions consisted of a 16/8-h day/night light cycle achieved with supplemental sodium iodide lighting. Plants were fertilized weekly with 0.5X Hoagland's solution.

Surface metabolite extractions

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

Bulk eggplant acylsugar extraction and purification for NMR analysis

Surface metabolites were bulk extracted from approximately 1000 *S. melongena* PI 555598 seedlings. Seeds were treated with 10% (v/v) bleach (Clorox, Oakland, CA, USA) for 10 min while being gently rocked at 24 rpm with a GyroMini nutating mixer and subsequently

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

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

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

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

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

cutting them at the base and placing them into two 2L beakers each containing 1L 100%

acetonitrile. Surface metabolites were extracted with gentle agitation with a metal spatula for five

min at room temperature. Plant material and sediment was removed by vacuum filtration through

a Büchner funnel lined with Whatman filter paper (MilliporeSigma). Solvent was removed *in*

vacuo by rotary evaporation and dried residue was dissolved in 20 mL of acetonitrile with 0.1%

formic acid. Solvent was removed again using a vacuum centrifuge and the dried residue was

dissolved in 1 mL of acetonitrile:water:formic acid (80:20:0.001). The semi-preparative LC

method is described in detail in Table S2.32 and is described in brief here. *S. melongena*

acylsugars were separated with a Waters 2795 HPLC (Waters Corporation, Milford, MA, USA)

equipped with an Acclaim 120 C18 HPLC column (4.6 x 150 mm, 5 μm; Thermo Fisher

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

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

was 1.5 mL/min, injection volume was 100 μL, and the column temperature was 40°C. Fractions

were collected automatically at 0.25 min intervals by a 2211 Superrac fraction collector (LKB

Bromma, Stockholm, Sweden) and assessed for acylsugar presence and purity by LC-MS

analysis, as described below. Column fractions were collected in the same tubes for each method

run which worked to pool the fractions between each run.

<u>LC-MS acylsugar analysis</u>

Acylsugars were analyzed by LC-MS with each method described below and in Tables S2.27-31. For each LC method, the mobile phase consisted of aqueous 10 mM ammonium formate, adjusted to pH 2.8 with formic acid (Solvent A) and 100% acetonitrile (Solvent B). The flow rate was maintained at 0.3 mL/min.

Acylsugar extracts were analyzed with a 22 min LC gradient using a Waters Acquity UPLC coupled to a Waters Xevo G2-XS QToF mass spectrometer (Waters Corporation, Milford, MA) equipped with electrospray ionization (ESI) operating in positive (ESI$^+$) and negative (ESI$^-$) modes. 10 µL acylsugar extracts were separated on an Acquity UPLC BEH C18 column (10 cm x 2.1 mm, 130 Å, 1.7 µm; Waters), kept at 40°C, using a binary solvent gradient. The 22 min linear gradient was as follows: 5% B at 0 min, 60% B at 2 min, 100% B at 16 min, held at 100% B until 20 min, 5% B at 20.01 min, held at 5% B until 22 min. Acylsugar extracts were analyzed with both ESI$^-$, and ESI$^+$. For ESI$^-$, the following parameters were used: capillary voltage, 2 kV; sampling cone voltage, 35 V; source temperature, 100°C; desolvation temperature 350°C; cone gas flow, 50 L/Hr; desolvation gas flow, 600 L/Hr. For ESI$^+$, the following parameters were used: capillary voltage, 3 kV; sampling cone voltage, 35 V; source temperature, 100°C; desolvation temperature 300°C; cone gas flow, 50 L/Hr; desolvation gas flow, 600 L/Hr. Acylsugars were fragmented in either MS^E or data-dependent acquisition (DDA) MS/MS modes as described previously (Lou et al., 2021; Lybrand et al., 2020). DDA survey and MS/MS functions acquired over m/z 50 to 1500 with scan times of 0.2 s. Ions selected for MS/MS were fragmented with a ramped collision energy where voltage varies based on ion mass. Collision energies followed a ramp with the voltage changing linearly for ions between the low mass setting (m/z 50, 15 to 30 V) and high mass setting (m/z 1500, 30 to 60 V). To increase mass

accuracy, lock mass correction was performed during data collection, with leucine enkephalin as the reference.

Semi-preparative LC fractions were analyzed with direct infusion and a 14-min LC gradient using a LC-20ADvp ternary pump (Shimadzu, Kyoto, Japan) coupled to a Waters Xevo G2-XS QToF mass spectrometer (Waters Corporation). The direct infusion method quickly screened fractions for acylsugar presence while the 14-min LC method tested fraction purity. 10 µL of acylsugar fractions were injected into an Ascentis Express C18 HPLC column (10 cm x 2.1 mm, 2.7 µm; Supelco, Bellefonte, PA, USA), kept at 40°C. The 14-min gradient was as 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 min, and held at 5% until 14 min. Fractions were analyzed under ESI[+] with the following parameters: capillary voltage, 3 kV; sampling cone voltage, 40 V; source temperature, 100°C; desolvation temperature 350°C; cone gas flow, 20 L/hr; desolvation gas flow, 500 L/hr. For direct infusion analysis, ions were acquired from *m/z* 50 to 1500 with a scan time of 0.1 s and one acquisition function with 0 V collision potential. For 14-min LC analysis, ions were acquired from *m/z* 50 to 1200 with a scan time of 0.1 s and three acquisition functions with different collision potentials (0, 25, 60 V). Lock mass calibration referenced to the leucine enkephalin [M+H][+] ion was applied during data acquisition.

Enzyme assays products were analyzed with a 7-min gradient using Waters Acquity UPLC coupled to a Waters Xevo G2-XS QToF mass spectrometer (Waters Corporation, Milford, MA) equipped with electrospray ionization (ESI) operating in negative (ESI[-]) mode. Reaction products were separated with an Ascentis Express C18 HPLC column (10 cm x 2.1 mm, 2.7 µm; Supelco), kept at 40°C, using a binary solvent gradient. The 7 min linear gradient was as follows: 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,

held at 5% B until 7 min. The ESI- parameters described for acylsugar extract analysis were used. Ions were acquired from *m/z* 50 to 1200 with a scan time of 0.1 s and three acquisition functions with different collision potentials (0, 25, 60 V). Lock mass calibration referenced to the leucine enkephalin [M+H]- ion was applied during data acquisition.

For coelution analysis between enzymatically- and plant-produced AI3:16 and AI4:18, a 24-min linear gradient was used with an Acquity UPLC BEH C18 column (10 cm x 2.1 mm, 130 Å, 1.7 µm; Waters), kept at 40ºC, on the same instrument used for enzyme assay analysis. The binary solvent, linear gradient was as follows: 5% B at 0 min, 60% B at 18 min, 100% B at 18.01 min, held at 100% B until 22 min, 5% B at 22.01 min, held at 5% B until 24 min. ESI- parameters and MS method from the above enzyme assay analysis were used.

Saponified acylsugars were analyzed using a Waters Acquity UPLC coupled to a Waters Xevo G2-XS QToF mass spectrometer (Waters Corporation, Milford, MA). 10 µL of saponified acylsugars were injected into either an Acquity UPLC BEH C18 column (10 cm x 2.1 mm, 130 Å, 1.7 µm; Waters) or an Acquity UPLC BEH Amide column (10 cm x 2.1 mm, 130 Å, 1.7 µm; Waters), kept at 40°C. Samples were analyzed on the C18 column with the 22-min method described above. To detect free sugars resulting from saponification, samples were analyzed on the BEH amide column with a 9 min method which was as follows: 95% B at 0 min 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. Both methods used the ESI- parameters and MS acquisition parameters described for enzyme assays.

Acylsugar annotation

We inferred acylsugar structures utilizing negative and positive mode MS and MS/MS collision-induced dissociation (CID) as previously described (Ghosh et al., 2014a; Hurney, 2018; Leong et al., 2019; Lou et al., 2021; Lybrand et al., 2020). We employed modified confidence

criteria developed by Lou and coworkers (Lou et al., 2021). Putative acylsugars only meeting criterion A were annotated as low confidence. Putative acylsugars meeting criteria A-D were annotated as medium confidence. We also added two additional criteria, G and H, described below. Acylsugars meeting criteria A-F and either G or H were annotated with highest confidence. Confidence levels for each reported acylsugar are specified in their species-specific annotation table (Tables S2.1-26).

(G) **NMR structural determination.** Acylsugars from *S. melongena* were further characterized by NMR experiments in which sugar core relative stereochemistry, acyl chain positions, and acyl chain branching were resolved.

(H) Coelution with NMR characterized compounds. Acylsugars with the same acyl chains and sugar core size were analyzed by coelution with *S. melongena* acylsugars. If acylsugars coeluted with an NMR characterized peak and shared the same acyl chains and fragmentation patterns, they met this criterion and were annotated with high confidence. This criterion was employed for acylsugars from species other than *S. melongena*.

Acylsugar metabolomic analysis

To collect accurate formate adduct masses and relative abundances from putative acylsugars, negative mode LC-MS raw data from metabolite extracts were analyzed with Progenesis QI software v3.0 (Waters, Milford, MA, USA) using retention time alignment, peak detection, adduct grouping, and deconvolution. Lock mass correction was performed during data collection. We used the following analysis parameters: peak picking sensitivity, default; retention time range, 0-22 min; adducts detected, M+FA-H, M+Cl, M-H, 2M-H, and 2M+FA-H. Average percent peak abundance was calculated for each putative acylsugar by dividing its raw

abundance by the sum of the total acylsugar raw abundance and then averaging that value across all samples for a species.

NMR analysis of acylsugars

Purified acylsugars were dissolved in acetonitrile-d_3 (99.96 atom % D; MilliporeSigma, Burlington, MA, USA) and transferred to Kontes NMR tubes (MilliporeSigma, Burlington, MA, USA). Samples were analyzed at the Michigan State University Max T. Rogers NMR Core with a Varian Inova 600 MHz spectrometer (Agilent , Santa Clara, CA, USA) equipped with a Nalorac 5 mm PFG switchable probe, a DirectDrive2 500 MHz spectrometer (Agilent, Santa Clara, CA, USA) equipped with a OneNMR probe, Bruker Avance NEO 800 MHz spectrometer (Bruker, Billerica, MA, USA) equipped with a 5 mm helium cryogenic HCN probe, or a Bruker Avance NEO 600 MHz spectrometer equipped with a 5 mm nitrogen cryogenic HCN Prodigy probe. All acylsugars were analyzed with a series of 1D (^{1}H, and ^{13}C) and 2D (gCOSY, TOCSY, gHSQC, gHMBC, gH2BC, J-resolved) experiments. Resulting spectra were referenced to acetonitrile-d_3 ($^\delta$H = 1.94 and $^\delta$C = 118.70 ppm). NMR spectra were processed and analyzed with MestReNova 12.0 software (MestreLab, A Coruña, Spain). For full NMR metadata, see Tables S2.47-50.

Sugar core composition analysis

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

ammonium hydroxide for 15 min at room temperature. Addition of 50 µL 20 mg/mL NaBH$_4$ and incubation at 100°C for 10 min converted aldoses and ketoses to polyols. Excess sodium borohydride was quenched with 100 µL 1M trifluoroacetic acid. Two volumes of methanol were added, and the sample was dried down by a stream of nitrogen gas. Dried down reaction products were redissolved in 200 µL of methanolic 0.5 M HCl and heated at 100°C for 15 min in a closed ½ dram vial, 12 x 35 mm (Kimble Chase, Vineland, NJ, USA). Solvent was removed by a stream of nitrogen gas. To acetylate the sugar cores, residue was dissolved in 200 µL of pyridine:acetic anhydride (1:1) and incubated at 100°C for 30 min. Acetylated polyols were then dried down *in vacuo* in a vacuum centrifuge and redissolved in 100 µL hexane for GC-MS analysis.

To determine the composition of disaccharide sugar cores, saponified sugars were hydrolyzed prior to derivatization. Dried down sugars were redissolved in 98 µL water in a ½ dram vial, 12 x 35 mm (Kimble Chase). After addition of 102 µL of 88% formic acid, the vial was sealed and heated at 100°C for 15 hours in a heat block. Hydrolyzed sugars were dried down *in vacuo* with a vacuum centrifuge and derivatized to alditol acetates following the method described above.

Hydroxyl acyl chain stereochemistry analysis

We derivatized and prepared hydroxyacyl chains as previously described (Jenske and Vetter, 2007). The commercial standards (3*R*)-OH-tetradecanoate (Cayman Chemical, Ann Arbor, MI, USA) and (3*R*/*S*)-OH-tetradecanoate (Cayman Chemical, Ann Arbor, MI, USA) were first derivatized to their respective ethyl ester derivatives. 600 µg of each fatty acid was dissolved in 0.5 mL of 0.5 M ethanolic KOH and incubated at 80°C for 5 min. After the solution was cooled on ice, 1 mL of ethanolic BF$_3$ was added, and reactions were incubated at 80°C for

five min. Two-phase partitioning with hexane and saturated sodium chloride yielded ethyl ester derivatives in the hexane partition.

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

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

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

<u>GC-MS analysis</u>

All GC-MS analyses employed an Agilent 5890 GC and an Agilent 5975 single quadrupole MS equipped with a FactorFour VF-5ms column (30 m x 0.25 mm, 0.25 um; Agilent) and 10 m EZ-Guard column (Agilent, Santa Clara, CA, USA). Helium was used as the carrier gas with a constant flow rate of 1 mL/min. Electron energy was set at 70 eV. MS source and quadrupole were maintained at 230°C and 150°C, respectively. Parameters specific to each analysis are listed below.

For analysis of sugar core alditol acetate derivatives, we followed previously published GC-MS parameters (Sassaki et al., 2008). Inlet temperature was maintained at 275°C. GC oven temperature was held at 60°C for one min and then was ramped at a rate of 40°C/min to 180°C. Oven temperature was then ramped to 240°C at a rate of 5°C/min and held for three min. Total run time was 19 min. The MS detector transfer line was maintained at 280°C. Split ratio was 10:1 with a split flow rate of 10 mL/min. A three min solvent delay was used. MS data was collected in full scan mode, *m/z* 50-600.

For analysis of MPTA fatty acid ethyl ester derivatives, we followed previously published GC-MS parameters (Jenske and Vetter, 2007). GC oven temperature was held at 60°C 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 after which the oven temperature was ramped to 230°C at a rate of 2°C/min. After 230°C was held for nine min, the oven temperature was ramped to 300°C at 10°C/min and held for 7.5 min. Total run time was 55 min. A 5-min solvent delay was applied. The MS detector transfer line

was maintained at 280°C. Selective ion monitoring detected the ions *m/z* 189, 209, and 255 with a dwell time of 50 ms.

Sample collection and transcriptome sequencing

We sequenced 18 transcriptomes: six biological replicates each of trichomes isolated from hypocotyls (trichomes), hypocotyls stripped of their trichomes (stripped hypocotyls), and whole roots (Supplemental data file 1). We sampled trichomes and stripped hypocotyls from 7-day-old *S. melongena* 67/3 seedlings following methods developed for root hair cell isolation (Bucher et al., 1997) with modifications. We grew lawns of eggplant seedlings in soil flats as described above for bulk leaf surface metabolite extraction. At 7 days post germination, we removed roots and cotyledons from seedlings, then transferred hypocotyls to liquid nitrogen in a plastic 2 L Dewar flask. Frozen hypocotyls were gently stirred with a glass rod for 20 minutes to physically shear trichomes from hypocotyls. After confirming under a dissecting scope that a sample of three hypocotyls had been stripped bare, we filtered trichomes into a 2 L glass beaker by slowly pouring the contents of the Dewar flask through a 500 μm wire mesh sieve (MilliporeSigma, Burlington, MA, USA). To maximize trichome recovery, stripped hypocotyls were returned to the Dewar, rinsed with liquid nitrogen, and filtered through the 500 μm sieve six more times. Stripped hypocotyls were divided into six pre-weighed 50 mL conical tubes, quickly weighed, then transferred to -80ºC for storage. Trichomes were subsequently filtered through a 150 μm sieve (MilliporeSigma, Burlington, MA, USA) into a 500 mL beaker to increase sample purity then transferred to a 50 mL conical tube to allow the excess liquid nitrogen to evaporate. Finally, filtered trichomes were divided into six pre-weighed 2 mL screw-cap tubes (Dot Scientific, Inc., Burton, MI, USA), quickly weighed, then transferred to -80ºC for storage.

We extracted total RNA using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's directions, measured RNA concentration using a Qubit 1.0 instrument (Thermo Fisher Scientific, Waltham, MA, USA) with the RNA HS assay, and the samples were processed by the Michigan State University Research Technology Support Facility Genomics Core (East Lansing, MI, USA) for library preparation and high-throughput sequencing. The core checked RNA quality using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA), constructed sequencing libraries using an Illumina Stranded mRNA Prep kit (Illumina, Inc. San Diego, CA, USA), and sequenced all 18 libraries on a NovaSeq 6000 (Illumina, Inc. San Diego, CA) using a single S4 flow cell lane, producing 150-base pair paired-end reads (Supplemental datafile 1). The raw fastq files are available to download from the Sequence Read Archive database under BioProject PRJNA935765.

Transcriptome alignment and differential gene expression analysis

To prepare the 150-bp paired-end sequences for alignment, we employed Trimmomatic (Bolger et al., 2014) to trim adapters and low-quality bases, then filter reads shorter than 75 bp, removing an average of 3.7% of sequences (range: 3.3% - 4.2%). We mapped the resulting 75–150-bp (average length: 149 bp) paired-end RNAseq reads separately to the eggplant V4.1 and HQ reference genomes using STAR in two-pass mode, which enhances splice junction discovery and mapping sensitivity (Dobin et al., 2013; Dobin and Gingeras, 2015). Using this approach, an average of 79.2% (range: 60.1 - 87.1%) and 82.2% (range: 62.0 - 90.7%) of reads mapped to a unique genomic location in V4.1 and HQ, respectively (Supplemental datafile 1). We filtered the resulting transcriptome alignments according to best practices as defined by the Genome Analysis Toolkit (GATK) (DePristo et al., 2011; Van der Auwera et al., 2013). Briefly, we removed optical and PCR duplicates with MarkDuplicates from the Picard toolkit

(http://broadinstitute.github.io/picard), parsed reads into exon segments and removed intron-spanning bases using SplitNCigarReads from GATK (McKenna et al., 2010). Finally, we selected unique alignments by eliminating reads with a mapping quality score below Q60 with the view command from SAMtools (Li et al., 2009). To generate raw read counts, we used the HTSeq command, htseq-count with the nonunique parameter set to all (Anders et al., 2015).

Raw read counts generated by HTSeq-count were used to perform differential gene expression analysis in edgeR (Robinson et al., 2010). To restrict comparisons to expressed genes, only transcripts with at least one read count-per-million (CPM) in at least one sample were retained for further analysis. This filtering step removed 11,665 (33.4%) and 14,544 (39.8%) annotated genes in the V4.1 and HQ eggplant genomes, respectively, yielding 23,251 and 22,024 expressed transcripts for differential gene expression analysis. Next, we normalized transcript abundances across the 18 transcriptomes in our eggplant V4.1 and HQ alignments using the default trimmed mean of M-values (TMM) method in the calcNormFactors function, then performed multidimensional scaling (MDS) with the plotMDS function to compare global gene expression profiles (Figure S2.13). This showed that our samples cluster tightly by tissue (i.e., trichome, shaved hypocotyl, or root), with no obvious differences between V4.1 and HQ alignments. To test differences in gene expression across tissues, we implemented a generalized linear model (GLM) using a quasi-likelihood (QL) approach: we generated an experimental design matrix specifying the three tissues (i.e., trichomes, trichomeless hypocotyls, and roots) with the model.matrix function, then used the glmQLFit function to fit our data to a QL-GLM model. To identify genes with a log2 fold-change (FC) > 2 between tissues, we used the glmTreat function, which performs threshold hypothesis testing, a rigorous statistical approach that evaluates variance and magnitude to detect expression differences greater than the specified

value (e.g., log2 FC > 2), then applies false discovery rate (FDR) p-value corrections. Genes were classified as significantly differentially expressed between two tissues if log2 FC > 2 and FDR-corrected p-value < 0.05. We calculated absolute transcript abundance for all expressed genes as transcripts per million (TPM) with the calculateTPM function in scateR (McCarthy et al., 2017).

Phylogenetic analyses

Characterized ASAT enzymes fall into clade III of the BAHD family (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; Anthony L Schilmiller et al., 2015; Schilmiller et al., 2012). To identify ASAT candidates in eggplant, we searched for sequences containing the Pfam Transferase domain (PF02458) (Finn et al., 2010), associated with all characterized catalytically active BAHD proteins. The PF02458 HMM profile was obtained from the Pfam website (http://pfam.sanger.ac.uk/) and queried against the V4.1 and HQ eggplant proteomes using the hmmsearch tool from HMMER v3.2.1 (hmmer.org), revealing 106 and 108 putative BAHD sequences, respectively. Using MAFFT v7.471 in E-INS-i mode, we built multiple sequence alignments (MSA) of amino acid sequences from four sources: 1) V4.1 or HQ eggplant PF02458 hits, 2) published reference sequences for clades I-VII (Moghe et al., 2023), 3) characterized ASAT sequences from other Solanaceae species, and 4) the SaASAT3 and SaASAT3-L1 candidates from *S. aethiopicum* (Katoh and Standley, 2013). The E-INS-i algorithm implements local alignment with a generalized affine gap cost (Altschul, 1998), which aligns conserved regions (e.g., the BAHD transferase domain) and essentially ignores nonconserved regions. Phylogenetic reconstruction was performed using IQ-TREE v2.1.3 (Minh et al., 2020). The ModelFinder tool was implemented to identify the best maximum likelihood model for

estimating evolutionary relationships (Kalyaanamoorthy et al., 2017), leading to selection of

Jones-Taylor-Thornton (JTT)+F with seven rate categories and Q.plant+F with seven rate

categories for the V4.1 and HQ MSAs, respectively. Phylogenetic trees were inferred by

maximum likelihood using the chosen model, and branch support was obtained from 100,000

ultrafast bootstrap iterations (Hoang et al., 2018). The resulting BAHD phylogenies were

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

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

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

<u>BAHD acyltransferase cloning, expression, and purification</u>

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

Open reading frames from genes were either synthesized by Twist Biosciences (South San

Francisco, CA, USA) with or without codon-optimization for *E. coli* expression (Table S2.35) or

amplified from genomic DNA or cDNA with primers listed in Table S2.36. Q5 2X Hotstart

master mix (New England Biolabs, Ipswich, MA, USA) was used for cloning PCRs. The source

of cloned DNA and whether a gene was codon optimized is described in Table S2.35. Amplified

genes were purified by agarose gel electrophoresis and extraction with the Monarch DNA Gel

Extraction Kit (New England Biolabs, Ipswich, MA, USA). Both the synthesized genes and PCR

amplified genes were then inserted into a doubly digested BamHI/XhoI pET28b(+) through

Gibson assembly using the 2X Gibson Assembly Master Mix (New England Biolabs, Ipswich,

MA, USA) according to the manufacturer's instructions. The constructs using synthesized genes

were transformed into BL21(DE3) (MilliporeSigma, Burlington, MA, USA) and constructs with

PCR amplified genes were transformed into BL21 Rosetta(DE3) cells (MilliporeSigma,

Burlington, MA, USA). Constructs were verified with colony PCR and Sanger sequencing using

T7 terminator and promoter primers (Table S2.36). Sanger sequencing was completed by the Michigan State University Research Technology Support Facility Genomics Core (East Lansing, MI, USA).

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

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

<u>Enzyme assays</u>

Enzyme assays were conducted in 100 mM sodium phosphate buffer at pH 6. For forward assays, acetyl-CoA (MilliporeSigma, Burlington, MA, USA) was added to a final concentration of 0.1 mM, and for reverse assays free CoA (MilliporeSigma, Burlington, MA, USA) was added to a final concentration of 1 mM. Purified AI3:16 and AI4:18 substrates were dried down using a vacuum centrifuge and redissolved in ethanol:water:formic acid (1:1:0.001). One microliter of the prepared acylsugars were used as acyl acceptors. Six microliters of enzyme

were added to a final volume of 60 µL. For negative controls, 6 µL of enzyme that was heat inactivated at 95°C for 10 minutes was substituted in place of untreated enzyme. Assays were incubated at 30°C for 30 minutes after which 120 µL of acetonitrile:isopropanol:formic acid (1:1:0.001) with 1.5 µM telmisartan (MilliporeSigma, Burlington, MA, USA) stop solution was added. Reactions were then spun at 17,000 x *g* for 10 minutes to remove precipitate. Supernatant was placed in autosampler vials and analyzed by LC-MS.

RT-PCR of *S. aethiopicum* BAHDs

We employed semi-quantitative RT-PCR to test *S. aethiopicum* BAHD expression in glandular trichomes with cDNA and genomic DNA (gDNA). Total RNA was isolated with the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany), including an on-column DNase digestion (Qiagen, Hilden, Germany), from shaved hypocotyls from accession *S. aethiopicum* PI 666075 and glandular trichomes from accessions PI 666075 and Grif 14165 isolated as described above for *S. melongena* glandular trichomes. RNA was quantified with a Nanodrop 2000c instrument (Thermo Fisher Scientific, Waltham, MA, USA). We synthesized cDNA using 10 ng of RNA and SuperScript III Reverse Transcriptase (Invitrogen, Waltham, MA, USA). gDNA was isolated with the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) from young leaf tissue collected from mature *S. aethiopicum* PI 666075 plants. PCR reactions (25 µL) were set up with GoTaq Green Master Mix (Promega, Madison, WI, USA), 200 nM of forward and reverse primers (Table S2.36), and 1 µL of cDNA or gDNA. PCR was performed under these conditions: 2 min at 95°C followed by 22, 30, and 35 cycles of 30s at 95°C, 30s at 58°C, and 1 min at 72°C.

We identified the putative *S. aethiopicum* orthologs of the *S. melongena* candidate BAHDs by querying *S. aethiopicum* annotated transcripts (Song et al., 2019) with *S. melongena* candidate ASAT DNA sequences using BLASTn. SaASAT3 and SaASAT3-L1 (GenBank

accession: OQ547782) were not annotated in the *S. aethiopicum* genome and were identified by querying *S. aethiopicum* scaffolds with BLASTn and then identifying open reading frames with Geneious software v9.1.8 (Dotmatics, Boston, MA, USA). The *S. aethiopicum* ef1α gene was identified by querying a putative *S. melongena* ef1α, SMEL4.1_06g005890, against the *S. aethiopicum* scaffolds.

S. melongena acyl chain composition analysis

S. melongena acyl chain composition was assessed with a previously developed esterification and GC-MS methods (Fan et al., 2020; Ning et al., 2015; Schenck et al., 2022). Acyl chain identifies were determined through authentic reference standards for nC8, nC10, nC12, and nC14 fatty acid ethyl esters (MilliporeSigma). The QuanLynx function of MassLynx v4.1 (Waters Corporation) integrated peaks from extracted ion chromatograms. Extracted ion chromatograms for the medium acyl chains iC8, nC8, iC10, nC10, iC12, nC12, iC14, and nC14 were generated for the *m/z* of 88 with a mass window of *m/z* of 0.50. Extracted ion chromatograms for the short chains iC4, aiC5, and iC5 were generated for with the *m/z* values of 71, 102, and 101, respectively, with a mass window of *m/z* of 0.50.

Accession numbers

Sequence data in this article can be found in the GenBank/EMBL data libraries under accession number SaASAT3-L1; OQ547782.

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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. https://doi.org/10.1016/j.foodres.2015.11.021

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. 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. https://doi.org/10.1021/jf960322f

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. https://doi.org/10.1021/acs.jnatprod.7b01069

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

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

Denoeud, F., Carretero-Paulet, L., Dereeper, A., Droc, G., Guyot, R., Pietrella, M., Zheng, C., Alberti, A., Anthony, F., Aprea, G., Aury, J.-M., Bento, P., Bernard, M., Bocs, S., Campa, C., Cenci, A., Combes, M.-C., Crouzillat, D., Da Silva, C., Daddiego, L., De Bellis, F., Dussert, S., Garsmeur, O., Gayraud, T., Guignon, V., Jahn, K., Jamilloux, V., Joët, T., Labadie, K., Lan, T., Leclercq, J., Lepelley, M., Leroy, T., Li, L.-T., Librado, P., Lopez, L., Muñoz, A., Noel, B., Pallavicini, A., Perrotta, G., Poncet, V., Pot, D., Priyono, null, Rigoreau, M., Rouard, M., Rozas, J., Tranchant-Dubreuil, C., VanBuren, R., Zhang, Q., Andrade, A.C., Argout, X., Bertrand, B., de Kochko, A., Graziosi, G., Henry, R.J., Jayarama, null, Ming, R., Nagai, C., Rounsley, S., Sankoff, D., Giuliano, G., Albert, V.A., Wincker, P., Lashermes, P., 2014. The coffee genome provides insight into the

convergent evolution of caffeine biosynthesis. Science 345, 1181–1184. https://doi.org/10.1126/science.1255274

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. https://doi.org/10.1038/ng.806

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. https://doi.org/10.1093/bioinformatics/bts635

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

Dutartre, L., Hilliou, F., Feyereisen, R., 2012. Phylogenomics of the benzoxazinoid biosynthetic pathway of Poaceae: gene duplications and origin of the Bx cluster. BMC Evol. Biol. 12, 64. https://doi.org/10.1186/1471-2148-12-64

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. https://doi.org/10.1016/j.pbi.2019.03.005

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. https://doi.org/10.1038/s41467-017-02045-7

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

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, 1–26. https://doi.org/10.7554/eLife.56717

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. https://doi.org/10.1039/D2NP00003B

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. https://doi.org/10.1104/pp.77.3.567

Gagnon, E., Hilgenhof, R., Orejuela, A., McDonnell, A., Sablok, G., Aubriot, X., Giacomin, L., Gouvêa, Y., Bragionis, T., Stehmann, J.R., Bohs, L., Dodsworth, S., Martine, C., Poczai,

P., Knapp, S., Särkinen, T., 2022. Phylogenomic discordance suggests polytomies along the backbone of the large genus Solanum. Am. J. Bot. 109, 580–601. https://doi.org/10.1002/ajb2.1827

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

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

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

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

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

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

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

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

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

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

135

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

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

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

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

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

Kruse, L.H., Bennett, A.A., Mahood, E.H., Lazarus, E., Park, S.J., Schroeder, F., Moghe, G.D., 2022. Illuminating the lineage-specific diversification of resin glycoside acylsugars in the morning glory (Convolvulaceae) family using computational metabolomics. Hortic. Res. 9, uhab079. https://doi.org/10.1093/hr/uhab079

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Moore, B.M., Wang, P., Fan, P., Leong, B., Schenck, C.A., Lloyd, J.P., Lehti-Shiu, M.D., Last, R.L., Pichersky, E., Shiu, S.-H., 2019. Robust predictions of specialized metabolism genes through machine learning. Proc. Natl. Acad. Sci. U.S.A. 116, 2344–2353. https://doi.org/10.1073/pnas.1817074116

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

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

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

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

Ning, J., Moghe, G.D., Leong, B., Kim, J., Ofner, I., Wang, Z., Adams, C., Jones, A.D., Zamir, D., Last, R.L., 2015. A feedback-insensitive isopropylmalate synthase affects acylsugar

composition in cultivated and wild tomato. Plant Physiol. 169, 1821–1835. https://doi.org/10.1104/pp.15.00474

Panchy, N., Lehti-Shiu, M., Shiu, S.-H., 2016. Evolution of Gene Duplication in Plants. Plant Physiol. 171, 2294–2316. https://doi.org/10.1104/pp.16.00523

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

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

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

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

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

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

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

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

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

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

Schilmiller, A.L., Moghe, G.D., Fan, P., Ghosh, B., Ning, J., Jones, A.D., Last, R.L., 2015. Functionally divergent alleles and duplicated Loci encoding an acyltransferase contribute

to acylsugar metabolite diversity in Solanum trichomes. Plant Cell 27, 1002–17. https://doi.org/10.1105/tpc.15.00087

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

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

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

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

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

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

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

Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K.V., Altshuler, D., Gabriel, S., DePristo, M.A., 2013. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Curr. Protoc. Bioinforma. 43, 11.10.1-11.10.33. https://doi.org/10.1002/0471250953.bi1110s43

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

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

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

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

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

Zizovic, I., Stamenić, M., Orlović, A., Skala, D., 2007. Supercritical carbon dioxide extraction of essential oils from plants with secretory ducts: mathematical modelling on the micro-scale. J. Supercrit. Fluids 39, 338–346. https://doi.org/10.1016/j.supflu.2006.03.009
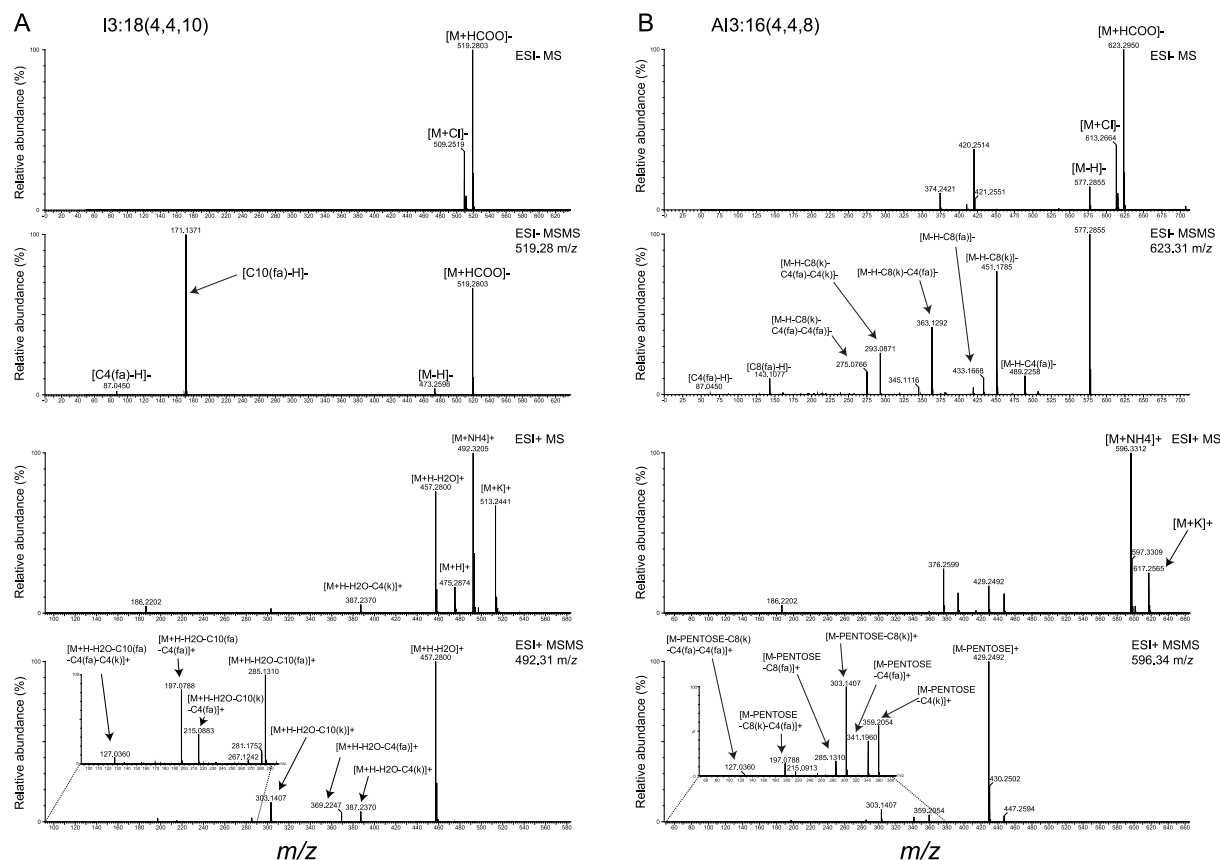
Figure S2.1. Annotation of I3:18(4,4,10) and AI3:16(4,4,8) using negative and positive mode MS and CID MS/MS fragmentation. (A) I3:18(4,4,10) negative and positive mode MS and MS/MS fragmentation. (B) AI3:16(4,4,8) negative and positive MS and MS/MS fragmentation. For each compound, ESI- MS (top panels) display the formate adduct accurate mass which was used to determine chemical formulas. ESI- MS/MS (second from the top panels) exhibit acyl chain carboxylate fragment ions for both compounds and stepwise loss of acyl chains for only AI3:16. ESI+ MS/MS exhibit stepwise losses of acyl chains for both compounds. Positive mode CID of AI3:16 also produces a fragment ion (*m/z* 429.2492) corresponding to the neutral loss of the pentose moiety. This indicates that all acyl chains reside on the hexose ring. The acylsugar LC-MS annotation method is described in the Materials and Methods. fa = fatty acid; k = ketene.
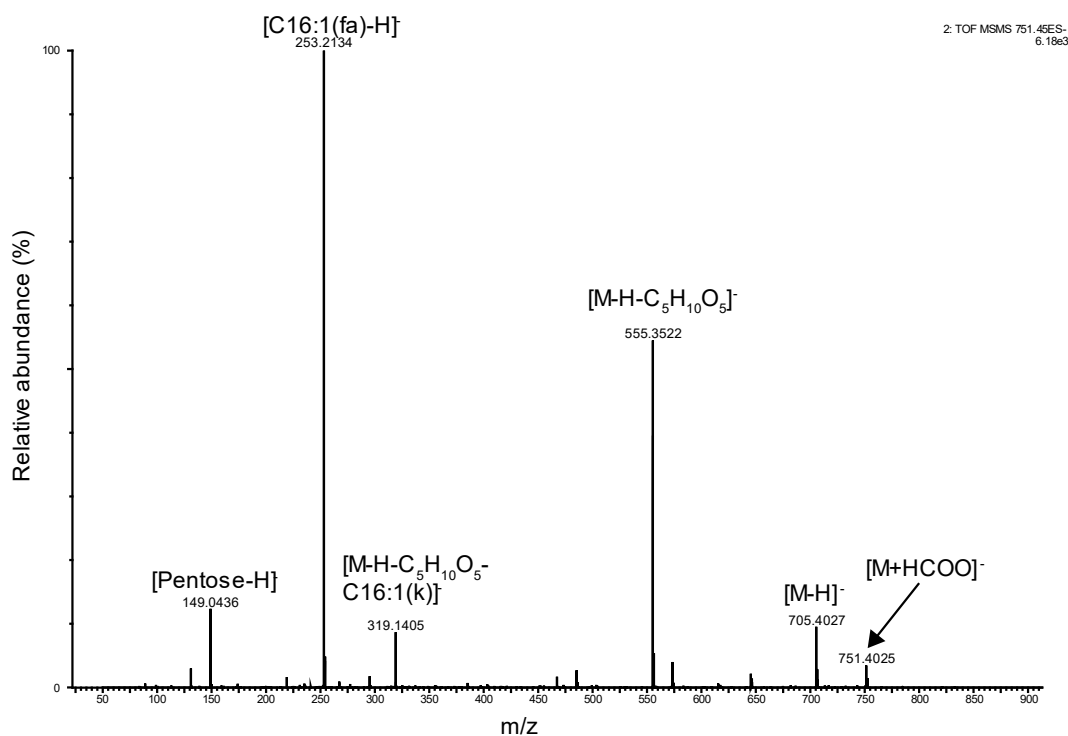
Figure S2.2. Negative mode MS/MS fragmentation of H3:22(4,4,16-O-p) from *S. sisymbriifolium* results in neutral loss of a pentose moiety. This neutral loss of a sugar group ($C_5H_{10}O_5$ plus formic acid) from the [M+formate]$^-$ ion *m/z* 751.4025 to form *m/z* 555.3522 is not observed with negative CID of other acyldisaccharides (Figures S2.1 and S2.8) suggesting an unusual glycosidic linkage on the acyl chain. fa = fatty acid; k = ketene; 16-O-p = pentosylated 16 carbon acyl chain.
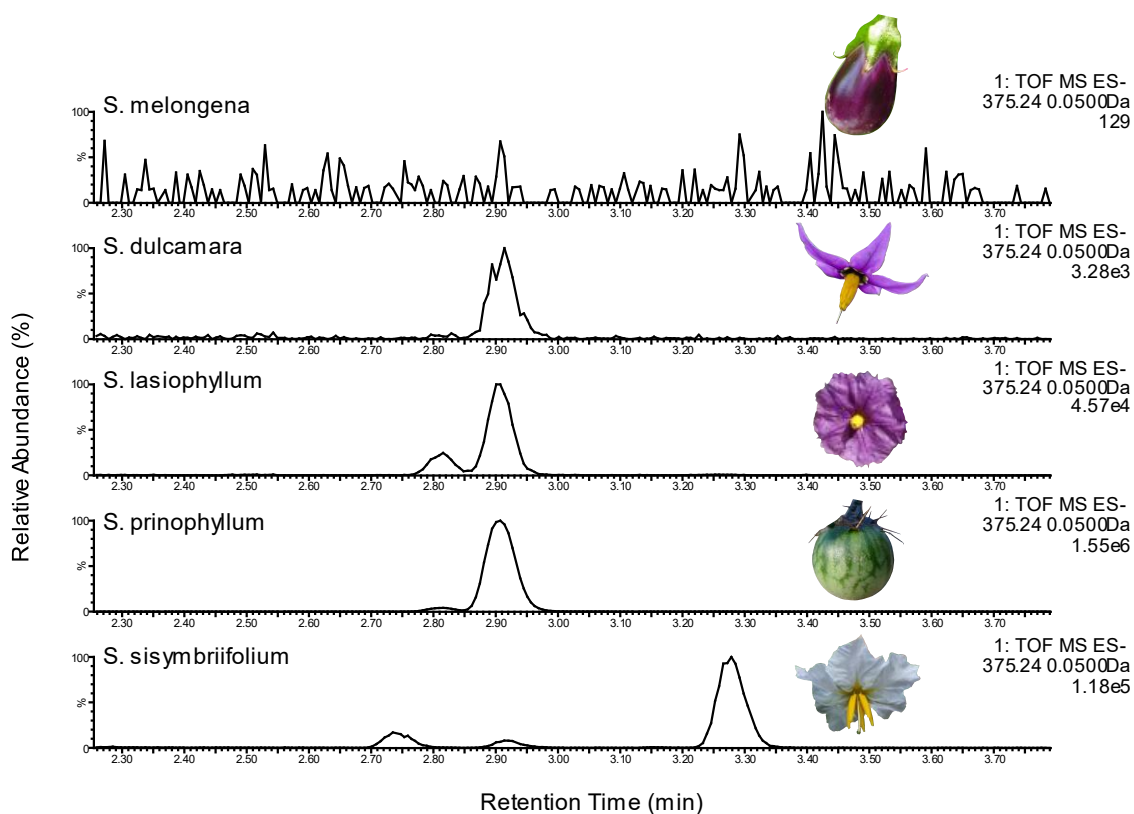
Figure S2.3. Acylsugar saponification produces free glycosylated fatty acids in four *Solanum* species supporting the identification of glycohydroxyacyl hexoses. All traces are extracted ion LC/MS chromatograms for pentosylated hydroxymyristic acid, *m/z* 375.24±0.05 (retention time 2.90 min) from saponified leaf surface extracts. *S. melongena* does not contain detectable glycosylated fatty acids and acts as a negative control. Vertical scale is normalized to the largest signal within the displayed region (values in the upper right of each chromatogram).
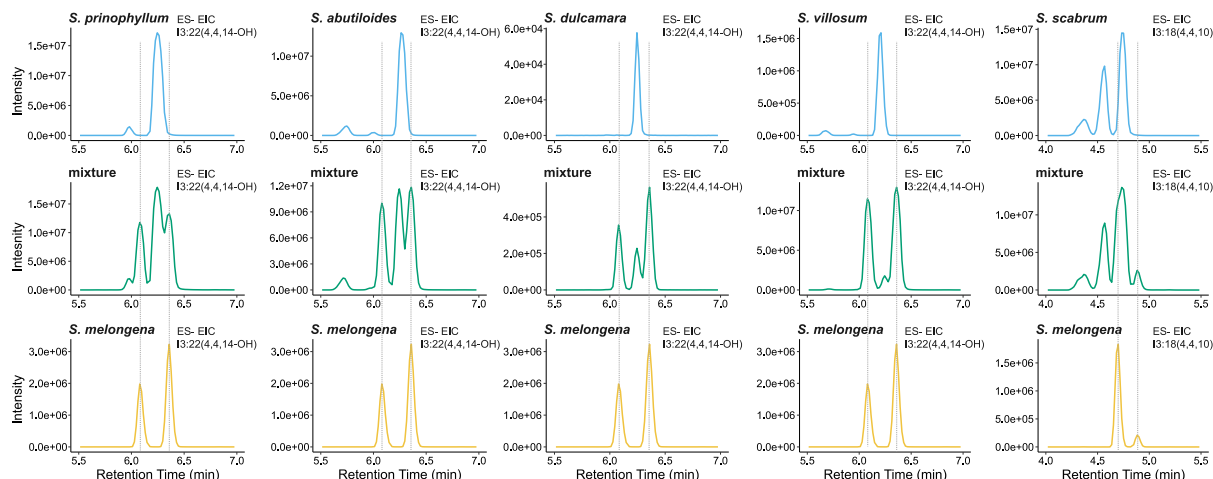
Figure S2.4. Coelution analysis reveals that five *Solanum* species accumulate acylsugar positional or branching isomers that differ from *S. melongena* acylsugars. Extracted ion chromatograms were generated for combined signals of *m/z* 591.34 and 519.28 corresponding to I3:22(4,4,14-OH) and I3:18(4,4,10) [M+formate]⁻ adducts. The yellow traces (bottom row) represent *S. melongena*, the blue traces represent the five other *Solanum* species (*S. prinophylum, S. abutiloides, S. dulcamara, S. villosum*, and *S. scabrum*), and the middle row green traces represent acylsugar mixtures between *S. melongena* and the species in the trace below. The mixed samples allow for corrections of retention time drift that may occur between samples.
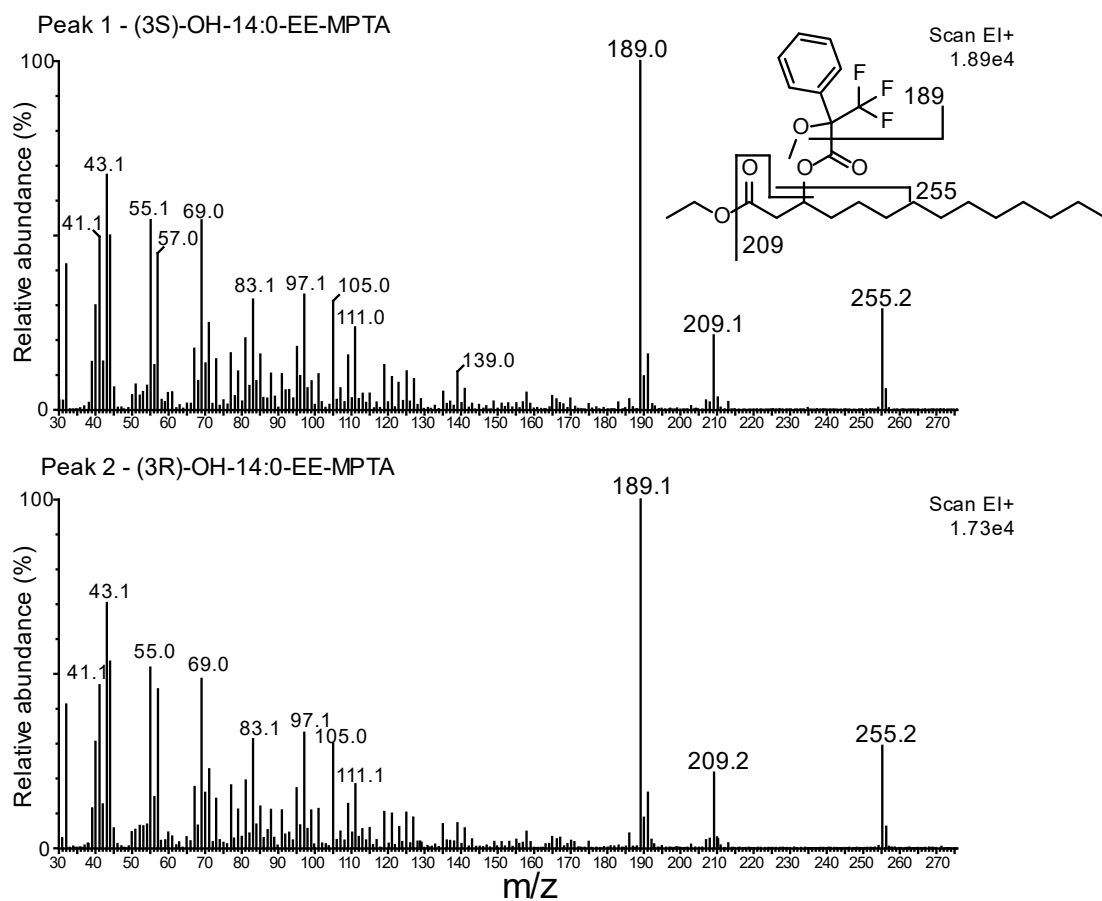
Figure S2.5. GC-MS full scan mass spectra of 3-OH-14:0 ethyl ester stereoisomers as their MPTA derivatives. The mass spectra contain ions at *m/z* 189, 209, and 255 corresponding to the expected fragmentation of the derivatized fatty acid. EE = ethyl ester. Peak numbers correspond to the peaks displayed in Figure 2.3.

	AI3:16(4,4,8)	AI4:18(2,4,4,8)
Solanum rigidum		
Solanum incanum	+	+
Solanum cerasiferum		
Solanum campylacanthum		
Solanum umtuma		
Solanum lichtesteinii	+	+
Solanum linnaeanum	+	+
Solanum aureitomentosum		
Solanum melongena	+	+
Solanum insanum		
Solanum lanzae		
Solanum usambarense		
Solanum agnewiorum		
Solanum trilobatum		
Solanum supinum		
Solanum polhilii		
Solanum anguivi	+	+
Solanum aethiopicum	+	-
Solanum violaceum		
Solanum macrocarpon		
Solanum dasyphyllum		
Solanum glabratum		
Solanum richardii	+	+

Figure S2.6. Phylogenetic analysis of Eggplant Clade and Anguivi Grade species reveals *S. aethiopicum* is the only analyzed acylsugar-producing species to not accumulate detectable AI4:18(2,4,4,8). AI3:16(4,4,8) and AI4:18(2,4,4,8) presence and absence in surface extracts as detected by LC-MS are plotted on a phylogeny of the Eggplant Clade and Anguivi Grade. Species that either do not accumulate detectable acylsugars (*S. macrocarpon* and *S. virginianum)* or were not analyzed did not have AI3:16 and AI4:18 presence plotted. The phylogeny was modified from a previously published version (Aubriot et al., 2018). Full acylsugar profiles are detailed Tables 2.1, S2.2-7.
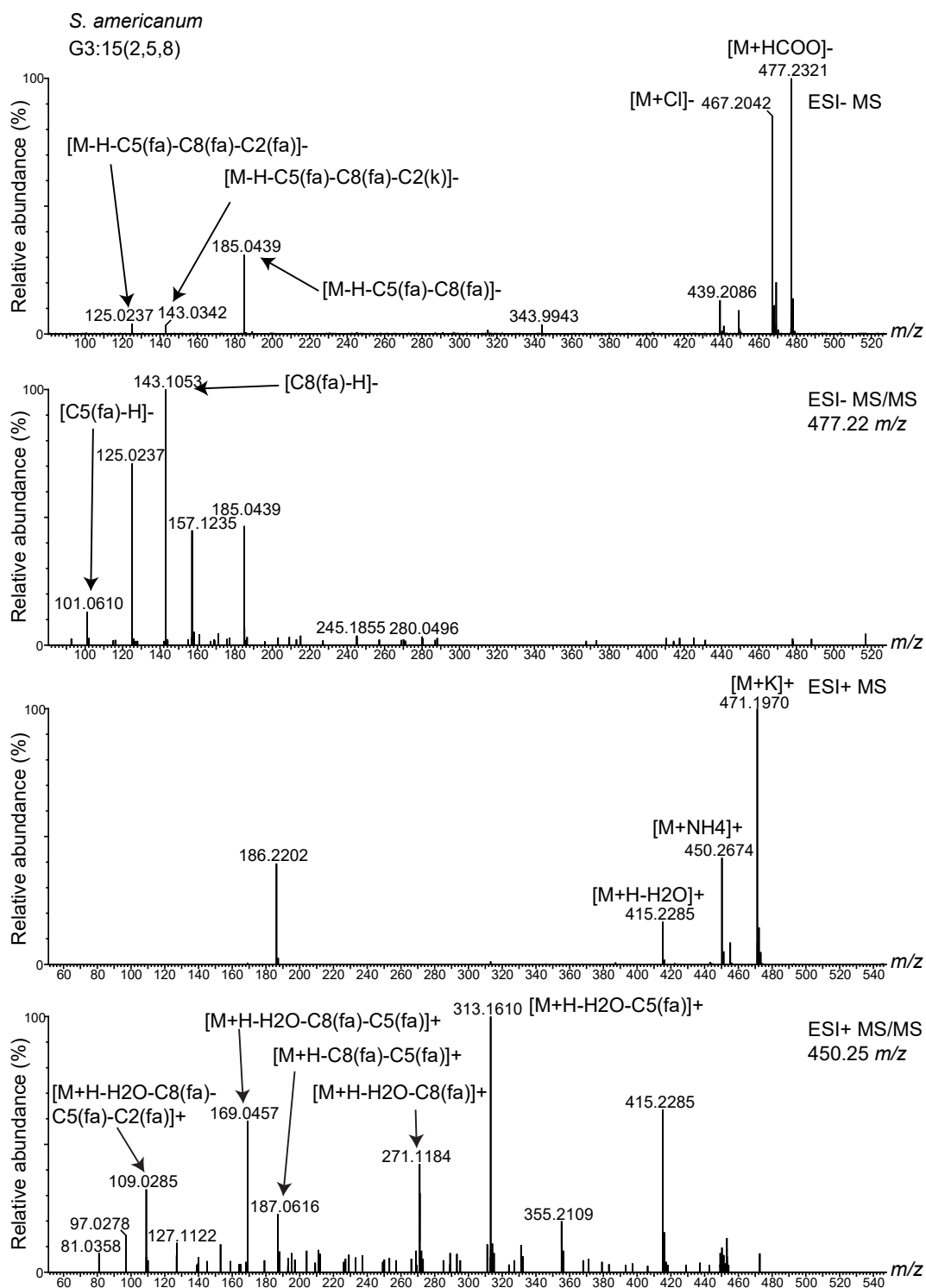
Figure S2.7. Negative and positive mode MS and CID MS/MS fragmentation of G3:15(2,5,8) from *S. americanum*. Acylglucoses fragment characteristically in negative mode MS and MS/MS functions producing fragment ions corresponding to stepwise acyl chain loss. This is in stark contrast to acylinositol negative mode fragmentation which does not produce

148

Figure S2.7. (cont'd)

major fragment ions corresponding to stepwise acyl chain loss (Figure S2.1). This fragmentation difference distinguishes the two sugar cores and enables their annotation by LC-MS. G3:15 formate and ammonium adduct ions were selected by data-dependent acquisition software and fragmented with a ramped collision energy detailed in the Materials and Methods. fa = fatty acid; k = ketene.
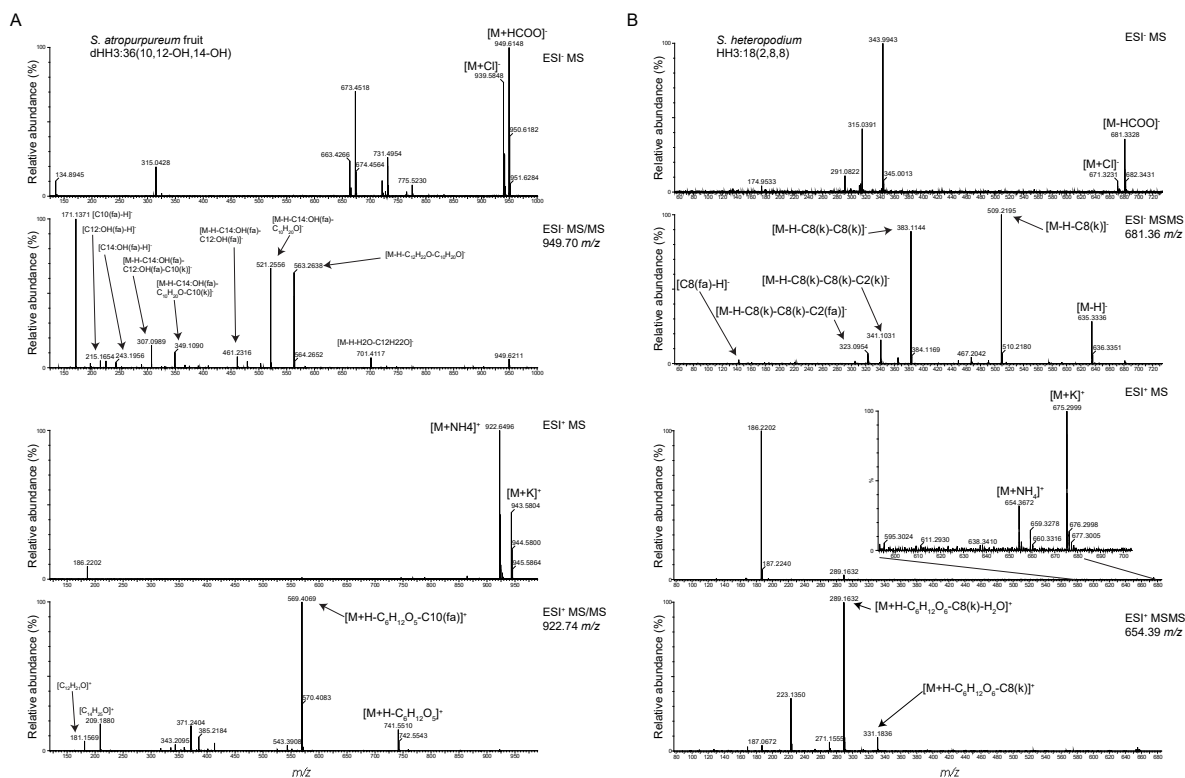
Figure S2.8. Negative and positive mode MS and CID MS/MS fragmentation of dHH3:36(10,12-OH,14-OH) from *S. atropurpureum* fruit and HH3:18(2,8,8) from *S. heteropodium*. Formate and ammonium adduct ions were selected by data-dependent acquisition software and fragmented with a ramped collision energy detailed in the Materials and Methods. Under negative mode CID MS/MS, each compound produces fragment ions corresponding to stepwise acyl chain loss down to a free sugar core fragment ion. In contrast, under positive mode CID MS/MS each compound produces fragment ions corresponding to neutral loss of a sugar moiety. (A) dHH3:36(10,12-OH,14-OH) MS and MS/MS fragmentation. Positive mode CID produces a fragment ion (*m/z* 569.4069) corresponding to a neutral loss of a deoxyhexose acylated with a ten carbon acyl chain. dHH = deoxyhexose-hexose (B) HH3:18(2,8,8) MS and MS/MS fragmentation. Positive mode CID produces a fragment ion (*m/z* 331.1836) corresponding to a neutral loss of a hexose acylated with an eight carbon acyl chain. HH = hexose-hexose; fa = fatty acid; k = ketene.
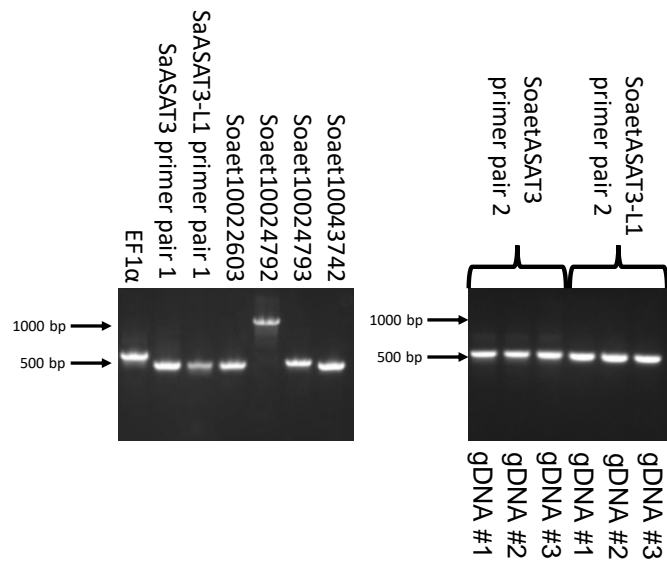
Figure S2.9. *S. aethiopicum* **RT-PCR primers validated with gDNA controls.** Soaet10024792 primers amplified a region containing an intron resulting in the larger amplicon length.
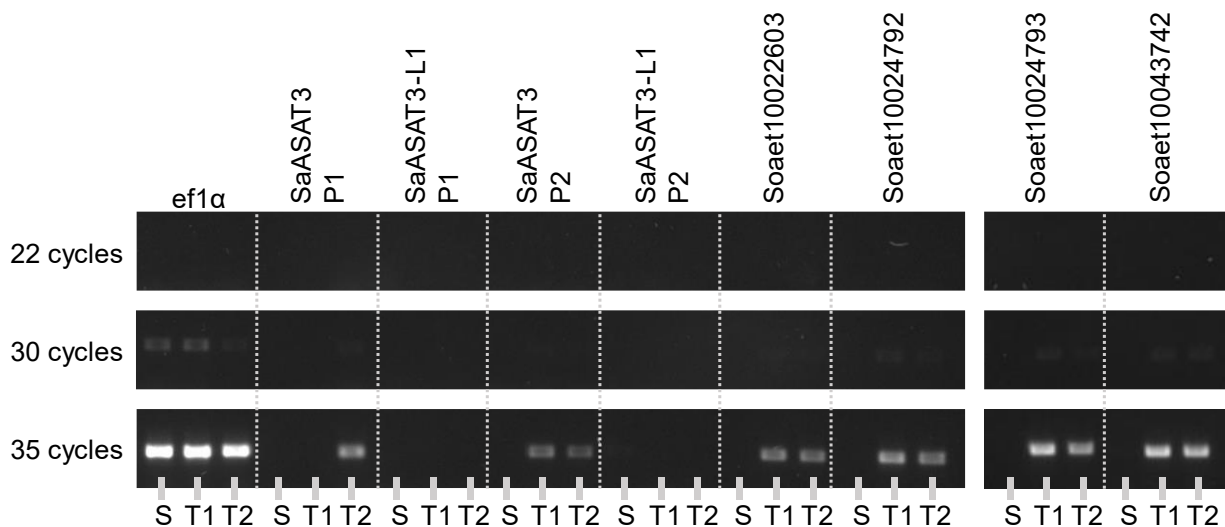
Figure S2.10. Semi-quantitative RT-PCR analysis of *S. aethiopicum* BAHD expression in glandular trichomes. *Elongation factor1α* (EF1α) was used as a positive control. P1 = primer pair 1; P2 = primer pair 2; S = shaved hypocotyl; T1 = PI 666075 glandular trichomes; T2 = Grif 14165 glandular trichomes.
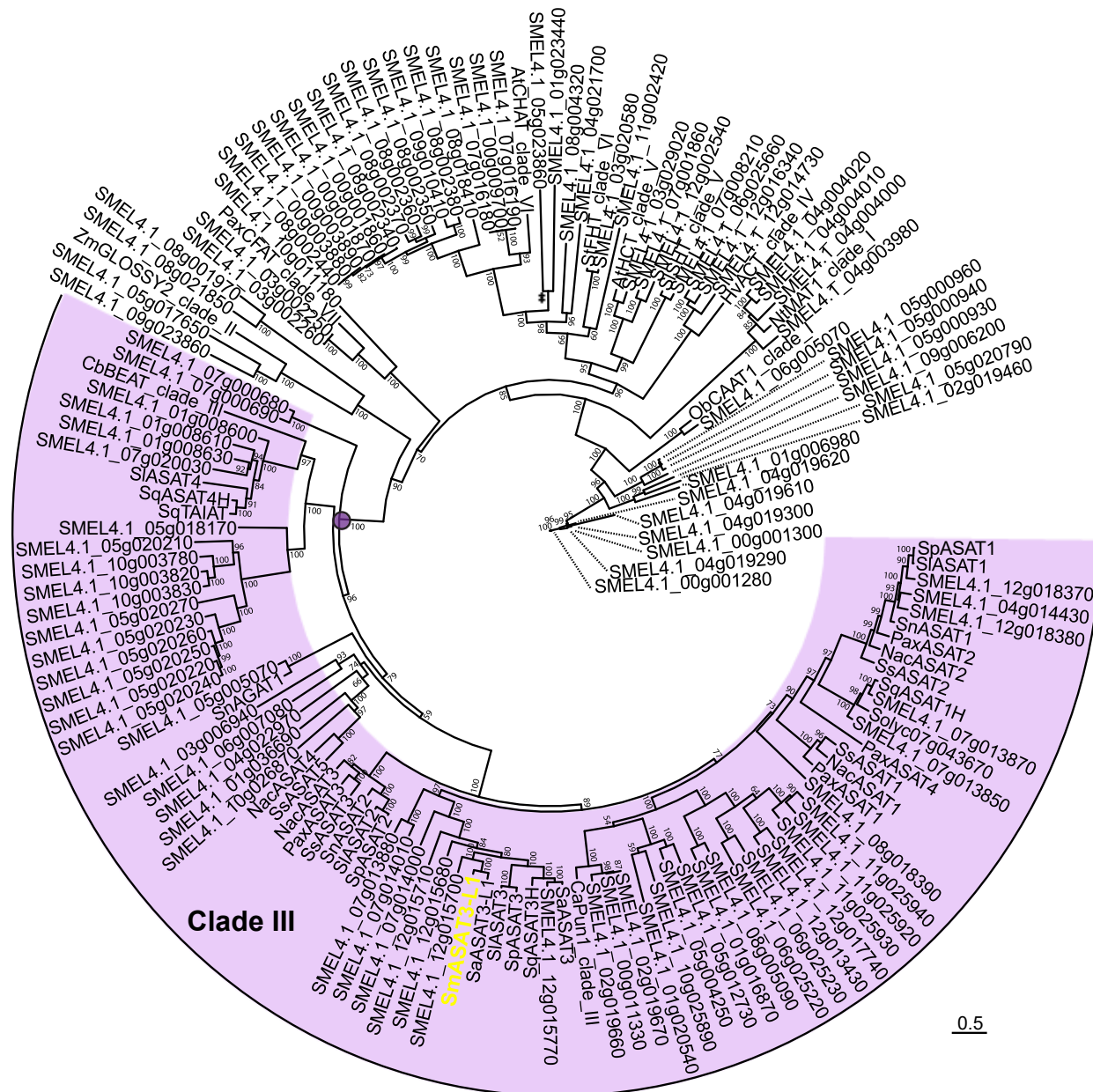
Figure S2.11. *S. melongena* **BAHD tree.** Shown is a phylogeny including 106 predicted BAHDs (PF002458) in the eggplant Smel_V4.1 reference genome, published reference BAHD sequences for clades I-VII, characterized ASAT sequences from other Solanaceae species, and the SaASAT3 and SaASAT3-L1 candidates from *S. aethiopicum*. The maximum likelihood tree was inferred from amino acid sequences using the Jones-Taylor-Thornton algorithm with seven rate categories in IQ-TREE v2.1.3. Bootstrap support calculated from 100,000 ultrafast bootstrap replicates and values >=95 indicate strong support. Clade III BAHDs are highlighted in purple, with a purple dot marking the node. Branch lengths represent substitution rates as denoted by scale bar.
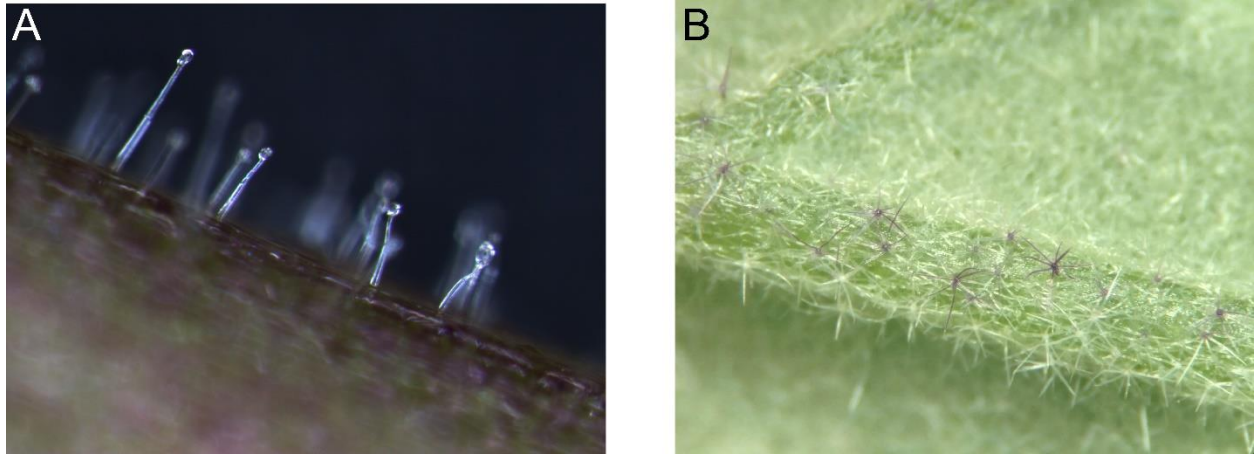
Figure S2.12. *S. melongena* **produces single stalked glandular trichomes on seedling above ground tissue but stellate nonglandular trichomes on non-seedling above ground tissue.** (A) Close up photo of a *S. melongena* hypocotyl displaying glandular trichomes with similar morphology to the acylsugar-producing *S. lycopersicum* Type I/IV trichomes (Luckwill, 1943; Schilmiller et al., 2012). (B) Close up photo of a *S. melongena* leaf from a reproductive stage plant displaying non-glandular stellate trichomes. Leaf surface metabolite extracts from this tissue do not have detectable acylsugars by LC-MS.
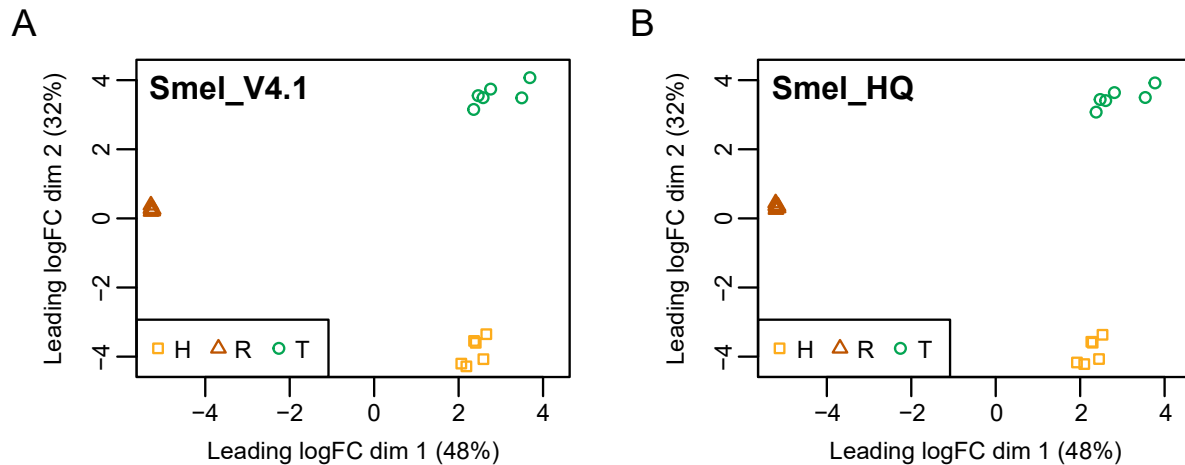
Figure S2.13. MDS plot of *S. melongena* RNAseq data. Genome-wide expression patterns across eggplant trichomes, trichomeless hypocotyls, and roots. Depicted is a multidimensional scaling (MDS) plot demonstrating that our 18 eggplant RNAseq samples cluster tightly by tissue identity. Distance between points illustrates expression differences between pairs of samples, calculated as leading (i.e., largest absolute) log2 fold-change (FC) in two dimensions, dim 1 and dim 2. Along the y-axis (dim 2), samples are separated into three tissue-specific groups. Along the x-axis (dim 1), samples are separated into two groups representing aerial (trichomes and trichomeless hypocotyls) and subterranean (roots) tissues, respectively. Colors and shapes represent the three tissues. T, trichomes (green circles); H, trichomeless hypocotyls (yellow squares); R, roots (brown triangles).
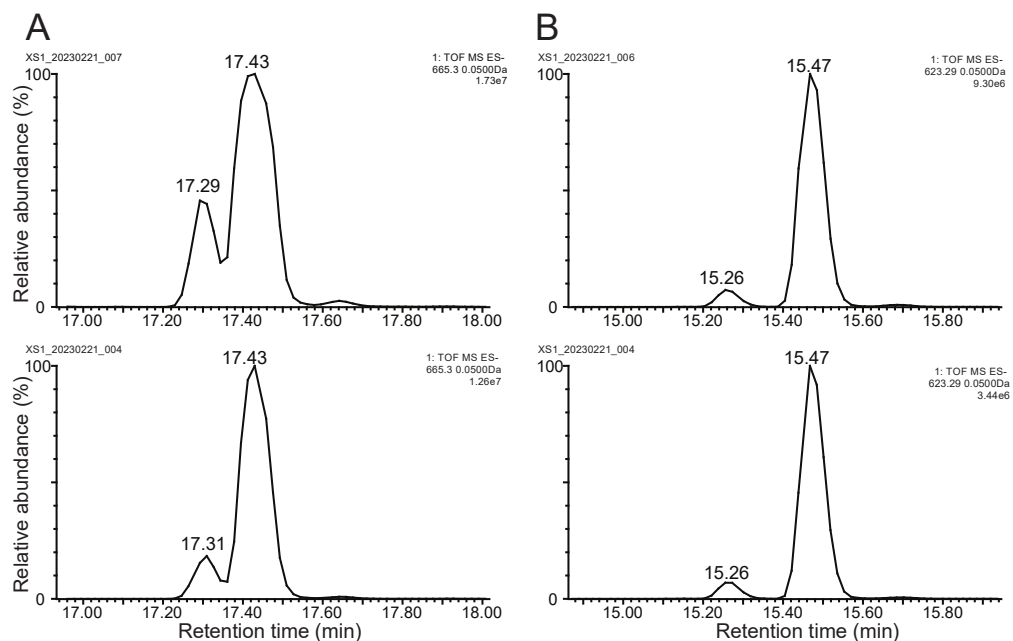
Figure S2.14. SmASAT3-L1 forward and reverse assay produced AI4:18 and AI3:16, respectively, coelute with plant produced AI4:18 and AI3:16. (A) Forward assay produced AI4:18 (top) coelutes with AI4:18 from a S. melongena 555598 leaf surface extract (bottom). The extracted ion chromatograms display the formate adduct of AI4:18, *m/z* 665.30. (B) Reverse assay produced AI3:16 (top) coelutes with AI3:16 from a S. melongena 555598 leaf surface extract (bottom). The extracted ion chromatograms display the formate adduct of AI3:16, *m/z* 623.20.

Table S2.1. Sugar cores and acyl chains annotated in *Solanum* Clade II, VANAns, and DulMo acylsugars. All acylsugars were annotated from leaf surface extracts except for the combined *S. acerifolium* and *S. atropurprueum* fruit data. Table includes S. nigrum and *S. quitoense* acylsugar traits from published reports (Hurney, 2018; Leong et al., 2020; Lou et al., 2021). Acylsugar profiles are described further in Tables S2.1-26.

Major Clade	Minor Clade	Species	Sugar core Mono-	Sugar core Di-	Acyl chain Non-modified	Acyl chain hydroxylated	Acyl chain unsaturated
Clade II	EHS	*Solanum melongena*	I	PH	2,4,5,8,10,12,14	12,14	
		Solanum incanum	I	PH	2,4,5,8,10,12,14	12,14	
		Solanum aethiopicum	I	PH	4,5,8,10,12,14	12,14	
		Solanum anguivi	I	PH	2,4,5,8,10,12,14	12,14	
		Solanum lichtensteinii	I	PH	2,4,5,8,10,12,14	12,14	
		Solanum linnaeanum	I	PH	2,4,5,8,10,12,14	12,14	
		Solanum richardii	I	PH, HH	2,4,5,8,10,12,14	12,14	
		Solanum heteropodium	I	HH	2,4,5,8,10,12,14	12,14	
		Solanum sejuntum	I		2,4,5,8,9,10,12	12,14	
		Solanum melanospermum	I		4,5,8,9,10,11,12	14	
		Solanum dioicum	I		4,5,8,9,10,11,12	14	
		Solanum prinophyllum	I	PH	4,5,10,12,14	12,14	
		Solanum lasiophyllum	I	PH	2,4,8,10,12,14	12,14	
	Torva	*Solanum torvum*	I		2,4,5,10,11,12, 13,14,16,17,18		5:1,18:1,18:2,18:3
	Sisymbriifolium	*Solanum sisymbriifolium*	I		4,5	14,15,16	
	Lasiocarpa	*Solanum quitoense*	I	PH, HH, GlNac-I	2,10,12		
	Acanthophora	*Solanum capsioides*	I		6,8,9,10,12	12,14	
		Solanum acerifolium/ atropurpureum	I	PH	2,5,6,8,10	12,14,15,16	
		Solanum acerifolium/ atropurpureum fruit	I	dHH	8,9,10,11	10,12,13,14	
		Solanum mammosum	I		4,5,6,8	6,12,14,16	
	Brevantherum	*Solanum abutiloides*	I		2,3,4,5,12,14,16	12,14	
VANAns	Archaeasolanum	*Solanum laciniatum*	I		6,8		

157

Table S2.1. (cont'd)

DulMo	Dulcamaroid	*Solanum duclamara*	I		4	14	
	Morelloid	*Solanum nigrum*	I, G		2,4,5,8,9,10		
		Solanum americanum	G	PH	2,4,5,6,8,9,10		
		Solanum scabrum	I, G		4,5,8,9,10	10	
		Solanum villosum	I, G	HH	2,4,5,10	12,14,16	

Table S2.2. Summary of annotated acylsugars detected in *S. incanum* trichome extracts. PH = pentose-hexose; AI = arabinose-inositol; I = inositol. RT = retention time; Annotation method and confidence level criteria are described within the Methods. *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over two samples. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acyldisaccharides							
PH3:14(2,4,8)	medium	2.23	C25H42O13	595.2607	595.2621	2.4	0.207
PH4:16(2,2,4,8)	medium	2.54	C27H44O14	637.2713	637.2720	1.1	1.44
PH3:16(4,4,8)	medium	2.61	C27H46O13	623.2920	623.2919	-0.2	15.3
AI3:16(4,4,8)	high	2.66	C27H46O13				
PH3:17(4,5,8)	medium	2.96	C28H48O13	637.3077	637.3083	1.0	0.159
AI4:18(2,4,4,8)	high	3.33	C29H48O14	665.3026	665.3026	0.0	63.1
PH3:18(4,4,10)	medium	3.52	C29H50O13	651.3233	651.3239	0.9	0.00728
PH3:18(4,4,10)	medium	3.61	C29H50O13	651.3233	651.3231	-0.3	0.151
PH4:19(2,4,5,8)	medium	3.76	C30H50O14	679.3183	679.3186	0.5	0.284
PH4:19(2,4,5,8)	medium	3.83	C30H50O14				
PH3:20(4,4,12)	medium	5.05	C31H54O13	679.3546	679.3554	1.2	0.000206
PH3:20(4,4,12)	medium	5.23	C31H54O13	679.3546	679.3552	0.9	0.0318
Acylhexoses							
I3:20(4,4,12-OH)	medium	4.16	C26H46O10	563.3073	563.3071	-0.4	0.0301
I3:20(4,4,12-OH)	medium	4.32	C26H46O10	563.3073	563.3068	-0.9	0.0450
I3:20(2,4,14-OH)	medium	4.68	C26H46O10	563.3073	563.3082	1.6	0.208
I3:18(4,4,10)	high	4.70	C24H42O9	519.2811	519.2802	-1.7	6.21
I3:18(4,4,10)	medium	4.88	C24H42O9	519.2811	519.2794	-3.3	0.164
I3:20(2,4,14-OH)	medium	4.90	C26H46O10	563.3073	563.3077	0.7	0.131
I3:22(4,4,14-OH)	high	6.06	C28H50O10	591.3386	591.3385	-0.2	5.83
I3:22(4,4,14-OH)	high	6.34	C28H50O10	591.3386	591.3380	-1.0	2.74
I3:20(4,4,12)	high	6.71	C26H46O9	547.3124	547.3119	-1.0	2.85
I3:22(4,4,14)	high	9.34	C28H50O9	575.3437	575.3435	-0.3	0.900
I3:22(4,4,14)	high	9.68	C28H50O9	575.3437	575.3433	-0.7	0.174

Table S2.3. Summary of annotated acylsugars detected in *S. anguivi* trichome extracts. PH = pentose-hexose; AI = arabinose-inositol; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over six samples. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	m/z acc	m/z exp	Δm(ppm)	Average acylsugar abundance (%)
Acyldisaccharides							
PH4:16(2,2,4,8)	medium	2.54	C27H44O14	637.2713	637.2720	1.1	2.73
PH3:16(4,4,8)	medium	2.58	C27H46O13	623.2920	623.2919	0.97	16.8
AI3:16(4,4,8)	high	2.63	C27H46O13				
PH3:17(4,5,8)	medium	2.93	C28H48O13	637.3077	637.3083	1.0	0.0921
AI4:18(2,4,4,8)	high	3.33	C29H48O14	665.3026	665.3026	0.0	59.8
PH4:19(2,4,5,8)	medium	3.70	C30H50O14	679.3183	679.3186	1.1	0.147
PH4:19(2,4,5,8)	medium	3.77	C30H50O14				
PH3:20(4,4,12)	medium	4.98	C31H54O13	679.3546	679.3554	1.2	0.00021
PH3:20(4,4,12)	medium	5.15	C31H54O13	679.3546	679.3552	0.9	0.0262
Acylhexoses							
I3:20(4,4,12-OH)	medium	4.14	C26H46O10	563.3073	563.3071	-0.4	0.023
I3:20(4,4,12-OH)	medium	4.32	C26H46O10	563.3073	563.3068	-0.9	0.0461
I3:18(4,4,10)	high	4.68	C24H42O9	519.2811	519.2802	-1.7	6.35
I3:20(2,4,14-OH)	medium	4.68	C26H46O10	563.3073	563.3082	1.6	0.393
I3:18(4,4,10)	medium	4.88	C24H42O9	519.2811	519.2794	-3.3	0.0955
I3:20(2,4,14-OH)	medium	4.90	C26H46O10	563.3073	563.3077	0.7	0.132
I3:22(4,4,14-OH)	high	5.98	C28H50O10	591.3386	591.3385	-0.2	5.78
I3:22(4,4,14-OH)	high	6.24	C28H50O10	591.3386	591.3380	-1.0	4.37
I3:20(4,4,12)	high	6.71	C26H46O9	547.3124	547.3119	-1.0	2.53
I3:22(4,4,14)	high	9.25	C28H50O9	575.3437	575.3435	-0.3	0.594
I3:22(4,4,14)	high	9.56	C28H50O9	575.3437	575.3433	-0.7	0.134

Table S2.4. Summary of annotated acylsugars detected in *S. aethiopicum* trichome extracts. Annotation method and confidence level criteria are described within the Methods. PH = pentose-hexose; AI = arabinose-inositol; I = inositol. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over 25 samples. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acyldisaccharides							
PH3:16(4,4,8)	medium	2.61	C27H46O13	623.2920	623.2919	1.1	47.8
AI3:16(4,4,8)	high	2.66	C27H46O13				
PH3:18(4,4,10)	medium	3.49	C29H50O13	651.3233	651.3239	0.9	3.47
PH3:18(4,4,10)	medium	3.59	C29H50O13	651.3233	651.3231	-0.3	3.35
PH3:20(4,4,12)	medium	4.98	C31H54O13	679.3546	679.3554	1.2	0.191
PH3:20(4,4,12)	medium	5.15	C31H54O13	679.3546	679.3552	0.9	0.179
Acylhexoses							
I3:16(4,4,8)	medium	3.28	C22H38O9	491.2498	491.2492	-1.3	1.29
I3:20(4,4,12-OH)	medium	4.09	C26H46O10	563.3073	563.3071	-0.4	0.590
I3:20(4,4,12-OH)	medium	4.27	C26H46O10	563.3073	563.3068	-0.9	1.44
I3:18(4,4,10)	high	4.70	C24H42O9	519.2811	519.2802	-1.7	15.3
I3:18(4,4,10)	medium	4.87	C24H42O9	519.2811	519.2794	-3.3	0.596
I3:22(4,4,14-OH)	high	6.01	C28H50O10	591.3386	591.3385	-0.2	15.8
I3:20(4,8,8)	medium	6.21	C26H46O9	547.3124	547.3122	-0.3	0.369
I3:22(4,4,14-OH)	high	6.26	C28H50O10	591.3386	591.338	-1.0	5.73
I3:20(4,4,12)	high	6.84	C26H46O9	547.3124	547.3119	-1.0	1.81
I3:20(4,4,12)	medium	7.11	C26H46O9	547.3124	547.3127	0.6	0.0694
I3:22(4,8,10)	medium	8.58	C28H50O9	575.3437	575.3439	0.3	0.261
I3:22(4,4,14)	high	9.41	C28H50O9	575.3437	575.3435	-0.3	1.71
I3:22(4,4,14)	high	9.75	C28H50O9	575.3437	575.3433	-0.7	0.139

Table S2.5. Summary of annotated acylsugars detected in *S. lichtensteinii* trichome extracts. PH = pentose-hexose; AI = arabinose-inositol; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over four samples. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acyldisaccharides							
PH3:14(2,4,8)	medium	2.23	C25H42O13	595.2607	595.2621	2.4	1.03
PH4:16(2,2,4,8)	medium	2.58	C27H44O14	637.2713	637.2720	1.1	4.64
AI3:16(4,4,8)	high	2.66	C27H46O13	623.2920	623.2919	-0.2	6.35
PH3:17(4,5,8)	medium	2.94	C28H48O13	637.3077	637.3083	1.0	0.0256
AI4:18(2,4,4,8)	high	3.29	C29H48O14	665.3026	665.3026	0.0	31.6
PH3:18(4,4,10)	medium	3.48	C29H50O13	651.3233	651.3239	0.9	2.13
PH3:18(4,4,10)	medium	3.59	C29H50O13	651.3233	651.3231	-0.3	4.47
PH4:19(2,4,5,8)	medium	3.76	C30H50O14	679.3183	679.3186	0.5	0.0866
PH3:20(4,4,12)	medium	5.21	C31H54O13	679.3546	679.3554	1.2	0.0609
PH3:20(4,4,12)	medium	5.45	C31H54O13	679.3546	679.3552	0.9	0.00228
Acylhexoses							
I3:20(4,4,12-OH)	medium	4.16	C26H46O10	563.3073	563.3071	-0.4	0.00268
I3:20(4,4,12-OH)	medium	4.32	C26H46O10	563.3073	563.3068	-0.9	0.0794
I3:20(2,4,14-OH)	medium	4.68	C26H46O10	563.3073	563.3082	1.6	0.459
I3:18(4,4,10)	high	4.70	C24H42O9	519.2811	519.2802	-1.7	30.4
I3:18(4,4,10)	medium	4.88	C24H42O9	519.2811	519.2794	-3.3	0.448
I3:20(2,4,14-OH)	medium	4.92	C26H46O10	563.3073	563.3077	0.7	0.527
I3:22(4,4,14-OH)	high	6.08	C28H50O10	591.3386	591.3385	-0.2	3.00
I4:22(4,4,4,10)	medium	6.15	C28H48O10	589.3229	589.3227	-0.3	4.37
I3:20(4,8,8)	medium	6.23	C26H46O9	547.3124	547.3122	-0.3	0.0428
I4:22(4,4,4,10)	medium	6.34	C28H48O10	589.3229	589.3237	1.3	0.0209
I3:22(4,4,14-OH)	high	6.36	C28H50O10	591.3386	591.3380	-1.0	8.85
I3:20(4,4,12)	high	6.84	C26H46O9	547.3124	547.3119	-1.0	0.567
I3:20(4,4,12)	medium	7.13	C26H46O9	547.3124	547.3127	0.6	0.125
I3:22(4,8,10)	medium	8.58	C28H50O9	575.3437	575.3439	0.3	0.0877
I3:22(4,4,14)	high	9.41	C28H50O9	575.3437	575.3435	-0.3	0.345
I3:22(4,4,14)	high	9.76	C28H50O9	575.3437	575.3433	-0.7	0.232

Table S2.6. Summary of annotated acylsugars detected in *S. linnaeanum* trichome extracts.
PH = pentose-hexose; AI = arabinose-inositol; I = inositol. Annotation method and confidence
level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical
monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) =
mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis
QI and was averaged over seven samples. Acylsugars are sorted by number of sugar moieties and
then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acyldisaccharides							
PH3:14(2,4,8)	medium	2.23	C25H42O13	595.2607	595.2621	2.4	2.28
PH4:16(2,2,4,8)	medium	2.58	C27H44O14	637.2713	637.2720	1.1	1.64
AI3:16(4,4,8)	high	2.66	C27H46O13	623.2920	623.2919	-0.2	26.8
PH3:17(4,5,8)	medium	2.94	C28H48O13	637.3077	637.3083	1.0	0.365
AI4:18(2,4,4,8)	high	3.29	C29H48O14	665.3026	665.3026	0.0	56.4
PH4:19(2,4,5,8)	medium	3.76	C30H50O14	679.3183	679.3186	0.5	0.385
Acylhexoses							
I3:20(4,4,12-OH)	medium	4.16	C26H46O10	563.3073	563.3071	-0.4	0.00207
I3:20(4,4,12-OH)	medium	4.32	C26H46O10	563.3073	563.3068	-0.9	0.0248
I3:20(2,4,14-OH)	medium	4.68	C26H46O10	563.3073	563.3082	1.6	0.0482
I3:18(4,4,10)	high	4.70	C24H42O9	519.2811	519.2802	-1.7	4.71
I3:18(4,4,10)	medium	4.88	C24H42O9	519.2811	519.2794	-3.3	0.235
I3:20(2,4,14-OH)	medium	4.92	C26H46O10	563.3073	563.3077	0.7	0.216
I3:22(4,4,14-OH)	high	6.10	C28H50O10	591.3386	591.3385	-0.2	0.763
I3:20(4,8,8)	medium	6.23	C26H46O9	547.3124	547.3122	-0.3	0.371
I3:22(4,4,14-OH)	high	6.36	C28H50O10	591.3386	591.3380	-1.0	4.37
I3:20(4,4,12)	high	6.84	C26H46O9	547.3124	547.3119	-1.0	1.13
I3:20(4,4,12)	medium	7.13	C26H46O9	547.3124	547.3127	0.6	0.0103
I3:22(4,8,10)	medium	8.58	C28H50O9	575.3437	575.3439	0.3	0.000535
I3:22(4,4,14)	high	9.41	C28H50O9	575.3437	575.3435	-0.3	0.108
I3:22(4,4,14)	high	9.76	C28H50O9	575.3437	575.3433	-0.7	0.130

Table S2.7. Summary of annotated acylsugars detected in *S. richardii* trichome extracts. PH = pentose-hexose; AI = arabinose-inositol; I = inositol. Annotation method and confidence level criteria are described within the Methods section. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over two samples. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm (ppm)	Average acylsugar abundance (%)
Acyldisaccharides							
HH3:16(4,4,8)	medium	2.39	C28H48O14	653.3026	653.3024	-0.3	0.284
PH3:21(4,4,13-O)	medium	2.48	C32H54O14	707.3496	707.3477	-2.7	0.0361
PH3:16(4,4,8)	medium	2.61	C27H46O13	623.2920	623.2916	-0.6	4.49
AI3:16(4,4,8)	high	2.66	C27H46O13				
HH4:18(2,4,4,8)	medium	2.87	C30H50O15	695.3132	695.3104	-4.0	0.407
PH3:17(4,5,8)	medium	2.89	C28H48O13	637.3077	637.3091	2.2	2.51
HH3:18(4,4,10)	medium	3.09	C30H52O14	681.3339	681.3330	-1.3	0.653
AI4:18(2,4,4,8)	high	3.24	C29H48O14	665.3026	665.3047	3.2	24.0
PH3:20(4,4,12-OH)	medium	3.46	C31H54O14	695.3496	695.3482	-2.0	0.449
HH3:19(4,5,10)	medium	3.47	C31H54O14				
PH3:18(4,4,10)	medium	3.49	C29H50O13	651.3233	651.3205	-4.3	15.9
PH3:18(4,4,10)	medium	3.61	C29H50O13	651.3233	651.3205	-4.3	0.482
PH4:19(2,4,5,8)	medium	3.70	C30H50O14	679.3183	679.3146	-5.4	3.75
PH3:19(4,5,10)	medium	4.02	C30H52O13	665.3390	665.3364	-3.9	4.98
PH3:19(4,5,10)	medium	4.10	C30H52O13	665.3390	665.3364	-3.9	0.156
PH4:20(4,4,4,8)	medium	4.18	C31H52O14	693.3339	693.3358	2.7	1.27
PH4:20(2,4,4,10)	medium	4.55	C31H52O14	693.3339	693.3358	2.7	5.63
PH3:20(4,8,8)	medium	4.57	C31H54O13	679.3582	679.3520	-9.1	0.842
PH4:21(4,4,5,8)	medium	4.78	C32H54O14	707.3496	707.3531	4.9	0.0192
PH4:21(4,4,5,8)	medium	4.85	C32H54O14	707.3496	707.3531	4.9	0.0465
PH3:22(4,4,14-OH)	medium	4.85	C33H58O14	723.3809	723.3819	1.4	0.458
PH3:20(4,4,12)	medium	5.05	C31H54O13	679.3546	679.3520	-3.8	0.317
PH3:22(4,4,14-OH)	medium	5.07	C33H58O14	723.3809	723.3819	1.4	0.369
PH4:21(2,4,5,10)	medium	5.20	C32H54O14	707.3496	707.3477	-2.7	0.544
PH3:20(4,4,12)	medium	5.27	C31H54O13	679.3546	679.3574	4.1	0.0756
PH3:21(5,8,8)	medium	5.28	C32H56O13	693.3703	693.3682	-3.0	0.199
PH3:21(4,5,12)	medium	5.84	C32H56O13	693.3703	693.3682	-3.0	0.0945
PH3:21(4,5,12)	medium	6.08	C32H56O13	693.3703	693.3735	4.6	0.0213
PH3:22(4,8,10)	medium	6.48	C33H58O13	707.3859	707.3858	-0.1	0.791
PH3:22(4,8,10)	medium	6.71	C33H58O13	707.3859	707.3858	-0.1	0.0243

Table S2.7. (cont'd)

PH3:22(4,4,14)	medium	7.31	C33H58O13	707.3859	707.3858	-0.1	0.280
PH3:22(4,4,14)	medium	7.62	C33H58O13	707.3859	707.3858	-0.1	0.0837
Acylhexoses							
I3:20(4,4,12-OH)	medium	4.16	C26H46O10	563.3073	563.3071	-0.4	0.186
I3:20(4,4,12-OH)	medium	4.32	C26H46O10	563.3073	563.3068	-0.9	1.06
I3:20(2,4,14-OH)	medium	4.68	C26H46O10	563.3073	563.3082	1.6	0.00899
I3:18(4,4,10)	high	4.70	C24H42O9	519.2811	519.2802	-1.7	12.0
I3:18(4,4,10)	medium	4.88	C24H42O9	519.2811	519.2794	-3.3	0.193
I3:20(2,4,14-OH)	medium	4.90	C26H46O10	563.3073	563.3077	0.7	0.00280
I3:19(4,5,10)	medium	5.30	C25H44O9	533.2967	533.2962	-0.9	2.71
I3:22(4,4,14-OH)	high	6.01	C28H50O10	591.3386	591.3385	-0.2	3.27
I3:20(4,8,8)	medium	6.23	C26H46O9	547.3124	547.3122	-0.3	0.820
I3:22(4,4,14-OH)	high	6.28	C28H50O10	591.3386	591.3380	-1.0	4.98
I3:20(4,4,12)	high	6.84	C26H46O9	547.3124	547.3119	-1.0	1.03
I3:23(4,5,14-OH)	medium	6.86	C29H52O10	605.3542	605.3545	0.4	0.279
I3:23(4,5,14-OH)	medium	6.98	C29H52O10	605.3542	605.3549	1.1	0.0493
I3:20(4,4,12)	medium	7.13	C26H46O9	547.3124	547.3127	0.6	0.0675
I3:23(4,5,14-OH)	medium	7.14	C29H52O10	605.3542	605.3543	0.2	0.437
I3:23(4,5,14-OH)	medium	7.27	C29H52O10	605.3542	605.3551	1.4	0.105
I3:22(4,8,10)	medium	8.58	C28H50O9	575.3437	575.3439	0.3	1.46
H3:22	low	8.86	C28H50O9	575.3437	575.3447	1.7	0.0232
I3:22(4,4,14)	high	9.41	C28H50O9	575.3437	575.3435	-0.3	1.29
I3:23(5,8,10)	medium	9.49	C29H52O9	589.3593	589.3587	-1.0	0.255
I3:22(4,4,14)	high	9.75	C28H50O9	575.3437	575.3433	-0.7	0.0984
H3:23	low	9.76	C29H52O9	589.3593	589.3594	0.2	0.00569
I3:26(4,8,14-OH)	medium	10.21	C32H58O10	647.4012	647.4017	0.7	0.101
I3:23(4,5,14)	medium	10.30	C29H52O9	589.3593	589.3596	0.5	0.0538
H3:23	low	10.46	C29H52O9	589.3593	589.3596	0.5	0.0173
I3:26(4,8,14-OH)	medium	10.54	C32H58O10	647.4012	647.4021	1.3	0.217
H3:24	low	11.04	C30H54O9	603.3750	603.3762	2.0	0.0465
H3:24	low	11.11	C30H54O9	603.3750	603.3745	-0.8	0.0394

Table S2.8. Summary of annotated acylsugars detected in *S. heteropodium* trichome extracts. HH = hexose-hexose; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Hexose	Hexose	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acyldisaccharides									
HH3:18(2,8,8)	2 8	8	medium	3.29	C30H52O14	681.3339	681.3350	1.6	4.40
HH3:18(2,8,8)	2 8	8	medium	3.35	C30H52O14				
Acylhexoses									
I2:15(5,10)			medium	3.04	C21H38O8	463.2549	463.2542	-1.6	3.83
I3:19(4,5,10)			medium	5.43	C25H44O9	533.2967	533.2961	-1.1	47.3
I3:22(5,5,12-OH)			medium	5.68	C28H50O10	591.3386	591.3421	5.9	0.0311
I3:20(5,5,10)			medium	6.34	C26H46O9	547.3124	547.3122	-0.3	32.4
I3:23(4,5,14-OH)			medium	7.06	C29H52O10	605.3542	605.3549	1.2	0.296
I3:21(4,5,12)			medium	7.24	C27H48O9	561.3280	561.3297	3.1	0.610
I3:21(4,5,12)			medium	7.77	C27H48O9	561.3280	561.3284	0.7	3.97
I3:24(5,5,14-OH)			medium	8.02	C30H54O10	619.3699	619.3728	4.8	0.0987
I3:22(5,5,12)			medium	8.80	C28H50O9	575.3437	575.3446	1.5	2.04
H3:23			low	9.74	C29H52O9	589.3593	589.3610	2.8	0.0636
I3:23(4,5,14)			medium	10.43	C29H52O9	589.3593	589.3615	3.8	3.09
I3:24(5,5,14)			medium	11.40	C30H54O9	603.3750	603.3764	2.3	2.02

Table S2.9. Summary of annotated acylsugars detected in *S. sejunctum* trichome extracts.
PH = pentose-hexose; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acyldisaccharides							
HH4:22(2,4,4,12-OH)	medium	3.61	C34H58O16	767.3707	767.3769	8.1	0.593
HH4:23(2,4,5,12-OH)	medium	4.14	C35H60O16	781.3863	781.3919	7.2	1.19
HH4:24(2,5,5,12-OH)	medium	4.85	C36H62O16	795.4020	795.4096	9.6	0.488
HH4:24(2,4,4,14-OH)	medium	5.23	C36H62O16	795.4020	795.4080	7.5	2.29
HH4:25(2,4,5,14-OH)	medium	6.01	C37H64O16	809.4176	809.4241	8.1	2.43
HH4:25(2,4,5,14-OH)	medium	6.10	C37H64O16	809.4176	809.4251	9.2	1.05
Acylhexoses							
I2:12(4,8)	medium	2.19	C18H32O8	421.2079	421.2071	-1.9	0.572
H2:13	low	2.38	C19H34O8	435.2230	435.2226	-1.1	1.62
I2:14(5,9)	medium	2.58	C22H40O8	449.2392	449.2384	-1.7	0.420
I2:14(4,10)	medium	2.73	C20H36O8	449.2392	449.2384	-1.8	1.93
H2:15	low	3.04	C21H38O8	463.2543	463.2542	-0.3	0.438
I3:16(4,4,8)	medium	3.22	C22H38O9	491.2498	491.2493	-1.0	3.26
I3:16(4,4,8)	medium	3.31	C22H38O9	491.2498			
I3:16(4,4,8)	medium	3.41	C22H38O9	491.2498	491.2484	-2.9	0.654
I3:17(4,5,8)	medium	3.69	C23H40O9	505.2654	505.2656	0.3	0.180
I3:17(4,4,9)	medium	3.79	C23H40O9	505.2654	505.2649	-0.9	16.5
I3:17(4,4,9)	medium	3.92	C23H40O9	505.2654	505.2648	-1.2	0.339
I3:18(4,5,9)	medium	4.37	C24H42O9	519.2811	519.2804	-1.4	9.84
I3:18(4,4,10)	high	4.75	C24H42O9	519.2811	519.2810	-0.1	13.9
H3:21:[O1]	low	4.92	C27H48O10	577.3229	577.3254	4.3	0.767
I3:19(5,5,9)	medium	5.10	C25H44O9	533.2967	533.2984	3.3	0.169
H3:19	low	5.30	C25H44O9	533.2967	533.3000	6.2	0.0507
I3:19(4,5,10)	medium	5.47	C25H44O9	533.2967	533.2961	-1.1	13.8
I3:22(4,4,14-OH)	medium	5.71	C28H50O10	591.3386	591.3421	5.9	0.187
I3:20(4,5,11)	medium	6.11	C26H46O9	547.3124	547.3174	9.2	0.0317
I3:22(4,4,14-OH)	high	6.23	C28H50O10	591.3386	591.3395	1.6	6.86
I3:20(5,5,10)	medium	6.34	C26H46O9	547.3124	547.3122	-0.3	0.626
I3:20(4,8,8)	medium	6.52	C26H46O9	547.3124	547.3131	1.2	0.452
I3:20(4,8,8)	medium	6.63	C26H46O9	547.3124	547.3119	-0.9	0.231
I3:20(4,4,12)	high	6.91	C26H46O9	547.3124	547.3129	0.9	0.992
I3:23(4,5,14-OH)	medium	7.12	C29H52O10	605.3542	605.3549	1.2	10.2

Table S2.9. (cont'd)

H3:21	low	7.22	C27H48O9	561.3280	561.3297	3.1	0.0689
I3:21(5,8,8)	medium	7.34	C27H48O9	561.3280	561.3292	2.1	0.388
I3:21(5,8,8)	medium	7.44	C27H48O9	561.3280	561.3280	0.1	1.20
I3:21(4,5,12)	medium	7.53	C27H48O9	561.3280			
I3:21(4,5,12)	medium	7.77	C27H48O9	561.3280	561.3284	0.7	0.521
I3:21(4,5,12)	medium	7.83	C27H48O9	561.3280			
I3:24(5,5,14-OH)	medium	8.09	C30H54O10	619.3699	619.3728	4.8	2.14
I3:22(5,8,9)	medium	8.29	C28H50O9	575.3437	575.3449	2.1	0.311
I3:22(5,8,9)	medium	8.42	C28H50O9	575.3437	575.3453	2.9	0.291
I3:22(5,8,9)	medium	8.66	C28H50O9	575.3437	575.3491	9.3	0.0206
H3:24:[O1]	low	8.75	C30H54O10	619.3699	619.3776	12.4	0.0323
I3:22(4,8,10)	medium	8.88	C28H50O9	575.3437	575.3446	1.5	1.07
I3:22(4,8,10)	medium	9.05	C28H50O9	575.3437	575.3450	2.2	0.340
I3:22(4,8,10)	medium	9.53	C28H50O9	575.3437	575.3469	5.5	0.132
I3:23(5,8,10)	medium	9.73	C29H52O9	589.3593	589.3610	2.8	0.575
I3:23(5,8,10)	medium	9.88	C29H52O9	589.3593	589.3600	1.1	0.423
I3:23(4,9,10)	medium	9.96	C29H52O9	589.3593	589.3602	1.5	0.296
I3:23(4,9,10)	medium	10.43	C29H52O9	589.3593	589.3615	3.8	0.195
I3:23(4,9,10)	medium	10.51	C29H52O9	589.3593			

Table S2.10. Summary of annotated acylsugars detected in *S. melanospermum* trichome extracts. I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acylhexoses							
I2:12(4,8)	medium	2.19	C18H32O8	421.2079	421.2071	-1.9	1.19
H2:13	low	2.36	C19H34O8	435.2230	435.2226	-1.1	3.67
I2:14(4,10)	medium	2.58	C20H36O8	449.2392	449.2384	-1.7	0.980
I2:14(4,10)	medium	2.73	C20H36O8	449.2392	449.2384	-1.8	3.82
H2:15	low	3.04	C21H38O8	463.2543	463.2542	-0.3	0.850
I3:16(4,4,8)	medium	3.31	C22H38O9	491.2498	491.2493	-1.0	4.95
I3:16(4,4,8)	medium	3.41	C22H38O9	491.2498	491.2484	-2.9	2.74
I3:17(4,5,8)	medium	3.69	C23H40O9	505.2654	505.2656	0.3	0.282
I3:17(4,4,9)	medium	3.79	C23H40O9	505.2654	505.2649	-0.9	23.0
I3:17(4,5,8)	medium	3.90	C23H40O9	505.2654	505.2648	-1.2	1.25
I3:18(4,5,9)	medium	4.38	C24H42O9	519.2811	519.2804	-1.4	11.7
I3:18(4,4,10)	high	4.75	C24H42O9	519.2811	519.2810	-0.1	14.5
H3:21:[O1]	low	4.92	C27H48O10	577.3229	577.3254	4.3	0.0992
I3:19(5,5,9)	medium	5.10	C25H44O9	533.2967	533.2984	3.3	0.200
I3:19(5,5,9)	medium	5.30	C25H44O9	533.2967	533.3000	6.2	0.0972
I3:19(4,5,10)	medium	5.47	C25H44O9	533.2967	533.2961	-1.1	14.1
H3:22:[O1]	low	5.71	C25H44O9	591.3386	591.3421	5.9	0.0242
I3:20(4,5,11)	medium	6.13	C26H46O9	547.3124	547.3174	9.2	0.0651
I3:22(4,4,14-OH)	high	6.23	C28H50O10	591.3386	591.3395	1.6	0.952
I3:20(5,5,10)	medium	6.34	C26H46O9	547.3124	547.3122	-0.3	0.882
I3:20(4,8,8)	medium	6.52	C26H46O9	547.3124	547.3131	1.2	0.827
I3:20(4,8,8)	medium	6.63	C26H46O9	547.3124	547.3119	-0.9	1.04
I3:20(4,4,12)	high	6.91	C26H46O9	547.3124	547.3129	0.9	1.02
I3:23(4,5,14-OH)	medium	7.12	C29H52O10	605.3542	605.3549	1.2	1.42
I3:21(5,8,8)	medium	7.22	C27H48O9	561.3280	561.3297	3.1	0.124
I3:21(4,8,9)	medium	7.34	C27H48O9	561.3280	561.3292	2.1	0.555
I3:21(4,8,9)	medium	7.46	C27H48O9	561.3280	561.3280	0.1	2.75
I3:21(4,5,12)	medium	7.52	C27H48O9	561.3280	561.3267	-2.3	
I3:21(4,5,12)	medium	7.77	C27H48O9	561.3280	561.3284	0.7	0.528
I3:21(4,5,12)	medium	7.83	C27H48O9	561.3280	561.3267	-2.3	
I3:24(5,5,14-OH)	medium	8.07	C30H54O10	619.3699	619.3728	4.8	0.251
I3:22(5,8,9)	medium	8.29	C28H50O9	575.3437	575.3449	2.1	0.339

Table S2.10. (cont'd)

I3:22(5,8,9)	medium	8.42	C28H50O9	575.3437	575.3453	2.9	0.595
I3:22(5,8,9)	medium	8.64	C28H50O9	575.3437	575.3491	9.3	0.0313
I3:22(4,8,10)	medium	8.90	C28H50O9	575.3437	575.3446	1.5	1.48
I3:22(4,8,10)	medium	9.05	C28H50O9	575.3437	575.3450	2.2	1.21
I3:22(4,8,10)	medium	9.53	C28H50O9	575.3437	575.3469	5.5	0.0664
I3:23(5,8,10)	medium	9.73	C29H52O9	589.3593	589.3610	2.8	0.594
I3:23(5,8,10)	medium	9.89	C29H52O9	589.3593	589.3600	1.1	0.702
I3:23(5,8,10)	medium	9.99	C29H52O9	589.3593	589.3602	1.5	0.960
I3:23(4,9,10)	medium	10.41	C29H52O9	589.3593	589.3615	3.8	0.0883
I3:23(4,9,10)	medium	10.49	C29H52O9				

Table S2.11. Summary of annotated acylsugars detected in *S. lasiophyllum* trichome extracts. PH = pentose-hexose; AI = arabinose-inositol; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acylsugars with two sugar groups							
PH3:16	low	2.66	C27H46O13	623.2920	623.2928	1.3	0.00437
PH3:20:[O1]	low	2.94	C31H54O14	695.3496	695.3484	-1.7	0.0804
H3:20(4,4,12-O-P)	medium	3.04	C31H54O14	695.3496	695.3496	-0.1	17.5
AI4:18(2,4,4,8)	high	3.29	C29H48O14	665.3026	665.3027	0.2	0.0519
H3:21(4,5,12-O-P)	medium	3.42	C32H56O14	709.3652	709.3644	-1.1	11.9
PH3:18(4,4,10)	medium	3.72	C29H50O13	651.3233	651.3225	-1.2	0.262
PH3:22:[O1]	low	3.96	C33H58O14	723.3809	723.3826	2.3	0.0254
H3:22:1(4,4,14-O-P)	medium	4.14	C33H58O14	723.3809	723.3814	0.7	0.210
H3:22:1(4,4,14-O-P)	medium	4.31	C33H58O14	723.3809	723.3803	-0.8	1.20
PH3:23:[O1]	low	4.77	C34H60O14	737.3965	737.3961	-0.5	0.102
PH3:23:[O1]	low	4.97	C34H60O14	737.3965	737.3948	-2.3	0.496
PH3:20(4,4,12)	medium	5.23	C31H54O13	679.3546	679.3543	-0.4	0.0678
PH3:20(4,4,12)	medium	5.47	C31H54O13	679.3546	679.3535	-1.6	2.78
PH3:21(4,5,12)	medium	6.01	C32H56O13	693.3703	693.3704	0.2	0.0502
PH3:21(4,5,12)	medium	6.28	C32H56O13	693.3703	693.3682	-3.0	3.10
PH3:22	low	7.26	C33H58O13	707.3859	707.3885	3.6	0.00458
PH3:22(4,4,14)	medium	7.55	C33H58O13	707.3859	707.3887	3.9	0.00894
PH3:22(4,4,14)	medium	7.89	C33H58O13	707.3859	707.3896	5.2	0.0176
PH3:23	low	8.52	C34H60O13	721.4010	721.4040	4.1	0.0104
PH3:23	low	8.85	C34H60O13	721.4010	721.4055	6.1	0.0157
Acylhexoses							
H3:20:1	low	3.04	C26H44O9	545.2962	545.2958	-0.7	0.0604
H3:19:[O1]	low	3.42	C25H44O10	549.2911	549.2910	-0.2	0.116
H3:20:[O1]	low	4.09	C26H46O10	563.3073	563.3091	3.1	0.0152
I3:20(4,4,12-OH)	medium	4.25	C26H46O10	563.3073	563.3062	-2.0	10.9
H3:18	low	4.57	C24H42O9	519.2811	519.2818	1.4	0.0123
I3:18(4,4,10)	high	4.75	C24H42O9	519.2811	519.2789	-4.3	2.42
H3:21:[O1]	low	4.70	C27H48O10	577.3229	577.3228	-0.1	0.0417
I3:21(4,5,12-OH)	medium	4.90	C27H48O10	577.3229	577.3201	-4.9	16.3
H3:19	low	5.27	C25H44O9	533.2967	533.2962	-0.9	0.0159
I3:19(4,5,10)	medium	5.47	C25H44O9	533.2967	533.2953	-2.7	3.89
I3:22(5,5,12-OH)	low	5.65	C28H50O10	591.3386	591.3384	-0.4	0.0596

Table S2.11. (cont'd)

I3:22(4,4,14-OH)	high	5.98	C28H50O10	591.3386	591.3380	-1.1	0.0963
I3:22(4,4,14-OH)	high	6.24	C28H50O10	591.3386	591.3364	-3.7	2.66
H3:20	low	6.31	C26H46O9	547.3124	547.3143	3.5	0.00880
H3:20	low	6.65	C26H46O9	547.3124	547.3106	-3.3	0.117
I3:20(4,4,12)	high	6.94	C26H46O9	547.3124	547.3115	-1.7	11.3
I3:23(4,5,14-OH)	medium	6.79	C29H52O10	605.3542	605.3541	-0.2	0.112
I3:23(4,5,14-OH)	medium	7.09	C29H52O10	605.3542	605.3530	-2.0	1.43
I3:21(4,5,12)	medium	7.52	C27H48O9	561.3280	561.3270	-1.8	0.130
I3:21(4,5,12)	medium	7.81	C27H48O9	561.3280	561.3273	-1.3	12.1
H3:22	low	8.83	C28H50O9	575.3437	575.3454	3.0	0.0155
H3:22	low	9.25	C28H50O9	575.3437	575.3438	0.2	0.0365
H3:22	low	9.6	C28H50O9	575.3437	575.3441	0.6	0.0878
H3:23	low	10.16	C29H52O9	589.3593	589.3612	3.2	0.0238
H3:23	low	10.51	C29H52O9	589.3593	589.3598	0.8	0.0882

Table S2.12. Summary of annotated acylsugars detected in *S. prinophyllum* trichome extracts. PH = pentose-hexose; AI = arabinose-inositol; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over three samples. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acylsugars with two sugar groups							
PH3:16	low	2.73	C27H46O13	623.2920	623.2906	-2.2	0.545
PH4:18	low	3.09	C29H48O14	665.3026	665.3060	5.2	0.000629
PH4:18	low	3.29	C29H48O14	665.3026	665.3008	-2.8	0.263
H3:20(4,4,12-O-P)	medium	3.04	C31H54O14	695.3496	695.3479	-2.4	3.17
PH3:20:[O1]	low	2.91	C31H54O14	695.3496	695.3481	-2.1	0.00561
PH3:20:[O1]	low	3.16	C31H54O14	695.3496	695.3539	6.1	0.00359
PH3:20:[O1]	low	3.26	C31H54O14	695.3496	695.3529	4.7	0.00365
PH3:20:[O1]	low	3.41	C31H54O14	695.3496	695.3473	-3.3	0.0759
H3:21(4,5,12-O-P)	medium	3.44	C32H56O14	709.3652	709.3628	-3.3	0.342
PH3:21:[O1]	low	3.58	C32H56O14	709.3652	709.3660	1.1	0.00176
PH3:21:[O1]	low	3.72	C32H56O14	709.3652	709.3667	2.1	0.0147
PH3:18(4,4,10)	medium	3.61	C29H50O13	651.3233	651.3224	-1.4	0.814
PH3:18(4,4,10)	medium	3.72	C29H50O13	651.3233	651.3233	0.0	2.13
HH3:22:[O1]	low	3.66	C34H60O15	753.3914	753.3915	0.2	0.0227
H3:22(4,4,14-O-H)	medium	3.81	C34H60O15	753.3914	753.3916	0.3	2.79
H3:22(4,4,14-O-P)	medium	4.14	C33H58O14	723.3809	723.3812	0.4	1.03
H3:22(4,4,14-O-P)	medium	4.32	C33H58O14	723.3809	723.3813	0.6	26.6
PH3:22:[O1]	low	4.52	C33H58O14	723.3809	723.3813	0.5	0.0691
PH3:22:[O1]	low	4.93	C33H58O14	723.3809	723.3809	-0.1	0.0230
HH3:23:[O1]	low	4.39	C35H62O15	767.4065	767.4082	2.2	0.0734
H4:24(4,4,4,12-O-P)	medium	4.53	C35H60O15	765.3914	765.3905	-1.2	0.399
PH3:23:[O1]	low	4.77	C34H60O14	737.3965	737.3969	0.6	0.0224
H3:23(4,5,14-O-P)	medium	4.86	C34H60O14	737.3965	373.3950	-2.0	3.71
H3:23(4,5,14-O-P)	medium	4.99	C34H60O14				
PH3:20(4,4,12)	medium	5.23	C31H54O13	679.3546	679.3534	-1.8	0.0507
PH3:20(4,4,12)	medium	5.47	C31H54O13	679.3546	679.3530	-2.4	2.28
H4:26(4,4,4,14-O-P)	medium	6.31	C37H64O15	793.4227	793.4241	1.8	0.0528
H4:26(4,4,4,14-O-P)	medium	6.56	C37H64O15	793.4227	793.4224	-0.3	2.74

173

Table S2.12. (cont'd)

H4:27(4,4,5,14-O-P)	medium	7.31	C38H66O15	807.4384	807.4380	-0.5	0.559
H4:27(4,4,5,14-O-P)	medium	7.42	C38H66O15	807.4384	807.4381	-0.4	0.0593
PH3:22(4,4,14)	medium	7.55	C33H58O13	707.3859	707.3847	-1.7	0.144
PH3:22(4,4,14)	medium	7.89	C33H58O13	707.3859	707.3848	-1.6	2.21
PH3:23	low	8.88	C34H60O13	721.4010	721.4012	0.2	0.0991
PH3:23	low	9.23	C34H60O13	721.4010	721.4043	4.5	0.00364
Acylhexoses							
H4:17	low	3.36	C23H38O10	519.2442	519.2452	1.9	0.00412
H4:17	low	3.62	C23H38O10	519.2442	519.2429	-2.4	0.495
H5:19	low	3.72	C25H42O11	563.2704	563.2710	1.1	0.0123
H3:20:[O1]	low	4.09	C26H46O10	563.3073	563.3091	3.1	0.00276
H4:18	low	4.14	C24H40O10	533.2598	533.2593	-0.9	0.669
I3:20(4,4,12-OH)	medium	4.25	C26H46O10	563.3073	563.3067	-1.1	4.23
H3:20:[O1]	low	4.57	C26H46O10	563.3073	563.3103	5.3	0.00220
H3:20:[O1]	low	4.73	C26H46O10	563.3073	563.3094	3.8	0.00130
I3:18(4,4,10)	medium	4.57	C24H42O9	519.2811	519.2793	-3.4	0.0608
I3:18(4,4,10)	medium	4.75	C24H42O9	519.2811	519.2792	-3.7	0.166
H3:18	low	4.82	C24H42O9	519.2811	519.2800	-2.1	0.000654
H4:19	low	4.90	C25H42O10	547.2755	547.2741	-2.6	0.104
I3:21(4,5,12-OH)	medium	4.85	C27H48O10	577.3229	577.3212	-3.0	0.292
H3:21:[O1]	low	5.10	C27H48O10	577.3229	577.3206	-3.9	0.00116
H5:20	low	5.35	C26H42O11	575.2704	575.2746	7.4	0.00229
H3:21:[O1]	low	5.35	C27H48O10	577.3229	577.3239	1.7	0.00318
H3:21:[O1]	low	5.45	C27H48O10	577.3229	577.3223	-1.0	0.00678
H3:19	low	5.27	C25H44O9	533.2962	533.2965	0.6	0.00792
H3:19	low	5.48	C25H44O9	533.2962	533.2955	-1.4	0.0139
H5:21	low	6.03	C27H44O11	589.2860	589.2847	-2.2	0.0454
H5:21	low	6.15	C27H44O11	589.2860	589.2858	-0.4	0.0146
I3:22(4,4,14-OH)	medium	5.98	C28H50O10	591.3386	591.3359	-4.6	1.10
I3:22(4,4,14-OH)	medium	6.24	C28H50O10	591.3386	591.3388	0.3	24.8
I3:20(4,4,12)	medium	6.65	C26H46O9	547.3124	547.3111	-2.4	0.0169
I3:20(4,4,12)	medium	6.94	C26H46O9	547.3124	547.3107	-3.1	1.23
H5:22	low	6.78	C28H46O11	603.3017	603.3004	-2.2	0.354
I3:23(4,5,14-OH)	medium	6.84	C29H52O10	605.3542	605.3510	-5.4	0.0488
I3:23(4,5,14-OH)	medium	7.13	C29H52O10	605.3542	605.3529	-2.2	4.39
H5:22	low	7.76	C29H48O11	617.3173	617.3167	-1.0	0.359
H3:24:[O1]	low	8.09	C30H54O10	619.3694	619.3694	0.0	0.0434
H3:21	low	7.79	C27H48O9	561.3275	561.3265	-1.8	0.0522
H3:21	low	8.24	C27H48O9	561.3275	561.3271	-0.7	0.0134
I3:22(4,4,14)	medium	9.25	C28H50O9	575.3437	575.3418	-3.4	0.947
I3:22(4,4,14)	medium	9.60	C28H50O9	575.3437	575.3424	-2.2	9.19
I3:23(4,5,14)	medium	10.16	C29H52O9	589.3593	589.3596	0.5	0.0258

174

Table S2.12. (cont'd)

I3:23(4,5,14)	medium	10.24	C29H52O9				
I3:23(4,5,14)	medium	10.50	C29H52O9	589.3593	589.3584	-1.6	0.917
I3:23(4,5,14)	medium	10.59	C29H52O9				
I3:23(4,5,14)	medium	10.94	C29H52O9	589.3593	589.3585	-1.4	0.0307
H3:24	low	11.57	C30H54O9	603.3744	603.3768	4.0	0.0141
H3:24	low	11.90	C30H54O9	603.3744	603.3781	6.1	0.00236
H3:24	low	12.03	C30H54O9	603.3744	603.3749	0.8	0.00432

Table S2.13. Summary of annotated acylsugars detected in *S. acerifolium* fruit surface metabolite extracts. dHH = deoxyhexose-hexose; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over three samples. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm (ppm)	Average acylsugar abundance (%)
Acyldisaccharides							
dHH3:30:[O2]	low	9.33	C42H76O15	865.5161	865.5191	3.5	0.00889
dHH3:30:[O2]	low	9.47	C42H76O15	865.5161	865.5189	3.3	0.0104
dHH3:30:[O2]	low	9.76	C42H76O15	865.5161	865.5192	3.6	0.00865
dHH3:30:[O2]	low	10.04	C42H76O15	865.5161	865.5199	4.4	0.00563
dHH3:34:[O3]	low	10.49	C46H84O16	937.5736	937.5772	3.8	0.0342
dHH3:28:[O1]	low	10.51	C40H72O14	821.4899	821.4919	2.5	0.0754
dHH3:31:[O2]	low	10.54	C43H78O15	879.5317	879.5347	3.4	0.0571
dHH3:31:[O2]	low	10.93	C43H78O15	879.5317	879.5352	4.0	0.0462
dHH3:31:[O2]	low	11.46	C43H78O15	879.5317	879.5345	3.2	0.0217
dHH3:31:[O2]	low	11.62	C43H78O15	879.5317	879.5359	4.7	0.00233
dHH3:29:[O1]	low	11.69	C41H74O14	835.5055	835.5074	2.3	0.205
dHH3:32:[O2]	low	11.75	C44H80O15	893.5474	893.5501	3.0	0.787
dHH3:32:[O2]	low	11.89	C44H80O15	893.5474	893.5503	3.3	0.0700
dHH3:32:[O2]	low	12.10	C44H80O15	893.5474	893.5503	3.3	0.624
dHH3:30(8,10,12-OH)	medium	12.52	C42H76O14	849.5217	849.5240	2.7	0.0289
dHH3:32:[O2]	low	12.58	C44H80O15	893.5474	893.5506	3.6	0.0411
dHH3:30(8,10,12-OH)	medium	12.75	C42H76O14	849.5217	849.5236	2.2	1.81
dHH3:30(8,10,12-OH)	medium	12.83	C42H76O14	849.5217	849.5236	2.2	
dHH3:33:[O2]	low	12.86	C45H82O15	907.5630	907.5658	3.1	0.809
dHH3:33:[O2]	low	13.10	C45H82O15	907.5630	907.5663	3.6	0.492
dHH3:34:[O3]	low	13.12	C46H84O16	937.5736	937.5783	5.0	0.0112
dHH3:30:[O1]	low	13.35	C42H76O14	849.5217	849.5259	4.9	0.00531
dHH3:33:[O2]	low	13.38	C45H82O15	907.5630	907.5666	4.0	0.0743
dHH3:33:[O2]	low	13.66	C45H82O15	907.5630	907.5667	4.0	0.244
dHH3:31:[O1]	low	13.58	C43H78O14	863.5374	863.5402	3.2	0.0553
dHH3:31(9,10,12-OH)	medium	13.83	C43H78O14	863.5374	863.5398	2.7	0.763
dHH3:31:[O1]	low	13.93	C43H78O14	863.5374	863.5404	3.4	0.0824
dHH3:34(10,12-OH,12-OH)	medium	13.94	C46H84O15	921.5792	921.5827	3.8	2.79
dHH3:34(10,12-OH,12-OH)	medium	14.12	C46H84O15	921.5792	921.5827	3.8	3.64

Table S2.13. (cont'd)

dHH3:34(10,12-OH,12-OH)	medium	14.39	C46H84O15	921.5792	921.5822	3.2	0.902
dHH3:31:[O1]	low	14.44	C43H78O14	863.5374	863.5412	4.5	0.0126
dHH3:32(10,10,12-OH)	medium	14.52	C44H80O14	877.5530	877.5554	2.7	0.144
dHH3:33:[O2]	low	14.54	C45H82O15	907.5630	907.5673	4.7	0.0188
dHH3:34(10,12-OH,12-OH)	medium	14.56	C46H84O15	921.5792	921.5820	3.0	0.418
dHH3:35(10,12-OH,13-OH)	medium	14.87	C47H86O15	935.5949	935.5978	3.1	1.67
dHH3:32:[O1]	low	14.59	C44H80O14	877.5530	877.5552	2.5	0.615
dHH3:32(10,10,12-OH)	medium	14.76	C44H80O14	877.5530	877.5554	2.7	2.17
dHH3:32(10,10,12-OH)	medium	14.91	C44H80O14	877.5530	877.5555	2.8	0.125
dHH3:35(10,12-OH,13-OH)	medium	15.06	C47H86O15	935.5949	935.5980	3.3	1.89
dHH3:32(10,10,12-OH)	medium	15.26	C44H80O14	877.5530	877.5553	2.7	0.0865
dHH3:35:[O2]	low	15.32	C47H86O15	935.5949	935.5984	3.7	0.154
dHH3:35:[O2]	low	15.49	C47H86O15	935.5949	935.5980	3.3	0.702
dHH3:36:[O2]	low	15.50	C45H82O14	891.5687	891.5716	3.2	0.491
dHH3:33(10,11,12-OH)	medium	15.63	C45H82O14	891.5687	891.5715	3.1	1.08
dHH3:36(10,12-OH,14-OH)	medium	15.74	C48H88O15	949.6105	949.6145	4.2	4.31
dHH3:36(10,12-OH,14-OH)	medium	15.89	C48H88O15	949.6105	949.6148	4.5	5.36
dHH3:31	low	16.07	C43H78O13	847.5419	847.5442	2.7	0.0769
dHH3:33:[O1]	low	16.10	C45H82O14	891.5687	891.5722	3.9	0.0542
dHH3:33:[O1]	low	16.18	C45H82O14	891.5687	891.5722	3.9	
dHH3:36:[O2]	low	16.17	C48H88O15	949.6105	949.6140	3.7	0.338
dHH3:36:[O2]	low	16.30	C48H88O15	949.6105	949.6136	3.3	0.311
Acylhexoses							
I2:20(8,12-OH)	medium	4.68	C26H48O9	549.3280	549.3287	1.3	0.00156
I2:20(8,12-OH)	medium	5.07	C26H48O9	549.3280	549.3324	8.0	0.0000280
I2:20(8,12-OH)	medium	5.36	C26H48O9	549.3280	549.3284	0.7	0.00164
I2:21(9,12-OH)	medium	5.62	C27H50O9	563.3437	563.3449	2.2	0.000590
I2:21(9,12-OH)	medium	6.39	C27H50O9	563.3437	563.3435	-0.4	0.00811
I2:22(10,12-OH)	medium	6.71	C28H52O9	577.3593	577.3592	-0.1	0.0353
I2:22(10,12-OH)	medium	7.22	C28H52O9	577.3593	577.3610	2.9	0.000368
I2:22(10,12-OH)	medium	7.59	C28H52O9	577.3593	577.3591	-0.3	0.117
I3:28(8,10-OH,10-OH)	medium	8.80	C34H62O11	691.4274	691.4288	2.1	0.00171
I3:28(8,10-OH,10-OH)	medium	9.07	C34H62O11	691.4274	691.4287	1.8	0.00257
H2:24:[O1]	low	9.27	C30H56O9	605.3901	605.3907	1.0	0.0128
I3:28(8,10-OH,10-OH)	medium	9.35	C34H62O11	691.4274	691.4281	1.0	0.00749
H3:26:[O1]	low	9.83	C32H58O10	647.4007	647.4013	1.1	0.0114
H2:22	low	9.90	C28H52O8	561.3639	561.3644	1.0	0.00318
H3:29:[O3]	low	9.96	C35H64O12	705.4425	705.4443	2.5	0.00203
H2:24:[O1]	low	10.23	C30H56O9	605.3901	605.3909	1.3	0.0120
H3:29:[O3]	low	10.24	C35H64O12	705.4425	705.4437	1.6	0.0104

177

Table S2.13. (cont'd)

H3:29:[O3]	low	10.38	C35H64O12	705.4425	705.4436	1.6	0.0124
H3:28:[O1]	low	11.85	C34H62O10	675.4325	675.4330	0.7	0.0142
I3:28(8,10,10-OH)	medium	12.19	C34H62O10	675.4325	675.4328	0.4	1.17
I3:28(8,8,12-OH)	medium	12.21	C34H62O10	675.4325	675.4328	0.4	
H3:31:[O2]	low	12.22	C37H68O11	733.4744	733.4750	0.8	0.0468
H3:31:[O2]	low	12.29	C37H68O11				
H4:33:[O3]	low	12.25	C39H70O13	791.4798	791.4830	4.0	0.00832
H3:34:[O3]	low	12.45	C40H74O12	791.5162	791.5167	0.6	0.00458
H3:31:[O2]	low	12.48	C37H68O11	733.4744	733.4749	0.6	0.529
I3:31(9,10-OH,12-OH)	medium	12.53	C37H68O11				
I3:31(9-OH,10,12-OH)	medium	12.68	C37H68O11	733.4744	733.4750	0.8	0.183
I3:34(10-OH,12-OH,12-OH)	medium	12.70	C40H74O12	791.5162	791.5174	1.5	0.739
H3:31:[O2]	low	12.88	C37H68O11	733.4744	733.4756	1.6	0.00716
H3:29:[O1]	low	12.88	C35H64O10	689.4481	689.4495	2.1	0.0132
H3:34:[O3]	low	12.90	C40H74O12	791.5162	791.5176	1.7	0.0346
I3:32(10,10-OH,12-OH)	medium	13.25	C38H70O11	747.4900	747.4908	1.1	0.121
I3:29(8,9,12-OH)	medium	13.28	C35H64O10	689.4481	689.4488	1.1	1.07
I3:29(8,9,12-OH)	medium	13.40	C35H64O10				
H3:29:[O1]	low	13.51	C35H64O10	689.4481	689.4493	1.7	0.0133
I3:32(10,10-OH,12-OH)	medium	13.53	C38H70O11	747.7900	747.4910	1.3	3.47
I3:32(10,10-OH,12-OH)	medium	13.61	C38H70O11				
I3:32(10,10-OH,12-OH)	medium	13.81	C38H70O11	747.4900	747.4910	1.3	0.433
H3:30:[O1]	low	13.86	C36H66O10	703.4638	703.4649	1.6	0.104
I3:30(9,10,10-OH)	medium	14.20	C36H66O10				
I3:30(8,10,12-OH)	medium	14.28	C36H66O10	703.4638	703.4651	1.9	3.37
H3:30:[O1]	low	14.40	C36H66O10				
H3:33:[O2]	low	14.24	C39H72O11	761.5057	761.5073	2.1	0.0356
I3:33(9,12-OH,12-OH)	medium	14.47	C39H72O11	761.5057	761.5067	1.3	2.83
H3:33:[O2]	low	14.60	C39H72O11				
I3:30(9,10,10-OH)	medium	14.52	C36H66O10	703.4638	703.4642	0.6	0.235
I3:33(9,10-OH,14-OH)	medium	14.71	C39H72O11	761.5057	761.5067	1.3	0.436
I3:31(9,10,12-OH)	medium	14.76	C37H68O10	717.4794	717.4808	2.0	0.0398
I3:31(9,10,12-OH)	medium	15.14	C37H68O10				
I3:31(8,11,12-OH)	medium	15.20	C37H68O10	717.4794	717.4803	1.3	4.22
H3:31:[O1]	low	15.34	C37H68O10				
H3:31:[O1]	low	15.47	C37H68O10				
I3:34(10,12-OH,12-OH)	medium	15.35	C40H74O11	775.5213	775.5237	3.1	13.8
I3:34(10,12-OH,12-OH)	medium	15.57	C40H74O11	775.5213	775.5228	2.0	3.86
I3:32(10,10,12-OH)	medium	15.61	C38H70O10	731.4951	731.4966	2.1	0.103
I3:32(10,10,12-OH)	medium	15.92	C38H70O10	731.4951	731.4967	2.2	4.64
I3:32(10,10,12-OH)	medium	16.03	C38H70O10	731.4951	731.4962	1.5	3.41
I3:35(10,12-OH,13-OH); I3:35(9,12-OH,14-OH); I3:35(11,12-OH,12-OH)	medium	16.17	C41H76O11	789.5370	789.5390	2.5	6.29

Table S2.13. (cont'd)

I3:32(10,10,12-OH)	medium	16.27	C38H70O10	731.4951	731.4960	1.2	0.127
I3:35(10,12-OH,13-OH); I3:35(11,12-OH,12-OH)	medium	16.37	C41H76O11	789.5370	789.5381	1.4	0.870
I3:33(10,11,12-OH)	medium	16.37	C39H72O10	745.5107	745.5095	-1.6	0.0392
H3:36:[O2]	low	16.44	C42H78O11	803.5526	803.5539	1.6	0.0906
I3:33(10,11,12-OH)	medium	16.65	C39H72O10	745.5107	745.5114	0.9	1.80
H3:36:[O2]	low	16.77	C42H78O11	803.5526	803.5538	1.5	6.28
I3:36(10,12-OH,14-OH)	medium	16.90	C42H78O11				
I3:33(10,11,12-OH)	medium	16.77	C39H72O10	745.5107	745.5111	0.6	0.767
I3:33(10,11,12-OH)	medium	17.00	C39H72O10	745.5107	745.5100	-0.9	0.0247
I3:36(10,12-OH,14-OH)	medium	17.06	C42H78O11	803.5526	803.5540	1.7	1.12
I3:34(10,10,14-OH)	medium	17.08	C40H74O10	759.5264	759.5285	2.7	0.0276
I3:34(10,12,12-OH)	medium	17.33	C40H74O10	759.5264	759.5273	1.1	1.70
I3:34(10,10,14-OH)	medium	17.46	C40H74O10	759.5264	759.5273	1.2	0.849
I3:37(11,12-OH,14-OH); I3:37(10,13-OH,14-OH)	medium	17.58	C43H80O11	817.5683	817.5699	1.9	0.896
I3:34(10,12,12-OH)	medium	17.71	C40H74O10	759.5264	759.5327	8.3	0.0111
H3:37:[O2]	low	17.81	C43H80O11	817.5683	817.5713	3.7	0.0343
H3:35:[O1]	low	18.06	C41H76O10	773.5415	773.5436	2.7	0.209
H3:38:[O2]	low	18.39	C44H82O11	831.5834	831.5869	4.2	0.0120

Table S2.14. Summary of annotated acylsugars detected in *S. atropurpureum* fruit surface metabolite extracts. dHH = deoxyhexose-hexose; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over three samples. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm (ppm)	Average acylsugar abundance (%)
Acyldisaccharides							
dHH3:30:[O2]	low	9.33	C42H76O15	865.5161	865.5191	3.5	0.00889
dHH3:30:[O2]	low	9.47	C42H76O15	865.5161	865.5189	3.3	0.0104
dHH3:30:[O2]	low	9.76	C42H76O15	865.5161	865.5192	3.6	0.00865
dHH3:30:[O2]	low	10.04	C42H76O15	865.5161	865.5199	4.4	0.00563
dHH3:34:[O3]	low	10.49	C46H84O16	937.5736	937.5772	3.8	0.0342
dHH3:28:[O1]	low	10.51	C40H72O14	821.4899	821.4919	2.5	0.0754
dHH3:31:[O2]	low	10.54	C43H78O15	879.5317	879.5347	3.4	0.0571
dHH3:31:[O2]	low	10.93	C43H78O15	879.5317	879.5352	4.0	0.0462
dHH3:31:[O2]	low	11.46	C43H78O15	879.5317	879.5345	3.2	0.0217
dHH3:31:[O2]	low	11.62	C43H78O15	879.5317	879.5359	4.7	0.00233
dHH3:29:[O1]	low	11.69	C41H74O14	835.5055	835.5074	2.3	0.205
dHH3:32:[O2]	low	11.75	C44H80O15	893.5474	893.5501	3.0	0.787
dHH3:32:[O2]	low	11.89	C44H80O15	893.5474	893.5503	3.3	0.0700
dHH3:32:[O2]	low	12.1	C44H80O15	893.5474	893.5503	3.3	0.624
dHH3:30(8,10,12-OH)	medium	12.52	C42H76O14	849.5217	849.5240	2.7	0.0289
dHH3:32:[O2]	low	12.58	C44H80O15	893.5474	893.5506	3.6	0.0411
dHH3:30(8,10,12-OH)	medium	12.75	C42H76O14	849.5217	849.5236	2.2	1.81
dHH3:30(8,10,12-OH)	medium	12.83	C42H76O14	849.5217	849.5236	2.2	
dHH3:33:[O2]	low	12.86	C45H82O15	907.5630	907.5658	3.1	0.809
dHH3:33:[O2]	low	13.10	C45H82O15	907.5630	907.5663	3.6	0.492
dHH3:34:[O3]	low	13.12	C46H84O16	937.5736	937.5783	5.0	0.0112
dHH3:30:[O1]	low	13.35	C42H76O14	849.5217	849.5259	4.9	0.00531
dHH3:33:[O2]	low	13.38	C45H82O15	907.5630	907.5666	4.0	0.0743
dHH3:33:[O2]	low	13.66	C45H82O15	907.5630	907.5667	4.0	0.244
dHH3:31:[O1]	low	13.58	C43H78O14	863.5374	863.5402	3.2	0.0553
dHH3:31(9,10,12-OH)	medium	13.83	C43H78O14	863.5374	863.5398	2.7	0.763
dHH3:31:[O1]	low	13.93	C43H78O14	863.5374	863.5404	3.4	0.0824
dHH3:34(10,12-OH,12-OH)	medium	13.94	C46H84O15	921.5792	921.5827	3.8	2.79
dHH3:34(10,12-OH,12-OH)	medium	14.12	C46H84O15	921.5792	921.5827	3.8	3.64
dHH3:34(10,12-OH,12-OH)	medium	14.39	C46H84O15	921.5792	921.5822	3.2	0.902

Table S2.14. (cont'd)

dHH3:31:[O1]	low	14.44	C43H78O14	863.5374	863.5412	4.5	0.0126
dHH3:32(10,10,12-OH)	medium	14.52	C44H80O14	877.5530	877.5554	2.7	0.144
dHH3:33:[O2]	low	14.54	C45H82O15	907.5630	907.5673	4.7	0.0188
dHH3:34(10,12-OH,12-OH)	medium	14.56	C46H84O15	921.5792	921.5820	3.0	0.418
dHH3:35(10,12-OH,13-OH)	medium	14.87	C47H86O15	935.5949	935.5978	3.1	1.67
dHH3:32:[O1]	low	14.59	C44H80O14	877.5530	877.5552	2.5	0.615
dHH3:32(10,10,12-OH)	medium	14.76	C44H80O14	877.5530	877.5554	2.7	2.17
dHH3:32(10,10,12-OH)	medium	14.91	C44H80O14	877.5530	877.5555	2.8	0.125
dHH3:35(10,12-OH,13-OH)	medium	15.06	C47H86O15	935.5949	935.5980	3.3	1.89
dHH3:32(10,10,12-OH)	medium	15.26	C44H80O14	877.5530	877.5553	2.7	0.0865
dHH3:35:[O2]	low	15.32	C47H86O15	935.5949	935.5984	3.7	0.154
dHH3:35:[O2]	low	15.49	C47H86O15	935.5949	935.5980	3.3	0.702
dHH3:36:[O2]	low	15.50	C45H82O14	891.5687	891.5716	3.2	0.491
dHH3:33(10,11,12-OH)	medium	15.63	C45H82O14	891.5687	891.5715	3.1	1.08
dHH3:36(10,12-OH,14-OH)	medium	15.74	C48H88O15	949.6105	949.6145	4.2	4.31
dHH3:36(10,12-OH,14-OH)	medium	15.89	C48H88O15	949.6105	949.6148	4.5	5.36
dHH3:31	low	16.07	C43H78O13	847.5419	847.5442	2.7	0.0769
dHH3:33:[O1]	low	16.10	C45H82O14	891.5687	891.5722	3.9	0.0542
dHH3:33:[O1]	low	16.18	C45H82O14	891.5687	891.5722	3.9	
dHH3:36:[O2]	low	16.17	C48H88O15	949.6105	949.6140	3.7	0.338
dHH3:36:[O2]	low	16.30	C48H88O15	949.6105	949.6136	3.3	0.311
Acylhexoses							
I2:20(8,12-OH)	medium	4.68	C26H48O9	549.3280	549.3287	1.3	0.00240
I2:20(8,12-OH)	medium	5.07	C26H48O9	549.3280	549.3324	8.0	0.000190
I2:20(8,12-OH)	medium	5.36	C26H48O9	549.3280	549.3284	0.7	0.00185
I2:21(9,12-OH)	medium	5.62	C27H50O9	563.3437	563.3449	2.2	0.000724
I2:21(9,12-OH)	medium	6.39	C27H50O9	563.3437	563.3435	-0.4	0.00690
I2:22(10,12-OH)	medium	6.71	C28H52O9	577.3593	577.3592	-0.1	0.0459
I2:22(10,12-OH)	medium	7.22	C28H52O9	577.3593	577.3610	2.9	0.00147
I2:22(10,12-OH)	medium	7.59	C28H52O9	577.3593	577.3591	-0.3	0.114
I3:28(8,10-OH,10-OH)	medium	8.80	C34H62O11	691.4274	691.4288	2.1	0.00667
I3:28(8,10-OH,10-OH)	medium	9.07	C34H62O11	691.4274	691.4287	1.8	0.00908
H2:24:[O1]	low	9.27	C30H56O9	605.3901	605.3907	1.0	0.0146
I3:28(8,10-OH,10-OH)	medium	9.35	C34H62O11	691.4274	691.4281	1.0	0.0236
H3:26:[O1]	low	9.83	C32H58O10	647.4007	647.4013	1.1	0.0312
H2:22	low	9.90	C28H52O8	561.3639	561.3644	1.0	0.00698
H3:29:[O3]	low	9.96	C35H64O12	705.4425	705.4443	2.5	0.00406
H2:24:[O1]	low	10.23	C30H56O9	605.3901	605.3909	1.3	0.00850
H3:29:[O3]	low	10.24	C35H64O12	705.4425	705.4437	1.6	0.0243
H3:29:[O3]	low	10.38	C35H64O12	705.4425	705.4436	1.6	0.0265
H3:28:[O1]	low	11.85	C34H62O10	675.4325	675.4330	0.7	0.0149

Table S2.14. (cont'd)

I3:28(8,10,10-OH)	medium	12.19	C34H62O10	675.4325	675.4328	0.4	1.97
I3:28(8,8,12-OH)	medium	12.21	C34H62O10	675.4325	675.4328	0.4	
H3:31:[O2]	low	12.22	C37H68O11	733.4744	733.4750	0.8	0.0632
H3:31:[O2]	low	12.29	C37H68O11				
H4:33:[O3]	low	12.25	C39H70O13	791.4798	791.4830	4.0	0.0230
H3:34:[O3]	low	12.45	C40H74O12	791.5162	791.5167	0.6	0.00353
H3:31:[O2]	low	12.48	C37H68O11	733.4744	733.4749	0.6	0.670
I3:31(9,10-OH,12-OH)	medium	12.53	C37H68O11				
I3:31(9-OH,10,12-OH)	medium	12.68	C37H68O11	733.4744	733.4750	0.8	0.220
I3:34(10-OH,12-OH,12-OH)	medium	12.70	C40H74O12	791.5162	791.5174	1.5	0.838
H3:31:[O2]	low	12.88	C37H68O11	733.4744	733.4756	1.6	0.00960
H3:29:[O1]	low	12.88	C35H64O10	689.4481	689.4495	2.1	0.00802
H3:34:[O3]	low	12.90	C40H74O12	791.5162	791.5176	1.7	0.0400
I3:32(10,10-OH,12-OH)	medium	13.25	C38H70O11	747.4900	747.4908	1.1	0.207
I3:29(8,9,12-OH)	medium	13.28	C35H64O10	689.4481	689.4488	1.1	1.45
I3:29(8,9,12-OH)	medium	13.40	C35H64O10				
H3:29:[O1]	low	13.51	C35H64O10	689.4481	689.4493	1.7	0.0144
I3:32(10,10-OH,12-OH)	medium	13.53	C38H70O11	747.4900	747.4910	1.3	4.45
I3:32(10,10-OH,12-OH)	medium	13.61	C38H70O11				
I3:32(10,10-OH,12-OH)	medium	13.81	C38H70O11	747.4900	747.4910	1.3	0.544
H3:30:[O1]	low	13.86	C36H66O10	703.4638	703.4649	1.6	0.0700
I3:30(9,10,10-OH)	medium	14.20	C36H66O10				
I3:30(8,10,12-OH)	medium	14.28	C36H66O10	703.4638	703.4651	1.9	4.23
H3:30:[O1]	low	14.40	C36H66O10				
H3:33:[O2]	low	14.24	C39H72O11	761.5057	761.5073	2.1	0.0449
I3:33(9,12-OH,12-OH)	medium	14.47	C39H72O11	761.5057	761.5067	1.3	2.33
H3:33:[O2]	low	14.60	C39H72O11				
I3:30(9,10,10-OH)	medium	14.52	C36H66O10	703.4638	703.4642	0.6	0.253
I3:33(9,10-OH,14-OH)	medium	14.71	C39H72O11	761.5057	761.5067	1.3	0.339
I3:31(9,10,12-OH)	medium	14.76	C37H68O10	717.4794	717.4808	2.0	0.0209
I3:31(9,10,12-OH)	medium	15.14	C37H68O10				
I3:31(8,11,12-OH)	medium	15.20	C37H68O10	717.4794	717.4803	1.3	4.44
H3:31:[O1]	low	15.34	C37H68O10				
H3:31:[O1]	low	15.47	C37H68O10				
I3:34(10,12-OH,12-OH)	medium	15.35	C40H74O11	775.5213	775.5237	3.1	10.7
I3:34(10,12-OH,12-OH)	medium	15.57	C40H74O11	775.5213	775.5228	2.0	3.36
I3:32(10,10,12-OH)	medium	15.61	C38H70O10	731.4951	731.4966	2.1	0.0746
I3:32(10,10,12-OH)	medium	15.92	C38H70O10	731.4951	731.4967	2.2	6.14
I3:32(10,10,12-OH)	medium	16.03	C38H70O10	731.4951	731.4962	1.5	3.42
I3:35(10,12-OH,13-OH); I3:35(9,12-OH,14-OH); I3:35(11,12-OH,12-OH)	medium	16.17	C41H76O11	789.5370	789.5390	2.5	3.76
I3:32(10,10,12-OH)	medium	16.27	C38H70O10	731.4951	731.4960	1.2	0.164

Table S2.14. (cont'd)

I3:35(10,12-OH,13-OH); I3:35(11,12-OH,12-OH)	medium	16.37	C41H76O11	789.5370	789.5381	1.4	0.557
I3:33(10,11,12-OH)	medium	16.37	C39H72O10	745.5107	745.5095	-1.6	0.0156
H3:36:[O2]	low	16.44	C42H78O11	803.5526	803.5539	1.6	0.0595
I3:33(10,11,12-OH)	medium	16.65	C39H72O10	745.5107	745.5114	0.9	1.91
H3:36:[O2]	low	16.77	C42H78O11	803.5526	803.5538	1.5	5.37
I3:36(10,12-OH,14-OH)	medium	16.90	C42H78O11				
I3:33(10,11,12-OH)	medium	16.77	C39H72O10	745.5107	745.5111	0.6	0.539
I3:33(10,11,12-OH)	medium	17.00	C39H72O10	745.5107	745.5100	-0.9	0.0175
I3:36(10,12-OH,14-OH)	medium	17.06	C42H78O11	803.5526	803.5540	1.7	0.883
I3:34(10,10,14-OH)	medium	17.08	C40H74O10	759.5264	759.5285	2.7	0.0237
I3:34(10,12,12-OH)	medium	17.33	C40H74O10	759.5264	759.5273	1.1	1.84
I3:34(10,10,14-OH)	medium	17.46	C40H74O10	759.5264	759.5273	1.2	0.625
I3:37(11,12-OH,14-OH); I3:37(10,13-OH,14-OH)	medium	17.58	C43H80O11	817.5683	817.5699	1.9	0.534
I3:34(10,12,12-OH)	medium	17.71	C40H74O10	759.5264	759.5327	8.3	0.0137
H3:37:[O2]	low	17.81	C43H80O11	817.5683	817.5713	3.7	0.0143
H3:35:[O1]	low	18.06	C41H76O10	773.5415	773.5436	2.7	0.155
H3:38:[O2]	low	18.39	C44H82O11	831.5834	831.5869	4.2	0.00855

Table S2.15. Summary of annotated acylsugars detected in *S. acerifolium* trichome extracts. PH = pentose-hexose; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over two samples. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	m/z acc	m/z exp	Δm (ppm)	Average acylsugar abundance (%)
Acyldisaccharides							
PH3:16	low	2.21	C27H46O13	623.2915	623.2966	8.2	0.0143
PH3:17	low	2.28	C28H48O13	637.3071	637.3080	1.4	1.45
PH3:18	low	2.50	C33H50O10	651.3381	651.3326	-8.4	0.00475
PH4:18	low	2.66	C29H48O14	665.3021	665.3074	8.0	0.0468
PH4:19(2,5,6,6)	medium	2.81	C30H50O14	679.3183	679.3183	-0.1	8.82
Acylhexoses							
H3:17	low	3.16	C23H40O9	505.2649	505.2661	2.4	0.00823
H2:18:[O1]	low	3.41	C24H44O9	521.2962	521.2984	4.3	0.272
H3:17	low	3.79	C23H40O9	505.2649	505.2670	4.1	0.0456
H2:19:[O1]	low	4.05	C25H46O9	535.3118	535.3146	5.2	0.253
H2:19:[O1]	low	4.20	C25H46O9	535.3118	535.3132	2.6	0.00233
I3:18(6,6,6)	medium	4.49	C24H42O9	519.2811	519.2802	-1.7	29.2
I2:20(6,14-OH)	medium	4.92	C26H48O9	549.3280	549.3271	-1.6	46.7
H3:19	low	5.35	C25H44O9	533.2967	533.2965	-0.4	3.29
I2:21(5,14-OH)	medium	5.80	C27H50O9	563.3437	563.3467	5.3	0.156
I2:21(6,15-OH)	medium	5.98	C27H50O9	563.3437	563.3462	4.5	0.246
I3:20(6,6,8)	medium	6.30	C26H46O9	547.3124	547.3127	0.5	1.62
I3:20(6,6,8)	medium	6.38	C26H46O9				
I2:22(6,16-OH)	medium	7.19	C28H52O9	577.3593	577.3588	-0.8	7.76
H3:21	low	7.42	C27H48O9	561.3275	561.3317	7.6	0.106

Table S2.16. Summary of annotated acylsugars detected in *S. atropurpureum* trichome extracts. PH = pentose-hexose; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acyldisaccharides							
PH3:16	low	2.21	C27H46O13	623.2915	623.2966	8.2	0.0143
PH3:17	low	2.28	C28H48O13	637.3071	637.3080	1.4	1.45
PH3:18	low	2.50	C33H50O10	651.3381	651.3326	-8.4	0.00475
PH4:18	low	2.66	C29H48O14	665.3021	665.3074	8.0	0.0468
PH4:19(2,5,6,6)	medium	2.81	C30H50O14	679.3183	679.3183	-0.1	8.82
Acylhexoses							
H3:17	low	3.16	C23H40O9	505.2649	505.2661	2.4	0.00823
H2:18:[O1]	low	3.41	C24H44O9	521.2962	521.2984	4.3	0.272
H3:17	low	3.79	C23H40O9	505.2649	505.2670	4.1	0.0456
H2:19:[O1]	low	4.05	C25H46O9	535.3118	535.3146	5.2	0.253
H2:19:[O1]	low	4.20	C25H46O9	535.3118	535.3132	2.6	0.00233
I3:18(6,6,6)	medium	4.49	C24H42O9	519.2811	519.2802	-1.7	29.2
I2:20(6,14-OH)	medium	4.92	C26H48O9	549.3280	549.3271	-1.6	46.7
H3:19	low	5.35	C25H44O9	533.2967	533.2965	-0.4	3.29
I2:21(5,14-OH)	medium	5.80	C27H50O9	563.3437	563.3467	5.3	0.156
I2:21(6,15-OH)	medium	5.98	C27H50O9	563.3437	563.3462	4.5	0.246
I3:20(6,6,8)	medium	6.30	C26H46O9	547.3124	547.3127	0.5	1.62
I3:20(6,6,8)	medium	6.38	C26H46O9				
I2:22(6,16-OH)	medium	7.19	C28H52O9	577.3593	577.3588	-0.8	7.76
H3:21	low	7.42	C27H48O9	561.3275	561.3317	7.6	0.106

Table S2.17. Summary of annotated acylsugars detected in *S. torvum* trichome extracts. I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over eight samples. Acylsugars are sorted by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acylhexoses							
H3:17:1	low	3.37	C23H38O9	503.2498	503.2506	1.7	0.00119
I3:17:1(2,5:1,10)	medium	3.48	C23H38O9	503.2498	503.2482	-3.1	0.143
I2:16(2,14)	medium	3.48	C22H40O8	477.2705	477.2699	-1.3	0.000158
I2:17:1(5:1,12)	medium	3.82	C23H40O8	489.2705	489.2671	-7.0	0.822
I2:16(2,14)	medium	3.90	C22H40O8	477.2705	477.2679	-5.5	0.00411
H3:17	low	3.97	C23H40O9	505.2649	505.2628	-4.1	0.0400
I3:18:1(2,5:1,11)	medium	4.02	C24H40O9	517.2654	517.2625	-5.6	0.154
H2:16	low	4.07	C22H40O8	477.2705	477.2665	-8.4	0.00973
H3:18:1	low	4.20	C24H40O9	517.2600	517.2638	7.4	0.000868
I2:16(2,14)	medium	4.24	C22H40O8	477.2705	477.2689	-3.3	0.0408
I2:17(5,12)	medium	4.39	C23H42O8	491.2862	491.2851	-2.3	0.323
I2:18:1(5:1,13)	medium	4.45	C24H42O8	503.2862	503.2852	-2.0	0.0603
I3:18(2,5,11)	medium	4.62	C24H42O9	519.2811	519.2802	-1.7	0.0571
I2:18:1(5:1,13)	medium	4.65	C24H42O8	503.2862	503.2859	-0.5	0.00213
I3:19:1(2,5:1,12)	medium	5.08	C25H42O9	531.2811	531.2800	-2.2	3.43
H2:18	low	5.12	C24H44O8	505.3013	505.3010	-0.6	0.0223
I3:18(2,4,12)	medium	5.13	C24H42O9	519.2811	519.2805	-1.2	0.0377
H2:18	low	5.35	C24H44O8	505.3013	505.3037	4.8	0.000392
I3:19:1(2,5:1,12)	medium	5.48	C25H42O9	531.2811	531.2798	-2.4	0.0943
I2:19:1(5:1,14)	medium	5.70	C25H44O8	517.3018	517.3002	-3.2	7.6
I2:18(2,16)	medium	5.71	C24H44O8	505.3018	505.3004	-2.7	0.189
H2:18	low	5.75	C24H44O8	505.3018	505.2999	-3.7	0.00340
I3:19(2,5,12)	medium	5.86	C25H44O9	533.2967	533.2955	-2.3	2.44
I3:19	low	6.21	C25H44O9	533.2967	533.2963	-0.7	0.0213
I2:18(2,16)	medium	6.34	C24H44O8	505.3018	505.3012	-1.1	0.0858
I2:19(5,14)	medium	6.52	C25H46O8	519.3175	519.3173	-0.4	3.33
I2:23:3(5:1,18:2)	medium	6.89	C29H48O8	569.3331	569.3329	-0.4	0.187
I3:25:4(2,5:1,18:3)	medium	7.11	C31H48O9	609.3280	609.3270	-1.6	0.288
I3:21:1(2,5:1,14)	medium	7.37	C27H46O9	559.3124	559.3114	-1.7	16.4
I3:21:1(2,5:1,14)	medium	7.42	C27H46O9				
H3:20	low	7.85	C26H46O9	547.3124	547.3126	0.4	0.663
I3:21:1(2,5:1,14)	medium	7.89	C27H46O9	559.3124	559.3118	-1.1	1.83
I3:25:3(2,5,18:3)	medium	7.97	C31H50O9	611.3437	611.3425	-1.9	0.0341

Table S2.17. (cont'd)

I2:21:1(5:1,16)	medium	8.20	C27H48O8	545.3331	545.3311	-3.6	3.85
I2:20(4,16)	medium	8.27	C26H48O8	533.3331	533.3322	-1.7	0.109
I3:21(2,5,14)	medium	8.38	C27H48O9	561.3280	561.3269	-1.9	12.3
I3:25:3(2,5:1,18:2)	medium	8.64	C31H50O9	611.3437	611.3430	-1.1	1.54
I3:21(2,5,14)	medium	8.77	C27H48O9	561.3280	561.3268	-2.1	0.304
I2:20(2,18)	medium	9.01	C26H48O8	533.3331	533.3332	0.2	0.215
I3:25:3	low	9.03	C31H50O9	611.3437	611.3447	1.7	0.0362
H3:21	low	9.20	C27H48O9	561.3280	561.3291	1.9	0.0119
I2:21(5,16)	medium	9.23	C27H50O8	547.3488	547.3477	-2.1	0.653
H3:22	low	9.41	C28H50O9	575.3437	575.3425	-2.1	0.0247
H5:23	low	9.53	C29H48O10	601.3224	601.3216	-1.4	0.0806
I3:25:2(2,5,18:2)	medium	9.61	C31H52O9	613.3593	613.3581	-2.0	0.175
H3:22	low	9.76	C28H50O9	575.3437	575.3426	-1.9	0.0242
I3:23:1(2,5:1,16)	medium	10.06	C29H50O9	587.3437	587.3437	0.1	15.5
I3:22(2,4,16)	medium	10.23	C28H50O9	575.3437	575.3427	-1.7	0.739
I3:23:1(2,5:1,16)	medium	10.52	C29H50O9	587.3437	587.3425	-2.1	2.45
I3:25:2(2,5:1,18:1)	medium	10.52	C31H52O9	613.3593	613.3580	-2.1	1.42
I3:22(2,4,16)	medium	10.56	C28H50O9	575.3437	575.3442	0.8	0.0256
I2:23:1(5:1,18)	medium	10.77	C29H52O8	573.3644	573.3639	-0.8	1.37
I3:25:2(2,5:1,18:1)	medium	10.89	C31H52O9	613.3593	613.3545	-7.9	0.0757
I3:24:1(2,4,18:1)	medium	10.92	C30H52O9	601.3593	601.3604	1.7	0.00330
I3:24:1(2,5:1,17)	medium	11.04	C30H52O9	601.3593	601.3588	-0.9	0.0669
I3:23(2,5,16)	medium	11.11	C29H52O9	589.3593	589.3588	-0.9	5.67
H4:26	low	11.27	C32H56O10	645.3850	645.3848	-0.3	0.0191
I3:24:1(2,5:1,17)	medium	11.37	C30H52O9	601.3593	601.3594	0.2	0.215
I3:23(2,5,16)	medium	11.44	C29H52O9	589.3593	589.3579	-2.4	0.311
H3:25:1	low	11.54	C31H54O9	615.3744	615.3765	3.3	0.190
H3:23	low	11.67	C29H52O9	589.3593	589.3556	-6.2	0.0136
I3:24:1(2,5:1,17)	medium	11.79	C30H52O9	601.3593	601.3599	1.0	0.0212
I4:25:1	low	12.00	C32H56O9	629.3542	629.3531	-1.7	0.539
I2:23(5,18)	medium	12.00	C29H54O8	575.3801	575.3794	-1.2	0.242
H3:24	low	12.05	C30H54O9	603.3744	603.3755	1.8	0.00517
H3:25:1	low	12.12	C31H54O9	615.3744	615.3695	-8.0	0.0282
H4:25	low	12.18	C31H54O10	631.3694	631.3722	4.5	0.00570
H3:25:1	low	12.29	C31H54O9	615.3744	615.3746	0.3	0.0483
I3:24(2,4,18)	medium	12.40	C30H54O9	603.3750	603.3746	-0.7	0.0115
I3:25:1(2,5:1,18)	medium	12.60	C31H54O9	615.3750	615.3743	-1.1	8.35
I3:24(2,4,18)	medium	12.81	C30H54O9	603.3750	603.3758	1.4	0.626
H4:25	low	12.83	C31H54O10	631.3694	631.3695	0.3	0.0244
I3:25:1(2,5:1,18)	medium	12.98	C31H54O9	615.3750	615.3739	-1.7	2.56
I3:25(2,5,18)	medium	13.56	C31H56O9	617.3906	617.3911	0.8	1.74
I3:25(2,5,18)	medium	13.78	C31H56O9	617.3906	617.3913	1.1	0.0999

Table S2.18. Summary of annotated acylsugars detected in *S. scabrum* trichome extracts. H = hexose; G = glucose; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over six samples. Acylsugars are sorted by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm (ppm)	Average acylsugar abundance (%)
Acylhexoses							
I3:18:OH(4,4,10-OH)	medium	2.86	C24H42O10	535.2760	535.2780	3.8	0.0251
G2:13(4,9)	medium	2.86	C19H34O8	435.2236	435.2248	2.8	0.0166
I3:18:OH(4,4,10-OH)	medium	2.94	C24H42O10	535.2760	535.2768	1.5	0.0712
G2:13(4,9)	medium	2.94	C19H34O8	435.2236	435.2235	-0.3	0.116
I3:18:OH(4,4,10-OH)	medium	3.02	C24H42O10	535.2760	535.2760	0.0	0.168
G2:13(4,9)	medium	3.04	C19H34O8	435.2236	435.2233	-0.6	0.192
H2:14	low	3.04	C20H36O8	449.2392	449.2421	6.4	0.00312
I3:16(4,4,8)	medium	3.06	C22H38O9	491.2498	491.2497	-0.1	0.150
G2:13(4,9)	medium	3.14	C19H34O8	435.2236	435.2233	-0.6	0.0890
I3:16(4,4,8)	medium	3.22	C22H38O9	491.2498	491.2497	-0.1	3.58
H2:13	low	3.24	C19H34O8	435.2236	435.2237	0.2	0.00505
H4:18:[O1]	low	3.27	C25H42O11	563.2704	563.2726	4.0	0.0772
I3:16(4,4,8)	medium	3.31	C22H38O9	491.2498	491.2498	-0.1	7.44
G2:14(4,10)	medium	3.46	C20H36O8	449.2392	449.2388	-1.0	1.37
H3:13	low	3.49	C20H34O9	463.2179	463.2217	8.1	0.0219
G2:14(4,10)	medium	3.61	C20H36O8	449.2392	449.2389	-0.7	3.20
I3:17(4,5,8)	medium	3.69	C23H40O9	505.2654	505.2652	-0.3	3.11
G2:14(4,10)	medium	3.76	C20H36O8	449.2392	449.2391	-0.2	1.14
I3:17(4,5,8)	medium	3.77	C23H40O9	505.2654	505.2658	0.8	11.5
I3:17(4,4,9)	medium	3.94	C23H40O9	505.2654	505.2652	-0.5	2.95
G2:15(5,10)	medium	3.99	C21H38O8	463.2549	463.2545	-0.9	1.52
G2:15(5,10)	medium	4.18	C21H38O8	463.2549	463.2544	-1.1	4.32
I3:18(4,4,10)	medium	4.37	C24H42O9	519.2811	519.2806	-0.9	3.33
G2:15(5,10)	medium	4.38	C21H38O8	463.2549	463.2543	-1.2	2.54
I3:18(4,4,10)	medium	4.57	C24H42O9	519.2811	519.2812	0.1	9.13
G3:16(2,4,10)	medium	4.70	C22H38O9	491.2498	491.2510	2.5	0.0287
I3:18(4,4,10)	medium	4.73	C24H42O9	519.2811	519.2804	-1.3	15.8
G3:16(2,4,10)	medium	4.87	C22H38O9	491.2498	491.2492	-1.1	0.562
H3:21:[O1]	low	4.90	C27H48O10	577.3229	577.3237	1.3	0.143
I3:19(4,5,10)	medium	5.07	C25H44O9	533.2967	533.2967	-0.1	0.221
G3:16(2,4,10)	medium	5.08	C22H38O9	491.2498	491.2494	-0.8	0.464

H4:19	low	5.10	C25H42O10	547.2760	547.2757	-0.5	1.97
H4:19	low	5.27	C25H42O10	547.2760	547.2760	0.0	3.27
I3:19(4,5,10)	medium	5.27	C25H44O9	533.2967	533.2969	0.4	8.48
I3:19(4,5,10)	medium	5.47	C25H44O9	533.2967	533.2967	0.0	18.0
H4:19	low	5.48	C25H42O10	547.2760	547.2763	0.5	0.270
G3:17(2,5,10)	medium	5.48	C23H40O9	505.2654	505.2687	6.6	0.0125
G3:17(2,5,10)	medium	5.68	C23H40O9	505.2654	505.2654	-0.1	0.378
G3:17(2,5,10)	medium	5.89	C23H40O9	505.2654	505.2651	-0.7	0.473
I4:20	low	6.01	C26H44O10	561.2916	561.2918	0.4	0.276
I4:20	low	6.08	C26H44O10	561.2916	561.2919	0.5	0.227
I3:20(5,5,10)	medium	6.11	C26H46O9	547.3124	547.3123	-0.2	0.484
I3:20(5,5,10)	medium	6.34	C26H46O9	547.3124	547.3120	-0.7	1.72
I4:20(2,4,4,10)	medium	6.34	C26H44O10	561.2916	561.2913	-0.5	1.47
I3:20(5,5,10)	medium	6.56	C26H46O9	547.3124	547.3123	-0.1	0.330
I4:20(2,4,4,10)	medium	6.58	C26H44O10	561.2916	561.2916	0.0	3.45
G2:18(8,10)	medium	6.67	C24H44O8	505.3018	505.3030	2.3	0.0517
H3:20	low	6.91	C26H46O9	547.3124	547.3124	0.0	0.108
G2:18(8,10)	medium	6.94	C24H44O8	505.3018	505.3018	0.1	0.153
I4:21(2,4,5,10)	medium	7.20	C27H46O10	575.3073	575.3072	-0.1	1.76
G2:18(8,10)	medium	7.26	C24H44O8	505.3018	505.3023	0.9	0.102
I4:21(2,4,5,10)	medium	7.46	C27H46O10	575.3073	575.3074	0.2	5.40
H4:22	low	8.20	C28H48O10	589.3229	589.3247	3.0	0.0549
H4:22	low	8.45	C28H48O10	589.3229	589.3237	1.4	0.222
H4:22	low	8.68	C28H48O10	589.3229	589.3274	7.7	0.0189
H4:22	low	8.72	C28H48O10	589.3229			
H4:22	low	9.06	C28H48O10	589.3229	589.3277	8.1	0.00728

Table S2.19. Summary of annotated acylsugars detected in *S. villosum* trichome extracts.
HH = hexose-hexose; G = glucose; I = inositol. Annotation method and confidence level criteria
are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic
formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass
measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and
was averaged over six samples. Acylsugars are sorted by number of sugar moieties and then by
elution order.

Name	Confidence level	RT (min)	Chemical formula	m/z acc	m/z exp	Δm(ppm)	Average acylsugar abundance (%)
Acyldisaccharides							
HH4:22(2,4,4,12-OH)	medium	3.61	C34H58O16	767.3707	767.3722	2.0	3.07
HH3:22:[O1]	low	3.99	C34H60O15	753.3914	753.3972	7.7	0.00373
HH4:23(2,4,5,12-OH)	medium	4.12	C35H60O16	781.3863	781.3881	2.3	3.04
HH4:23(2,4,5,12-OH)	medium	4.19	C35H60O16				
HH3:22(4,4,14-OH)	medium	4.35	C34H60O15	753.3914	753.3937	3.1	0.0268
HH4:24(2,5,5,12-OH)	medium	4.82	C36H62O16	795.4020	795.4037	2.2	0.907
HH4:24(2,4,4,14-OH)	medium	5.01	C36H62O16	795.4020	795.4045	3.1	0.0509
HH4:24(2,4,4,14-OH)	medium	5.21	C36H62O16	795.4020	795.4036	2.0	3.25
HH4:25(2,4,5,14-OH)	medium	5.98	C37H64O16	809.4176	809.4196	2.5	2.07
HH4:25(2,4,5,14-OH)	medium	6.06	C37H64O16	809.4176	809.4196	2.4	0.900
HH4:26(2,5,5,14-OH)	medium	6.93	C38H66O16	823.4333	823.4355	2.7	0.690
Acylhexoses							
G2:14(4,10)	medium	3.46	C20H36O8	449.2392	449.2384	-1.7	9.06
G2:14(4,10)	medium	3.61	C20H36O8	449.2392	449.2385	-1.6	13.4
G2:14(4,10)	medium	3.76	C20H36O8	449.2392	449.2383	-2.1	2.70
G2:15(5,10)	medium	3.97	C21H38O8	463.2549	463.2542	-1.6	5.10
G2:15(5,10)	medium	4.19	C21H38O8	463.2549	463.2541	-1.7	8.85
I3:20(4,4,12-OH)	medium	4.24	C26H46O10	563.3073	563.3073	0.0	1.91
G2:15(5,10)	medium	4.37	C21H38O8	463.2549	463.2542	-1.5	1.61
G3:16(2,4,10)	medium	4.68	C22H38O9	491.2498	491.2493	-1.1	0.484
G3:16(2,4,10)	medium	4.85	C22H38O9	491.2498	491.2490	-1.6	4.68
I3:21(4,5,12-OH)	medium	4.90	C27H48O10	577.3229	577.3226	-0.5	4.62
G3:16(2,4,10)	medium	5.05	C22H38O9	491.2498	491.2492	-1.2	1.13
G3:17(2,5,10)	medium	5.45	C23H40O9	505.2654	505.2657	0.7	0.0569
G3:17(2,5,10)	medium	5.63	C23H40O9	505.2654	505.2649	-1.0	2.25
I3:22(5,5,12-OH)	medium	5.68	C28H50O10	591.3386	591.3388	0.4	0.659
G3:17(2,5,10)	medium	5.86	C23H40O9	505.2654	505.2659	0.9	0.315
I3:22(4,4,14-OH)	medium	5.95	C28H50O10	591.3386	591.3394	1.4	0.145
I3:22(4,4,14-OH)	medium	6.21	C28H50O10	591.3386	591.3387	0.1	12.8
I3:23(4,5,14-OH)	medium	6.79	C29H52O10	605.3542	605.3556	2.3	0.0467
I3:23(4,5,14-OH)	medium	7.09	C29H52O10	605.3542	605.3542	0.0	14.0

Table S2.19. (cont'd)

I3:24(5,5,14-OH)	medium	8.04	C30H54O10	619.3699	619.3701	0.2	2.06
I3:24(4,4,16-OH)	medium	8.70	C30H54O10	619.3699	619.3708	1.4	0.0296

Table S2.20. Summary of annotated acylsugars detected in *S. mammosum* trichome extracts. H = hexose; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over six samples.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm (ppm)	Average acylsugar abundance (%)
Acylhexoses							
I3:16(4,6,6-OH)	medium	2.28	C22H38O10	507.2447	507.2439	-1.6	0.0311
I3:16(4,6,6-OH)	medium	2.36	C22H38O10	507.2447	507.2439	-1.6	0.159
H3:17:[O2]	low	2.45	C23H40O10	521.2598	521.2613	2.8	0.00197
I3:14(4,4,6)	medium	2.46	C20H34O9	463.2185	463.2170	-3.2	0.912
I3:14(4,4,6)	medium	2.53	C20H34O9				
H3:17:[O2]	low	2.59	C23H40O10	521.2598	521.2613	2.9	0.00722
H3:18:[O1]	low	2.66	C24H42O10	535.2760	535.2746	-2.6	0.0659
I3:15(4,5,6)	medium	2.71	C21H36O9	477.2341	477.2326	-3.2	0.211
I3:18(6,6,6-OH)	medium	2.87	C24H42O10	535.2760	535.2740	-3.7	1.27
I3:18(6,6,6-OH)	medium	2.96	C24H42O10				
I3:16(4,6,6)	medium	3.08	C22H38O9	491.2498	491.2476	-4.5	6.93
I3:16(4,6,6)	medium	3.13	C22H38O9				
H2:18:[O2]	low	3.13	C24H44O10	537.2911	537.2908	-0.5	0.00314
I3:16(4,6,6)	medium	3.22	C22H38O9	491.2498	491.2474	-4.9	1.42
H2:18:[O2]	low	3.29	C24H44O10	537.2911	537.2893	-3.4	0.0610
I3:16(4,6,6)	medium	3.33	C22H38O9	491.2498	491.2472	-5.3	0.317
H2:18:[O2]	low	3.42	C24H44O10	537.2911	537.2912	0.2	0.0155
H3:16	low	3.44	C22H38O9	491.2498	491.2482	-3.3	0.0866
I3:17(5,6,6)	medium	3.56	C23H40O9	505.2654	505.2647	-1.5	0.951
H2:18:[O2]	low	3.56	C24H44O10	537.2911	537.2920	1.7	0.0182
H3:20:[O1]	low	3.59	C26H46O10	563.3068	563.3092	4.3	0.00402
I3:17(5,6,6)	medium	3.66	C23H40O9	505.2654	505.2638	-3.1	0.394
H3:20:[O1]	low	3.66	C26H46O10	563.3068	563.3077	1.7	0.0312
H3:20:[O1]	low	3.72	C26H46O10				
H2:18:[O2]	low	3.66	C24H44O10	537.2911	537.2910	-0.1	0.0171
H3:17	low	3.77	C23H40O9	505.2654	505.2646	-1.5	0.0270
H2:18:[O2]	low	3.82	C24H44O10	537.2911	537.2911	-0.1	0.0844
H3:20:[O1]	low	3.85	C26H46O10	563.3068	563.3071	0.6	0.0140
H3:20:[O1]	low	3.96	C26H46O10	563.3068	563.3063	-0.8	0.0508
H2:18:[O2]	low	3.99	C24H44O10	537.2911	537.2904	-1.3	0.313
H3:20:[O1]	low	4.07	C26H46O10	563.3068	563.3055	-2.2	0.112
I3:18(6,6,6)	medium	4.09	C24H42O9	519.2811	519.2794	-3.2	7.78

Table S2.20. (cont'd)

H1:18:[O2]	low	4.12	C24H46O9	523.3118	523.3114	-0.8	0.0205
H3:20:[O1]	low	4.19	C26H46O10	563.3068	563.3067	-0.1	0.0287
I3:18(6,6,6)	medium	4.20	C24H42O9	519.2811	519.2800	-2.0	4.33
I3:18(6,6,6)	medium	4.35	C24H42O9	519.2811	519.2797	-2.6	2.33
I3:18(6,6,6)	medium	4.43	C24H42O9				
H1:18:[O2]	low	4.45	C24H46O9	523.3118	523.3113	-1.0	0.720
I3:18(6,6,6)	medium	4.55	C24H42O9	519.2811	519.2804	-1.3	0.943
I3:18(6,6,6)	medium	4.73	C24H42O9	519.2811	519.2797	-2.7	0.134
H3:19	low	4.79	C25H44O9	533.2962	533.2951	-2.0	0.0879
H2:20:[O2]	low	4.82	C26H48O10	565.3224	565.3214	-1.8	0.137
H3:24:[O2]	low	4.85	C30H54O11	635.3648	635.3685	5.9	0.00256
H3:19	low	4.93	C25H44O9	533.2962	533.2956	-1.1	0.0279
H3:24:[O2]	low	5.01	C30H54O11	635.3648	635.3679	4.9	0.00194
H3:19	low	5.03	C25H44O9	533.2962	533.2954	-1.5	0.0745
H3:19	low	5.10	C25H44O9				
H2:20:[O2]	low	5.05	C26H48O10	565.3224	565.3229	0.9	0.0218
H3:24:[O2]	low	5.20	C30H54O11	635.3648	635.3667	3.1	0.00373
H3:19	low	5.21	C25H44O9	533.2962	533.2957	-0.8	0.0537
H2:20:[O2]	low	5.23	C26H48O10	565.3224	565.3225	0.2	0.0932
H3:24:[O2]	low	5.28	C30H54O11	635.3648	635.3659	1.7	0.00490
H3:19	low	5.32	C25H44O9	533.2962	533.2962	0.1	0.0200
I3:24(4,6-OH,14-OH)	medium	5.41	C30H54O11	635.3648	635.3652	0.7	0.0180
H3:19	low	5.43	C25H44O9	533.2962	533.2973	2.0	0.00827
I3:24(4,6-OH,14-OH)	medium	5.55	C30H54O11	635.3648	635.3635	-2.1	0.0580
I3:20(6,6,8)	medium	5.69	C26H46O9	547.3124	547.3108	-2.9	1.02
H2:20:[O2]	low	5.71	C26H48O10	565.3224	565.3219	-0.9	0.317
I3:22(4,6,12-OH)	medium	5.73	C28H50O10	591.3386	591.3375	-1.9	0.438
I3:20(6,6,8)	medium	5.78	C26H46O9	547.3124	547.3109	-2.7	1.02
I3:20(6,6,8)	medium	5.84	C26H46O9				
I3:22(4,6,12-OH)	medium	5.90	C28H50O10	591.3386	591.3369	-2.8	0.204
I3:20(6,6,8)	medium	5.93	C26H46O9	547.3124	547.3105	-3.5	1.69
H2:20:[O2]	low	5.93	C26H48O10	565.3224	565.3212	-2.1	0.732
I3:20(6,6,8)	medium	6.03	C26H46O9	547.3124	547.3107	-3.1	1.31
I3:22(4,6,12-OH)	medium	6.10	C28H50O10	591.3386	591.3382	-0.7	0.0119
I3:20(6,6,8)	medium	6.13	C26H46O9	547.3124	547.3105	-3.5	1.03
H3:23:[O2]	low	6.31	C29H52O11	621.3122	621.3100	-3.6	0.0248
H3:20	low	6.52	C26H46O9	547.3124	547.3119	-1.0	0.0711
H3:23:[O2]	low	6.54	C29H52O11	621.3122	621.3065	-9.2	0.402
H3:20	low	6.74	C26H46O9	547.3124	547.3124	0.1	0.0123
H3:26:[O2]	low	6.74	C32H58O11	663.3961	663.3985	3.6	0.00455
I3:26(6,6-OH,14-OH)	medium	6.98	C32H58O11	663.3961	663.3953	-1.2	0.0455
H3:24:[O1]	low	7.01	C30H54O10	619.3699	619.3707	1.3	0.00225
H3:24:[O1]	low	7.14	C30H54O10	619.3699	619.3723	3.9	0.00812

Table S2.20. (cont'd)

I3:26(6,6-OH,14-OH)	medium	7.22	C32H58O11	663.3961	663.3951	-1.5	0.240
I3:26(6,6-OH,14-OH)	medium	7.39	C32H58O11	663.3961	663.3954	-1.1	0.294
I3:26(6,6-OH,14-OH)	medium	7.45	C32H58O11				
I3:24(6,6,12-OH)	medium	7.41	C30H54O10	619.3699	619.3687	-2.0	0.728
I3:26(6,6-OH,14-OH)	medium	7.62	C32H58O11	663.3961	663.3966	0.7	0.00643
I3:24(6,6,12-OH)	medium	7.67	C30H54O10	619.3699	619.3687	-1.9	0.943
I3:24(4,6,14-OH)	medium	7.70	C30H54O10	619.3699	619.3695	-0.6	0.787
I3:24(6,6,12-OH)	medium	7.77	C30H54O10				
I3:24(4,6,14-OH)	medium	7.94	C30H54O10	619.3699	619.3688	-1.7	2.84
I3:24(4,6,14-OH)	medium	8.00	C30H54O10				
I3:24(4,6,14-OH)	medium	8.25	C30H54O10	619.3699	619.3694	-0.9	8.80
I3:24(4,6,14-OH)	medium	8.45	C30H54O10	619.3699	619.3689	-1.7	1.10
H4:24:[O2]	low	8.53	C30H52O12	649.3435	649.3463	4.3	0.0131
H4:24:[O2]	low	8.68	C30H52O12	649.3435	649.3442	1.0	0.00273
I3:25(5,6,14-OH)	medium	8.91	C31H56O10	633.3855	633.3862	1.1	0.0998
H4:24:[O2]	low	9.01	C30H52O12	649.3435	649.3422	-2.0	0.601
I3:25(5,6,14-OH)	medium	9.12	C31H56O10	633.3855	633.3844	-1.7	0.514
I3:25(5,6,14-OH)	medium	9.23	C31H56O10				
I3:25(5,6,14-OH)	medium	9.43	C31H56O10	633.3855	633.3841	-2.2	0.0939
I3:26(6,6,14-OH)	medium	9.63	C32H58O10	647.4012	647.3994	-2.7	0.643
I3:26(6,6,14-OH)	medium	9.95	C32H58O10	647.4012	647.4005	-1.1	7.29
I3:26(6,6,14-OH)	medium	10.16	C32H58O10	647.4012	647.4003	-1.3	4.08
I3:26(6,6,14-OH)	medium	10.26	C32H58O10	647.4012	647.4011	-0.2	14.9
H3:24	low	10.30	C30H54O9	603.3744	603.3759	2.4	0.00245
I3:26(6,6,14-OH)	medium	10.46	C32H58O10	647.4012	647.4011	-0.1	8.24
H3:24	low	10.49	C30H54O9	603.3744	603.3742	-0.3	0.0595
I3:26(6,6,14-OH)	medium	10.67	C32H58O10	647.4012	647.3997	-2.3	1.36
H3:24	low	10.79	C30H54O9	603.3744	603.3730	-2.3	0.199
H3:27:[O1]	low	10.84	C33H60O10	661.4163	661.4147	-2.5	0.00106
H3:27:[O1]	low	11.00	C33H60O10	661.4163	661.4165	0.3	0.0338
I3:26(6,6,14-OH)	medium	11.02	C32H58O10	647.4012	647.4004	-1.2	0.335
H3:24	low	11.02	C30H54O9	603.3744	603.3742	-0.4	0.105
H3:27:[O1]	low	11.13	C33H60O10	661.4163	661.4175	1.7	0.0277
H3:24	low	11.22	C30H54O9	603.3744	603.3753	1.4	0.0311
H3:26:[O1]	low	11.24	C32H58O10	647.4012	647.4026	2.2	0.0177
H3:27:[O1]	low	11.32	C33H60O10	661.4163	661.4163	0.0	0.118
H3:27:[O1]	low	11.52	C33H60O10	661.4163	661.4172	1.3	0.0351
H3:24	low	11.54	C30H54O9	603.3744	603.3751	1.1	0.0286
H3:27:[O1]	low	11.70	C33H60O10	661.4163	661.4183	3.0	0.00867
H3:24	low	11.79	C30H54O9	603.3744	603.3791	7.7	0.00366
I3:28(6,8,14-OH)	medium	11.90	C34H62O10	675.4325	675.4324	-0.1	0.0360
H3:27:[O1]	low	11.98	C33H60O10	661.4163	661.4178	2.3	0.0124
I3:28(6,8,14-OH)	medium	12.18	C34H62O10	675.4325	675.4319	-0.9	0.497

194

Table S2.20. (cont'd)

H3:27:[O1]	low	12.20	C33H60O10	661.4163	661.4161	-0.2	0.00133
I3:28(6,8,14-OH)	medium	12.37	C34H62O10	675.4325	675.4318	-1.0	0.622
I3:28(6,8,14-OH)	medium	12.48	C34H62O10	675.4325	675.4313	-1.8	1.45
I3:28(6,8,14-OH)	medium	12.66	C34H62O10	675.4325	675.4315	-1.5	1.69
H3:26	low	12.73	C32H58O9	631.4057	631.4097	6.3	0.00272
I3:28(6,8,14-OH)	medium	12.77	C34H62O10	675.4325	675.4317	-1.1	1.31
I3:28(6,6,16-OH)	medium	12.78	C34H62O10				
I3:28(6,8,14-OH)	medium	12.86	C34H62O10	675.4325	675.4314	-1.7	0.275
H3:26	low	12.93	C32H58O9	631.4057	631.4078	3.3	0.0237
I3:28(6,6,16-OH)	medium	12.94	C34H62O10	675.4325	675.4319	-0.9	0.330
I3:28(6,8,14-OH)	medium	12.95	C34H62O10				
H3:29:[O1]	low	13.10	C35H64O10	689.4476	689.4518	6.1	0.000248
H3:26	low	13.16	C32H58O9	631.4057	631.4074	2.7	0.0283
H3:26	low	13.25	C32H58O9				
H3:29:[O1]	low	13.21	C35H64O10	689.4476	689.4431	-6.5	0.000322
H3:29:[O1]	low	13.41	C35H64O10	689.4476	689.4500	3.5	0.00342
H3:26	low	13.43	C32H58O9	631.4057	631.4091	5.3	0.00984
H3:29:[O1]	low	13.64	C35H64O10	689.4476	689.4484	1.2	0.00345
H3:26	low	13.64	C32H58O9	631.4057	631.4058	0.1	0.00120
H3:26	low	13.95	C32H58O9	631.4057	631.4076	2.9	0.00200
H3:30:[O1]	low	14.11	C36H66O10	703.4633	703.4660	3.9	0.0126
H3:30:[O1]	low	14.39	C36H66O10	703.4633	703.4641	1.2	0.0772
H3:30:[O1]	low	14.59	C36H66O10	703.4633	703.4639	0.9	0.174
H3:30:[O1]	low	14.82	C36H66O10	703.4633	703.4640	1.1	0.113

Table S2.21. Summary of annotated acylsugars detected in *S. abutiloides* trichome extracts. HH = hexose-hexose; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over six samples. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Hexose	Hexose			Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acyldisaccharides											
HH4:22(2,4,4,12-OH)	4	4	12-OH	2	medium	3.55	C34H58O16	767.3707	767.3713	0.7	4.43
HH3:22:[O1]					low	3.99	C34H60O15	753.3909	753.3934	3.3	0.0131
HH4:23(2,4,5,12-OH)	4	5	12-OH	2	medium	4.12	C35H60O16	781.3863	781.3873	1.3	7.24
HH4:23(2,4,5,12-OH)	4	5	12-OH	2	medium	4.18	C35H60O16				
HH3:22:[O1]					low	4.35	C34H60O15	753.3909	753.3923	1.8	0.100
HH4:24(2,5,5,12-OH)	5	5	12-OH	2	medium	4.85	C36H62O16	795.4020	795.4035	1.9	2.86
HH3:23:[O1]					low	4.99	C35H62O15	767.4065	767.4084	2.4	0.233
HH4:24(2,4,4,14-OH)	4	4	14-OH	2	medium	5.01	C36H62O16	795.4020	795.4037	2.1	0.122
HH3:23:[O1]					low	5.08	C35H62O15	767.4065	767.4085	2.6	0.110
HH4:24(2,4,4,14-OH)	4	4	14-OH	2	medium	5.23	C36H62O16	795.4020	795.4038	2.2	7.59
HH4:22(2,4,4,12)	4	4	12	2	medium	5.35	C34H58O15	751.3782	751.3772	-1.3	0.0624
HH4:22(2,4,4,12)	4	4	12	2	medium	5.75	C34H58O15				
HH4:25(2,4,5,14-OH)	4	5	14-OH	2	medium	5.75	C37H64O16	809.4176	809.4194	2.2	0.0615
HH3:24(5,5,14-OH)	5	5	14-OH		medium	5.82	C36H64O15	781.4227	781.4238	1.5	0.166
HH4:25(2,4,5,14-OH)	4	5	14-OH	2	medium	5.84	C37H64O16	809.4176	809.4195	2.3	0.0312
HH4:25(2,4,5,14-OH)	4	5	14-OH	2	medium	6.03	C37H64O16	809.4176	809.4190	1.7	7.48
HH4:25(2,4,5,14-OH)	4	5	14-OH	2	medium	6.10	C37H64O16	809.4176	809.4189	1.6	3.94
HH4:23					low	6.59	C35H60O15	765.3909	765.3928	2.5	0.0558
HH4:26(2,5,5,14-OH)	5	5	14-OH	2	medium	6.67	C38H66O16	823.4333	823.4359	3.2	0.00791
HH4:23					low	6.73	C35H60O15	765.3909	765.3935	3.5	0.0148
HH3:22					low	6.86	C34H60O14	737.3960	737.3980	2.7	0.0124

Table S2.21. (cont'd)

HH4:26(2,5,5,14-OH)	5	5	14-OH	2	medium	6.98	C38H66O16	823.4333	823.4347	1.7	4.10
HH4:24(2,5,5,12)	5	5	12	2	medium	7.58	C36H62O15	779.4071	779.4096	3.1	0.0122
HH4:24(2,5,5,12)	5	5	12	2	medium	7.67	C36H62O15				
HH3:23					low	7.76	C35H62O14	751.4116	751.4138	2.9	0.0299
HH4:24(2,4,4,14)	4	4	14	2	medium	7.89	C36H62O15	779.4071	779.4098	3.5	0.00798
HH3:23					low	7.92	C35H62O14	751.4116	751.4151	4.6	0.00580
HH4:24(2,4,4,14)	4	4	14	2	medium	8.22	C36H62O15	779.4071	779.4084	1.7	0.392
HH4:25					low	8.87	C37H64O15	793.4222	793.4263	5.2	0.00312
HH3:24					low	8.88	C36H64O14	765.4273	765.4299	3.5	0.0115
HH4:25					low	9.00	C37H64O15	793.4222	793.4275	6.7	0.00124
HH4:25					low	9.20	C37H64O15	793.4222	793.4240	2.3	0.329
HH4:25					low	9.33	C37H64O15	793.4222	793.4242	2.5	0.0986
HH3:24					low	9.55	C36H64O14	765.4273	765.4292	2.5	0.0408
HH4:26(2,5,5,14)	5	5	14	2	medium	10.32	C38H66O15	807.4384	807.4399	1.8	0.0982
HH3:25					low	10.51	C37H66O14	779.4429	779.4446	2.2	0.124
HH3:25					low	10.66	C37H66O14	779.4429	779.4455	3.4	0.0186
HH4:26(2,4,4,16)					medium	10.91	C38H66O15	807.4384	807.4393	1.2	0.487
HH3:26					low	11.62	C38H68O14	793.4586	793.4610	3.1	0.0592
HH4:27(2,4,5,16)					medium	11.89	C39H68O15	821.4540	821.4553	1.6	0.838
HH4:28(2,5,5,16)					medium	12.88	C40H70O15	835.4697	835.4715	2.2	0.166
Acylhexoses											
H3:20:[O1]					low	4.09	C26H46O10	563.3082	563.3087	0.8	0.00549
I3:20(4,4,12-OH)					medium	4.20	C26H46O10	563.3082	563.3070	-2.2	2.33
H3:21:[O1]					low	4.73	C27H48O10	577.3229	577.3254	4.4	0.00462
I3:21(4,5,12-OH)					medium	4.93	C27H48O10	577.3229	577.3229	0.0	5.10
H3:21:[O1]					low	5.20	C27H48O10	577.3229	577.3249	3.5	0.00201
I3:21(3,4,14-OH)					medium	5.47	C27H48O10	577.3229	577.3232	0.4	0.0295
I3:22(5,5,12-OH)					medium	5.71	C28H50O10	591.3386	591.3383	-0.5	1.68
I3:22(4,4,14-OH)					medium	5.96	C28H50O10	591.3386	591.3386	0.0	0.398
I3:22(4,4,14-OH)					medium	6.24	C28H50O10	591.3386	591.3382	-0.7	15.2

197

Table S2.21. (cont'd)

I3:23(4,5,14-OH)		medium	6.87	C29H52O10	605.3542	605.3541	-0.2	0.183
H3:20		low	6.96	C26H46O9	547.3118	547.3125	1.2	0.0286
I3:23(4,5,14-OH)		medium	7.17	C29H52O10	605.3542	605.3538	-0.7	20.8
I3:24(5,5,14-OH)		medium	7.85	C30H54O10	619.3699	619.3713	2.3	0.00615
H3:21		low	7.90	C27H48O9	561.3275	561.3288	2.3	0.0183
I3:24(5,5,14-OH)		medium	8.14	C30H54O10	619.3699	619.3703	0.7	7.98
I3:24(4,4,16-OH)		medium	8.83	C30H54O10	619.3699	619.3700	0.1	0.182
I3:22(4,4,14)		medium	8.97	C28H50O9	575.3437	575.3448	1.9	0.00372
I3:22(4,4,14)		medium	9.28	C28H50O9	575.3437	575.3452	2.6	0.00499
I3:22(4,4,14)		medium	9.63	C28H50O9	575.3437	575.3433	-0.7	0.369
H3:25:[O1]		low	9.83	C31H56O10	633.3850	633.3855	0.8	0.297
I3:23(4,5,14)		medium	10.19	C29H52O9	589.3593	589.3603	1.6	0.00181
I3:23(4,5,14)		medium	10.30	C29H52O9	589.3593	589.3607	2.4	0.00319
I3:23(4,5,14)		medium	10.55	C29H52O9	589.3593	589.3590	-0.5	0.789
I3:23(4,5,14)		medium	10.63	C29H52O9				
H3:26:[O1]		low	10.82	C32H58O10	647.4007	647.4018	1.8	0.0446
H3:24		low	11.32	C30H54O9	603.3750	603.3750	0.0	0.000515
I3:24(5,5,14)		medium	11.62	C30H54O9	603.3750	603.3750	0.0	0.201
H3:24		low	11.96	C30H54O9	603.3750	603.3782	5.4	0.000502
I3:24(4,4,16)		medium	12.29	C30H54O9	603.3750	603.3746	-0.7	0.832
H3:25		low	12.86	C31H56O9	617.3906	617.3928	3.6	0.000501
I3:25(4,5,16)		medium	13.08	C31H56O9	617.3906	617.3905	-0.2	1.73
I3:25(4,5,16)		medium	13.18	C31H56O9				
I3:26(5,5,16)		medium	14.01	C32H58O9	631.4063	631.4061	-0.3	0.422

Table S2.22. Summary of annotated acylsugars detected in *S. laciniatum* trichome extracts.
H = hexose; G = glucose; I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over two samples. Acylsugars are sorted by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acylhexoses							
I2:12(6,6)	medium	2.15	C18H32O8	421.2079	421.2066	-3.1	0.275
I2:12(6,6)	medium	2.28	C18H32O8	421.2079	421.2055	-5.6	3.48
G2:12(6,6)	medium	2.56	C18H32O8	421.2079	421.2061	-4.2	3.33
G2:12(6,6)	medium	2.63	C18H32O8	421.2079	421.2056	-5.5	1.07
I2:14(6,8)	medium	2.81	C20H36O8	449.2392	449.2364	-6.2	6.41
I2:14(6,8)	medium	2.87	C20H36O8	449.2392	449.2371	-4.6	10.3
G2:14(6,8)	medium	3.33	C20H36O8	449.2392	449.2371	-4.6	6.86
G2:14(6,8)	medium	3.42	C20H36O8	449.2392	449.2372	-4.4	8.35
G2:14(6,8)	medium	3.54	C20H36O8	449.2392	449.2377	-3.3	0.471
I2:16(8,8)	medium	3.70	C22H40O8	477.2705	477.2695	-2.0	0.423
I2:16(8,8)	medium	3.82	C22H40O8	477.2705	477.2686	-4.1	1.96
I2:16(8,8)	medium	3.96	C22H40O8	477.2705	477.2693	-2.5	1.26
I3:18(6,6,6)	medium	4.20	C24H42O9	519.2811	519.2806	-1.0	0.531
I3:18(6,6,6)	medium	4.32	C24H42O9	519.2811	519.2797	-2.6	8.23
I3:18(6,6,6)	medium	4.49	C24H42O9	519.2811	519.2793	-3.4	20.5
G2:16(8,8)	medium	4.53	C22H40O8	477.2705	477.2731	5.5	0.0945
G2:16(8,8)	medium	4.70	C22H40O8	477.2705	477.2711	1.3	0.505
I3:18(6,6,6)	medium	4.85	C24H42O9	519.2811	519.2825	2.8	0.327
G2:16(8,8)	medium	4.87	C22H40O8	477.2705	477.2730	5.2	0.327
I3:18(6,6,6)	medium	4.93	C24H42O9	519.2811	519.2804	-1.3	0.226
G3:18(6,6,6)	medium	5.65	C24H42O9	519.2811	519.2816	0.9	0.519
G3:18(6,6,6)	medium	5.82	C24H42O9	519.2811	519.2794	-3.2	1.96
I3:20(6,6,8)	medium	5.95	C26H46O9	547.3124	547.3114	-1.9	0.498
G3:18(6,6,6)	medium	6.11	C24H42O9	519.2811	519.2817	1.2	0.951
I3:20(6,6,8)	medium	6.18	C26H46O9	547.3124	547.3120	-0.8	1.23
H3:22:OH	low	6.24	C28H50O10	591.3386	591.3395	1.5	NA
G3:18(6,6,6)	medium	6.28	C24H42O9	519.2811	519.2798	-2.6	5.52
I3:20(6,6,8)	medium	6.30	C26H46O9	547.3124	547.3120	-0.8	0.751
I3:20(6,6,8)	medium	6.39	C26H46O9	547.3124	547.3109	-2.7	1.05
I3:20(6,6,8)	medium	6.56	C26H46O9	547.3124	547.3129	0.9	0.439
I3:20(6,6,8)	medium	6.67	C26H46O9	547.3124	547.3144	3.7	0.167
I3:20(6,6,8)	medium	6.81	C26H46O9	547.3124	547.3146	4.1	0.315

Table S2.22. (cont'd)

G3:20(6,6,8)	medium	7.77	C26H46O9	547.3124	547.3124	-0.1	0.640
G3:20(6,6,8)	medium	7.97	C26H46O9	547.3124	547.3121	-0.5	1.24
I3:22(6,8,8)	medium	8.04	C28H50O9	575.3437	575.3485	8.4	0.0682
G3:20(6,6,8)	medium	8.07	C26H46O9	547.3124	547.3134	1.8	0.301
I3:22(6,8,8)	medium	8.20	C28H50O9	575.3437	575.3430	-1.2	0.574
G3:20(6,6,8)	medium	8.29	C26H46O9	547.3124	547.3107	-3.1	1.96
I3:22(6,8,8)	medium	8.38	C28H50O9	575.3437	575.3426	-1.9	1.30
G3:20(6,6,8)	medium	8.52	C26H46O9	547.3124	547.3109	-2.7	3.70
I3:22(6,8,8)	medium	8.57	C28H50O9	575.3437	575.3425	-2.0	1.86

Table S2.23. Summary of annotated acylsugars detected in *S. dulcamara* trichome extracts.
I = inositol. Annotation method and confidence level criteria are described within the Methods.
RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp =
experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million.
Acylsugar abundance was reported by Progenesis QI and was averaged over six samples.
Acylsugars are sorted by number of sugar moieties.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Glycohydroxyacylhexoses							
H3:22(4,4,14-O-P)	medium	4.31	C33H58O14	723.3809	723.3777	-4.5	32.5
Acylhexoses							
I3:22(4,4,14-OH)	medium	6.24	C28H50O10	591.3386	591.3372	-2.4	67.5

Table S2.24. Summary of annotated acylsugars detected in *S. capsicoides* trichome extracts.
I = inositol. Annotation method and confidence level criteria are described within the Methods.
RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp =
experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million.
Acylsugar abundance was reported by Progenesis QI and was averaged over four samples.
Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acylhexoses							
H3:16:[O1]	low	2.81	C22H38O10	507.2447	507.2429	-3.6	0.945
H3:18:[O1]	low	3.89	C24H42O10	535.2760	535.2745	-2.8	0.228
H3:18:[O1]	low	4.05	C24H42O10	535.2760	535.2753	-1.4	10.4
H3:19:[O1]	low	4.64	C25H44O10	549.2911	549.2964	9.6	0.00628
H3:19:[O1]	low	4.72	C25H44O10	549.2911	549.2956	8.3	0.00459
I2:20(8,12-OH)	medium	5.21	C26H48O9	549.3280	549.3275	-0.9	0.0181
H3:20:[O1]	low	5.23	C26H46O10	563.3073	563.3076	0.5	0.00358
I2:20(8,12-OH)	medium	5.40	C26H48O9	549.3280	549.3269	-2.0	0.170
H3:20:[O1]	low	5.78	C26H46O10	563.3073	563.3070	-0.5	0.00217
H3:20:[O1]	low	6.06	C26H46O10	563.3073	563.3062	-2.0	0.891
H2:21:[O1]	low	6.12	C27H50O9	563.3431	563.3449	3.1	0.00696
H3:25:[O1]	low	6.65	C30H52O11	633.3492	633.3499	1.1	0.0120
H3:25:[O1]	low	6.93	C30H52O11	633.3492	633.3480	-1.9	1.21
H3:22	low	8.17	C28H50O9	575.3437	575.3450	2.3	0.00327
H3:22	low	8.40	C28H50O9	575.3437	575.3440	0.5	0.0386
H3:22	low	8.60	C28H50O9	575.3437	575.3430	-1.3	0.0511
H3:22	low	8.75	C28H50O9	575.3437	575.3440	0.6	0.0192
H3:26:[O1]	low	9.84	C32H58O10	647.4012	647.4023	1.8	0.0118
H3:26:[O1]	low	9.93	C32H58O10	647.4012	647.4009	-0.5	0.0690
I3:26(6,8,12-OH)	medium	10.08	C32H58O10	647.4012	647.3995	-2.6	0.727
I3:26(6,8,12-OH)	medium	10.30	C32H58O10	647.4012	647.4001	-1.7	1.68
H3:24	low	10.30	C30H54O9	603.3750	603.3760	1.6	0.00824
I3:26(6,8,12-OH)	medium	10.44	C32H58O10	647.4012	647.4003	-1.5	4.26
I3:24(8,8,8)	medium	10.54	C30H54O9	603.3750	603.3744	-1.0	0.164
H3:26:[O1]	low	10.67	C32H58O10	647.4012	647.4024	1.9	0.0185
I3:24(8,8,8)	medium	10.74	C30H54O9	603.3750	603.3745	-0.8	0.168
I3:24(8,8,8)	medium	10.82	C30H54O9	603.3750	603.3746	-0.7	0.383
I3:24(8,8,8)	medium	11.02	C30H54O9	603.3750	603.3746	-0.7	0.950
H3:28:[O1]	low	11.04	C33H60O10	661.4163	661.4176	2.0	0.0378
H3:28:[O1]	low	11.22	C33H60O10	661.4163	661.4166	0.5	0.153
H3:28:[O1]	low	11.32	C33H60O10	661.4163	661.4169	0.9	0.154
H3:28:[O1]	low	11.44	C34H62O10	675.4325	675.4342	2.6	0.0119

Table S2.24. (cont'd)

H3:28:[O1]	low	11.50	C33H60O10	661.4163	661.4176	2.0	0.0851
H3:28:[O1]	low	11.65	C34H62O10	675.4325	675.4327	0.3	0.0240
H3:28:[O1]	low	11.72	C33H60O10	661.4163	661.4209	6.9	0.00834
H3:28:[O1]	low	11.89	C34H62O10	675.4325	675.4325	0.0	0.160
I3:28(8,8,12-OH)	medium	12.12	C34H62O10	675.4325	675.4319	-0.9	5.56
I3:28(8,8,12-OH)	medium	12.38	C34H62O10	675.4325	675.4321	-0.6	22.5
I3:28(8,8,12-OH)	medium	12.58	C34H62O10	675.4325	675.4316	-1.4	26.9
H3:28:[O1]	low	12.78	C34H62O10	675.4325	675.4321	-0.5	0.173
H3:28:[O1]	low	12.93	C34H62O10	675.4325	675.4318	-1.1	0.390
H3:29:[O1]	low	12.98	C35H64O10	689.4481	689.4486	0.7	0.0478
I3:29(8,9,12-OH)	medium	13.14	C35H64O10	689.4481	######	4.0	0.964
I3:29(8,9,12-OH)	medium	13.23	C35H64O10				
I3:29(8,9,12-OH)	medium	13.34	C35H64O10	689.4481	######	4.0	2.61
I3:29(8,9,12-OH)	medium	13.40	C35H64O10				
H3:29:[O1]	low	13.71	C35H64O10	689.4481	689.4500	2.7	0.0743
I3:30(8,10,12-OH)	medium	14.19	C36H66O10	703.4638	703.4631	-1.0	0.189
I3:30(8,8,14-OH)	medium	14.24	C36H66O10	703.4638	703.4631	-1.0	1.03
I4:34(6,8,8,12-OH)	medium	14.31	C40H72O11	773.5057	773.5054	-0.4	0.327
I3:30(8,10,12-OH)	medium	14.42	C36H66O10	703.4638	703.4631	-0.9	2.40
I3:30(8,10,12-OH)	medium	14.50	C36H66O10	703.4638	######	0.3	2.68
I3:30(8,8,14-OH)	medium	14.50	C36H66O10				
I3:30(8,8,14-OH)	medium	14.57	C36H66O10				
I4:34(6,8,8,12-OH)	medium	14.51	C40H72O11	773.5057	773.5064	0.9	1.32
I4:34(6,8,8,12-OH)	medium	14.65	C40H72O11	773.5057	773.5067	1.3	0.763
I3:30(8,10,12-OH)	medium	14.65	C36H66O10	703.4638	######	1.0	3.46
I3:30(8,8,14-OH)	medium	14.75	C36H66O10				
H3:30:[O1]	low	15.04	C36H66O10	703.4638	703.4645	1.0	0.0685
H3:32:[O1]	low	15.92	C38H70O10	731.4951	731.4949	-0.3	0.221
I3:32(8,12,12-OH)	medium	16.07	C38H70O10	731.4951	731.4948	-0.4	0.494
I3:32(8,12,12-OH)	medium	16.22	C38H70O10	731.4951	731.4947	-0.5	0.777
I3:32(8,12,12-OH)	medium	16.30	C38H70O10	731.4951	731.4947	-0.5	1.54
I3:32(8,12,12-OH)	medium	16.47	C38H70O10	731.4951	731.4949	-0.3	2.34
H3:32:[O1]	low	16.72	C38H70O10	731.4951	731.4967	2.2	0.0316

Table S2.25. Summary of annotated acylsugars detected in *S. americanum* trichome extracts. PH = pentose-hexose. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over three samples. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Hexose			Pentose		Confidence level	RT (min)	Chemical formula	*m/z* acc	*m/z* exp	Δm(ppm)	Average acylsugar abundance (%)
Acyldisaccharides												
PH5:18(2,2,4,4,6)	2	4	6	2	4	medium	2.48	C29H46O15	679.2819	679.2800	-2.8	8.58
PH5:19(2,2,4,5,6)	2	5	6	2	4	medium	2.73	C30H48O15	693.2975	693.2964	-1.6	41.6
PH5:20(2,2,5,5,6)	2	5	6	2	5	medium	2.99	C31H50O15	707.3132	707.3127	-0.8	6.31
Acylhexoses												
H2:13						low	2.93	C19H34O8	435.2230	435.2220	-2.3	0.543
H2:13						low	3.06	C19H34O8	435.2230	435.2227	-0.8	0.439
G2:14(5,9)						medium	3.33	C20H36O8	449.2392	449.2397	1.2	0.00860
G2:14(5,9)						medium	3.48	C20H36O8	449.2392	449.2417	5.7	0.0187
G2:14(4,10)						medium	3.61	C20H36O8	449.2392	449.2384	-1.8	0.152
G2:14(4,10)						medium	3.76	C20H36O8	449.2392	449.2392	0.1	0.231
G3:15(2,5,8)						medium	3.77	C21H36O9	477.2341	477.2333	-1.6	2.17
H2:15						low	3.99	C21H38O8	463.2549	463.2572	5.0	0.0339
G3:15(2,5,8)						medium	4.02	C21H36O9	477.2341	477.2327	-3.0	6.22
G2:15(5,10)						medium	4.18	C21H38O8	463.2549	463.2534	-3.2	5.06
G2:15(5,10)						medium	4.39	C21H38O8	463.2549	463.2528	-4.4	4.79
G3:16(2,4,10)						medium	4.77	C22H38O9	491.2498	491.2482	-3.2	0.509
G3:16(2,4,10)						medium	5.07	C22H38O9	491.2498	491.2485	-2.6	1.29
G3:17(2,5,10)						medium	5.55	C23H40O9	505.2654	505.2636	-3.6	7.09
G3:17(2,5,10)						medium	5.91	C23H40O9	505.2654	505.2636	-3.6	15.0

Table S2.26. Summary of annotated acylsugars detected in *S. sisymbriifolium* trichome extracts. I = inositol. Annotation method and confidence level criteria are described within the Methods. RT = retention time; *m/z* acc = theoretical monoisotopic formate adduct mass; *m/z* exp = experimental formate adduct mass; Δm (ppm) = mass measurement error in parts per million. Acylsugar abundance was reported by Progenesis QI and was averaged over four samples. Acylsugars are sorted by number of sugar moieties and then by elution order.

Name	Hexose		Ring 2	Confidence level	RT (min)	Chemical formula	m/z acc	m/z exp	Δm (ppm)	Average acylsugar abundance (%)
Acylsugars with two sugar groups										
H22:3(4,4,14-O-h)	4	4	14-OH	medium	3.82	C34H60O15	753.3914	753.3945	4.1	0.00568
H22:3(4,4,14-O-p)	4	4	14-OH	medium	4.32	C33H58O14	723.3809	723.3822	1.8	0.0295
H22:3(4,4,14-O-p)	4	4	14-OH	medium	4.53	C33H58O14	723.3809	723.3813	0.5	0.272
H23:3(4,5,14-O-h)	4	5	14-OH	medium	5.01	C34H60O14	737.3965	737.3977	1.7	0.0440
H23:3(4,5,14-O-p)	4	5	14-OH	medium	5.23	C34H60O14	737.3965	737.3970	0.7	0.367
H24:3(4,4,16-O-p)	4	4	16-OH	medium	5.40	C35H62O14	751.4122	751.4135	1.7	0.000185
PH23:3[O1]				low	5.40	C34H60O14	737.3965	737.3989	3.3	0.00241
PH23:3[O1]				low	5.48	C34H60O14	737.3965	737.3989	3.2	0.00774
H24:3(4,4,16-O-p)	4	4	16-OH	medium	5.51	C35H62O14	751.4122	751.4137	2.0	0.00176
H24:3(4,4,16-O-p)	4	4	16-OH	medium	5.60	C35H62O14	751.4122	751.4136	1.9	0.00600
H24:3(4,4,16-O-p)	4	4	16-OH	medium	5.62	C36H64O15	781.4227	781.4244	2.2	0.133
H23:3(4,5,14-O-p)	4	5	14-OH	medium	5.75	C34H60O14	737.3965	737.3977	1.6	0.0733
H24:3(5,5,14-O-p)	5	5	14-OH	medium	5.80	C35H62O14	751.4122	751.4141	2.5	0.0206
H24:3(4,4,16-O-p)	4	4	16-OH	medium	6.01	C35H62O14				
H24:3(5,5,14-O-p)	5	5	14-OH	medium	6.03	C35H62O14	751.4122	751.4133	1.4	0.172
H24:3(4,5,15-O-p)	4	5	15-OH	medium	6.05	C35H62O14				
H24:3(4,4,16-O-p)	4	4	16-OH	medium	6.34	C35H62O14	751.4122	751.4130	1.1	1.39
H24:3(4,4,16-O-p)	4	4	16-OH	medium	6.65	C35H62O14	751.4122	751.4129	1.0	5.47
H25:3(5,5,15-O-p)	5	5	15-OH	medium	6.94	C36H64O14	765.4278	765.4300	2.9	0.0439
H25:3(4,5,16-O-p)	4	5	16-OH	medium	7.26	C36H64O14	765.4278	765.4293	2.0	1.97

Table S2.26. (cont'd)

H4:27(4,4,5,14-O-p)	4	4	14-OH	5	medium	7.34	C38H66O15	807.4384	807.4408	3.0	0.0241
H25:3(4,5,16-O-p)	4	5	16-OH		medium	7.45	C36H64O14	765.4278	765.4288	1.3	7.73
H25:3(4,5,16-O-p)	4	5	16-OH		medium	7.57	C36H64O14	765.4278			
H4:27(4,4,5,14-O-p)	4	4	14-OH	5	medium	7.50	C38H66O15	807.4384	807.4411	3.3	0.00673
H4:27(4,4,5,14-O-p)	4	4	14-OH	5	medium	7.64	C38H66O15	807.4384	807.4428	5.5	0.000622
PH4:26:[O1]					low	8.07	C37H64O15	793.4227	793.4260	4.2	0.222
PH4:26:[O1]					low	8.18	C37H64O15	793.4227	793.4264	4.7	0.0169
H4:28(4,5,5,14-O-p)	4	5	14-OH	5	medium	8.20	C39H68O15	821.4540	821.4557	2.1	0.0124
H3:26(5,5,16-O-p)	5	5	16-OH		medium	8.24	C37H66O14	779.4435	779.4446	1.4	0.782
H4:26(4,4,4,14-O-p)	4	4	14-OH	4	medium	8.35	C37H64O15	793.4227	793.4244	2.1	0.475
H3:26(5,5,16-O-p)	5	5	16-OH		medium	8.45	C37H66O14	779.4435	779.4451	2.0	4.02
H3:26(5,5,16-O-p)	5	5	16-OH		medium	8.53	C37H66O14	779.4435			
H4:27(4,4,5,14-O-p)	4	4	14-OH	5	medium	8.57	C38H66O15	807.4384	807.4421	4.5	0.0198
H4:26(4,4,4,14-O-p)					medium	8.64	C37H64O15	793.4227	793.4244	2.1	0.143
H4:27(4,4,5,14-O-p)	4	4	14-OH	5	medium	8.83	C38H66O15	807.4384	807.4416	4.0	0.00803
H3:26(5,5,16-O-p)	5	5	16-OH		medium	8.90	C37H66O14	779.4435	779.4458	3.0	0.0308
H4:28(4,5,5,14-O-p)	5	5	14-OH	4	medium	9.07	C39H68O15	821.4540	821.4560	2.4	0.737
PH4:29[O1]					low	9.10	C40H70O15	835.4697	835.4722	3.0	0.146
H4:27(4,4,5,14-O-p)	4	4	14-OH	5	medium	9.16	C38H66O15	807.4384	807.4406	2.7	1.87
H3:26(4,4,18-O-p)	5	5	18-OH		medium	9.23	C37H66O14	779.4435	779.4451	2.1	0.205
H4:27(4,4,5,14-O-p)	4	5	14-OH	4	medium	9.35	C38H66O15	807.4384	807.4399	1.8	1.18
H4:29(4,4,5,16-O-p)	4	4	16-OH	5	medium	9.45	C40H70O15	835.4697	835.4724	3.3	0.0353
H3:27(5,5,17-O-p)					medium	9.53	C38H68O14	793.4591	793.4604	1.6	0.106
H4:27(4,4,5,14-O-p)	4	4	14-OH	5	medium	9.63	C38H66O15	807.4384	807.4397	1.6	0.133
PH4:27[O1]					low	9.75	C38H66O15	807.4384	807.4409	3.1	0.0203
H4:29(4,4,5,16-O-p)	4	5	16-OH	4	medium	9.76	C40H70O15	835.4697	835.4718	2.5	2.88
PH4:27[O1]					low	9.84	C38H66O15	807.4384	807.4411	3.4	0.00800
PH3:27:[O1]					low	9.91	C38H68O14	793.4591	793.4613	2.8	0.0316
PH4:27:[O1]					low	9.95	C38H66O15	807.4384	807.4411	3.3	0.0143
H4:29(4,4,5,16-O-p)	4	5	16-OH	4	medium	9.96	C40H70O15	835.4697	835.4718	2.5	1.02
H4:28(4,5,5,14-O-p)	5	5	14-OH	5	medium	10.11	C39H68O15	821.4540	821.4563	2.8	2.86

Table S2.26. (cont'd)

H3:27(4,5,18-O-p)	5	5	18-OH		medium	10.23	C38H68O14	793.4591	793.4611	2.5	0.389
H4:28(4,5,5,14-O-p)	5	5	14-OH	5	medium	10.32	C39H68O15	821.4540	821.4565	3.0	1.16
PH4:30:[O1]					low	10.32	C41H72O15	849.4853	849.4886	3.8	0.0656
PH4:28:[O1]					low	10.63	C39H68O15	821.4540	821.4569	3.5	0.136
H4:30(4,5,5,16-O-p)	4	5	16-OH	5	medium	10.63	C41H72O15	849.4853	849.4878	3.0	2.47
H4:30(4,5,5,16-O-p)	4	5	16-OH	5	medium	10.72	C41H72O15	849.4853	849.4880	3.1	1.12
PH4:28:[O1]					low	10.76	C39H68O15	821.4540	821.4564	2.9	0.185
H4:30(4,5,5,16-O-p)	4	5	16-OH	5	medium	10.89	C41H72O15	849.4853	849.4876	2.7	0.452
H4:28(4,5,5,14-O-p)	5	5	14-OH	5	medium	10.94	C39H68O15	821.4540	821.4561	2.6	3.46
H4:29(5,5,5,14-O-p)	5	5	14-OH	5	medium	11.06	C40H70O15	835.4697	835.4716	2.2	1.64
H4:29(5,5,5,14-O-p)	5	5	14-OH	5	medium	11.26	C40H70O15	835.4697	835.4711	1.7	0.533
H4:29(4,4,5,16-O-p)	4	4	16-OH	5	medium	11.37	C40H70O15	835.4697	835.4711	1.7	0.375
H4:31(5,5,5,16-O-p)	4	5	16-OH	5	medium	11.49	C42H74O15	863.5010	863.5027	2.0	2.04
PH4:31:[O1]					low	11.56	C42H74O15				
H4:29(4,4,5,16-O-p)	4	4	16-OH	5	medium	11.69	C40H70O15	835.4697	835.4718	2.5	8.99
PH4:31:[O1]					low	11.77	C42H74O15	863.5010	863.5040	3.5	0.0933
H4:29(4,4,5,16-O-p)	4	5	16-OH	4	medium	11.89	C40H70O15	835.4697	835.4723	3.1	7.66
PH4:30:[O1]					low	11.96	C41H72O15	849.4853	849.4883	3.5	0.165
PH4:31:[O1]					low	12.12	C42H74O15	863.5010	863.5022	1.4	0.0229
H4:30(4,5,5,16-O-p)	4	5	16-OH	5	medium	12.25	C41H72O15	849.4853	849.4874	2.5	0.543
PH4:31:[O1]					low	12.35	C42H74O15	863.5010	863.5015	0.6	0.00472
H4:30(4,5,5,16-O-p)	4	5	16-OH	5	medium	12.53	C41H72O15	849.4853	849.4872	2.2	11.8
H4:30(4,5,5,16-O-p)	4	5	16-OH	5	medium	12.75	C41H72O15	849.4853	849.4878	2.9	6.60
H4:31(5,5,5,16-O-p)					medium	13.08	C42H74O15	863.5010	863.5036	3.1	0.139
H4:31(5,5,5,16-O-p)	5	5	16-OH	5	medium	13.38	C42H74O15	863.5010	863.5038	3.2	6.18
H4:31(5,5,5,16-O-p)	5	5	16-OH	5	medium	13.56	C42H74O15	863.5010	863.5034	2.7	2.76
PH4:31:[O1]					low	13.93	C42H74O15	863.5010	863.5039	3.4	0.272
H4:31(4,5,5,17-O-p)					medium	14.11	C42H74O15	863.5010	863.5040	3.4	0.150
Acylhexoses											
I3:22(4,4,14-OH)					high	6.24	C28H50O10	591.3382	591.3387	0.9	0.0334

Table S2.26. (cont'd)

I3:23(4,5,14-OH)	medium	7.14	C29H52O10	605.3542	605.3546	0.7	0.121
I3:24(4,4,16-OH)	medium	7.54	C30H54O10	619.3699	619.3713	2.2	0.00224
I3:24(5,5,14-OH)	medium	8.12	C30H54O10	619.3699	619.3701	0.4	0.112
I3:25:[O1]	low	8.22	C31H56O10	633.3855	633.3860	0.8	0.00561
I3:25(4,5,16-OH)	medium	8.33	C31H56O10	633.3855	633.3861	1.0	0.00424
I3:24(4,4,16-OH)	medium	8.45	C30H54O10	619.3699	619.3710	1.8	0.0161
I3:24(4,4,16-OH)	medium	8.79	C30H54O10	619.3699	619.3696	-0.4	1.90
I3:25(4,5,16-OH)	medium	9.12	C31H56O10	633.3855	633.3872	2.6	0.00658
I3:25:[O1]	low	9.25	C31H56O10	633.3855	633.3870	2.4	0.000808
I3:25(4,5,16-OH)	medium	9.43	C31H56O10	633.3855	633.3863	1.3	0.0264
I3:25(4,5,16-OH)	medium	9.78	C31H56O10	633.3855	633.3855	-0.1	2.36
I3:26(5,5,16-OH)	medium	10.43	C32H58O10	647.4012	647.4029	2.6	0.00809
I3:26(5,5,16-OH)	medium	10.78	C32H58O10	647.4012	647.4016	0.6	1.28
I3:26(5,5,16-OH)	medium	11.44	C32H58O10	647.4012	647.4018	1.0	0.0258

Table S2.27. 22 min C18 acylsugar analysis LC-MS method.

	Flow (mL/min)	%A	%B	
Initial	0.3	95	5	Column: Acquity UPLC BEH amide (10 cm x 2.1 mm, 130 Å, 1.7 μm)
1	0.3	40	60	A: 10 mM ammonium formate in water, pH 2.8
16	0.3	0	100	B: Acetonitrile
20	0.3	0	100	
20.01	0.3	95	5	
22	0.3	95	5	

Table S2.28. 14 min C18 LC-MS method for analysis of purified acylsugars.

	Flow (mL/min)	%A	%B	
Initial	0.3	95	5	Column: Ascentis Express C18 HPLC column (10 cm x 2.1 mm, 2.7 µm)
2	0.3	40	60	A: 10 mM ammonium formate in water, pH 2.8
10	0.3	0	100	B: Acetonitrile
12	0.3	0	100	
12.01	0.3	95	5	
14	0.3	95	5	

Table S2.29. 7 min C18 LC-MS method for analysis of enzyme assays.

	Flow (mL/min)	%A	%B	
Initial	0.3	95	5	Column: Ascentis Express C18 HPLC column (10 cm x 2.1 mm, 2.7 µm)
1	0.3	40	60	A: 10 mM ammonium formate in water, pH 2.8
5	0.3	0	100	B: Acetonitrile
6	0.3	0	100	
6.01	0.3	95	5	
7	0.3	95	5	

Table S2.30. 24 min C18 LC-MS method for analysis AI3:16 and AI4:18 coelution.

	Flow (mL/min)	%A	%B	
Initial	0.3	95	5	Column: Acquity UPLC BEH amide (10 cm x 2.1 mm, 130 Å, 1.7 μm)
18	0.3	40	60	A: 10 mM ammonium formate in water, pH 2.8
18.01	0.3	0	100	B: Acetonitrile
22	0.3	0	100	
22.01	0.3	95	5	
24	0.3	95	5	

Table S2.31. 9 min C18 LC-MS method for analysis saponified acylsugars.

	Flow (mL/min)	%A	%B	
Initial	0.3	5	95	Column: Acquity UPLC BEH amide (10 cm x 2.1 mm, 130 Å, 1.7 μm)
6	0.3	60	40	A: 10 mM ammonium formate in water, pH 2.8
7	0.3	95	5	B: Acetonitrile
7.01	0.3	5	95	
9	0.3	5	95	

Table S2.32. 54 min C18 LC-MS method for separation and purification of *S. melongena* acylsugars.

	Flow (mL/min)	%A	%B	
Initial	1	95	5	Column: Acclaim 120 C18 HPLC column (4.6 x 150 mm, 5 μm)
2	1	40	60	A: water + 0.1% Formic acid
40	1	20	80	B: Acetonitrile
42	1	0	100	
42.01	1	95	5	
44	1	95	5	

Table S2.33. Sugar core composition GC-MS method information.

Inlet temp		280°C
Transfer line temp		280°C
Helium gas flow rate		10 mL/min
Split ratio		10:1
	Temp (°C)	Temp gradient (°C/min)
initial	60	
1	60	0
4	180	40
16	240	5
19	240	0

Table S2.34. Hydroxylated acyl chain stereochemistry GC-MS method information.

	Temp (°C)	Temp gradient (°C/min)
Inlet temp	250°C	
Transfer line temp	280°C	
Helium gas flow rate	1 mL/min	
Split ratio	NA	
initial	60	
1.5	60	0
4.5	180	40
6.5	180	0
31.5	230	2
40.5	230	0
47.5	300	10
55	300	0

Table S2.35. *S. melongena* **ASAT candidate gene cloning information.**

Gene ID	Gene name	Source of cloned DNA	Species source for DNA sequence	Accession source for DNA sequence	Codon optimized (y/n)
SMEL_05g005070		synthesized gene fragment	*Solanum melongena*	67/3	Y
SMEL_06g025230		synthesized gene fragment	*Solanum melongena*	67/3	Y
SMEL_07g013870		synthesized gene fragment	*Solanum melongena*	67/3	N
SMEL_07g013880		cDNA	*Solanum melongena*	PI 555598	N
SMEL_08g013890		synthesized gene fragment	*Solanum melongena*	67/3	Y
SMEL_12g015770		gDNA	*Solanum melongena*	67/3	N
SMEL_12g015780	SmASAT3-L1	gDNA	*Solanum melongena*	PI 452123	N

Table S2.36. Oligonucleotides used in this study.

Sequence name	Sequence	Description
SMEL_07g013880_F	agcatgactggtggacagcaaatgggtcggATGGCTTCATCACAGATTCTATCAATCCAC	Sequence used for cloning into pET28b with BamHI/XhoI site
SMEL_07g013880_R	gccggatctcagtggtggtggtggtggtgcTTATATTGGGTGAGCAAACTCAAGGAGTTG	Sequence used for cloning into pET28b with BamHI/XhoI site
SMEL_07g013870_F	agcatgactggtggacagcaaatgggtcggATGGCTGCATCACGATTTGCTTTGATTTCC	Sequence used for cloning into pET28b with BamHI/XhoI site
SMEL_07g013870_R	gccggatctcagtggtggtggtggtggtgcTTAAAGACCCAAACTTGGAGAAGCAAATTC	Sequence used for cloning into pET28b with BamHI/XhoI site
SmASAT3_F	agcatgactggtggacagcaaatgggtcggATGGTAGCATCAAGAATTGTGTCTAAAAAG	Sequence used for cloning into pET28b with BamHI/XhoI site
SmASAT3_R	gccggatctcagtggtggtggtggtggtgcTCACAGACATGTAGTATCTTTGATTTTGAC	Sequence used for cloning into pET28b with BamHI/XhoI site
SmASAT3-L1_F	agcatgactggtggacagcaaatgggtcggATGGCATCATCAAGAATTATGTCTAGAAAG	Sequence used for cloning into pET28b with BamHI/XhoI site
SmASAT3-L1_R	gccggatctcagtggtggtggtggtggtgcTTATTCCGATGACCAACCAACCGGAGAAGC	Sequence used for cloning into pET28b with BamHI/XhoI site
Soaet10022603_RT_F	CCAAACCAACACCTCCAAAC	Primer for RT-PCR of Soaet10022603
Soaet10022603_RT_R	TCAACTCCACCATCATCATCTC	Primer for RT-PCR of Soaet10022603
Soaet10024792_RT_F	ACAAGGTTGCGGATGGATATAG	Primer for RT-PCR of Soaet10024792
Soaet10024792_RT_R	TGAACCTGTTGCGGAGTTT	Primer for RT-PCR of Soaet10024792
Soaet10024793_RT_F	CCATCTTCTCCAGTATCGTCTTT	Primer for RT-PCR of Soaet10024793
Soaet10024793_RT_R	ACATCATCGTCGTCCCTTTC	Primer for RT-PCR of Soaet10024793
Soaet10043742_RT_F	CGCGATAGGAGATGCAAGTAG	Primer for RT-PCR of Soaet10043742
Soaet10043742_RT_R	CGTCTCCCTAGCACATTCTTT	Primer for RT-PCR of Soaet10043742
Soaet_ASAT3_RT_1_F	GACGCCACGTGTCAAGAATA	Primer for RT-PCR of SaASAT3
Soaet_ASAT3_RT_1_R	CTTGCTAGGGTCACTCTTATGG	Primer for RT-PCR of SaASAT3
Soaet_ASAT3_RT_2_F	TCACCATTTGCCTTCTTCTACC	Primer for RT-PCR of SaASAT3
Soaet_ASAT3_RT_2_R	CGTGATTCTGGTGGAGCATTTA	Primer for RT-PCR of SaASAT3
Soaet_ASAT3L1_RT_1_F	CCTCACTCCTCCTTCACTTAGA	Primer for RT-PCR of SaASAT3-L1

Table S2.36. (cont'd)

Soaet_ASAT3L1_RT_1_R	TGCAGGAGATGGTTTGGAATC	Primer for RT-PCR of SaASAT3-L1
Soaet_ASAT3L1_RT_2_F	ACAATCGGGCGTCTTCAAA	Primer for RT-PCR of SaASAT3-L1
Soaet_ASAT3L1_RT_2_R	CGGAGAAGCAAACCGTAGAA	Primer for RT-PCR of SaASAT3-L1
Soaet10001586_EF1α_F	CTGACTGTGCTGTCCTGATTAT	Primer for RT-PCR of Soaet10001586 (EF1α)
Soaet10001586_EF1α_R	AGCTTCATGGTGCATCTCTAC	Primer for RT-PCR of Soaet10001586 (EF1α)
T7_promoter	TAATACGACTCACTATAGGG	Primer for Sanger sequencing and colony PCR of pET28b plasmids
T7_terminator	GCTAGTTATTGCTCAGCGG	Primer for Sanger sequencing and colony PCR of pET28b plasmids

Table S2.37. Plant material metadata. Surface metabolites from all listed accessions were analyzed by LC-MS. USDA-GRIN = United States Department of Agriculture - Germplasm Resources Information Network; CREA = Consiglio per la ricerca in agricoltura e l'analisi dell'economia agrarian.

Species	Accession	Name	Origin	Source	Major Clade	Acylsugars detected (y/n)
Solanum abutiloides			Mansfield, Missourri, United States	Baker Creek Heirloom Seeds	Clade II	y
Solanum acerifolium	PI 305325	GUADALUPE 3	Colombia	USDA-GRIN	Clade II	y
Solanum aethiopicum	Grif 14165	G-2563	Brazil	USDA-GRIN	Clade II	y
Solanum aethiopicum	PI 194166		Former Serbia and Montengro	USDA-GRIN	Clade II	y
Solanum aethiopicum	PI 247828	NSUA	Congo	USDA-GRIN	Clade II	y
Solanum aethiopicum	PI 374695		India	USDA-GRIN	Clade II	y
Solanum aethiopicum	PI 424859	W-1790	Brazil	USDA-GRIN	Clade II	y
Solanum aethiopicum	PI 636107	CGN 18558	United Kingdom	USDA-GRIN	Clade II	y
Solanum aethiopicum	PI 666075	CGN 17454	Japan	USDA-GRIN	Clade II	y
Solanum anguivi	PI 180485	10842	India	USDA-GRIN	Clade II	y
Solanum anguivi	PI 183357	11195	India	USDA-GRIN	Clade II	y
Solanum anguivi	PI 194789	11747	India	USDA-GRIN	Clade II	y
Solanum anguivi	PI 319855		Thailand	USDA-GRIN	Clade II	y
Solanum atropurpureum	PI 305320		Colombia	USDA-GRIN	Clade II	y
Solanum capsicoides	PI 196300	2930	Nicaragua	USDA-GRIN	Clade II	y
Solanum capsicoides	PI 370043		India	USDA-GRIN	Clade II	y
Solanum capsicoides	PI 390818	W-C 1203	Peru	USDA-GRIN	Clade II	y
Solanum carolinense			Bath Township, Michigan, United States	Beal Botanical Garden	Clade II	n

220

Table S2.37. (cont'd)
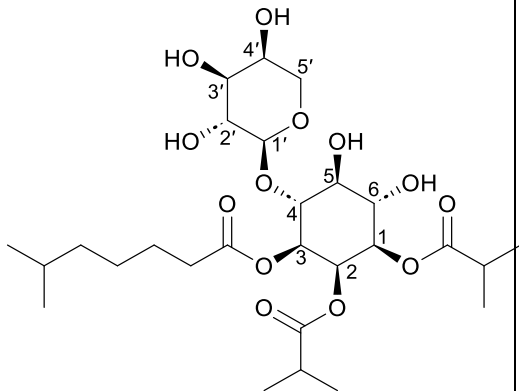
Solanum elaeagnifolium	PI 346963	1196	Mexico	USDA-GRIN	Clade II	n
Solanum incanum	PI 196043	9624	Ethiopia	USDA-GRIN	Clade II	y
Solanum incanum	PI 381155	PLB 294	Delhi, India	USDA-GRIN	Clade II	y
Solanum lasiophyllum	PI 678385	B and T World Seeds No. 16471		USDA-GRIN	Clade II	y
Solanum lichtensteinii	PI 645685	B and T World Seeds No. 442206		USDA-GRIN	Clade II	y
Solanum linnaeanum	PI 388846	WL-74	Italy	USDA-GRIN	Clade II	y
Solanum linnaeanum	PI 388847	WL-85	Italy	USDA-GRIN	Clade II	y
Solanum macrocarpon	PI 441915	BGH 841	Brazil	USDA-GRIN	Clade II	n
Solanum mammosum	PI 305323		Colombia	USDA-GRIN	Clade II	y
Solanum mammosum	PI 370045		India	USDA-GRIN	Clade II	y
Solanum mammosum	PI 413675	1738	Colombia	USDA-GRIN	Clade II	y
Solanum melanospermum			4216 Bing Bong Terminal NT, Australia	Dr. Christopher T. Martine, Bucknell University	Clade II	y
Solanum melongena	305E40		Montanaso Lombardo, Italy	CREA	Clade II	y
Solanum melongena	67/3		Montanaso Lombardo, Italy	CREA	Clade II	y
Solanum melongena	PI 441908	BGH 5008	Brazil	USDA-GRIN	Clade II	y
Solanum melongena	PI 491260	Tsakinik	Greece	USDA-GRIN	Clade II	y
Solanum melongena	PI 555598	LIAO JIAO 1 HAO	Beijing, China	USDA-GRIN	Clade II	y
Solanum melongena	PI 560904	SEVEN LEAVES	China	USDA-GRIN	Clade II	y
Solanum melongena	PI 639117	Grif 14479	Sri Lanka	USDA-GRIN	Clade II	y
Solanum melongena	PI 666079	Thai Green	Iowa, United States	USDA-GRIN	Clade II	y

221

Table S2.37. (cont'd)

Solanum melongena	PI 452123	Tonda di Manfredonia	Italy	USDA-GRIN	Clade II	not tested
Solanum prinophyllum				Dr. Joyce van Eck, Boyce Thompson Institute	Clade II	y
Solanum pseudocapsicum	Grif 16422	Red Giant		USDA-GRIN	Clade II	n
Solanum richardii	PI 500922		Zambia	USDA-GRIN	Clade II	y
Solanum rostratum	PI 420997	1300	Netherlands	USDA-GRIN	Clade II	n
Solanum rostratum	PI 675066	LBJWC-0110	United States	USDA-GRIN	Clade II	n
Solanum sejunctum			Barrk Sandstone Walk, Kakadu, NT, Australia	Dr. Christopher T. Martine, Bucknell University	Clade II	y
Solanum sibundoyensis			Santa Rosa, California, United States	Trade Winds Fruit	Clade II	n
Solanum sisymbriifolium	PI 358311		India	USDA-GRIN	Clade II	y
Solanum torvum			Fort Myers, Florida, United States	Top Tropicals	Clade II	y
Solanum virginianum	Grif 16940	B and T World Seeds No. 27142		USDA-GRIN	Clade II	n
Solanum virginianum	PI 381293	PLB 161	India	USDA-GRIN	Clade II	n
Solanum virginianum	PI 390213		Japan	USDA-GRIN	Clade II	n
Solanum americanum	PI 268152		United States	USDA-GRIN	DulMo	y
Solanum dulcamara	PI 643457	Sheffield's Seed Co. Lot No. 9804	Georgia, United States	USDA-GRIN	DulMo	y
Solanum retroflexum	PI 634755	WONDERBERRY	Wyoming, United States	USDA-GRIN	DulMo	n
Solanum retroflexum	PI 636106	WONDERBERRY	Wyoming, United States	USDA-GRIN	DulMo	n
Solanum scabrum				New York Botanical Garden	DulMo	y
Solanum villosum			Richfield, Minnesota, United States	Experimental Farm Network	DulMo	y
Solanum laciniatum	PI 337284		Hungary	USDA-GRIN	VANAns	y
Solanum laciniatum	PI 337310		New Zealand	USDA-GRIN	VANAns	y

Table S2.38. AI3:16(i4,i4,i8) chemical shifts and coupling constants.

	AI3:16(i4,i4,i8)
	Molecular Formula: $C_{27}H_{46}O_{13}$ Instrument: Agilent 500 MHz DDR2 NMR solvent: CD_3CN Fractions: 32-37 InChI Key: LLNCXHBBKCPKDQ-XSTJKNNJSA-N SMILES: O[C@H]1[C@H](O)[C@H](O)CO[C@H]1O[C@@H]2[C@@H](OC(CCCC(C)C)=O)[C@@H](OC(C(C)C)=O)[C@@H](OC(C(C)C)=O)[C@H](O)[C@H]2O

Carbon # (Group)	^1H (δ, ppm)	^{13}C (δ, ppm)
1(CH)	4.86 (dd, J = 2.91, 10.24 Hz)	70.92
-1(CO)		176.00
-2(CH)	2.61 (hept, J = 6.97 Hz)	33.80
-3,4(CH₃)	1.08 (d, J = 6.97 Hz)	18.02
	1.06 (d, J = 6.97 Hz)	18.24
2(CH)	5.50 (t, J = 2.93 Hz)	68.10 or 68.08[a]
-1(CO)		175.88
-2(CH)	2.47 (hept, J = 6.98 Hz)	33.69 or 33.64[a]
-3,4(CH₃)	1.20 (d, J = 6.97 Hz)	18.41, 18.40
3(CH)	4.99 (dd, J = 3.04, 10.22 Hz)	70.16
-1(CO)		172.71
-2(CH₂)	2.26 (m)	33.69 or 33.64[a]
-3(CH₂)	1.53 (m)	24.53
-4(CH₂)	1.29 (m)	26.57
-5(CH₂)	1.19 (m)	38.34
-6(CH)	1.55 (m)	27.57
-7,8(CH₃)	0.84 (d, J = 6.60 Hz)	21.89
4(CH)	3.85 (t, J = 10.00 Hz)	80.60
5(CH)	3.47 (t, J = 9.30 Hz)	72.49
6(CH)	3.78 (t, J = 9.85 Hz)	70.66
1'(CH)	4.26 (d, J = 7.01 Hz)	104.17
2'(CH)	3.41 (dd, J = 7.0, 9.36 Hz)	71.10
3'(CH)	3.47 (dd, J = 3.4, 9.3 Hz)	72.49
4'(CH)	3.73 (m)	68.10 or 68.08[a]
5'(CH₂)	3.88 (dd, J = 2.41, 12.58 Hz) 3.54 (dd, J = 1.54, 12.63 Hz)	66.13

a – ^{13}C signals not resolved in 2D spectra.
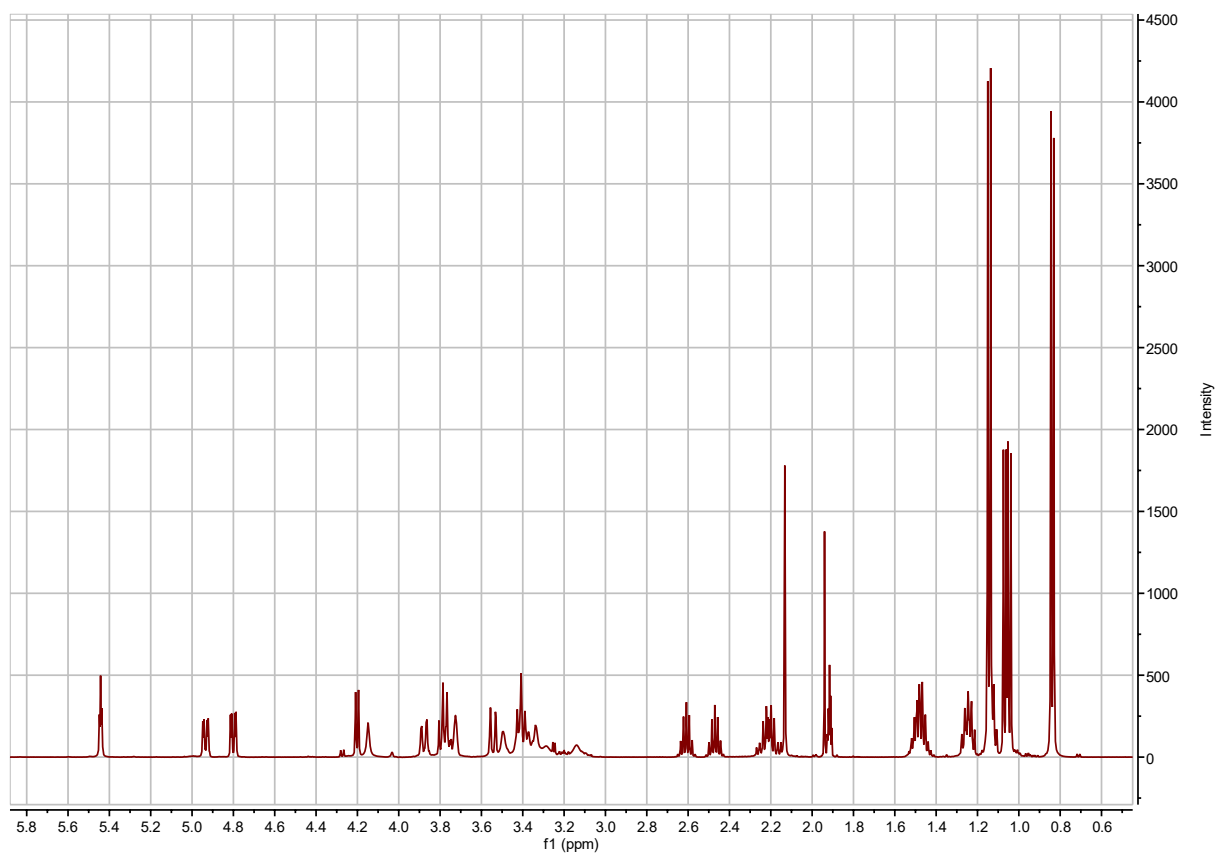
Figure S2.15. AI3:16(4,4,8) ¹H NMR.
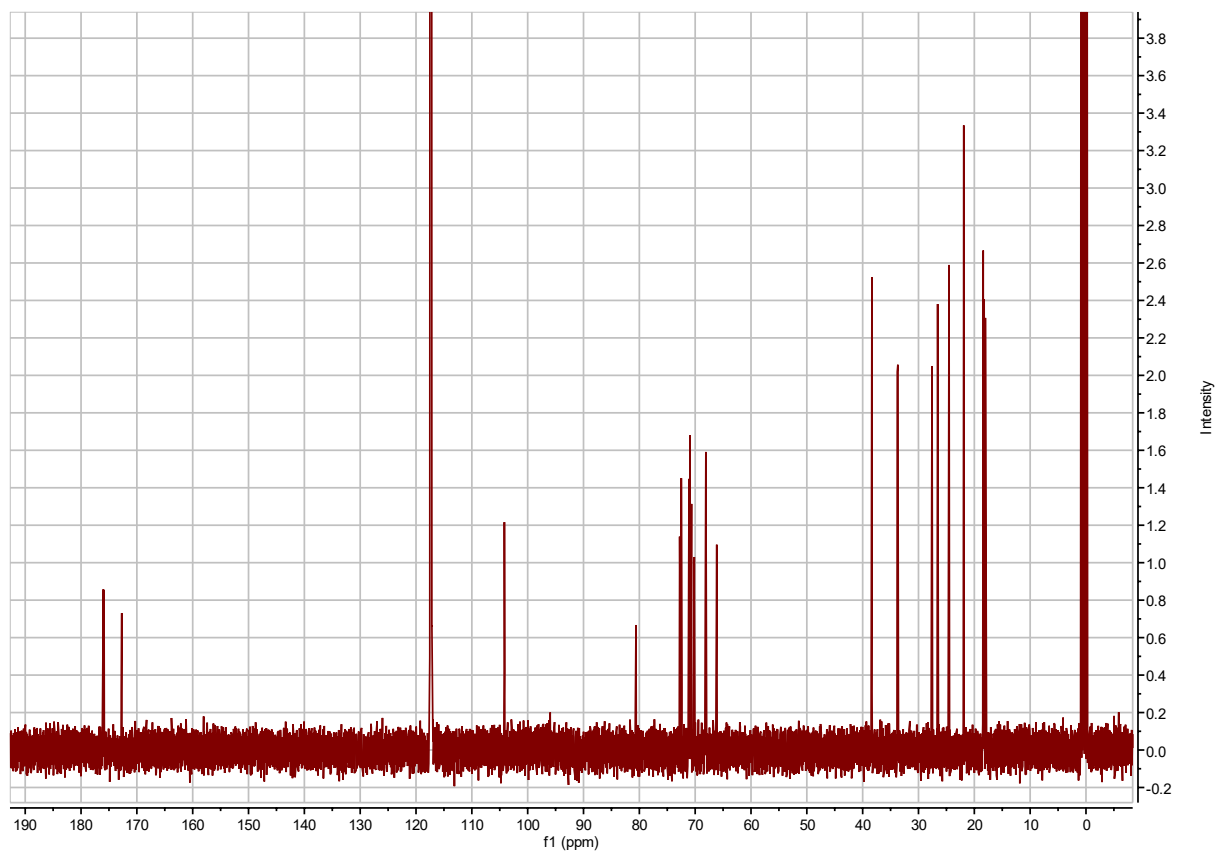
Figure S2.16. AI3:16(4,4,8) ^{13}C NMR.
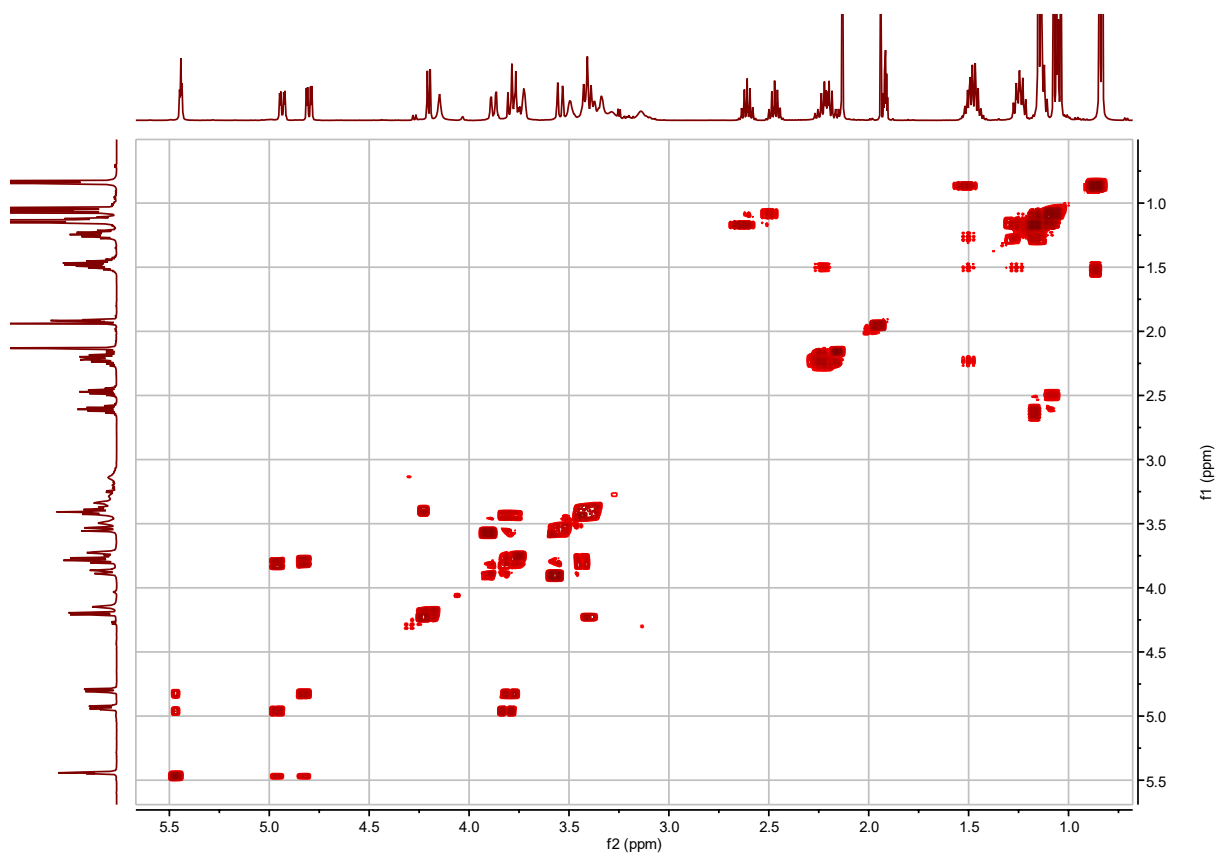
Figure S2.17. AI3:16(4,4,8) ¹H-¹H COSY.
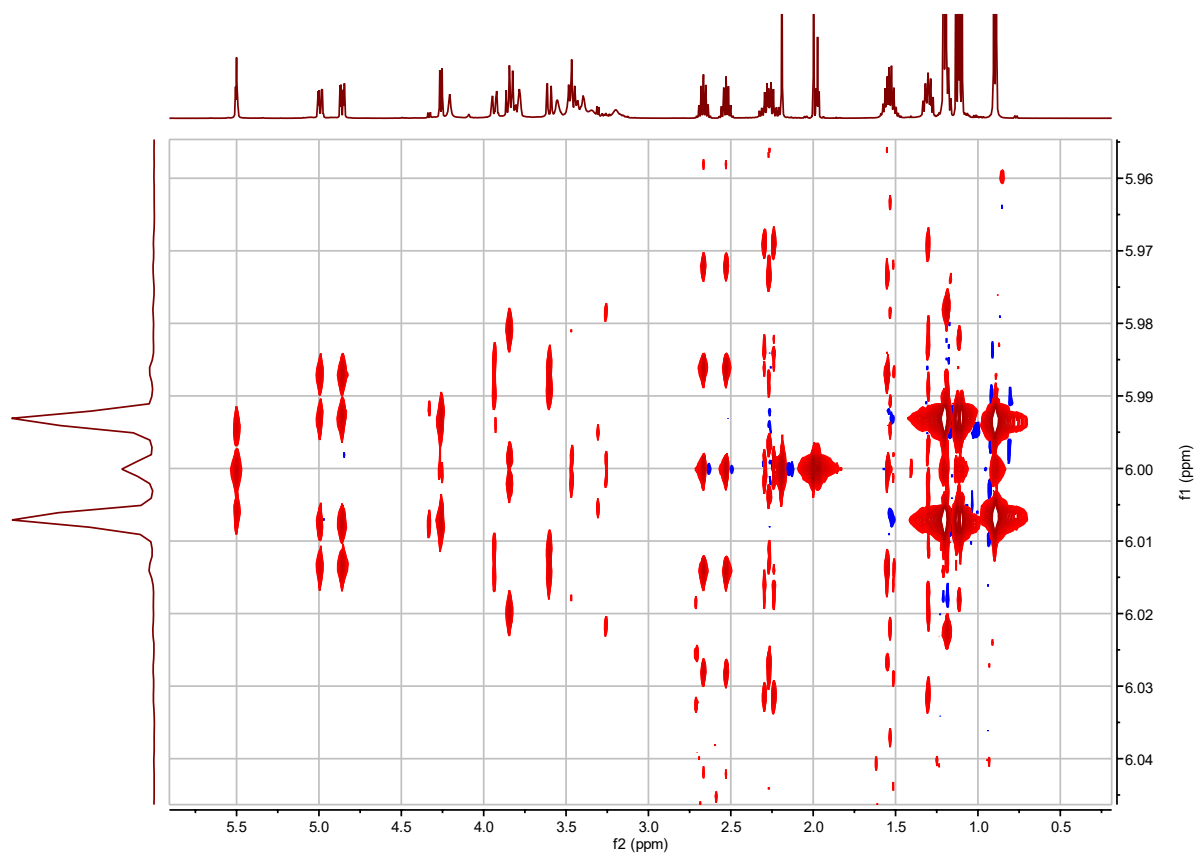
Figure S2.18. AI3:16(4,4,8) *J*-resolved.
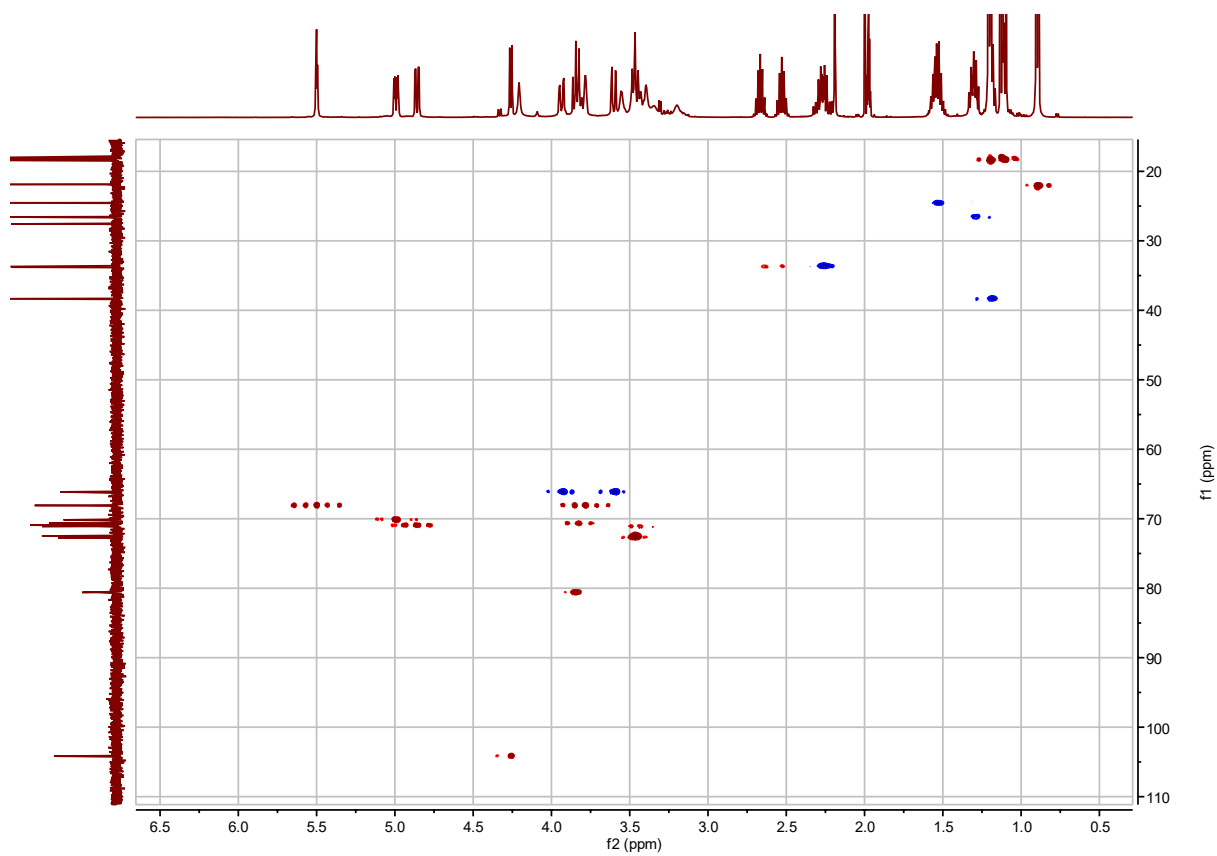
Figure S2.19. AI3:16(4,4,8) ¹H-¹³C HSQC.
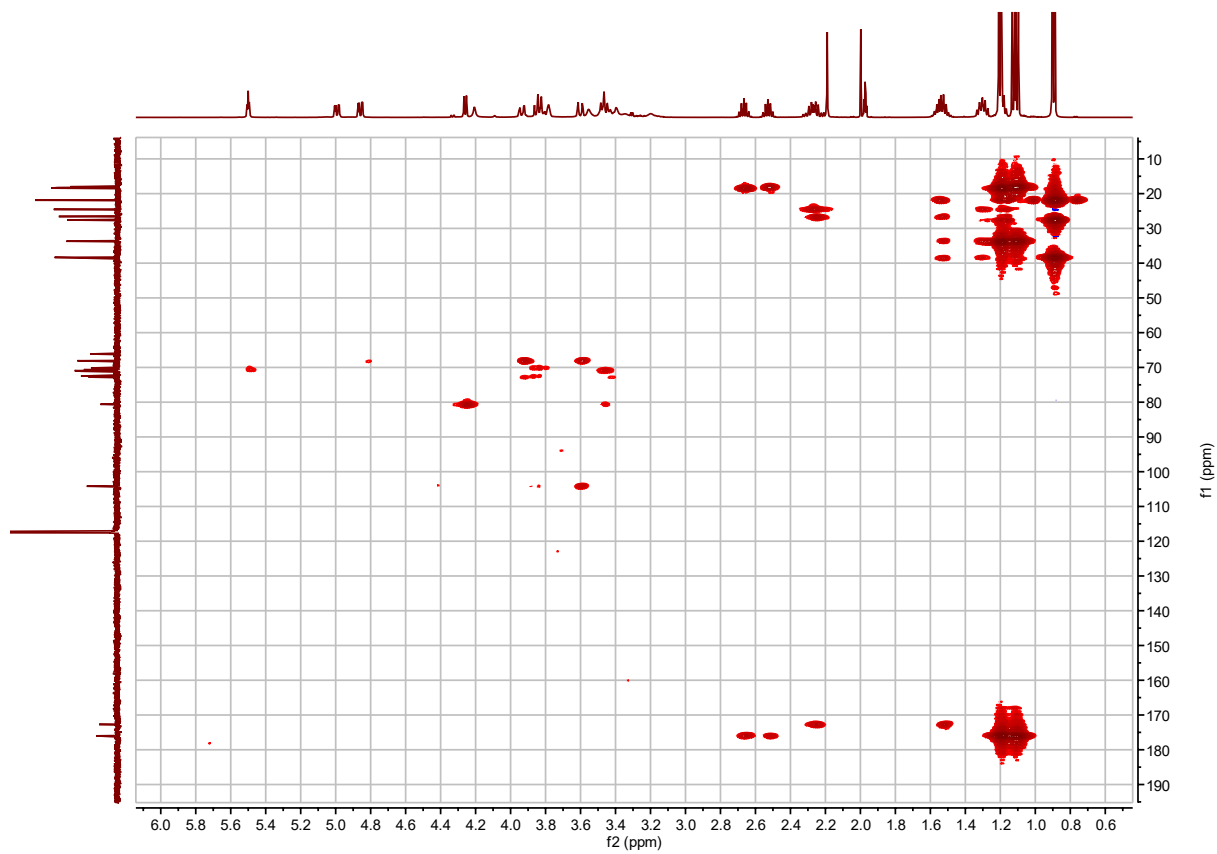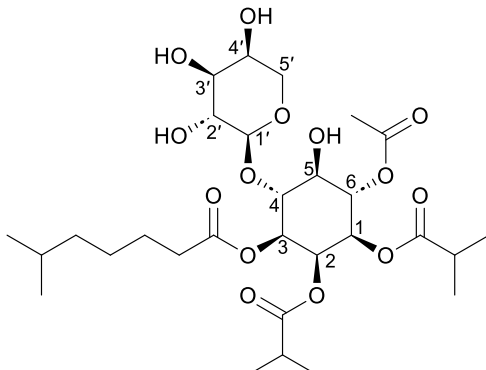
Figure S2.20. AI3:16(4,4,8) ¹H-¹³C HMBC with apodization optimized for correlations between acyl chain carbonyl carbon and C2 proton(s).
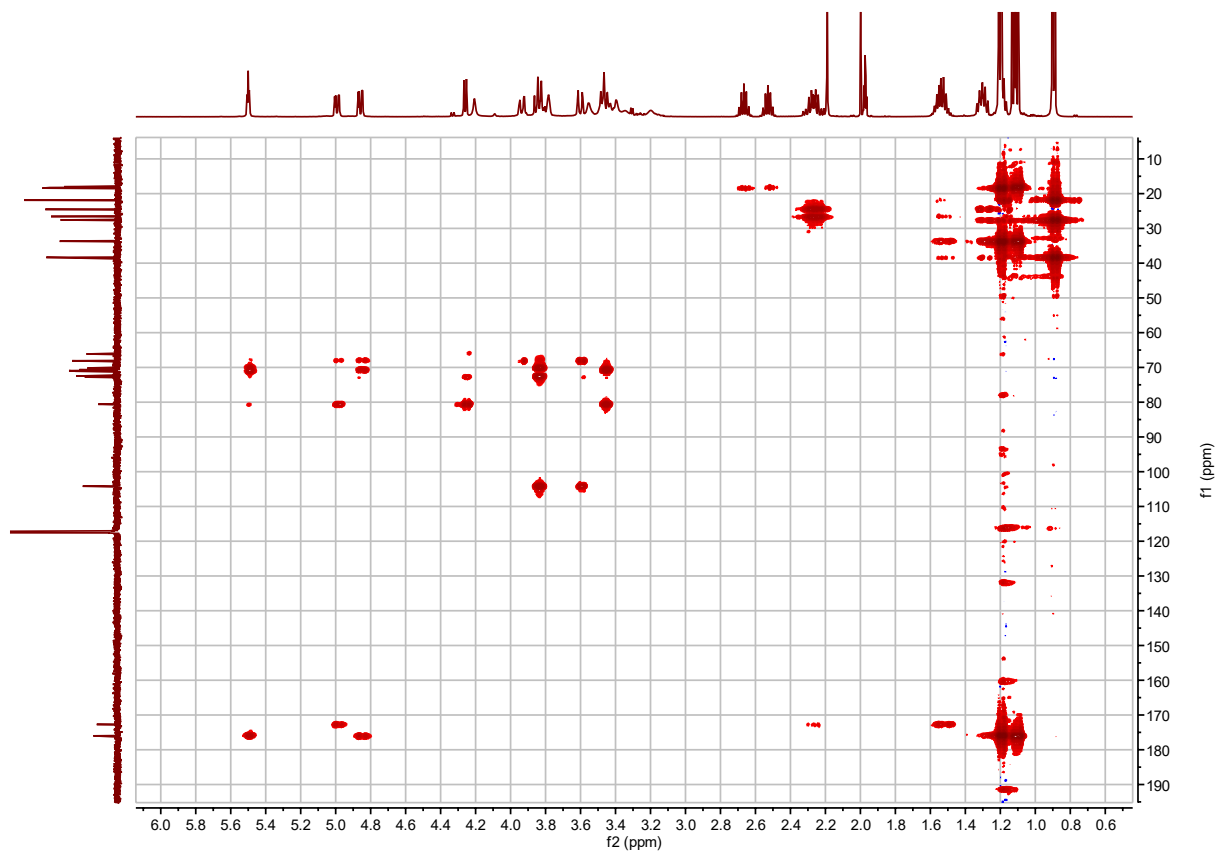
Figure S2.21. AI3:16(4,4,8) ^{1}H-^{13}C HMBC with apodization optimized for correlations between acyl chain carbonyl carbon and sugar ring protons.

Table S2.39. AI4:18(2,i4,i4,i8) chemical shifts and coupling constants.

	AI4:18(2,i4,i4,i8) Molecular Formula: $C_{29}H_{48}O_{14}$ Instrument: Agilent 500 MHz DDR2 and Varian 600 MHz Inova NMR solvent: CD_3CN Fractions: 44-52 InChI Key: JIKVNSLWQPFKKW-INTPKTGNSA-N SMILES: O[C@H]1[C@H](O)[C@H](O)CO[C@H]1O[C@@H]2[C@@H](OC(CCCCC(C)C)=O)[C@@H](OC(C(C)C)=O)[C@@H](OC(C(C)C)=O)[C@H](OC(C)=O)[C@H]2O

Carbon # (Group)	$^1H(\delta,\ ppm)$	$^{13}C\ (\delta,\ ppm)$
1(CH)	5.03 (dd, J = 2.9, 10.5 Hz)	68.88
-1(CO)		175.51
-2(CH)	2.45 (hept, J = 7.0 Hz)	33.67
-3,4(CH$_3$)	1.07 (d, J = 7.0 Hz)	17.85
2(CH)	5.54 (t, J = 2.97 Hz)	67.87
-1(CO)		175.79
-2(CH)	2.69 (hept, J = 6.94 Hz)	33.78
-3,4(CH$_3$)	1.21 (t, J = 6.80 Hz)	18.46
3(CH)	5.05 (dd, J = 2.9, 10.5 Hz)	69.72
-1(CO)		172.68
-2(CH$_2$)	2.29 (m)	33.60
-3(CH$_2$)	1.53 (m)	24.52
-4(CH$_2$)	1.30 (m)	26.54
-5(CH$_2$)	1.19 (m)	38.32
-6(CH)	1.55 (m)	27.56
-7,8(CH$_3$)	0.89 (d, J = 6.6 Hz)	21.88
4(CH)	3.94 (t, J = 9.7 Hz)	80.93
5(CH)	3.67 (t, J = 9.48 Hz)	70.22
6(CH)	5.35 (t, J = 10.10 Hz)	71.07
-1(CO)		169.75
-2(CH$_3$)	2.02 (s)	20.08
1′(CH)	4.25 (d, J = 6.98 Hz)	104.22
2′(CH)	3.43 (dd, J = 7.00, 9.36 Hz)	71.07
3′(CH)	3.47 (dd, J = 3.4, 9.3 Hz)	72.80
4′(CH)	3.78 (m)	68.12
5′(CH$_2$)	3.93 (dd, J = 2.5, 12.5 Hz) 3.61 (dd, J = 1.5, 12.7 Hz)	66.24

Figure S2.22. AI4:18(2,4,4,8) ¹H NMR.
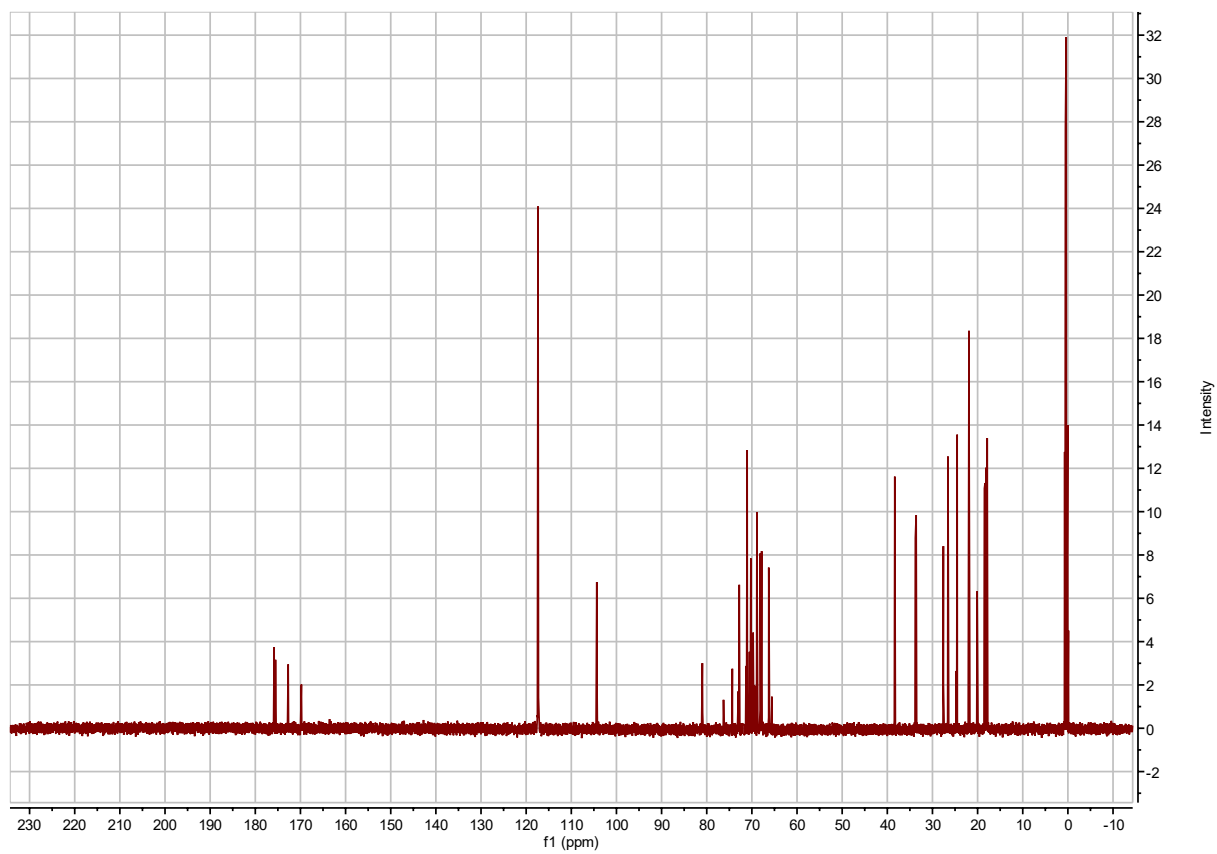
Figure S2.23. AI4:18(2,4,4,8) ^{13}C NMR.
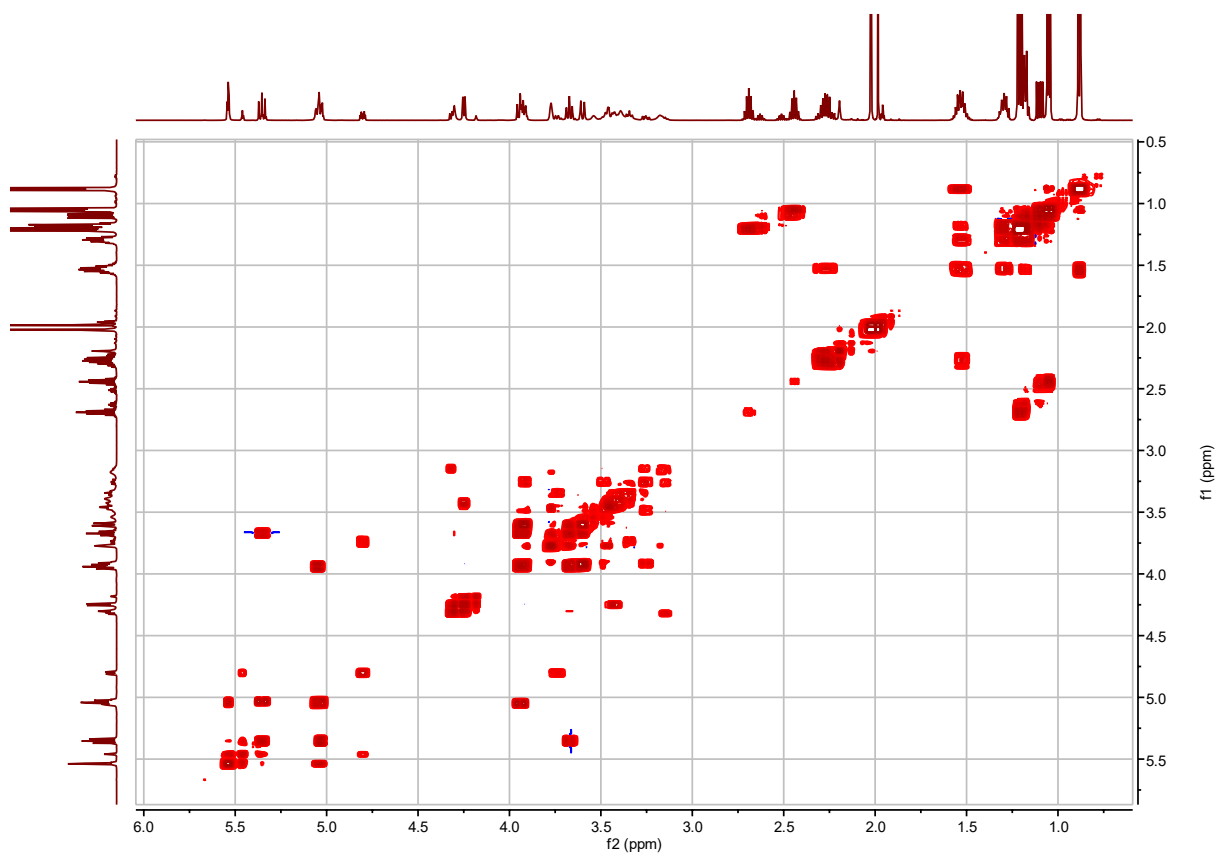
Figure S2.24. AI4:18(2,4,4,8) ¹H-¹H COSY.
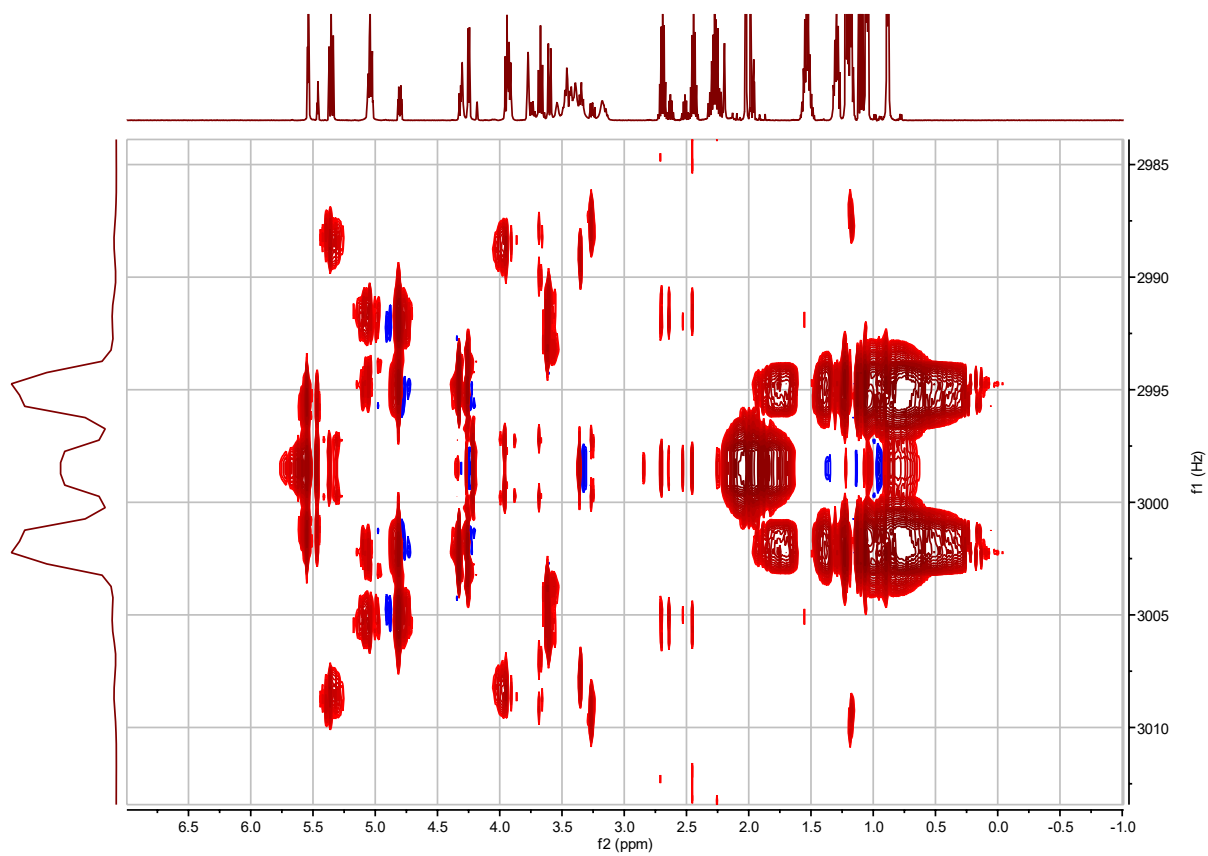
Figure S2.25. AI4:18(2,4,4,8) *J*-resolved.
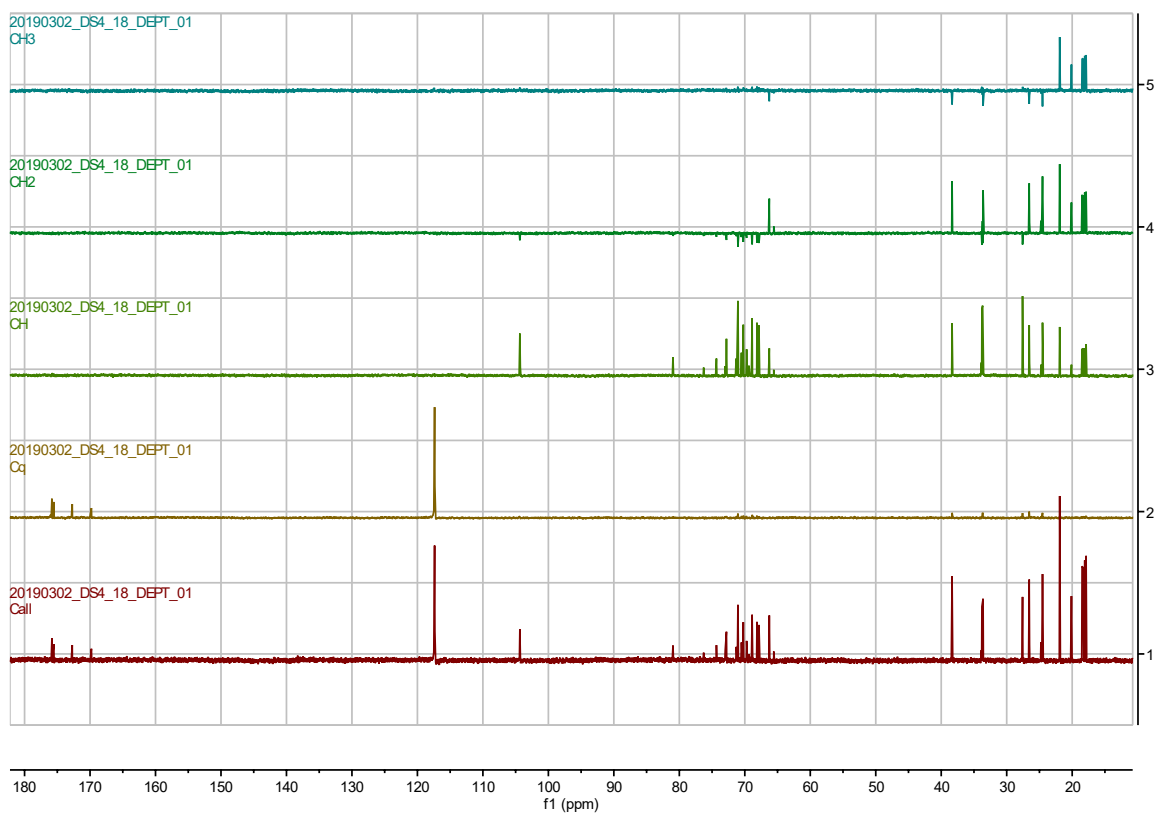
Figure S2.26. AI4:18(2,4,4,8) ^{13}C DEPT.
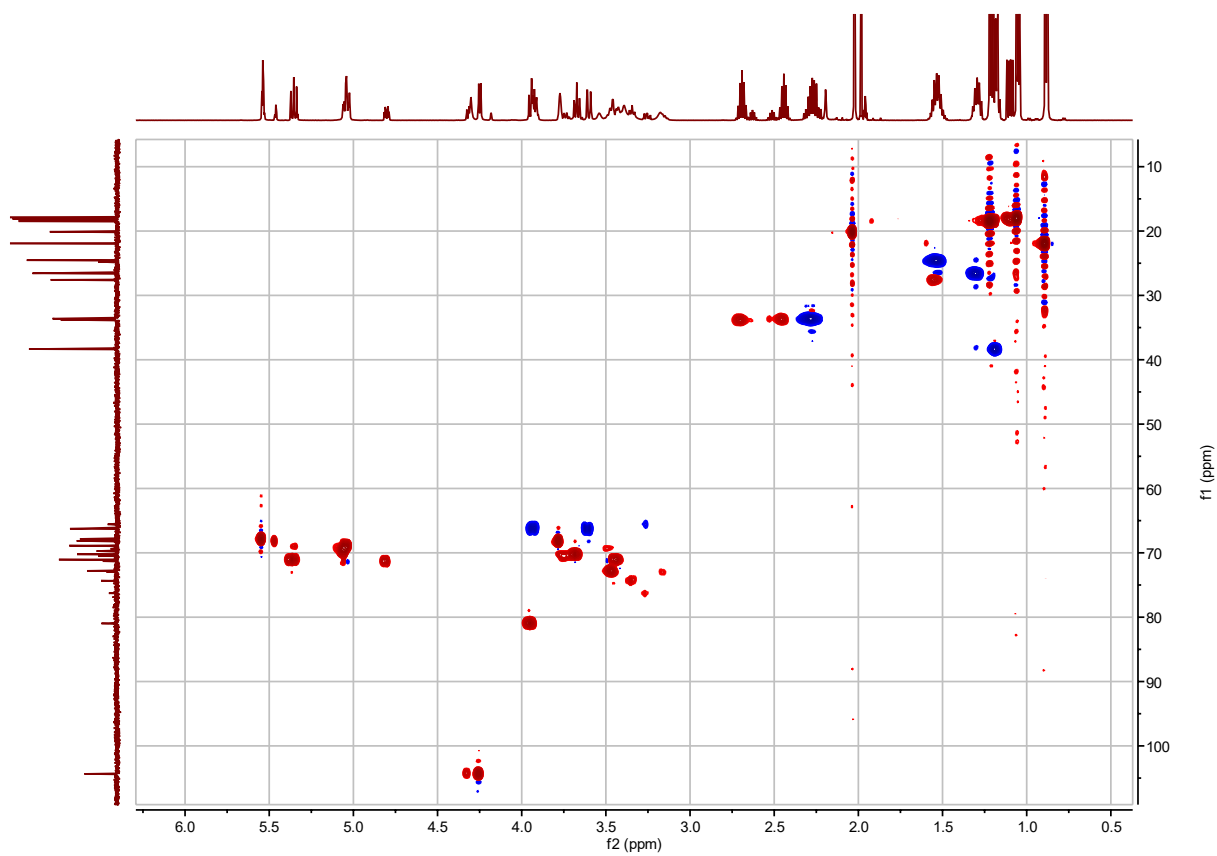
Figure S2.27. AI4:18(2,4,4,8) ^{1}H-^{13}C HSQC.
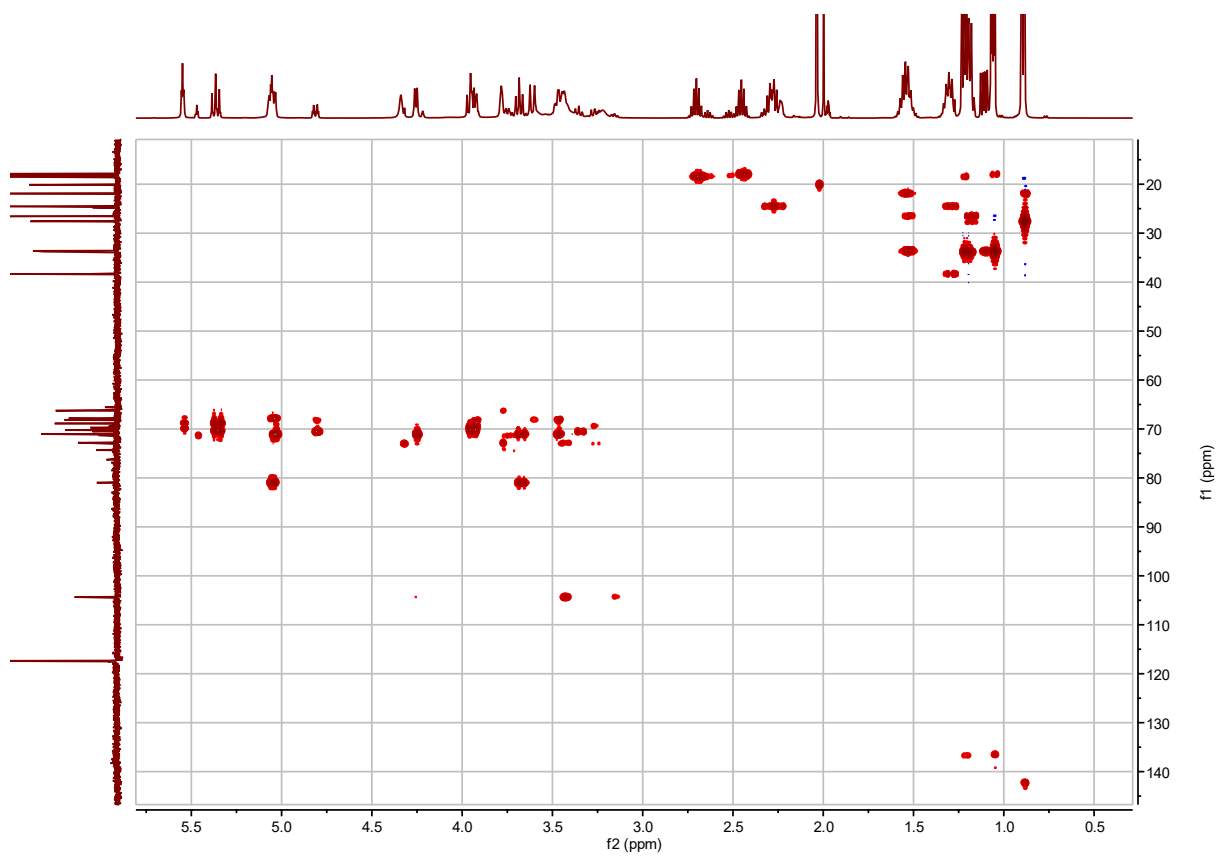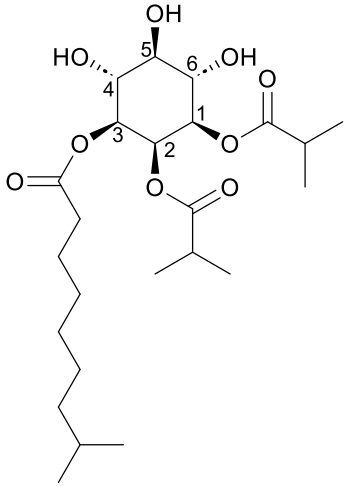
Figure S2.28. AI4:18(2,4,4,8) ^{1}H-^{13}C H2BC.
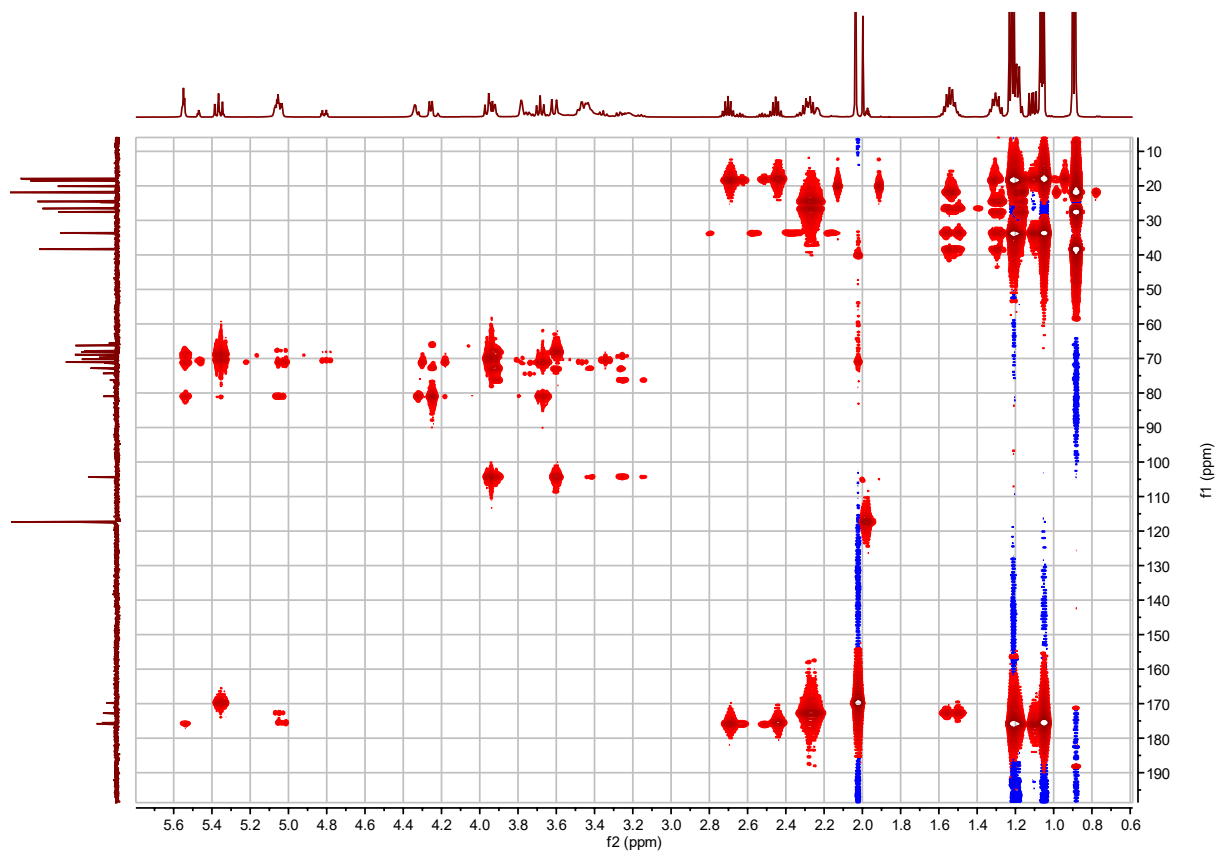
Figure S2.29. AI4:18(2,4,4,8) ^{1}H-^{13}C HMBC.

Table S2.40. I3:18(i4,i4,i10) Chemical shifts and coupling constants.

	I3:18(i4,i4,i10)
	Molecular Formula: $C_{24}H_{42}O_9$ Instrument: Agilent 500 MHz DDR2 NMR solvent: CD_3CN Fractions: 71-74 InChI Key: IRUZHFMTUQKAJJ-QANWUKQVSA-N SMILES: O[C@H]1[C@H](O)[C@@H](OC(CCCCCCC(C)C)=O)[C@@H](OC(C(C)C)=O)[C@@H](OC(C(C)C)=O)[C@@H]1O

Carbon # (Group)	^1H (δ, ppm)	^{13}C (δ, ppm)
1(CH)	4.80 (dd, J = 2.95, 7.15 Hz)	71.31 or 71.33[a]
-1(CO)		176.01
-2(CH)	2.52 (hept, J = 6.98 Hz)	33.68 or 33.70[a]
-3,4(CH$_3$)	1.09, 1.11 (d, J = 7.01 Hz)	18.01, 18.24
2(CH)	5.47 (t, J = 2.98 Hz)	68.16
-1(CO)		179.89
-2(CH)	2.64 (hept, J = 6.97 Hz)	33.85
-3,4(CH$_3$)	1.18 (d, J = 6.95 Hz)	18.35, 18.42
3(CH)	4.80 (dd, J = 2.95, 7.15 Hz)	71.31 or 71.33[a]
-1(CO)		172.80
-2(CH$_2$)	2.26 (m)	33.68 or 33.70[a]
-3(CH$_2$)	1.56 (m)	24.55
-4-6(CH$_2$)	1.36-1.25 (m)	29.27(5), 28.73(4), 26.96(6)
-7(CH$_2$)	1.18 (m)	38.70
-8(CH)	1.54 (hept, J = 6.7 Hz)	27.74
-9,10(CH$_3$)	0.89 (d, J = 6.6 Hz)	21.93
4(CH)	3.75 (t, J = 9.67 Hz)	70.49 or 70.51[a]
5(CH)	3.34 (t, J = 9.29 Hz)	70.49 or 70.51[a]
6(CH)	3.73 (t, J = 9.67 Hz)	74.33

a – ^{13}C signals not resolved in 2D spectra.
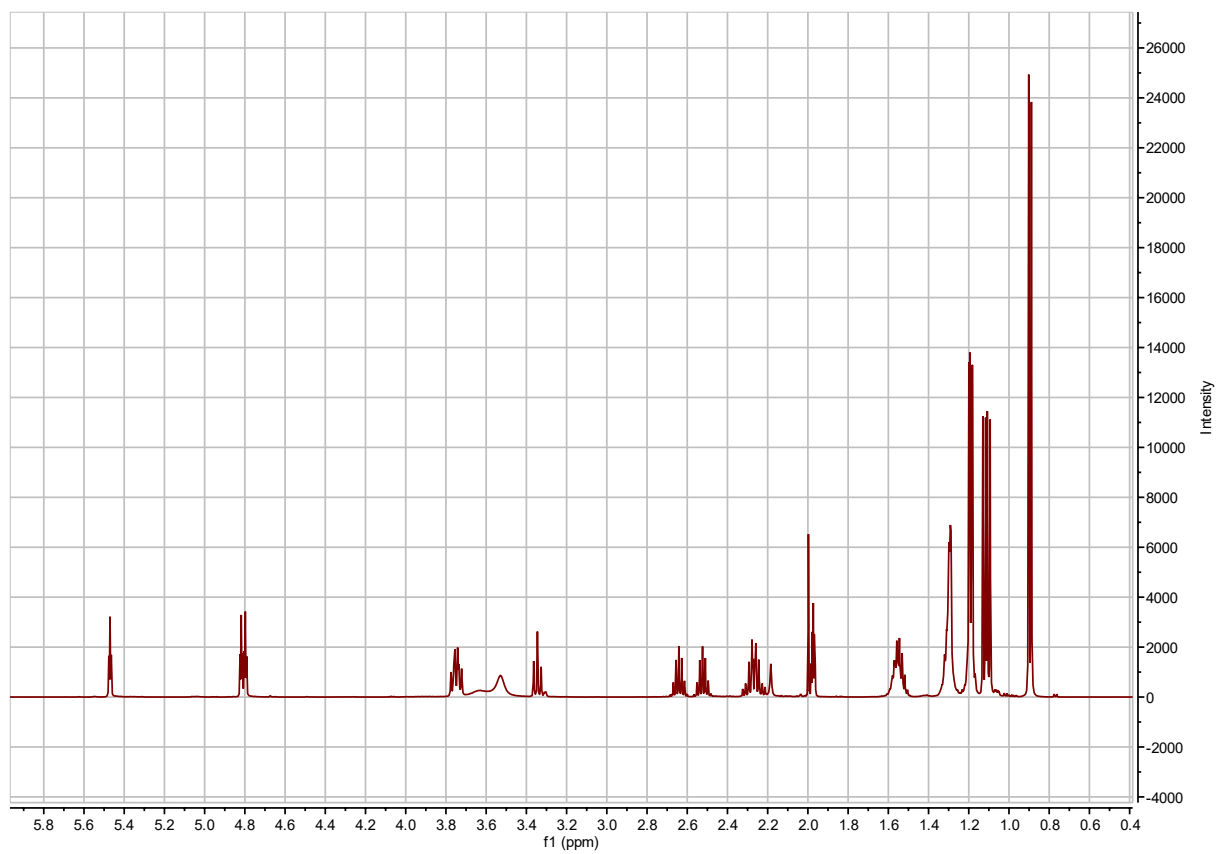
Figure S2.30. I3:18(i4,i4,i10) ^{1}H NMR.
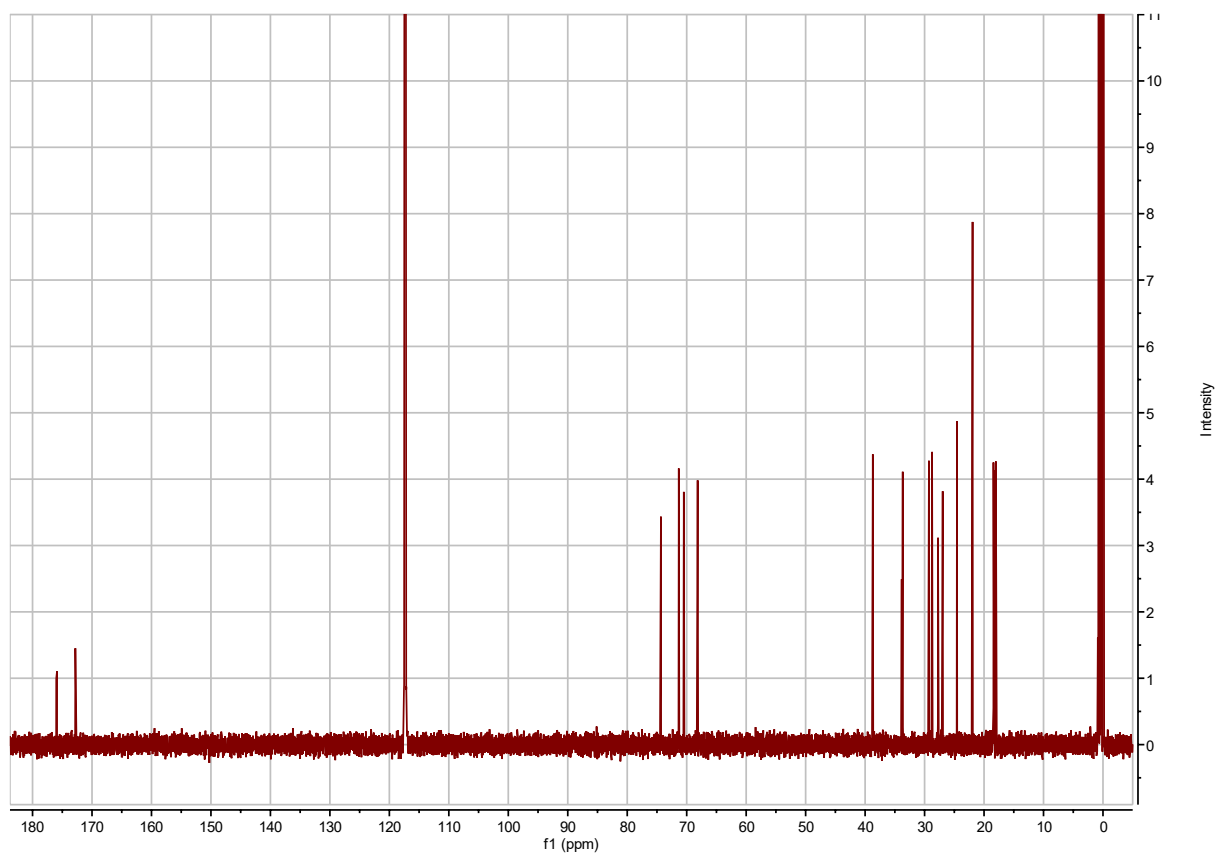
Figure S2.31. I3:18(i4,i4,i10) ¹³C NMR.
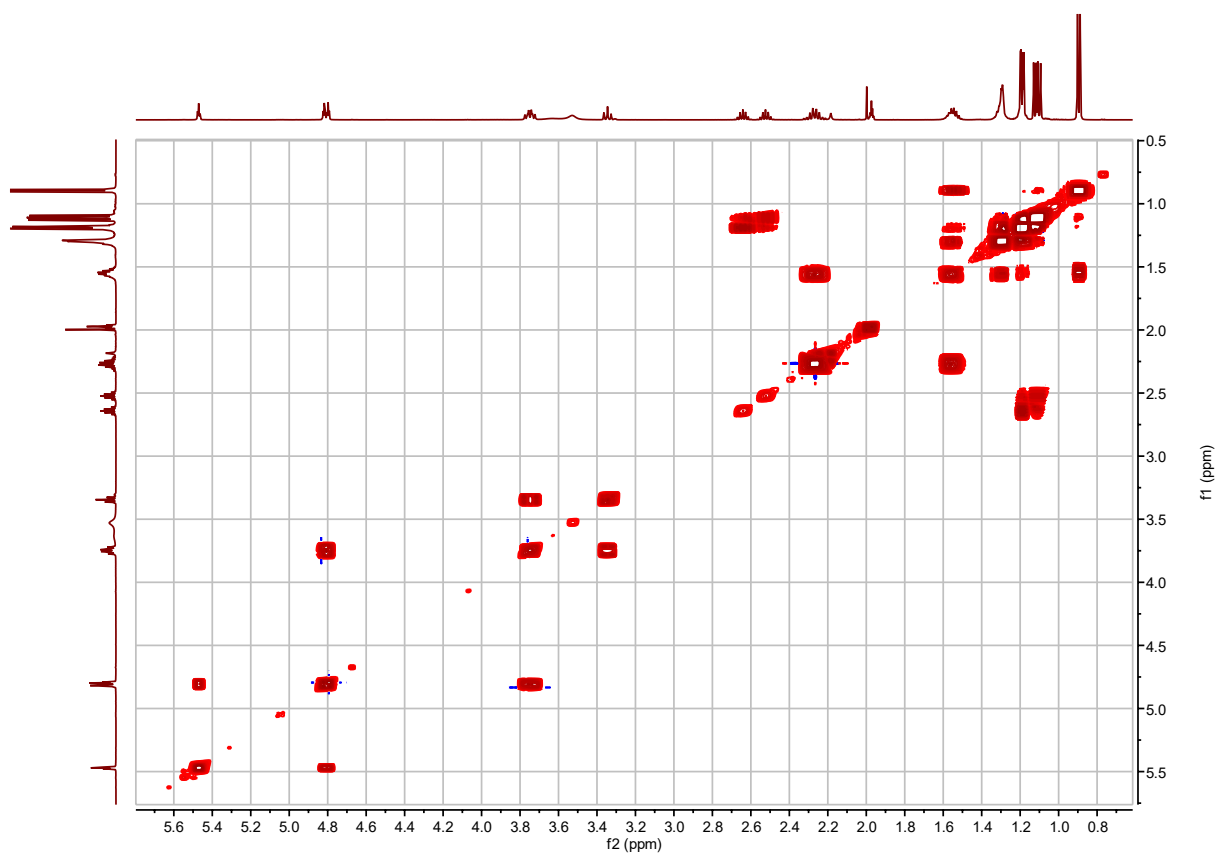
Figure S2.32. I3:18(i4,i4,i10) ¹H-¹H COSY.
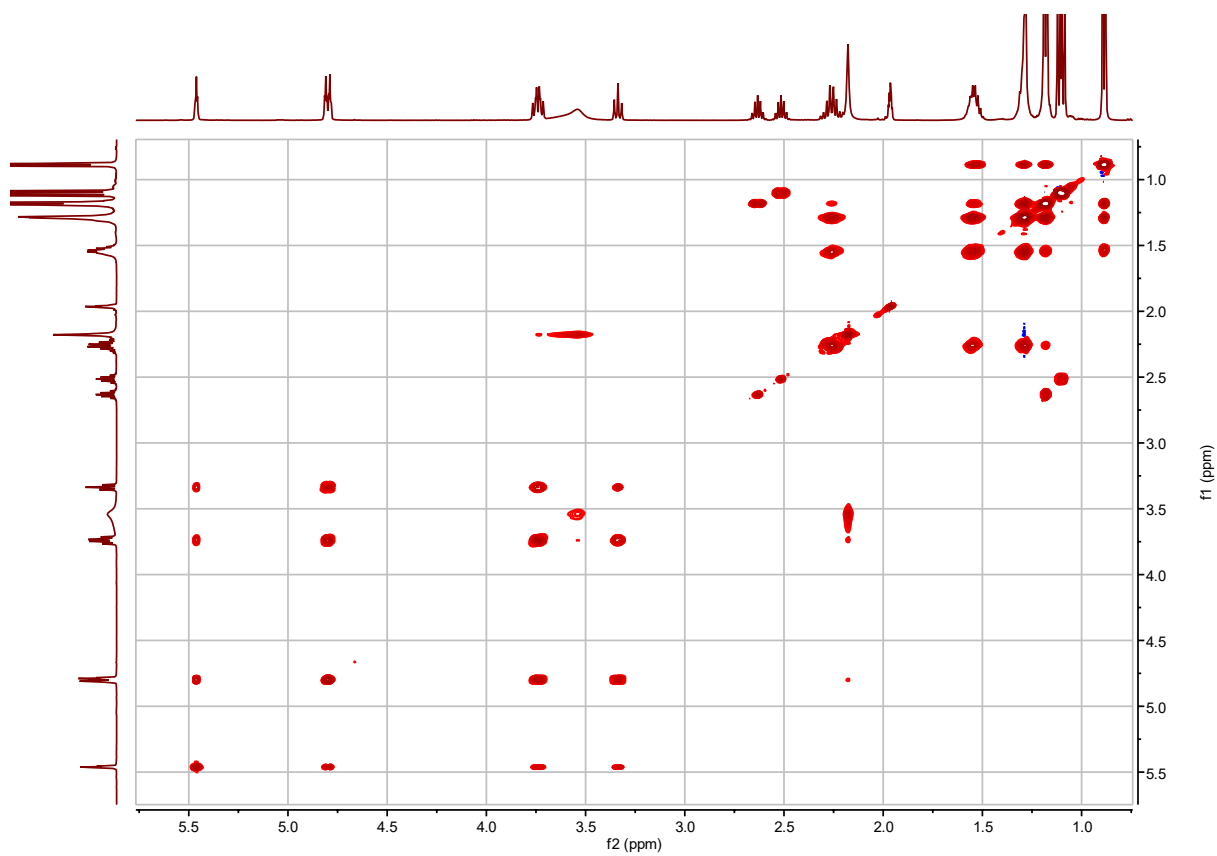
Figure S2.33. I3:18(i4,i4,i10) ^{1}H-^{1}H TOCSY.
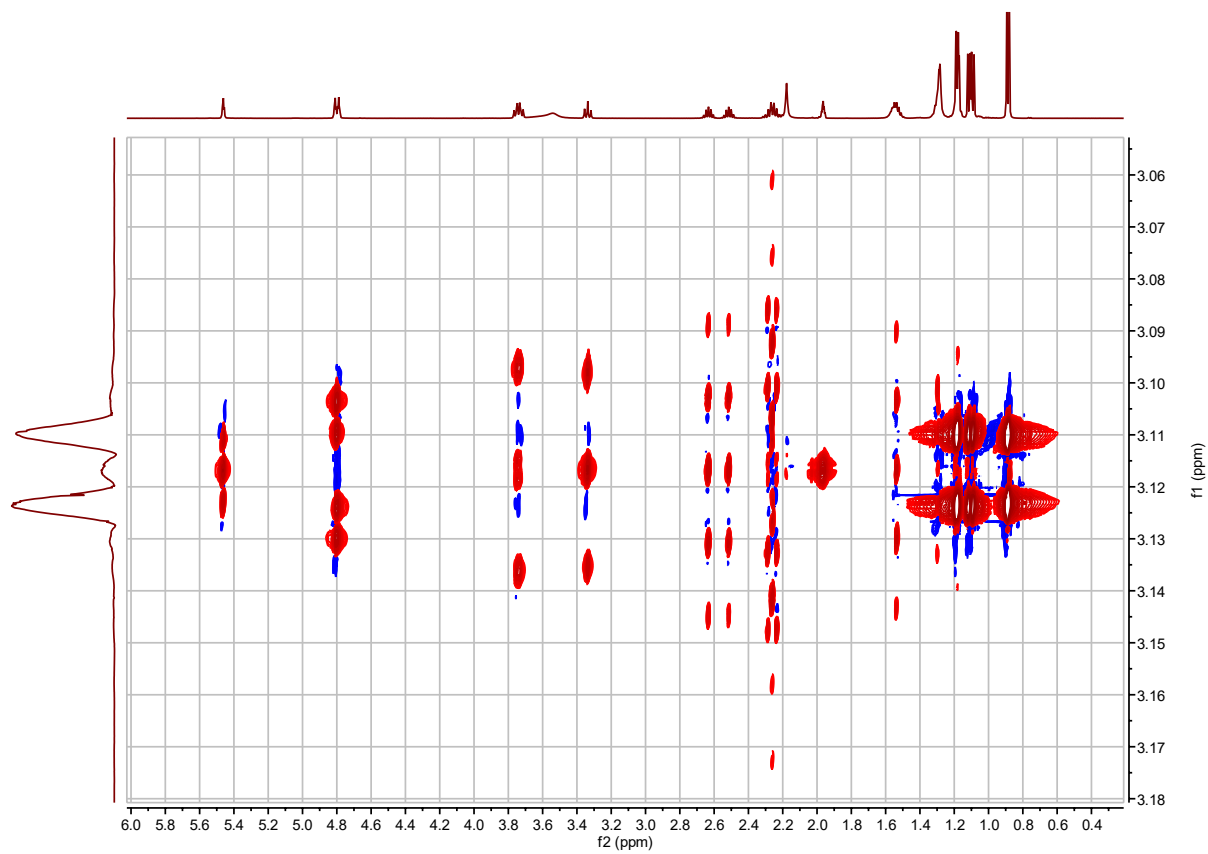
Figure S2.34. I3:18(i4,i4,i10) *J*-resolved.
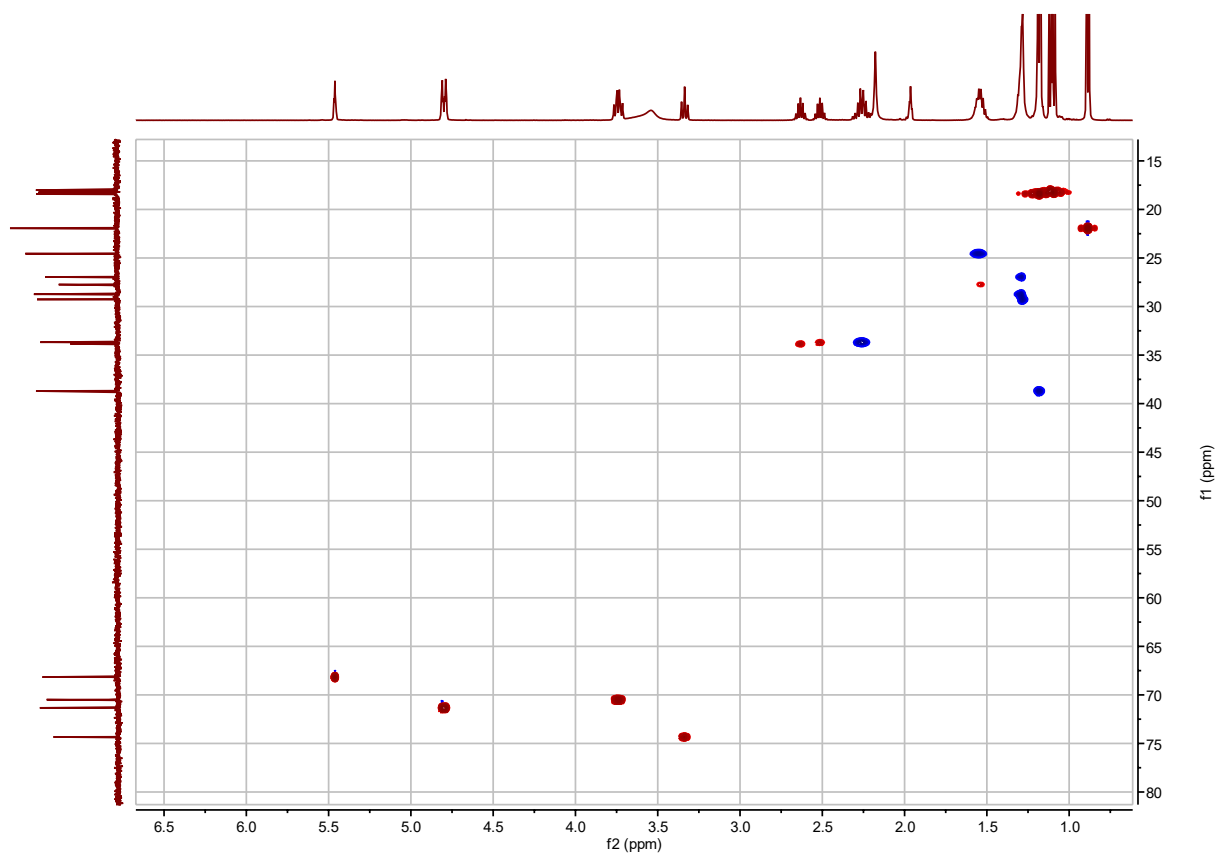
Figure S2.35. I3:18(i4,i4,i10) ^{1}H-^{13}C HSQC.
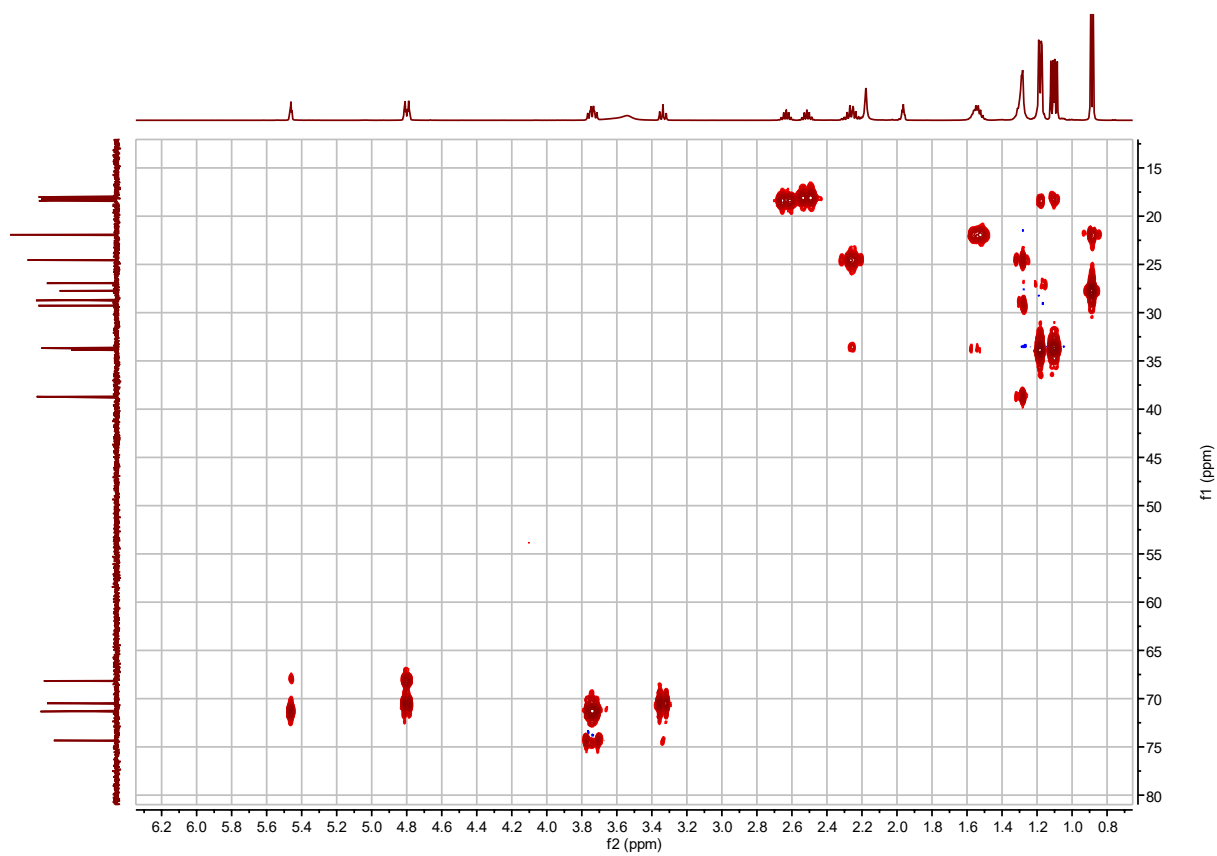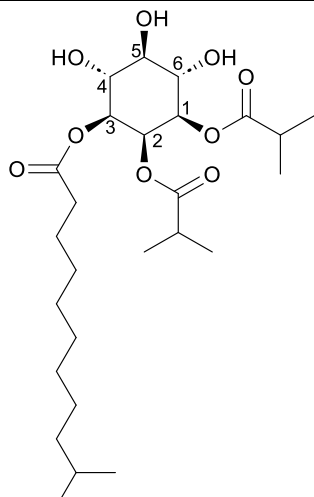
Figure S2.36. I3:18(i4,i4,i10) ^{1}H-^{13}C H2BC.
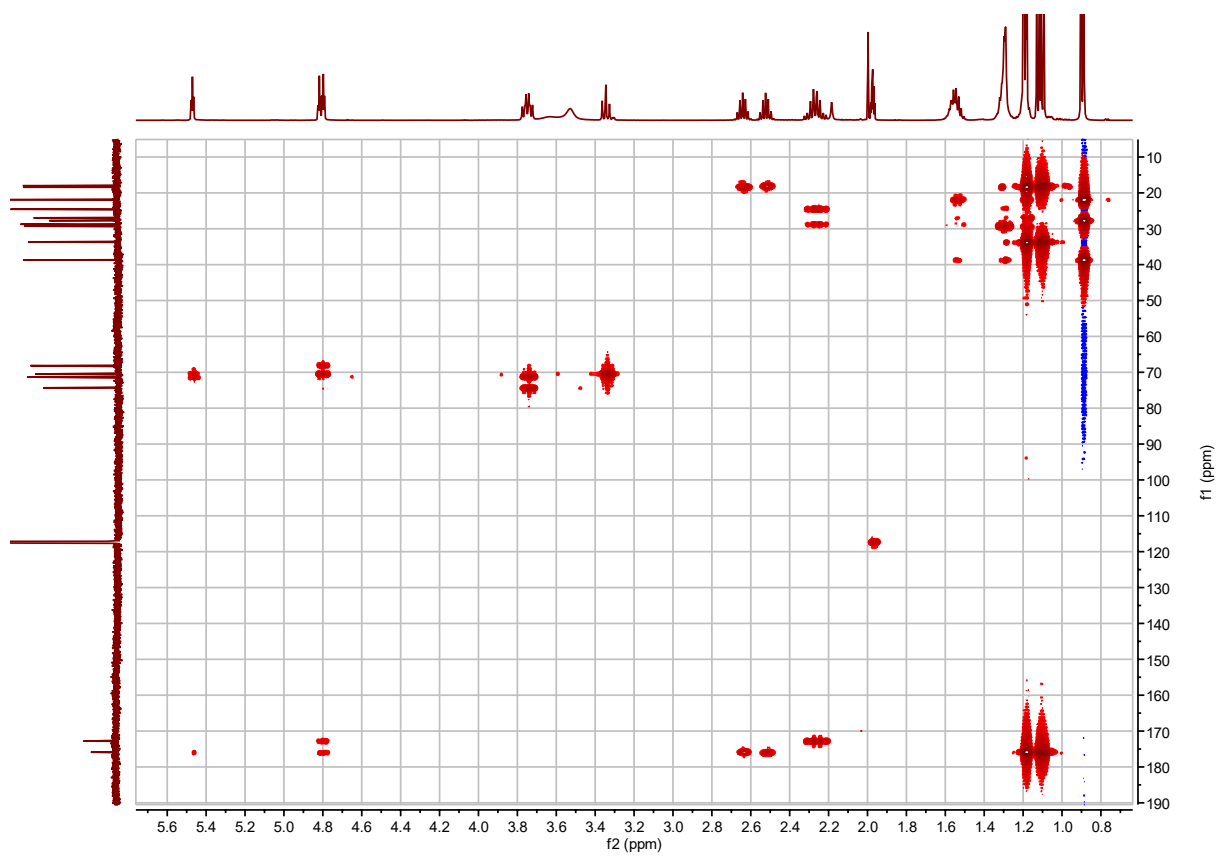
Figure S2.37. I3:18(i4,i4,i10) ^{1}H-^{13}C HMBC.

Table S2.41. I3:20(i4,i4,i12) chemical shifts and coupling constants.

	I3:20(i4,i4,i12)
	Molecular Formula: $C_{26}H_{46}O_9$
	Instrument: Agilent 500 MHz DDR2
	NMR solvent: CD_3CN
	Fractions: 105-107
	InChI Key: WFZVKWMBYUFKOS-WQRAYAPSSA-N
	SMILES: O[C@H]1[C@H](O)[C@@H](OC(CCCCCCCCC(C)C)=O)[C@@H](OC(C(C)C)=O)[C@@H](OC(C(C)C)=O)[C@@H]1O

Carbon # (Group)	^{1}H (δ, ppm)	^{13}C (δ, ppm)
1(CH)	4.80 or 4.79 (dd, J = 3.1, 10.1 Hz)	71.33 or 71.31[a]
-1(CO)		157.99
-2(CH)		33.69 or 33.67[a]
-3,4(CH$_3$)	2.51 (hept, J = 6.92 Hz) 1.11, 1.09 (d, J = 7.0 Hz)	18.21, 17.99
2(CH)	5.46 (t, J = 3.01 Hz)	68.18
-1(CO)		175.87
-2(CH)	2.62 (hept, J = 6.95 Hz)	33.84
-3,4(CH$_3$)	1.18 (d, J = 7.0 Hz)	18.39, 18.32
3(CH)	4.80 or 4.79 (dd, J = 3.1, 10.1 Hz)	71.33 or 71.31[a]
-1(CO)		172.77
-2(CH$_2$)		33.69 or 33.67[a]
-3(CH$_2$)	2.26 (m)	24.53
-4-8(CH$_2$)	1.54 (m) 1.33-1.27 (m)	28.68(4), 28.97(5), 29.18(6), 29.57 (7), 27.10(8)
-9(CH$_2$)		38.76
-10(CH)	1.19 (m)	27.73
-11,12(CH$_3$)	1.54 (m) 0.88 (d, J = 6.71 Hz)	21.91
4(CH)	3.75 or 3.73 (t, J = 9.9 Hz)	70.53 or 70.51[a]
5(CH)	3.33 (t, J = 9.30 Hz)	74.36
6(CH)	3.75 or 3.73 (t, J = 9.9 Hz)	70.53 or 70.51[a]

a – ^{13}C signals not resolved in 2D spectra.
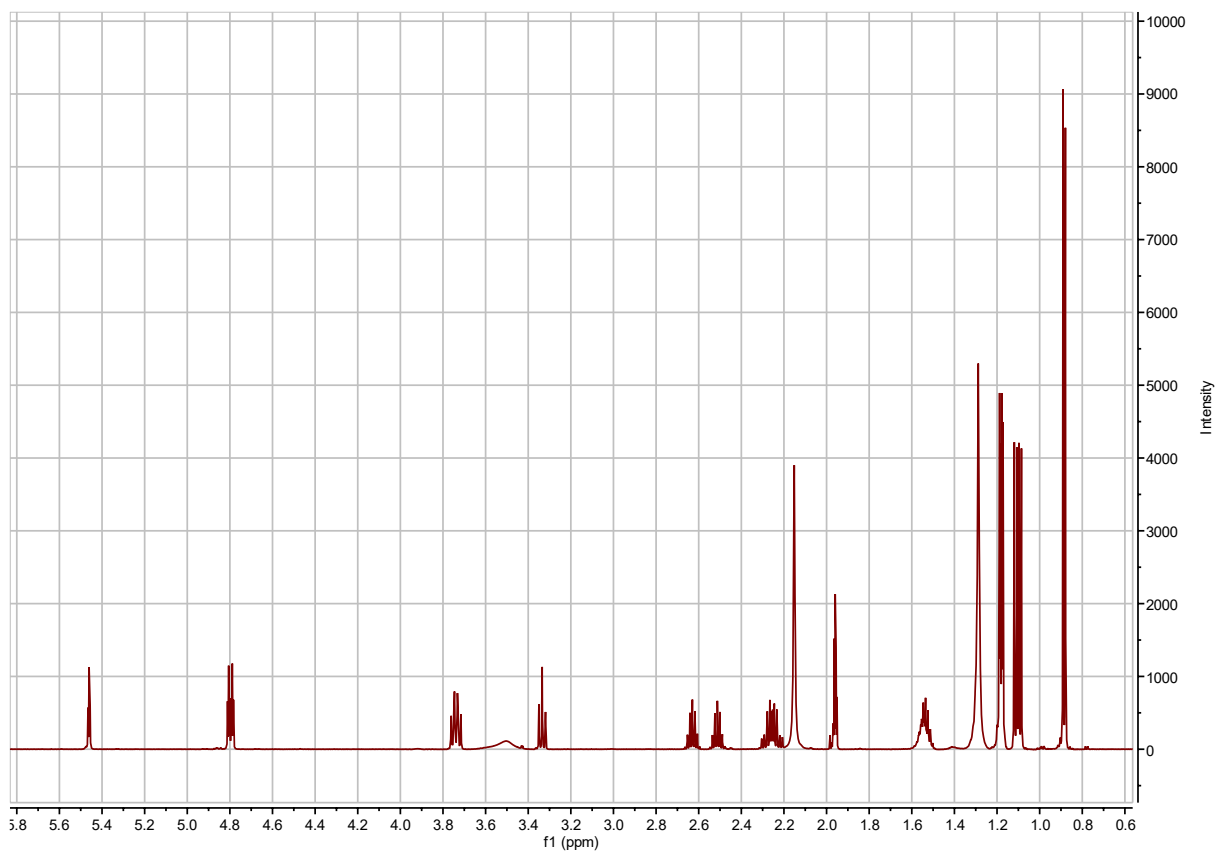
Figure S2.38. I3:20(i4,i4,i12) ¹H NMR.
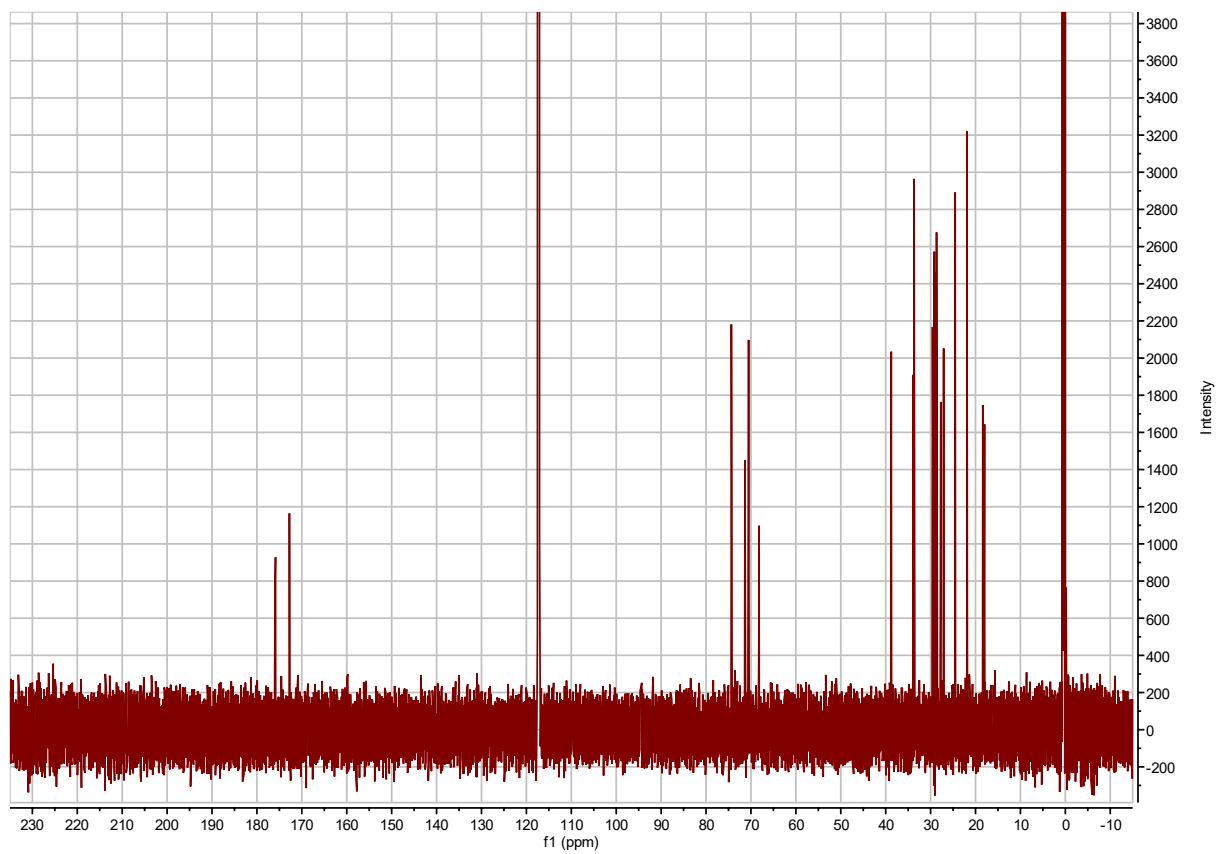
Figure S2.39. I3:20(i4,i4,i12) ¹³C NMR.
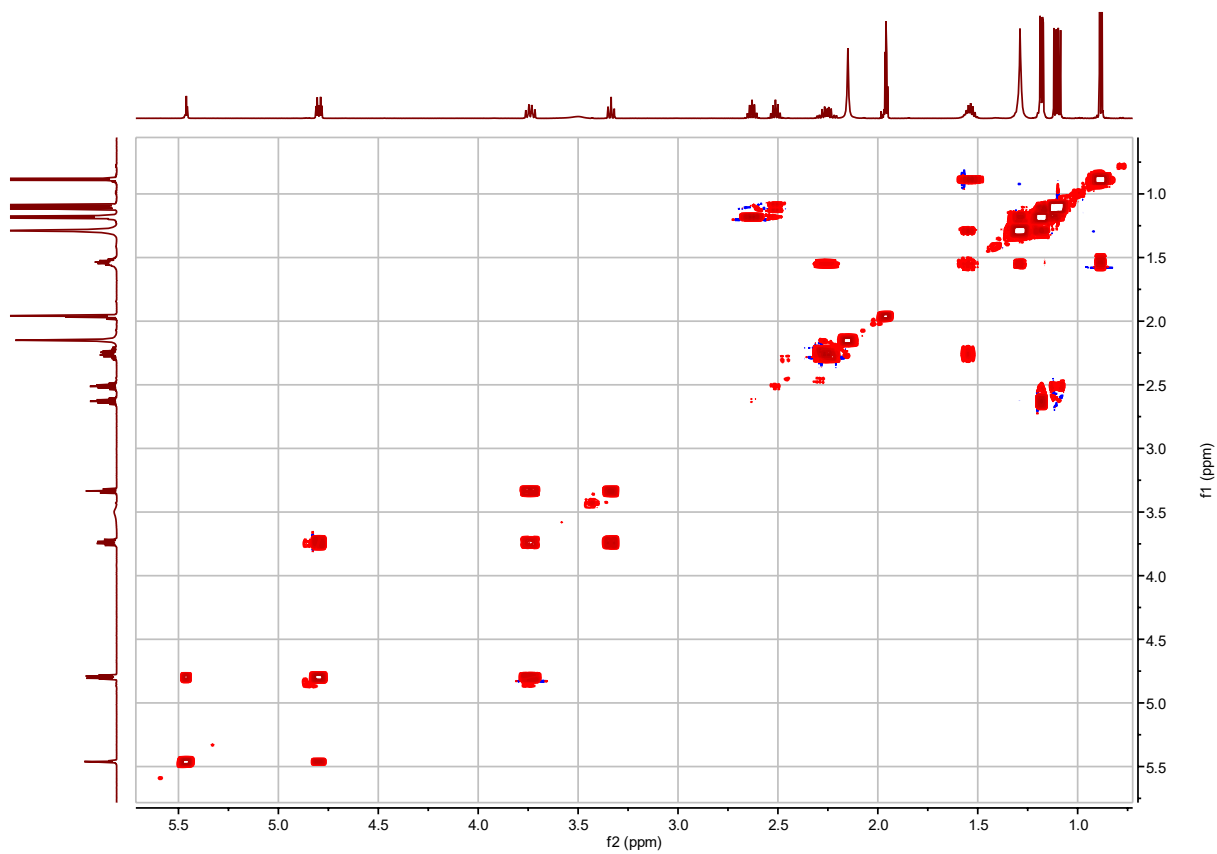
Figure S2.40. I3:20(i4,i4,i12) ^{1}H-^{1}H COSY.
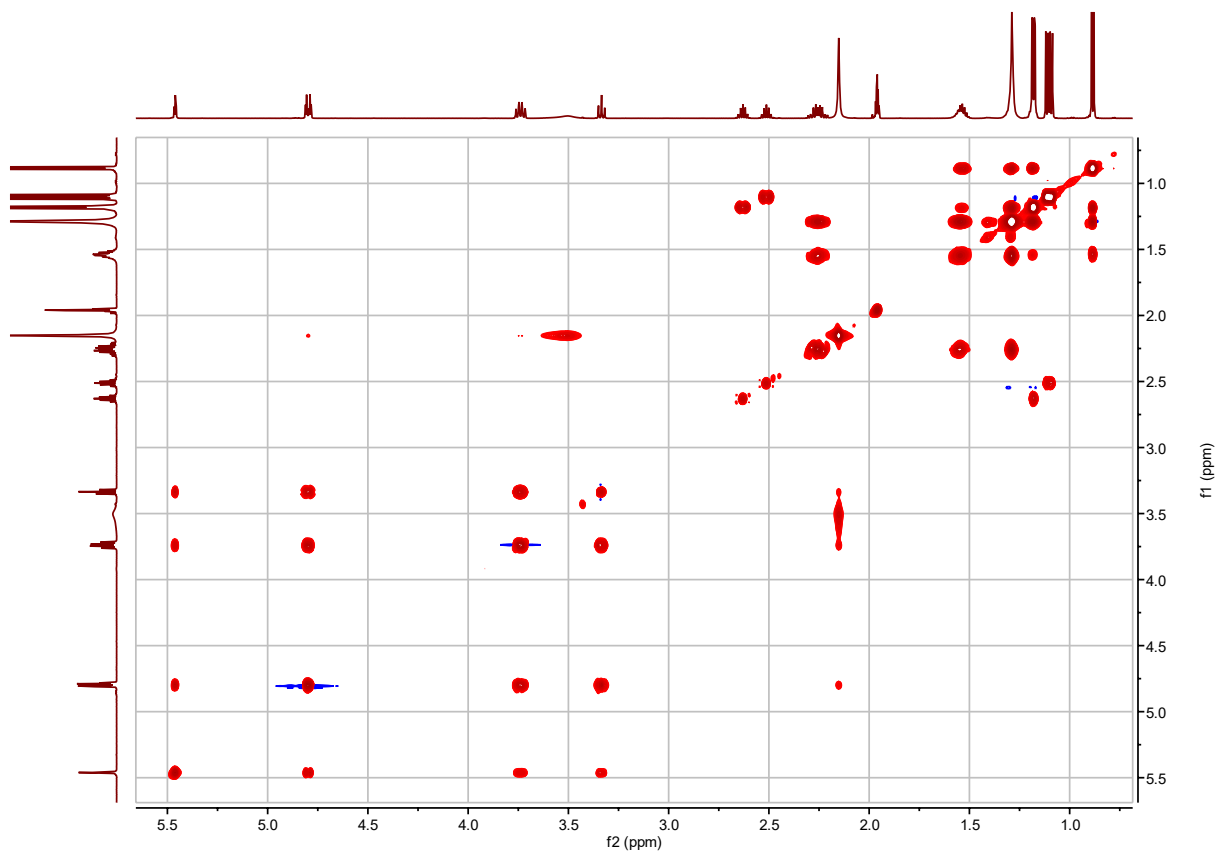
Figure S2.41. I3:20(i4,i4,i12) ¹H-¹H TOCSY.
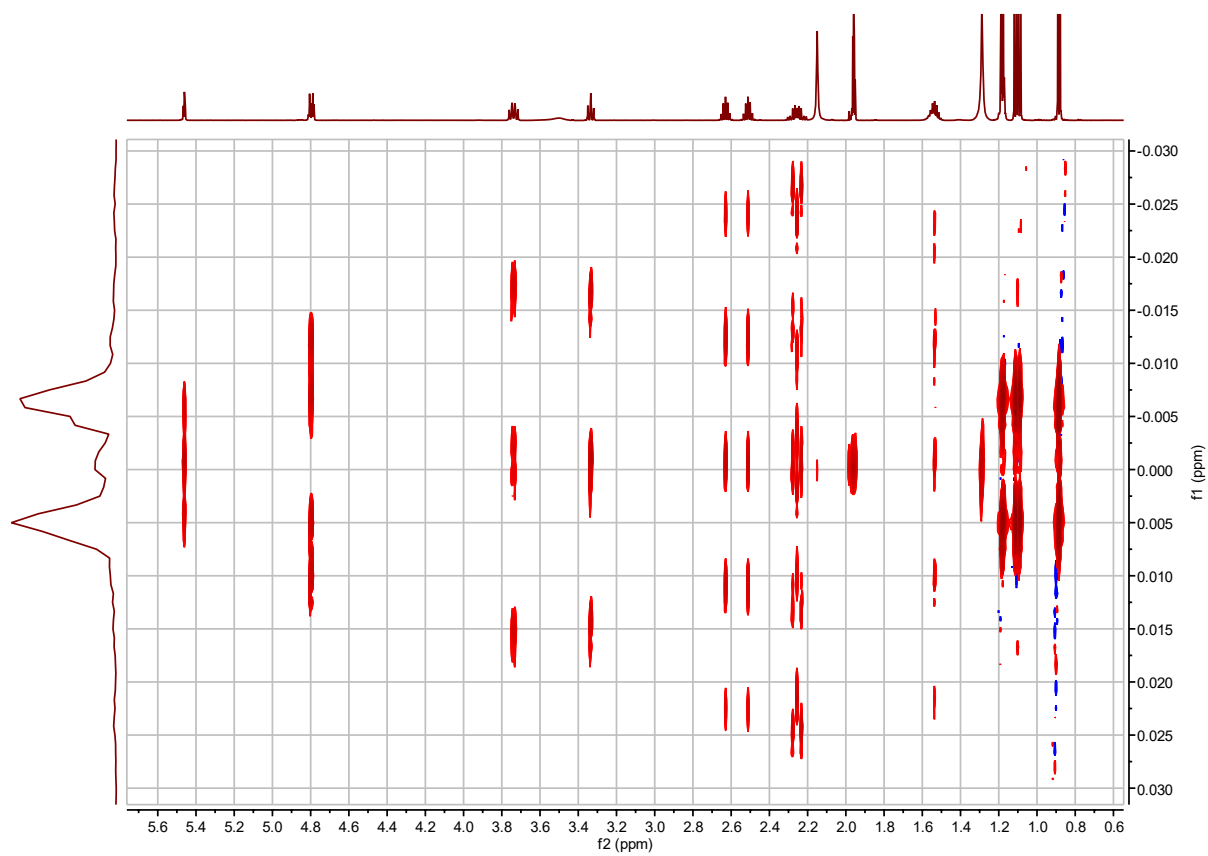
Figure S2.42. I3:20(i4,i4,i12) *J*-resolved.
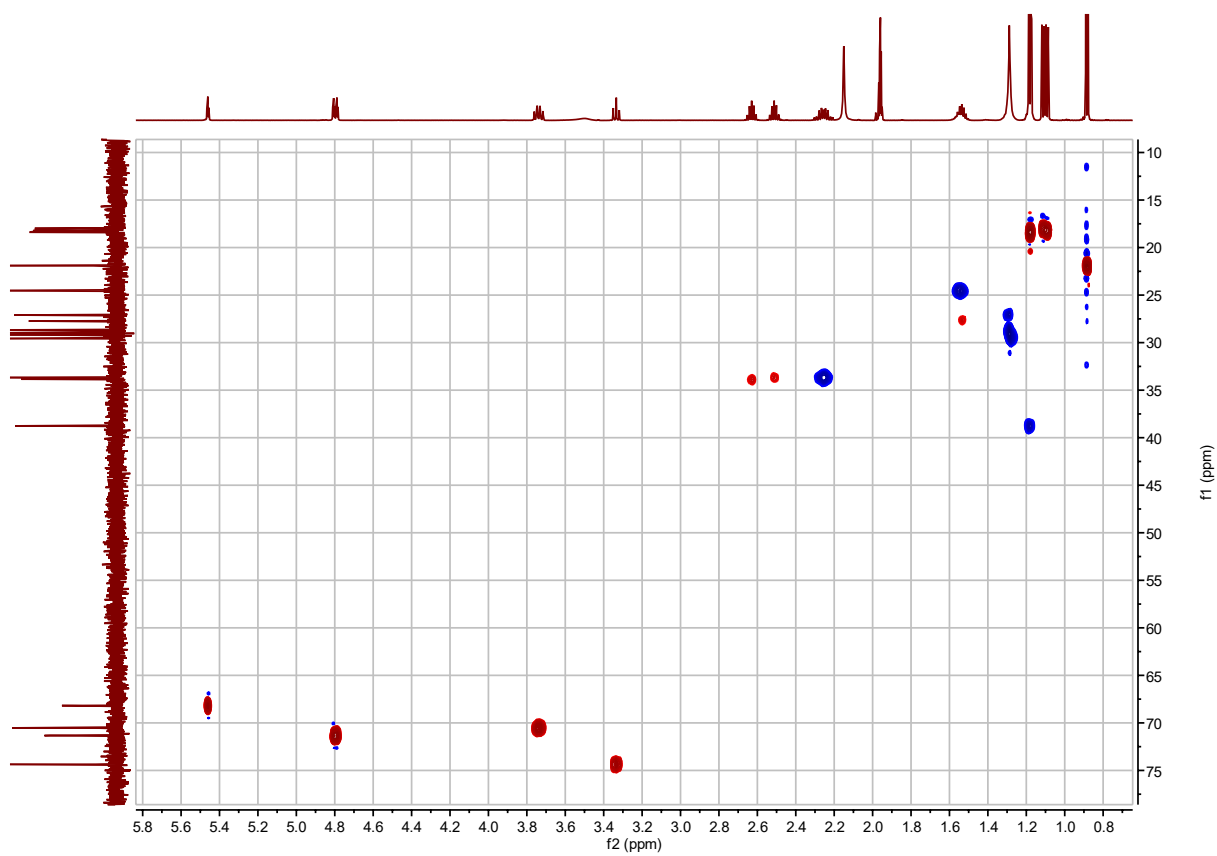
Figure S2.43. I3:20(i4,i4,i12) ^{1}H-^{13}C HSQC.
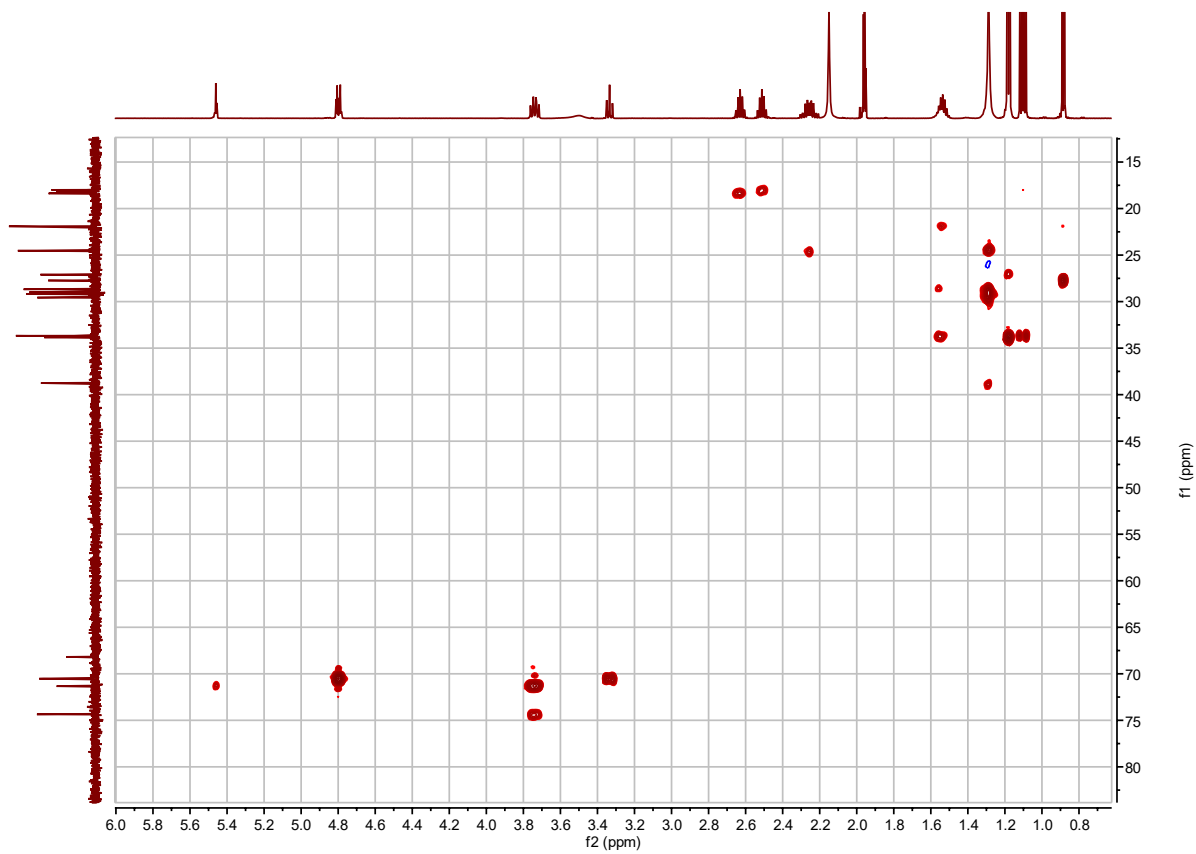
Figure S2.44. I3:20(i4,i4,i12) ¹H-¹³C H2BC.
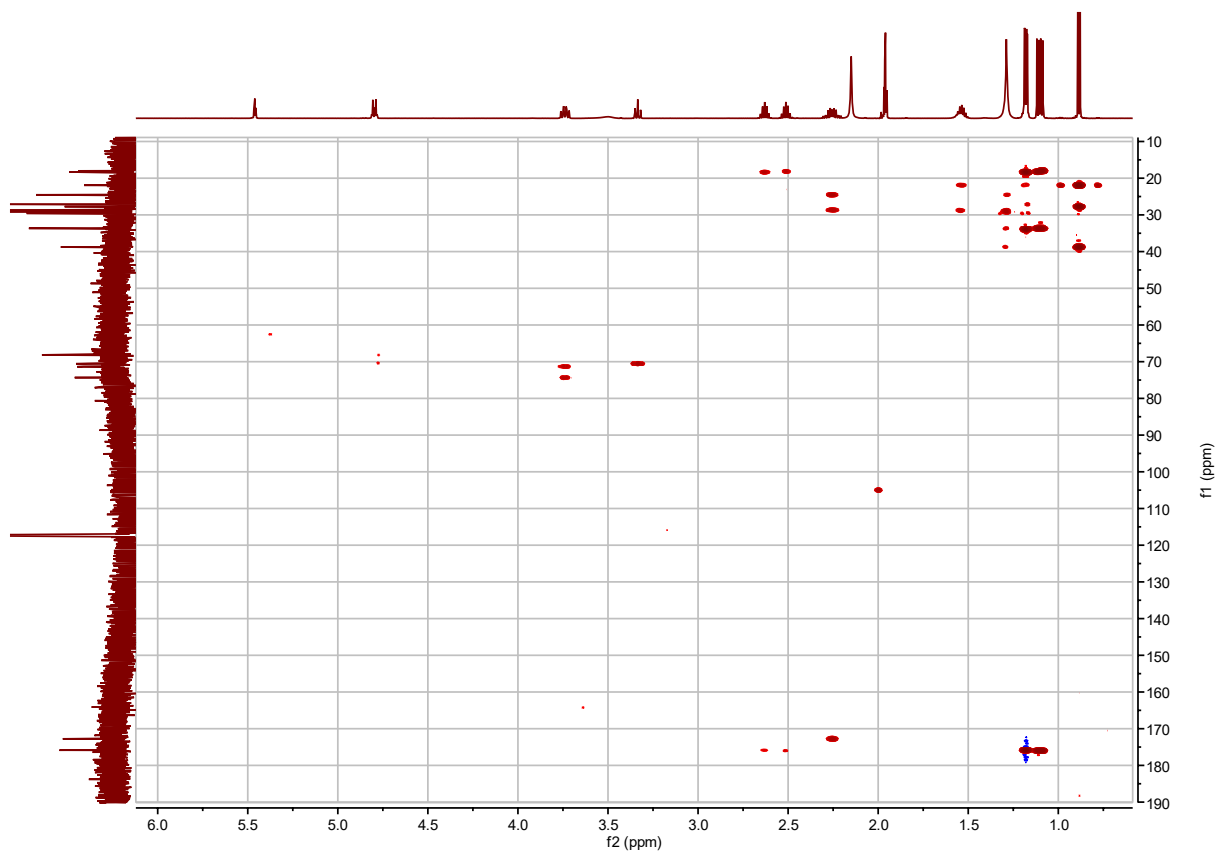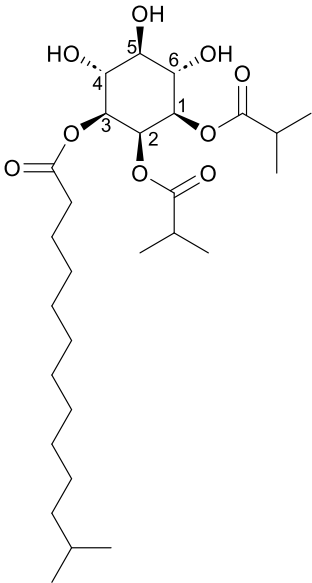
Figure S2.45. I3:20(i4,i4,i12) ¹H-¹³C HMBC with apodization optimized for correlations between acyl chain carbonyl carbon and C2 proton(s).
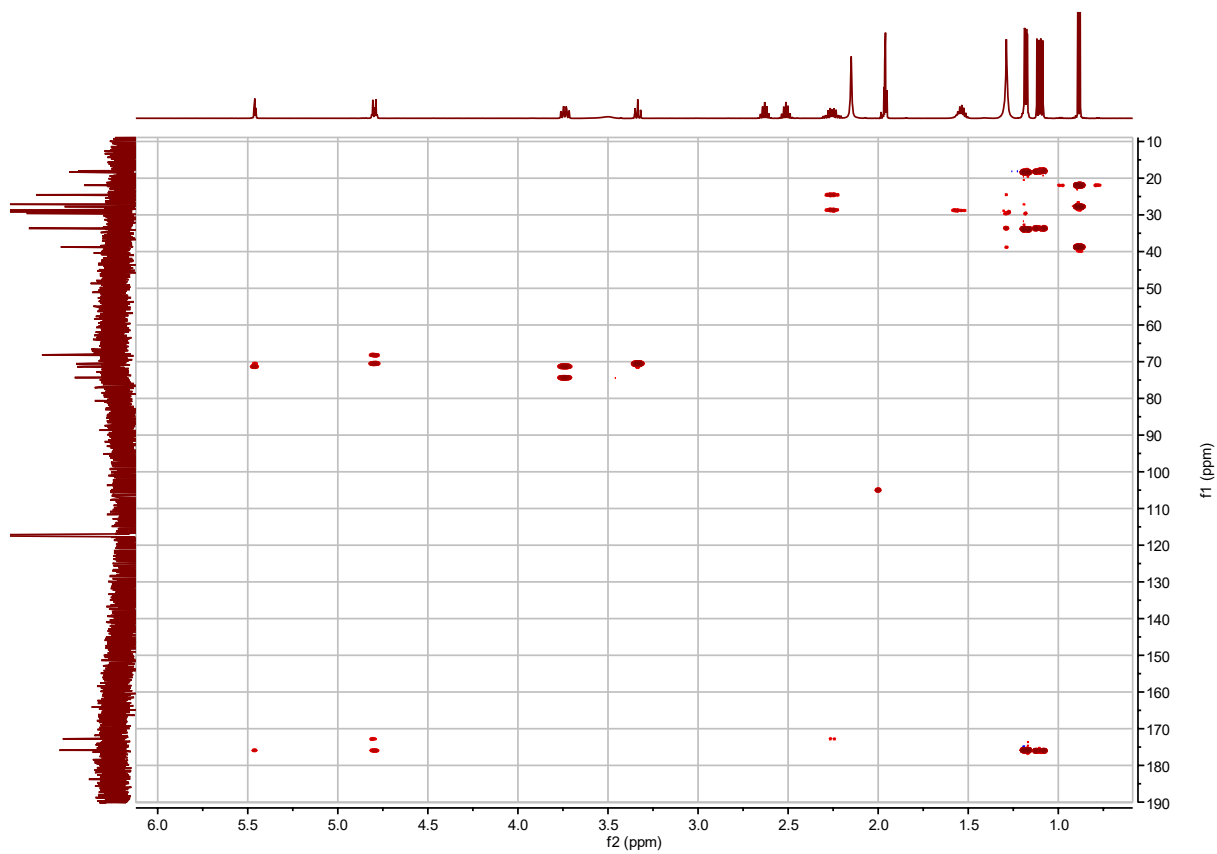
Figure S2.46. I3:20(i4,i4,i12) ¹H-¹³C HMBC with apodization optimized for correlations between acyl chain carbonyl carbon and sugar ring protons.

Table S2.42. I3:22(i4,i4,i14) chemical shifts and coupling constants.

<table>
<tr><td rowspan="3">

</td><td>I3:22(i4,i4,i14)</td></tr>
<tr><td>Molecular Formula: $C_{28}H_{50}O_9$</td></tr>
<tr><td>

Instrument: Varian 600 MHz Inova, Bruker Avance NEO 600 MHz NMR, Bruker Avance NEO 800 MHz

NMR solvent: CD_3CN

Fractions: 139-141

InChI Key: FPKCJTQXNJKTGE-KTKTTYTDSA-N

SMILES:
O[C@H]1[C@H](O)[C@@H](OC(CCCCCCCCCCC(C)C)=O)[C@@H](OC(C(C)C)=O)[C@@H](OC(C(C)C)=O)[C@@H]1O

</td></tr>
</table>

Carbon # (Group)	^1H (δ, ppm)	^{13}C (δ, ppm)
1(CH)	4.80 (dt, J = 3.25, 10.23 Hz)	71.34 or 71.31
-1(CO)		176.03 or 175.89[a]
-2(CH)	2.52 (hept, J = 6.96 Hz)	33.85
-3,4(CH_3)	1.11 (dd, J = 6.97, 13.05 Hz)	18.40 – 18.00
2(CH)	5.47 (t, J = 2.98 Hz)	68.18
-1(CO)		176.03 or 175.89[a]
-2(CH)	2.64 (hept, J = 6.96 Hz)	33.70
-3,4(CH_3)	1.19 (dd, J = 2.59, 6.96 Hz)	18.40 – 18.00
3(CH)	4.80 (dt, J = 3.25, 10.23 Hz)	71.34 or 71.31[a]
-1(CO)		172.81
-2(CH_2)	2.27 (m)	33.67
-3(CH_2)	1.55 (m)	24.53
-4(CH_2)	1.29 (m)	28.68
-5-10(CH_2)	1.30 (m)	29.63, 29.39, 29.36, 29.32, 29.16, 28.98
-11(CH_2)	1.19 (m)	38.78
-12(CH)	1.54 (m)	27.74
-13,14(CH_3)	0.89 (d, J = 6.64 Hz)	21.92
4(CH)	3.74 (td, J = 8.24, 9.85, 9.90 Hz)	70.53 or 70.50[a]
5(CH)	3.34 (t, J = 9.31 Hz)	74.36
6(CH)	3.74 (td, J = 8.24, 9.85, 9.90 Hz)	70.53 or 70.50[a]

a – ^{13}C signals not resolved in 2D spectra.
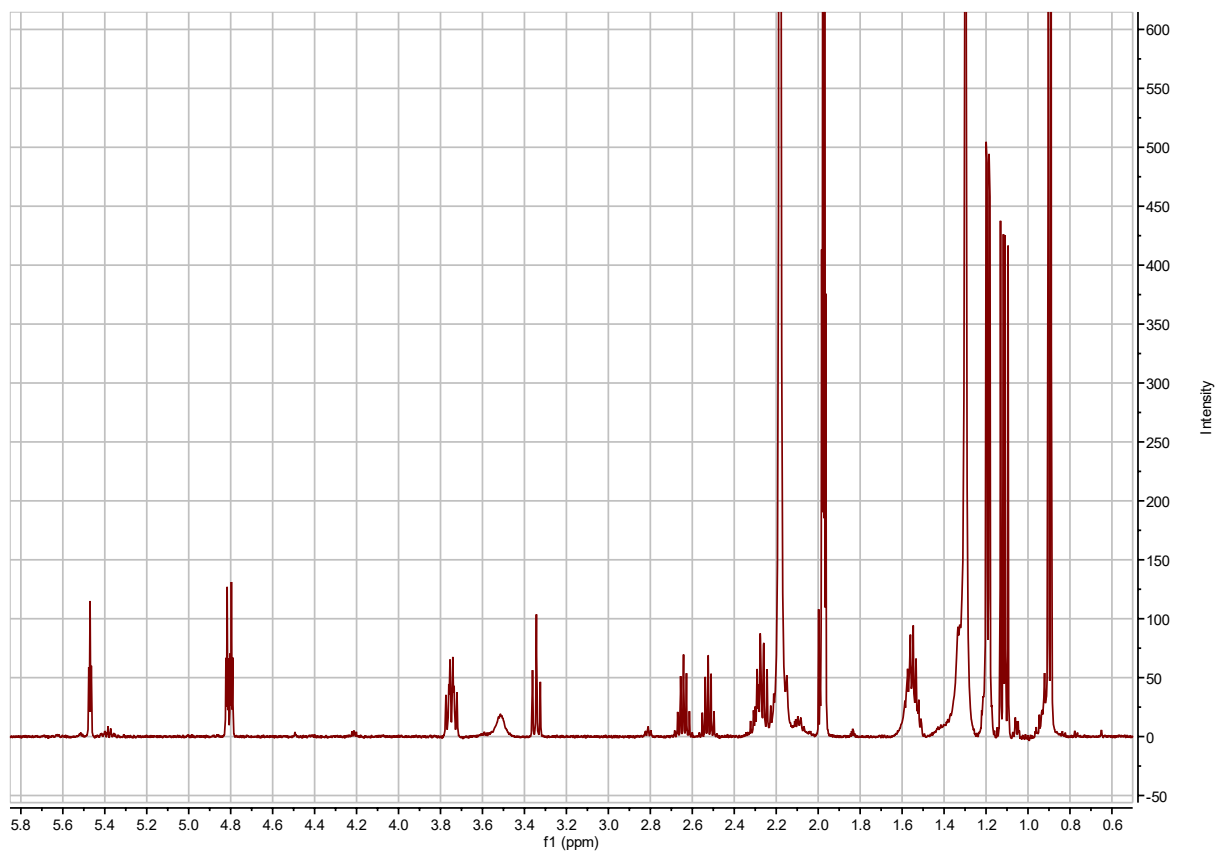
Figure S2.47. I3:22(i4,i4,i14) ^{1}H NMR.
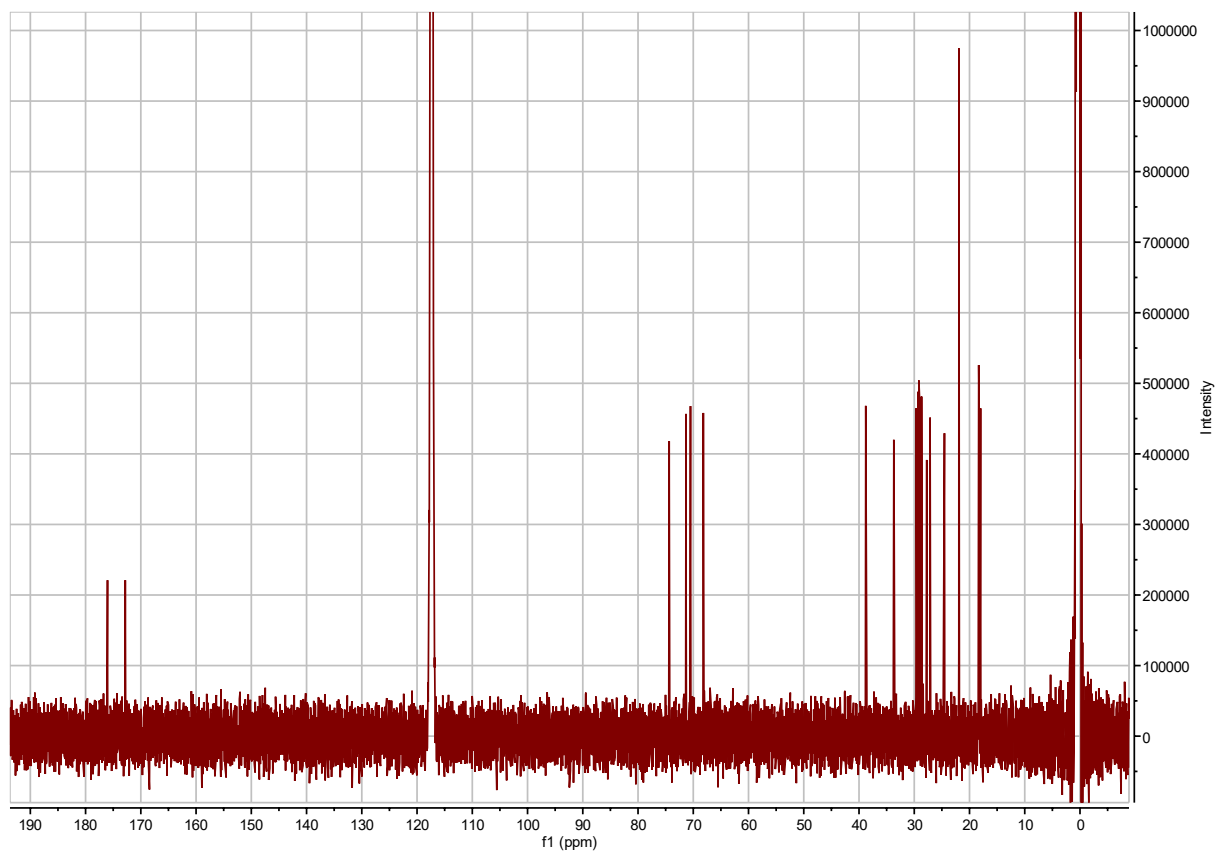
Figure S2.48. I3:22(i4,i4,i14) ^{13}C NMR.

Figure S2.49. I3:22(i4,i4,i14) ¹H-¹H COSY.
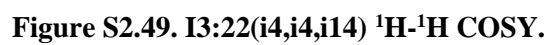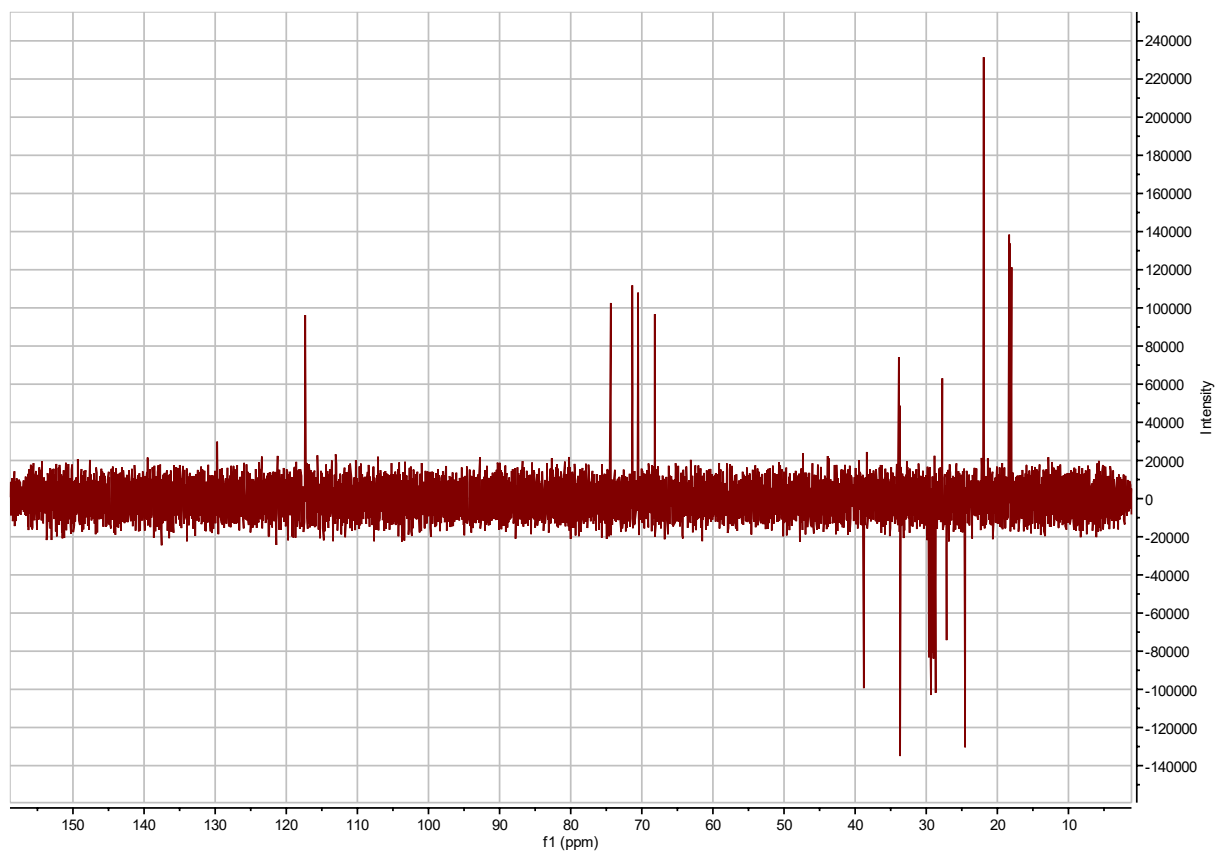
Figure S2.50. I3:22(i4,i4,i14) ^{13}C DEPT-Q.
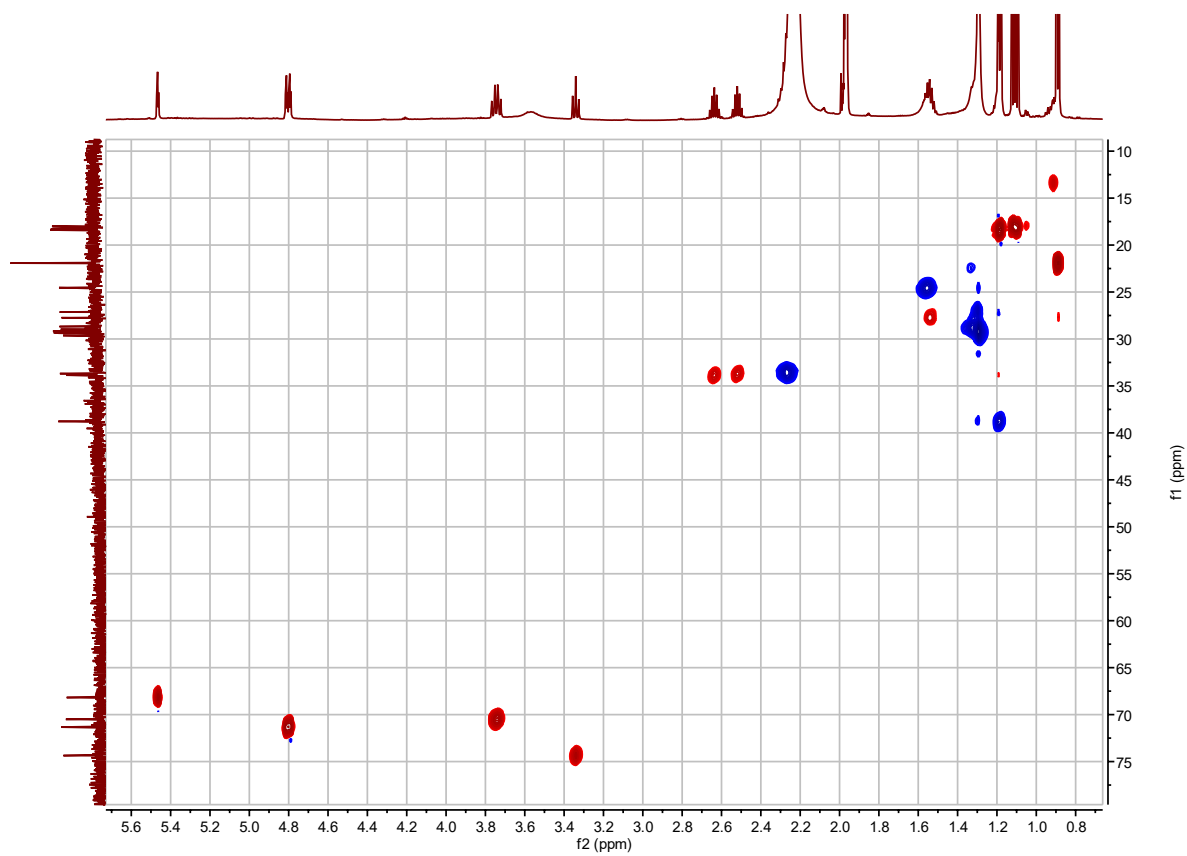
Figure S2.51. I3:22(i4,i4,i14) ^{1}H-^{13}C HSQC.
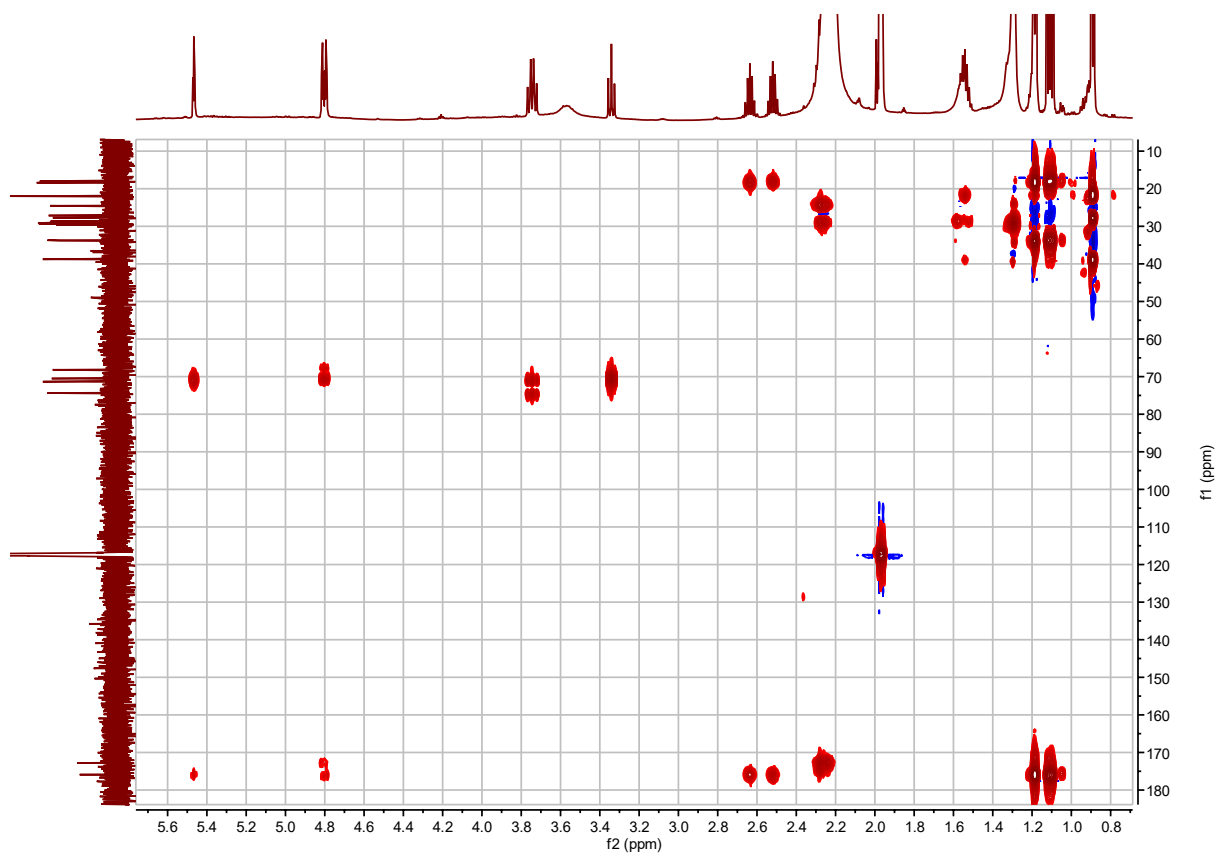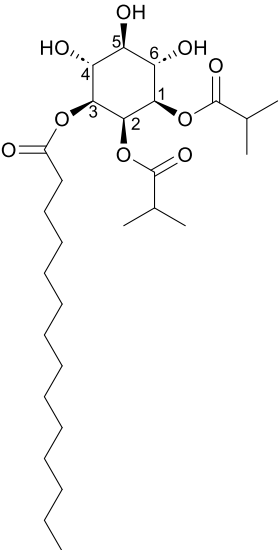
Figure S2.52. I3:22(i4,i4,i14) ^{1}H-^{13}C HMBC.

Table S2.43. I3:22(i4,i4,n14) chemical shifts and coupling constants.

	I3:22(i4,i4,n14)
	Molecular Formula: $C_{28}H_{50}O_9$ Instrument: Bruker Avance NEO 800 MHz NMR NMR solvent: CD_3CN Fractions: 144-146 InChI Key: ZHGFFHULNGWPKL-KTKTTYTDSA-N SMILES: O[C@H]1[C@H](O)[C@@H](OC(CCCCCCCCCCCCC)=O)[C@@H](OC(C(C)C)=O)[C@@H](OC(C(C)C)=O)[C@@H]1O

Carbon # (Group)	^1H (δ, ppm)	^{13}C (δ, ppm)
1(CH)	4.80 (ddd, J = 2.98, 4.60, 10.20 Hz)	71.89 or 71.87[a]
-1(CO)		176.56
-2(CH)	2.64 (hept, J = 6.97 Hz)	34.41
-3,4(CH₃)	1.18 (dd, J = 3.48, 6.98 Hz)	18.96, 18.90
2(CH)	5.46 (t, J = 2.97 Hz)	68.72
-1(CO)		176.43
-2(CH)	2.52 (hept, J = 6.97 Hz)	34.26 or 34.23[a]
-3,4(CH₃)	1.11 (dd, J = 6.98, 17.92 Hz)	18.79, 18.56
3(CH)	4.80 (ddd, J = 2.98, 4.60, 10.20 Hz)	71.89 or 71.87[a]
-1(CO)		173.34
-2(CH₂)	2.26 (m)	34.26 or 34.23[a]
-3(CH₂)	1.55 (m)	25.10
-4(CH₂)	1.29 (m)	32.22
-5-12(CH₂)	1.30 (m)	29.95, 29.93, 29.89, 29.73, 29.65, 29.55, 29.24
-13(CH₂)	1.33 (m)	22.97
-14(CH₃)	0.91 (m)	13.96
4(CH)	3.74 (q, J = 10.23 Hz)	71.07 or 71.05[a]
5(CH)	3.34 (t, J = 9.33 Hz)	74.90
6(CH)	3.74 (q, J = 10.23 Hz)	71.07 or 71.05[a]

a – ^{13}C signals not resolved in 2D spectra.
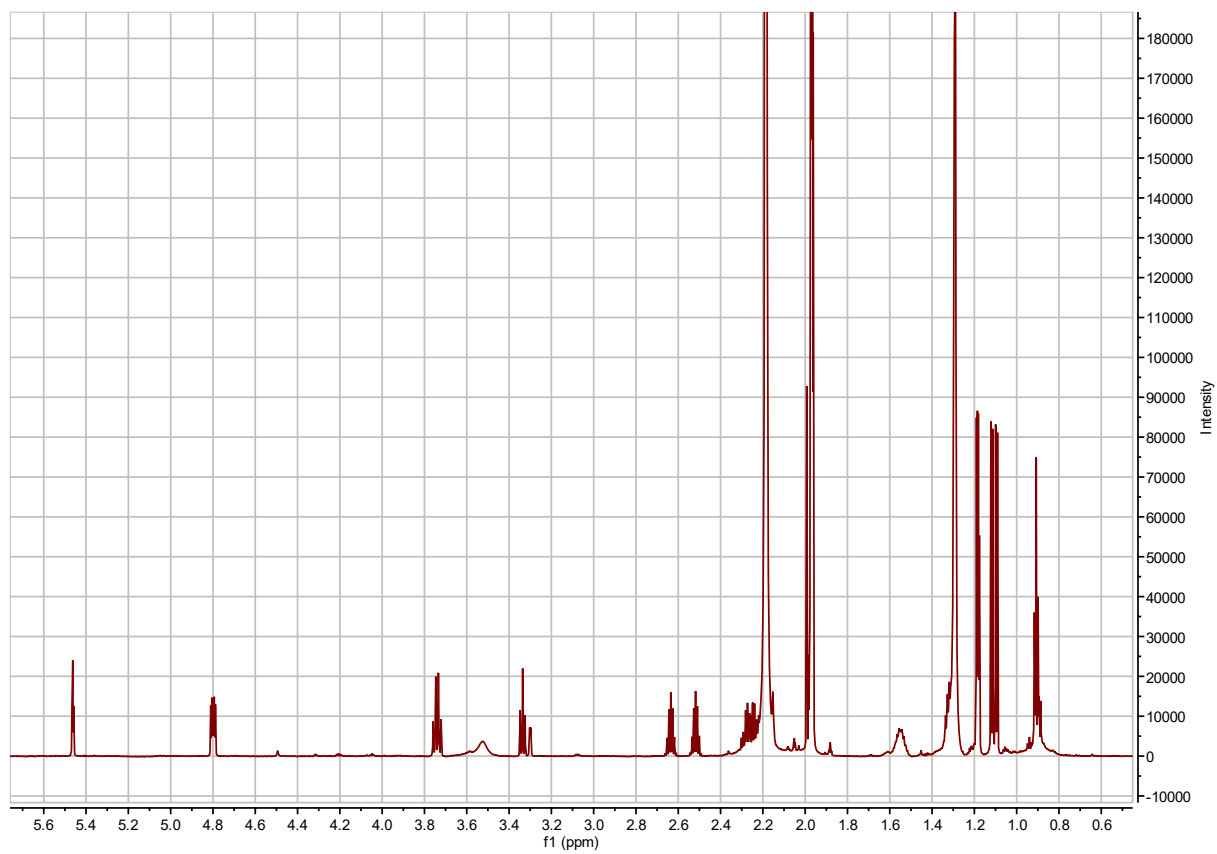
Figure S2.53. I3:22(i4,i4,n14) ¹H NMR.
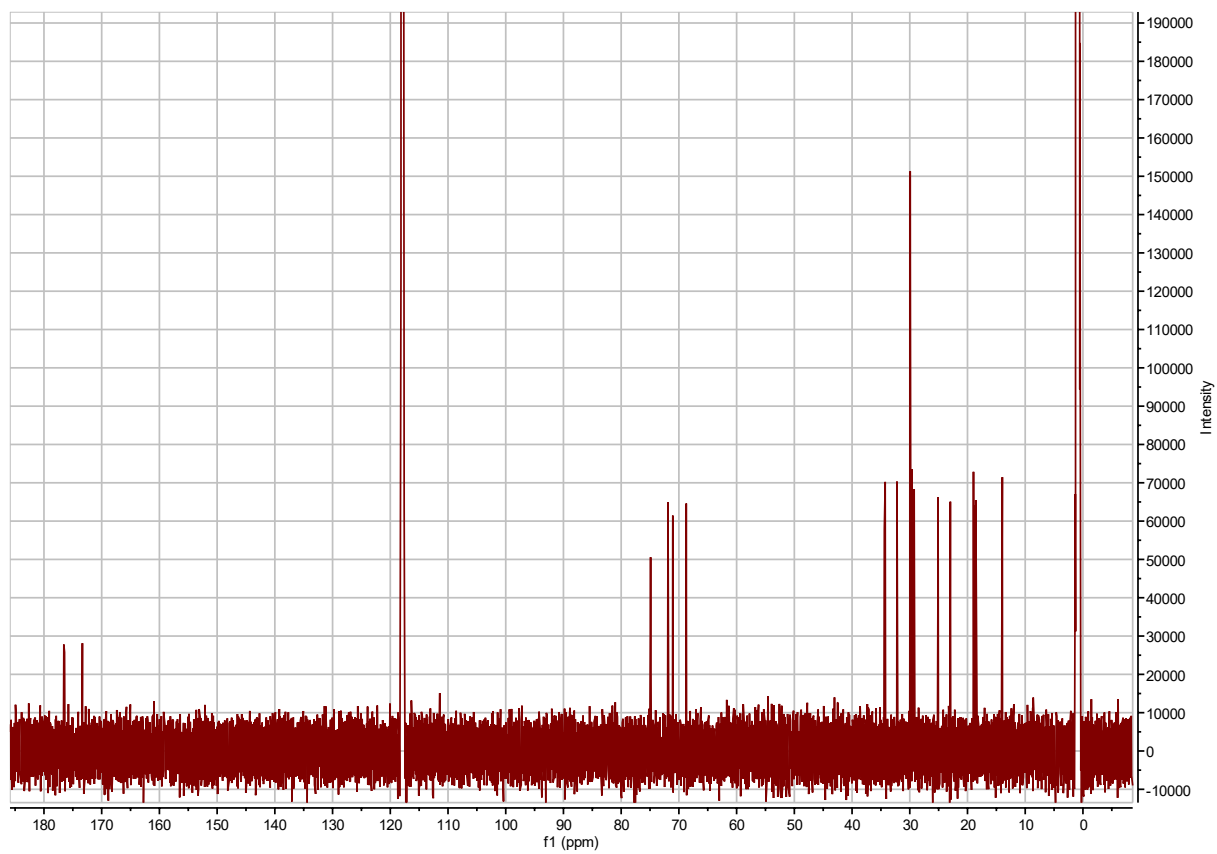
Figure S2.54. I3:22(i4,i4,n14) ¹³C NMR.
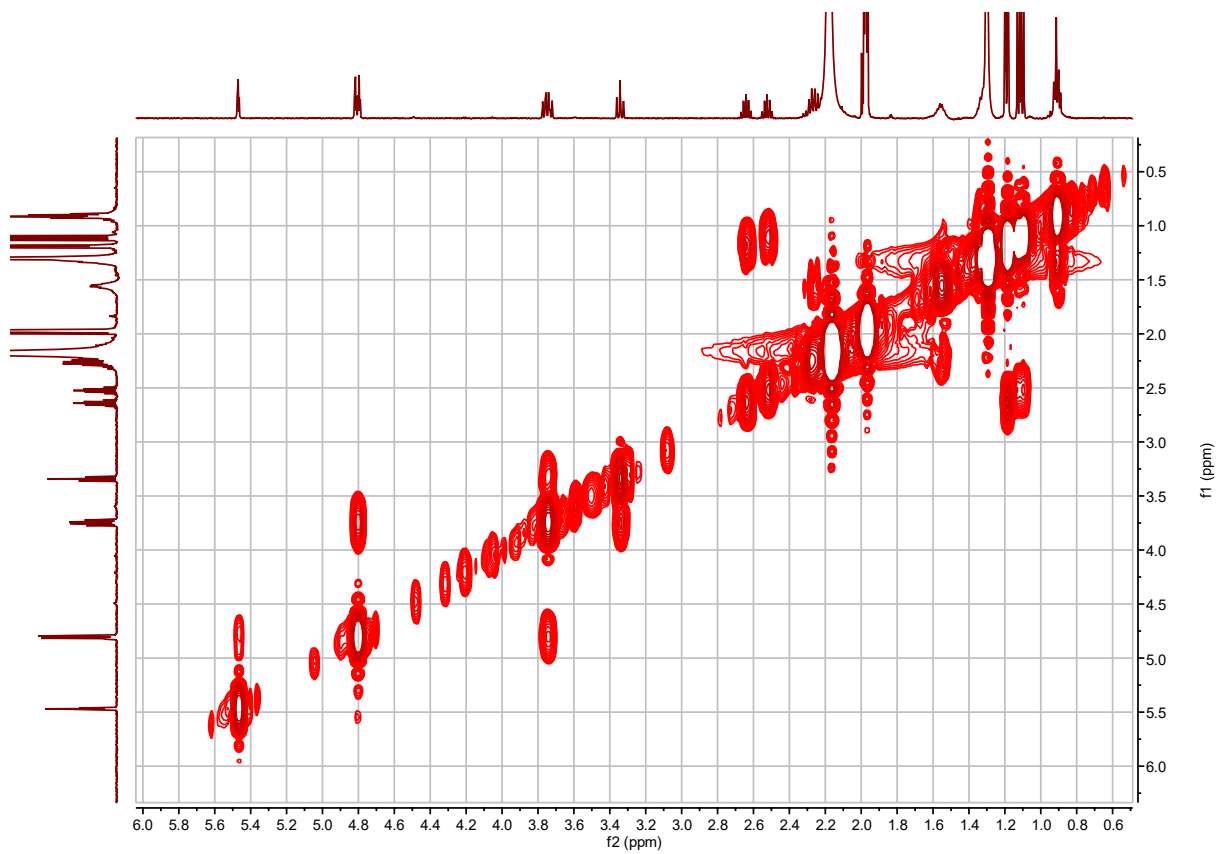
Figure S2.55. I3:22(i4,i4,n14) ¹H-¹H COSY.
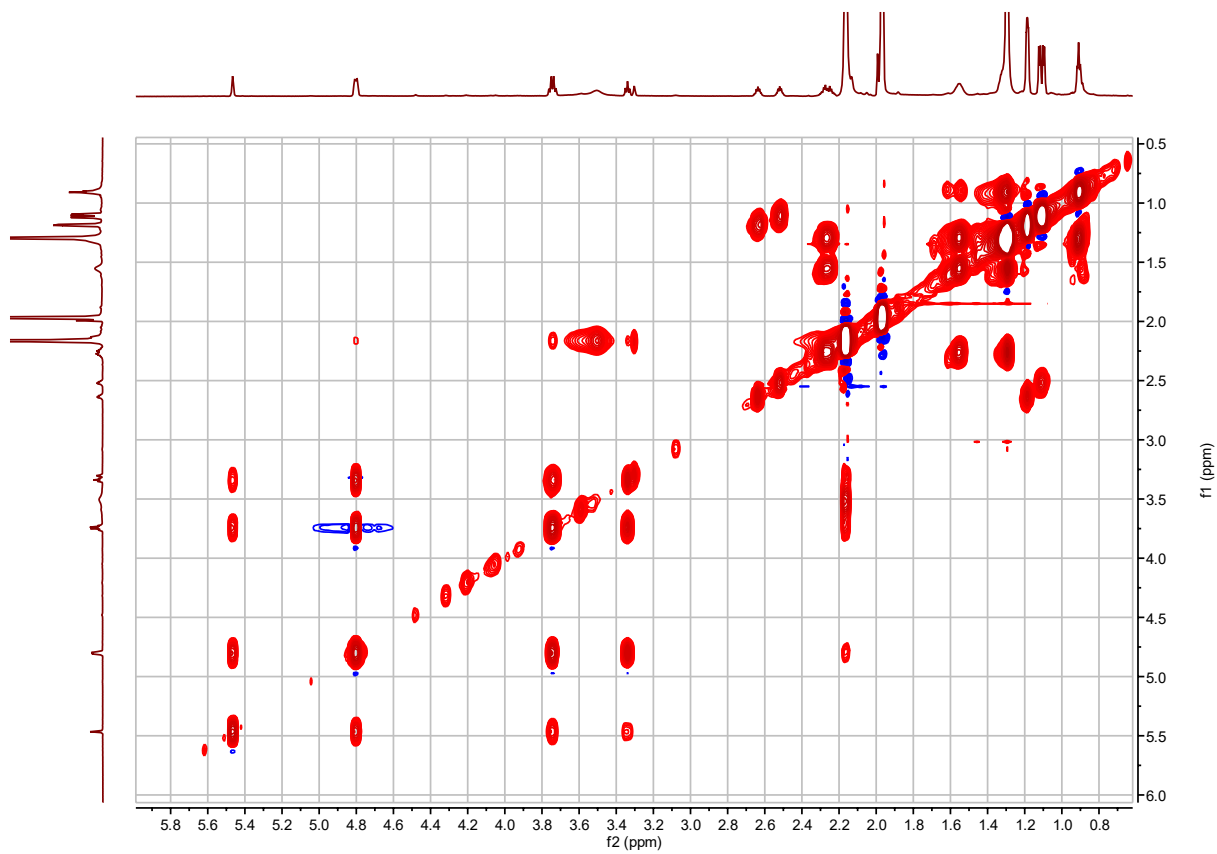
Figure S2.56. I3:22(i4,i4,n14) ^{1}H-^{1}H TOCSY.
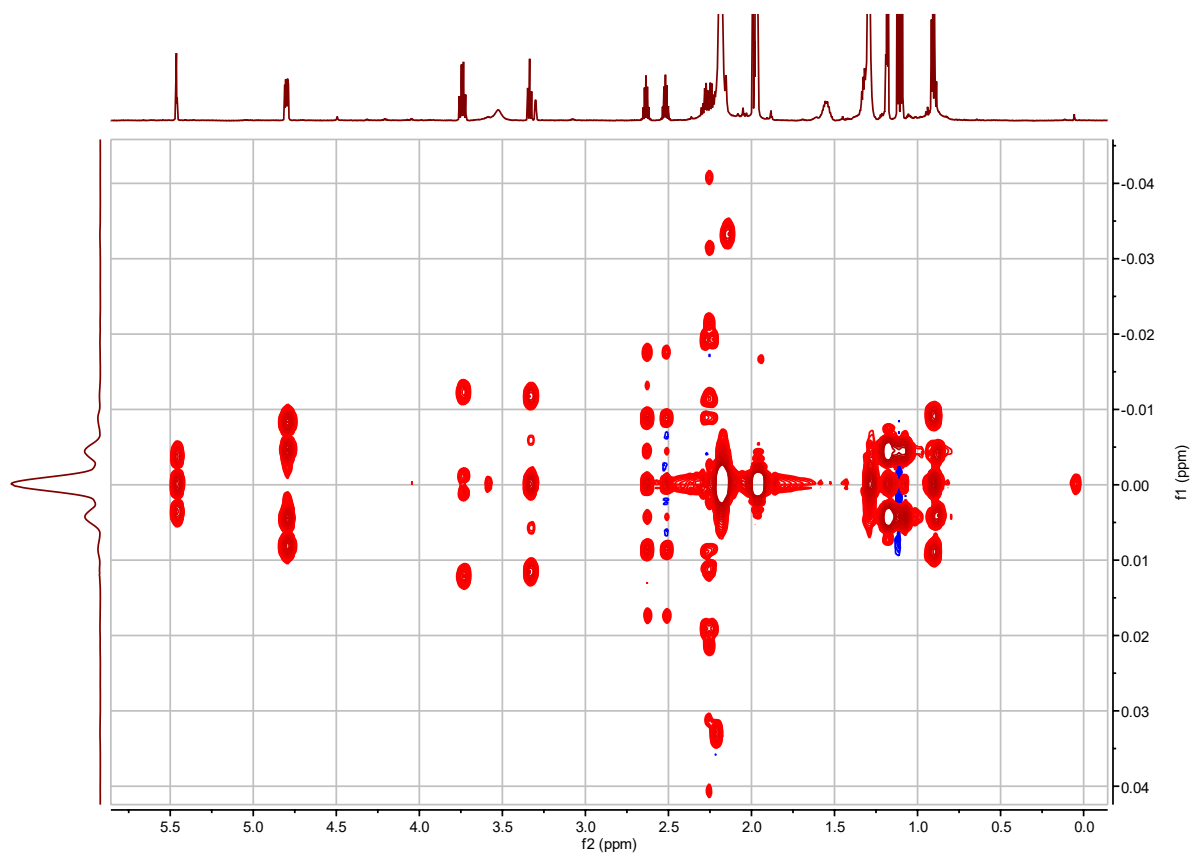
Figure S2.57. I3:22(i4,i4,n14) *J*-resolved.
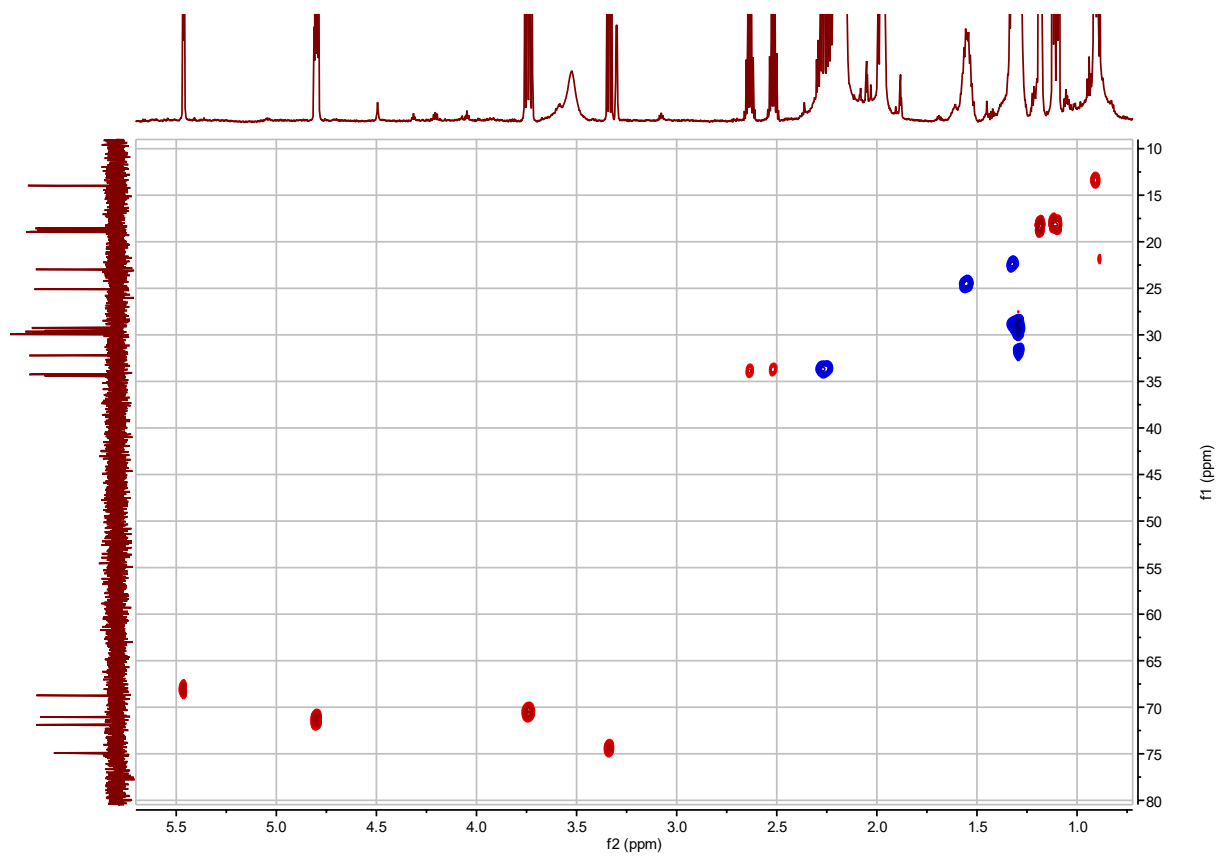
Figure S2.58. I3:22(i4,i4,n14) ^{1}H-^{13}C HSQC.
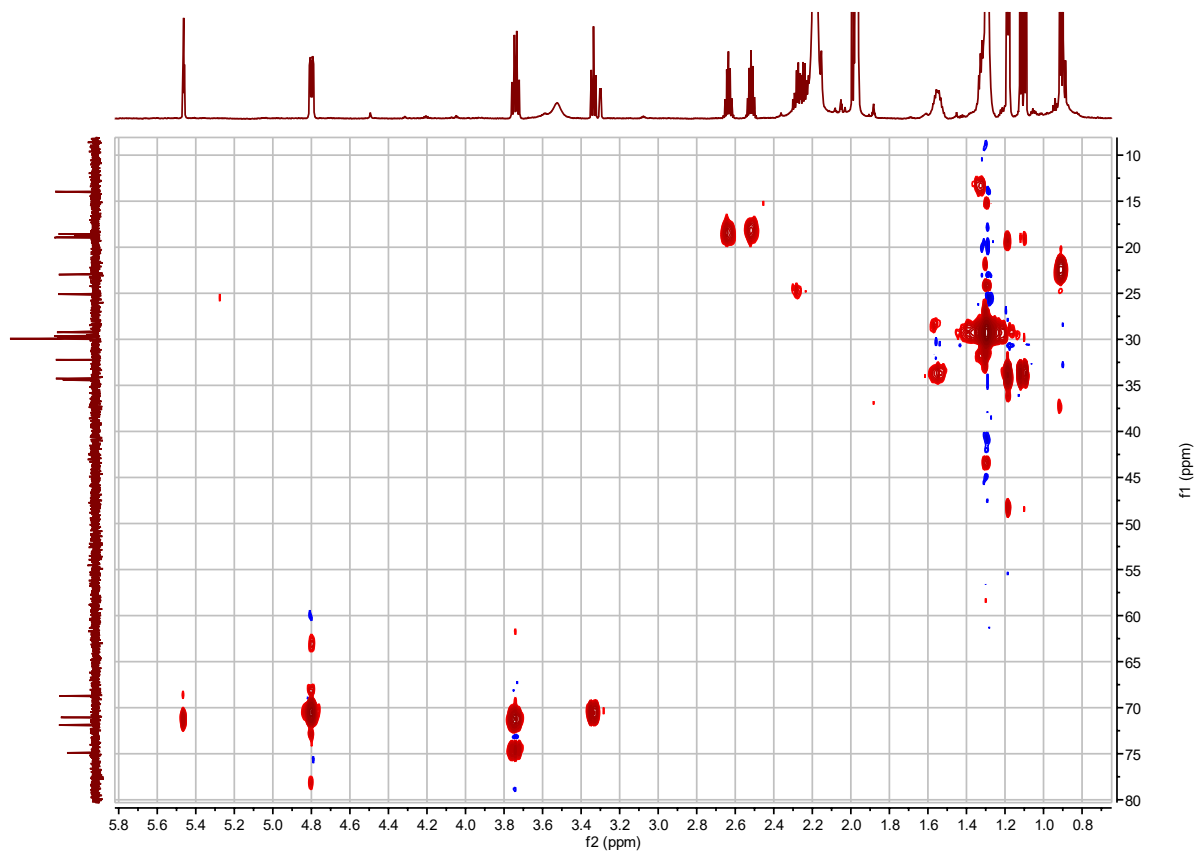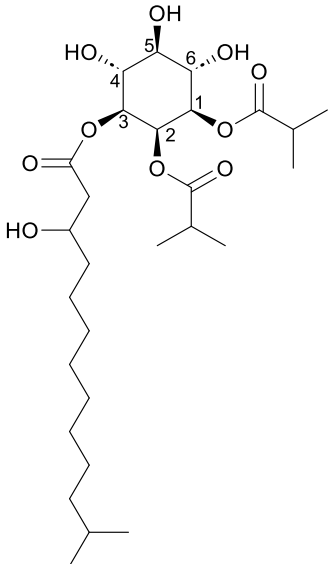
Figure S2.59. I3:22(i4,i4,n14) ¹H-¹³C H2BC.
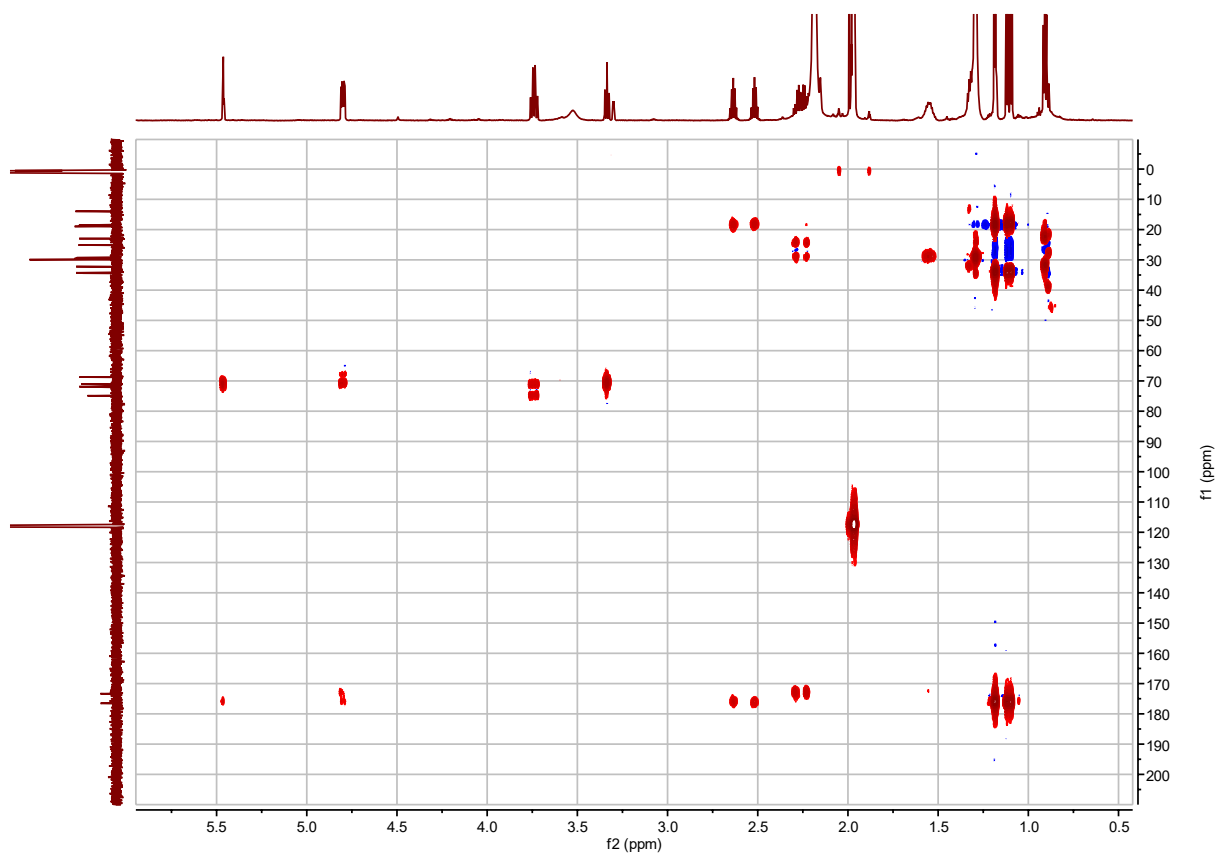
Figure S2.60. I3:22(i4,i4,n14) ^{1}H-^{13}C HMBC.

Table S2.44. I3:22(i4,i4,3-OH-i14) chemical shifts and coupling constants.

	I3:22(i4,i4,3-OH-i14)
	Molecular Formula: $C_{28}H_{50}O_{10}$ Instrument: Agilent 500 MHz DDR2 NMR solvent: CD_3CN Fractions: 95-97 InChI Key: REKBIFNRGXIEET-GBYYZRSBSA-N SMILES: O[C@H]1[C@H](O)[C@@H](OC(CC(O)CCCCCCCCC(C)C)=O)[C@@H](OC(C(C)C)=O)[C@@H](OC(C(C)C)=O)[C@@H]1O

Carbon # (Group)	^{1}H (δ, ppm)	^{13}C (δ, ppm)
1(CH)	4.82 (dd, J = 2.94, 10.21 Hz)	71.34
-1(CO)		175.98
-2(CH)	2.63 (hept, J = 7.04, 7.04, 6.96, 6.96, 6.96 Hz)	33.88
-3,4(CH₃)	1.18 (dd, J = 2.73, 6.97 Hz)	18.33, 18.40
2(CH)	5.48 (t, J = 2.97 Hz)	68.27
-1(CO)		176.01
-2(CH)	2.51 (hept, J = 7.00 Hz)	33.70
-3,4(CH₃)	1.11 (dd, J = 7.00, 12.57 Hz)	18.02, 18.22
3(CH)	4.86 (dd, J = 2.94, 10.18 Hz)	71.50
-1(CO)		171.16
-2(CH₂)	2.46 (dd, J = 4.25, 15.29 Hz) 2.30 (dd, J = 8.44, 15.32 Hz)	42.14
-3(CHOH)	3.92 (m)	67.75
-4(CH₂)	1.41 (m)	36.53
-5-9(CH₂)	1.29 (m)	25.25, 29.26, 29.32, 29.38, 29.65
-10(CH₂)	1.29 (m)	27.18
-11(CH₂)	1.18 (m)	38.7
-12(CH)	1.53 (hept, J = 6.72 Hz)	27.75
-13,14(CH₃)	0.89 (d, J = 6.64 Hz)	21.89
4(CH)	3.76 (td, J = 5.22, 9.71, 9.76 Hz)	70.66 or 70.54[a]
5(CH)	3.36 (t, J = 9.31 Hz)	74.15
6(CH)	3.76 (td, J = 5.22, 9.71, 9.76 Hz)	70.66 or 70.54[a]

a – ^{13}C signals not resolved in 2D spectra.
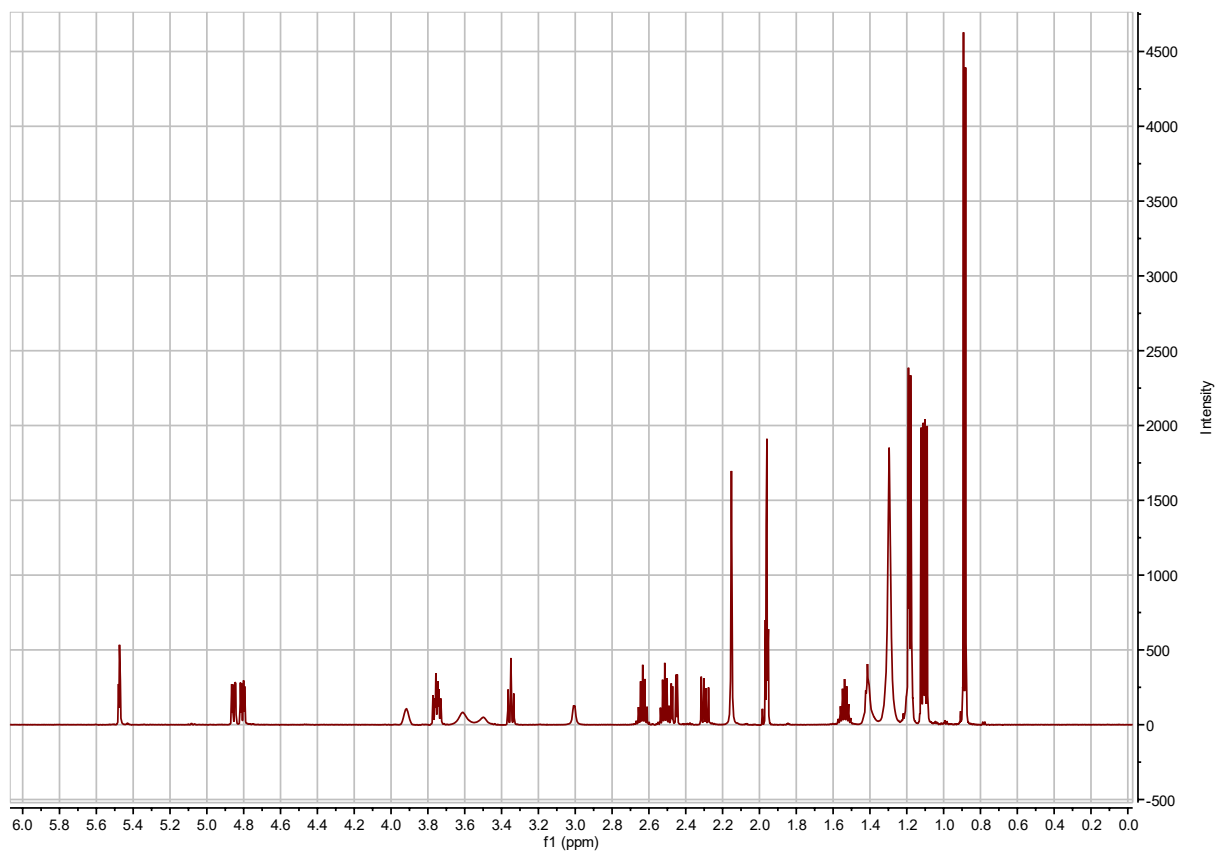
Figure S2.61. I3:22(i4,i4,3-OH-i14) ^{1}H NMR.
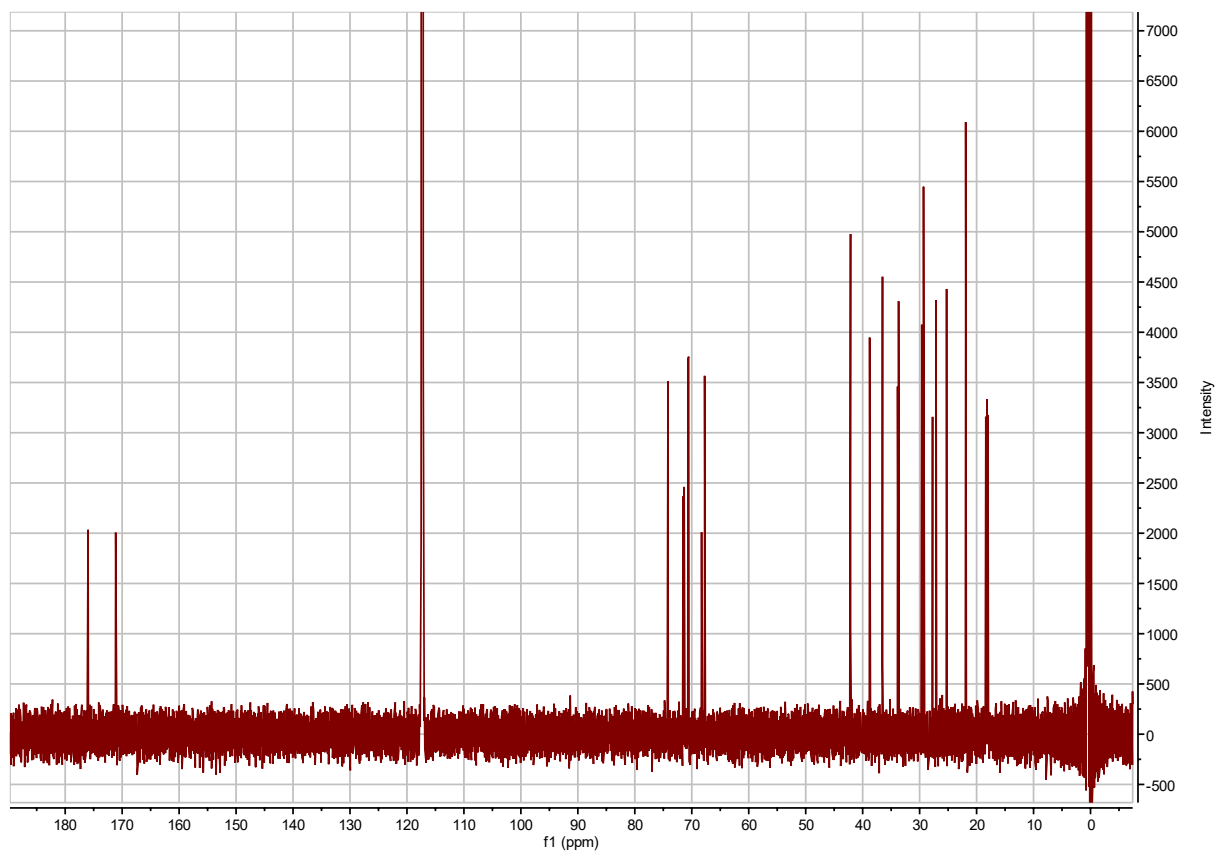
Figure S2.62. I3:22(i4,i4,3-OH-i14) ¹³C NMR.

Figure S2.63. I3:22(i4,i4,3-OH-i14) ^{1}H-^{1}H COSY.
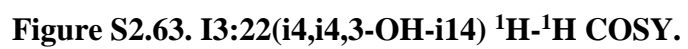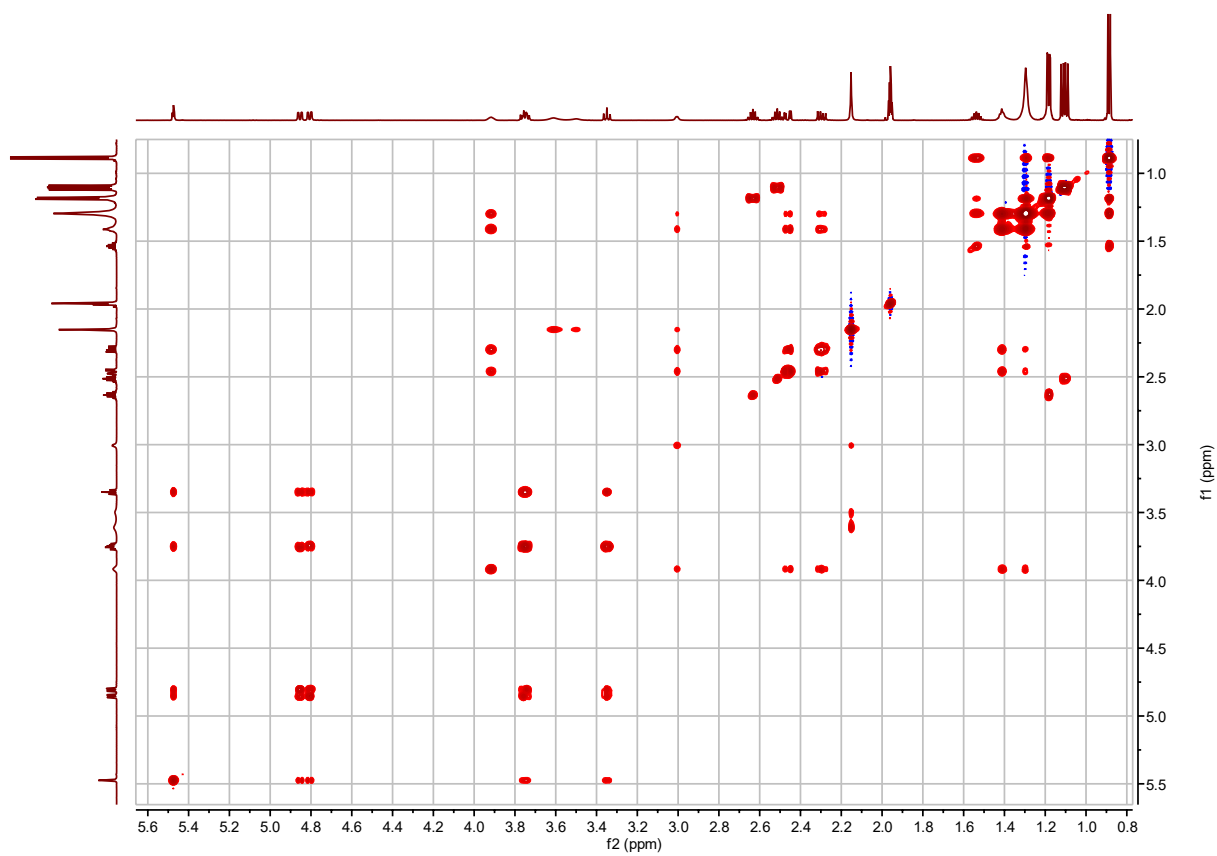
Figure S2.64. I3:22(i4,i4,3-OH-i14) ¹H-¹H TOCSY.
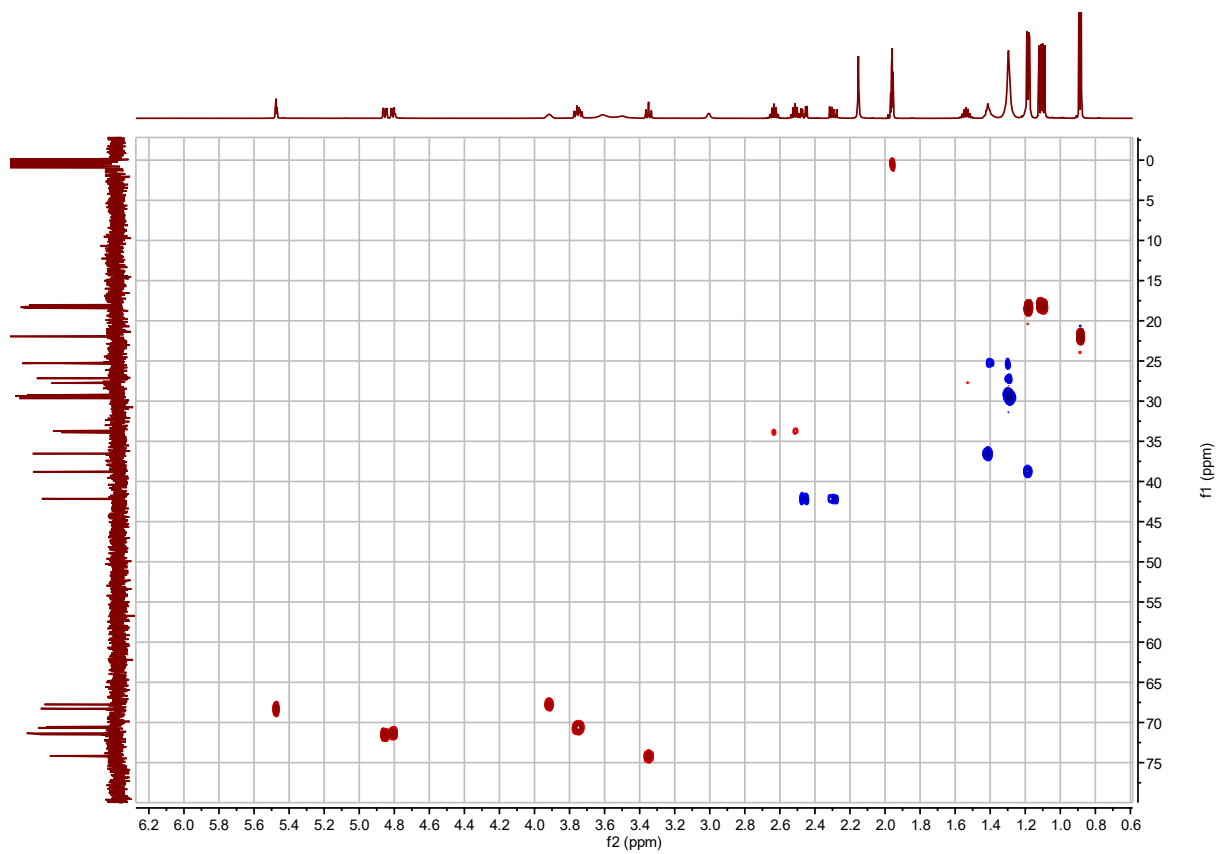
Figure S2.65. I3:22(i4,i4,3-OH-i14) ^{1}H-^{13}C HSQC.
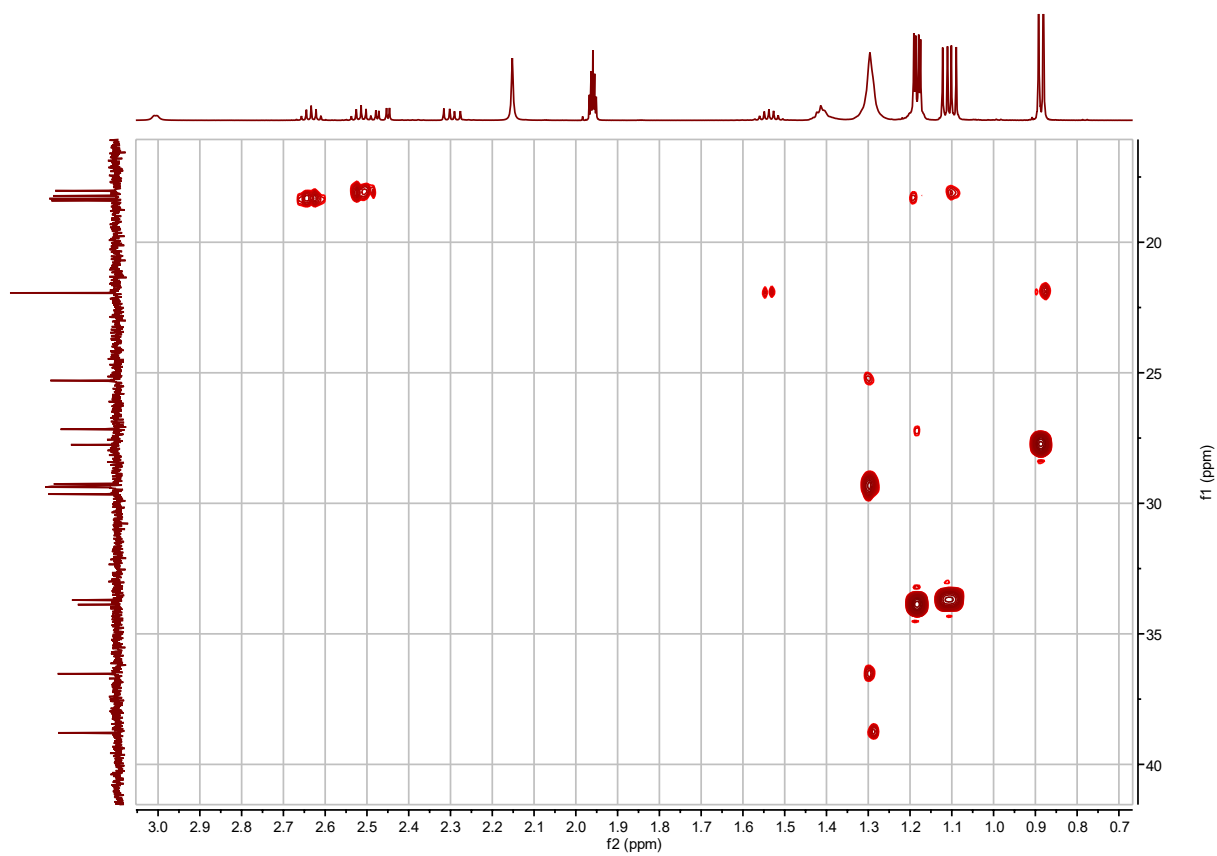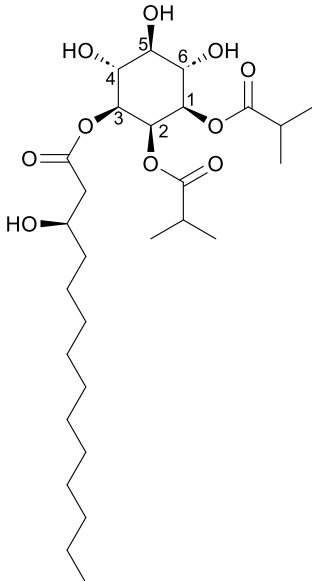
Figure S2.66. I3:22(i4,i4,3-OH-i14) ^{1}H-^{13}C H2BC.
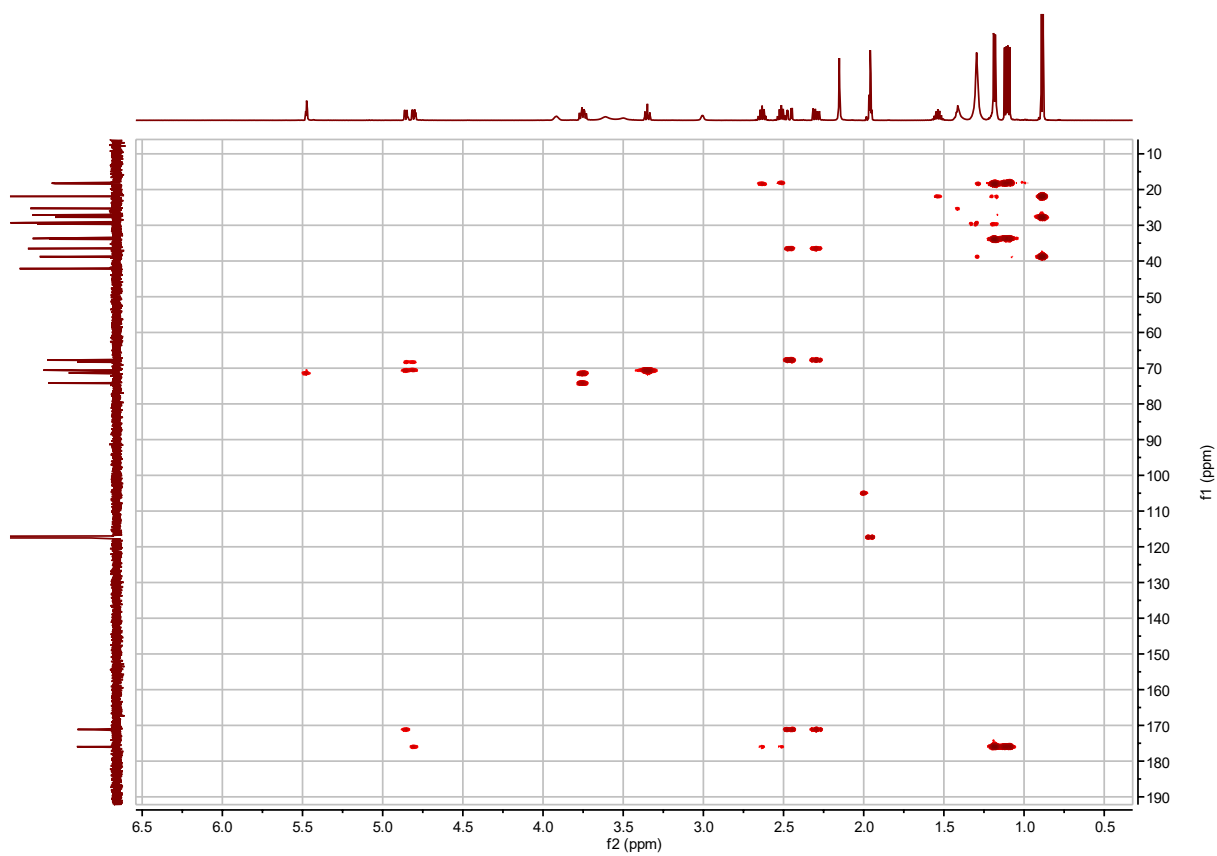
Figure S2.67. I3:22(i4,i4,3-OH-i14) ^{1}H-^{13}C HMBC.

Table S2.45. I3:22(i4,i4,(3*R*)-OH-n14) chemical shifts and coupling constants.

I3:22(i4,i4,(3*R*)-OH-n14)

Molecular Formula: $C_{28}H_{50}O_{10}$

Instrument: Agilent 500 MHz DDR2

NMR solvent: CD_3CN

Fractions: 101-103

InChI Key: KJJFWJBFXVGTFN-KGWGJSIPSA-N

SMILES:
O[C@H]1[C@H](O)[C@@H](OC(C[C@H](O)CCCCCCCCCC)=O)[C@@H](OC(C(C)C)=O)[C@@H](OC(C(C)C)=O)[C@@H]1O

Carbon # (Group)	^1H (δ, ppm)	^{13}C (δ, ppm)
1(CH)	4.82 (dd, J = 2.96, 10.21 Hz)	71.34
-1(CO)		175.98
-2(C_H)	2.63 (hept, J = 7.04, 7.04, 6.96, 6.96, 6.96 Hz)	33.88
-3,4(CH_3)	1.18 (dd, J = 2.73, 6.97 Hz)	18.33, 18.40
2(CH)	5.48 (t, J = 2.97 Hz)	68.27
-1(CO)		176.01
-2(CH)	2.51 (hept, J = 7.00 Hz)	33.71
-3,4(CH_3)	1.11 (d, J = 7.00 Hz)	18.02, 18.22
3(CH)	4.86 (dd, J = 2.95, 10.17 Hz)	71.50
-1(CO)		171.16
-2(CH_2)	2.30 (dd, J = 8.44, 15.32 Hz) 2.46 (dd, J = 4.25, 15.29 Hz)	42.14
-3(CHOH)	3.92 (m)	67.75
-4(CH_2)	1.42 (m)	36.53
-5-12(CH_2)	1.41-1.30 (m)	29.41, 29.38, 29.35, 29.25, 29.11, 25.30, 22.42
-13(CH_2)	1.30 (m)	31.67
-14(CH_3)	0.91(m)	13.42
4(CH)	3.76 (m)	70.54 or 70.66[a]
5(CH)	3.36 (t, J = 9.32 Hz)	74.16
6(CH)	3.76 (m)	70.54 or 70.66[a]

a – ^{13}C signals not resolved in 2D spectra.
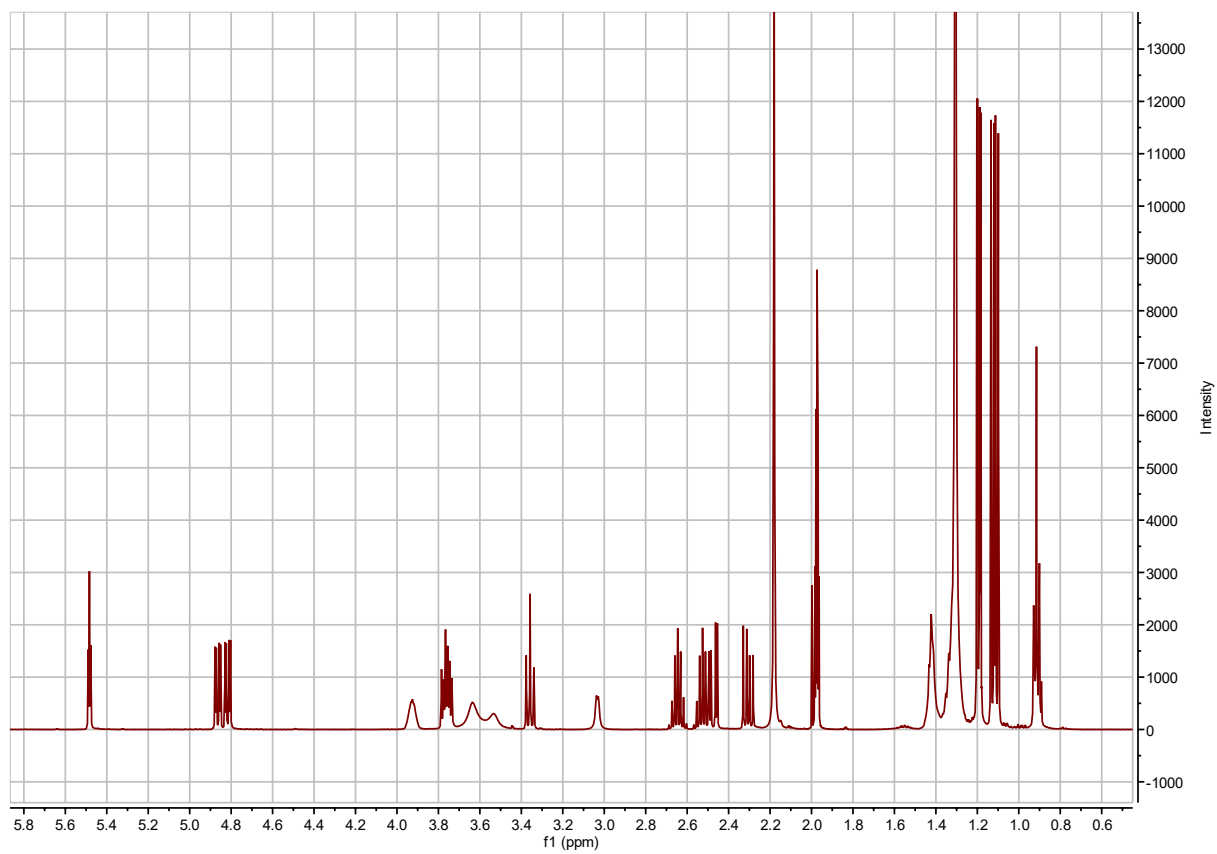
Figure S2.68. I3:22(i4,i4,3-OH-n14) ¹H NMR.
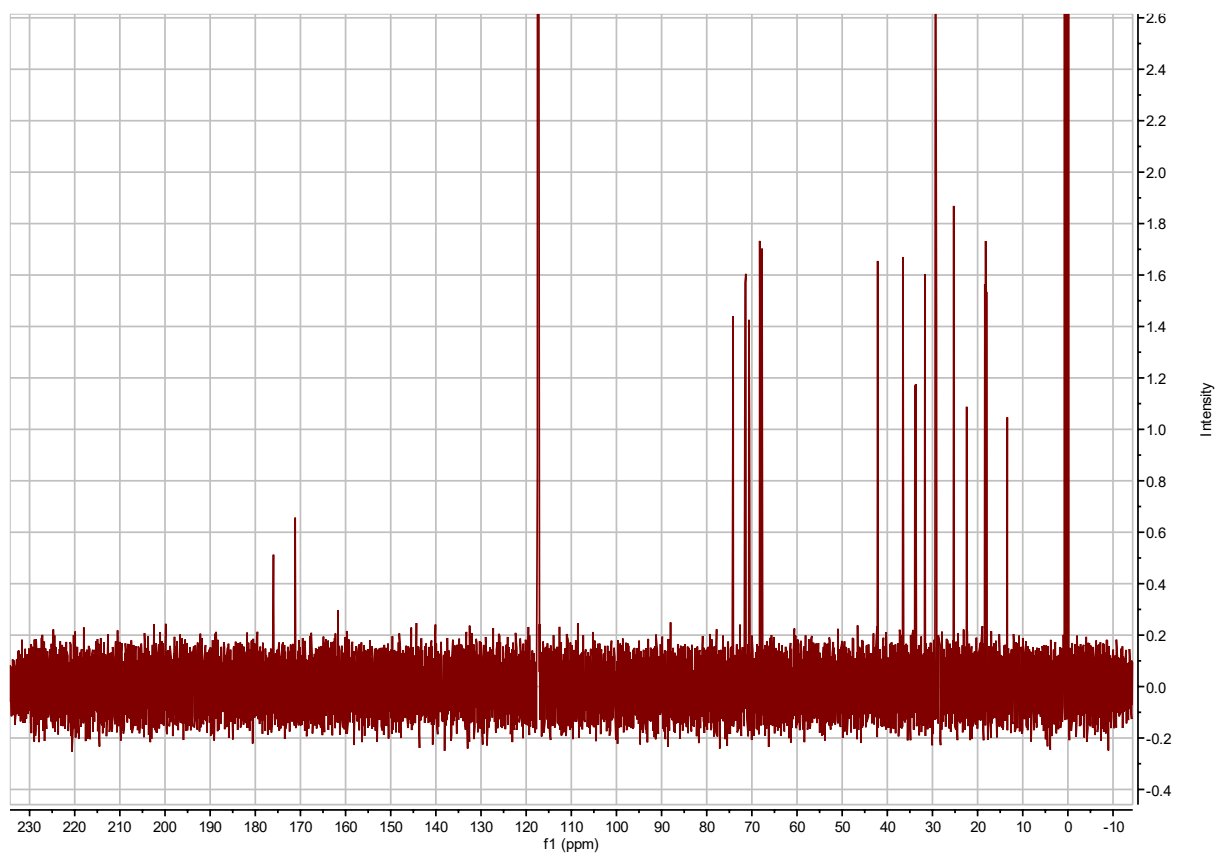
Figure S2.69. I3:22(i4,i4,3-OH-n14) ^{13}C NMR.
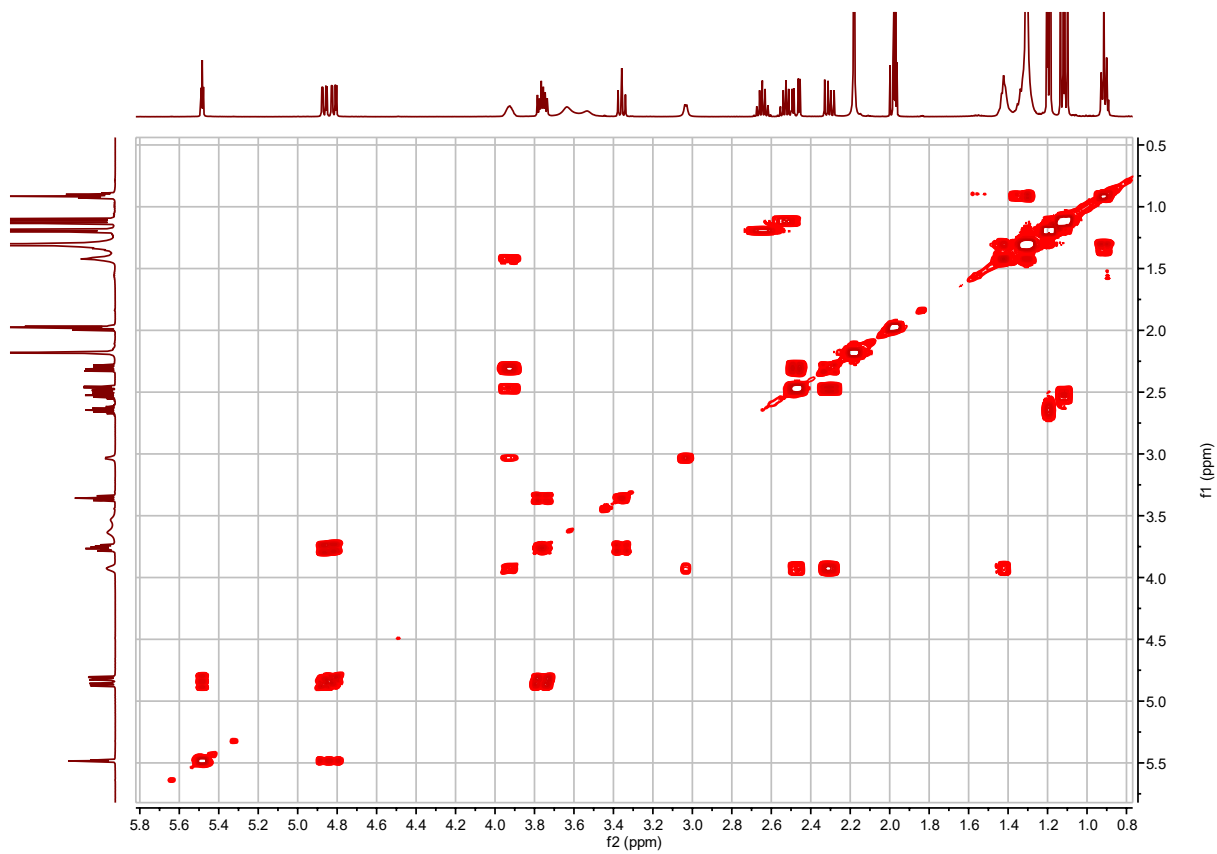
Figure S2.70. I3:22(i4,i4,3-OH-n14) ^{1}H-^{1}H COSY.

Figure S2.71. I3:22(i4,i4,3-OH-n14) ^{1}H-^{1}H TOCSY.
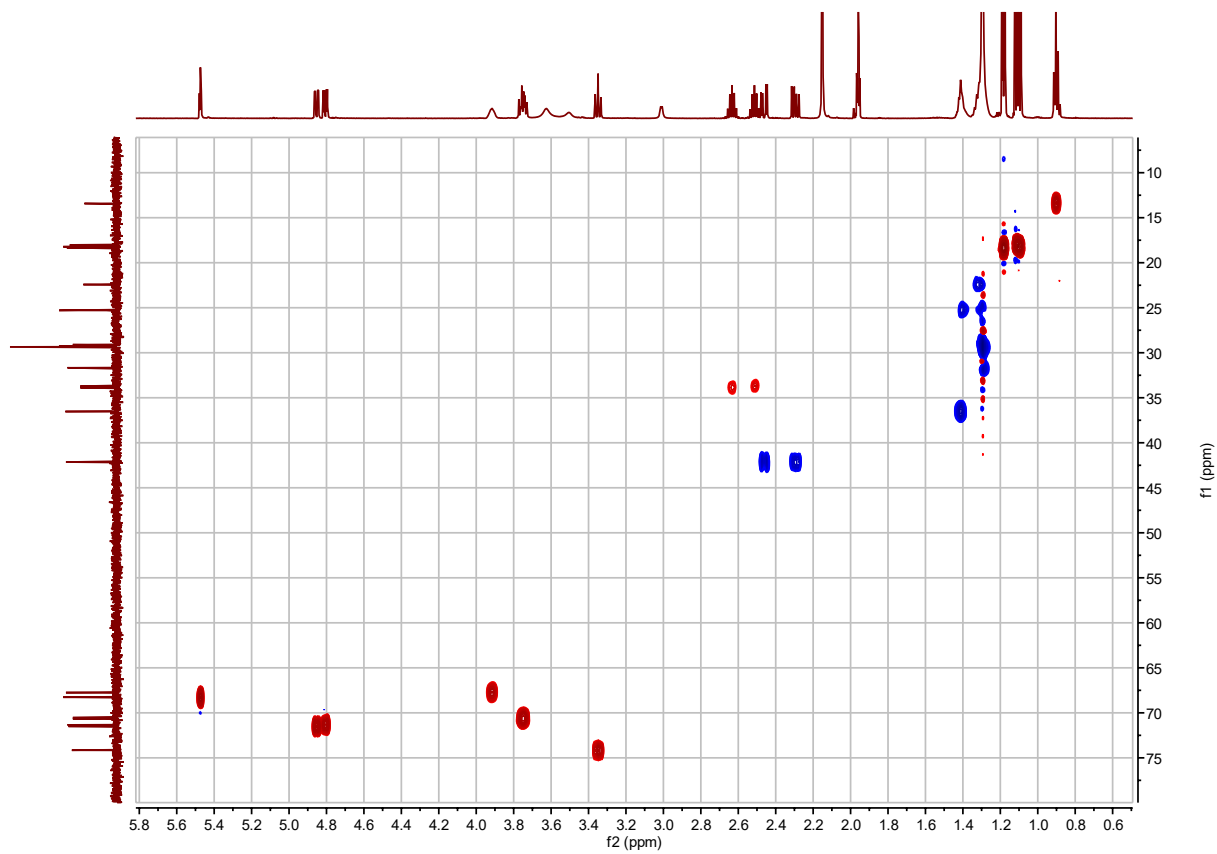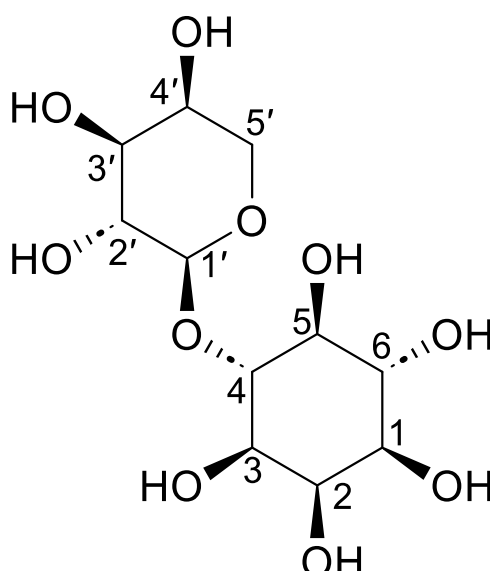
Figure S2.72. I3:22(i4,i4,3-OH-n14) ^{1}H-^{13}C HSQC.
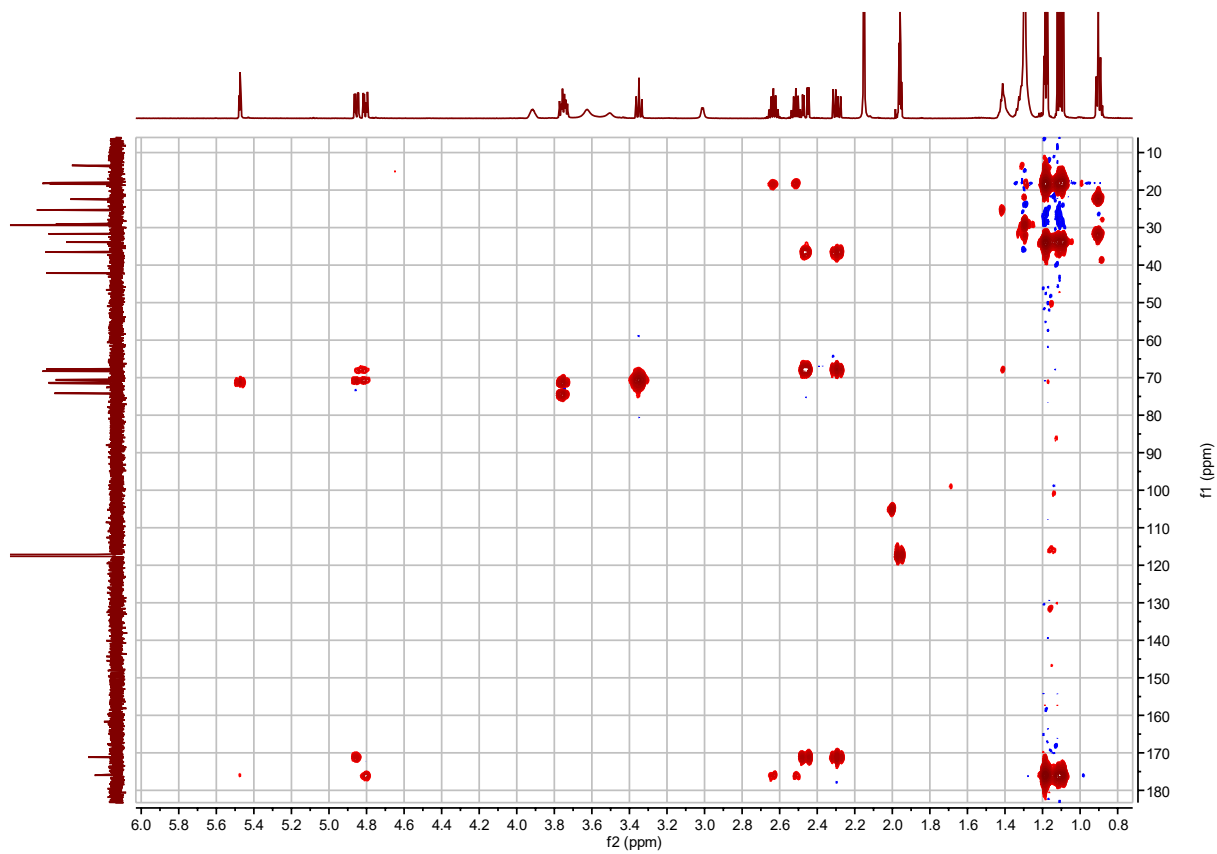
Figure S2.73. I3:22(i4,i4,3-OH-n14) ^{1}H-^{13}C HMBC.

Table S2.46. 4-*O*-β-arabinosyl-*myo*-inositol chemical shifts and coupling constants.

4-*O*-β-arabinosyl-*myo*-inositol derived from saponified AI4:18(2,4,4,8)

Molecular Formula: $C_{11}H_{20}O_{10}$

Instrument: Agilent 500 MHz DDR2

NMR solvent: D_2O

InChI Key: ZTUXXEBTGKCWOB-YFYAPIRNSA-N

SMILES:
O[C@H]1[C@@H]([C@@H](CO[C@H]1O[C@@H]2[C@H]([C@H]([C@H]([C@@H]([C@H]2O)O)O)O)O)O)O

Carbon # (Group)	^1H (δ, ppm)	^{13}C (δ, ppm)
1(CH)	3.50 (dd, J = 3.3, 9.7 Hz)	72.18 or 72.12[a]
2(CH)	3.86 (t, J = 2.95 Hz)	71.96
3(CH)	3.53 (dd, J = 2.9, 9.8 Hz)	70.69
4(CH)	3.60 (t, J = 10.1 Hz)	81.93
5(CH)	3.19 (t, J = 9.2 Hz)	72.48
6(CH)	3.44 (t, J = 10.4 Hz)	72.18 or 72.12[a]
1′(CH)	4.37 (d, J = 7.61 Hz)	103.72
2′(CH)	3.43 (t, J = 9.8 Hz)	71.18
3′(CH)	3.48 (dd, J = 1.7, 13.5 Hz)	72.18 or 72.12[a]
4′(CH)	3.75 (m)	68.19
5′(CH₂)	3.77 (dd, J = 2.38, 13.4 Hz); 3.48 (dd, J = 1.7, 13.4 Hz)	66.18

a – ^{13}C signals not resolved in 2D spectra.
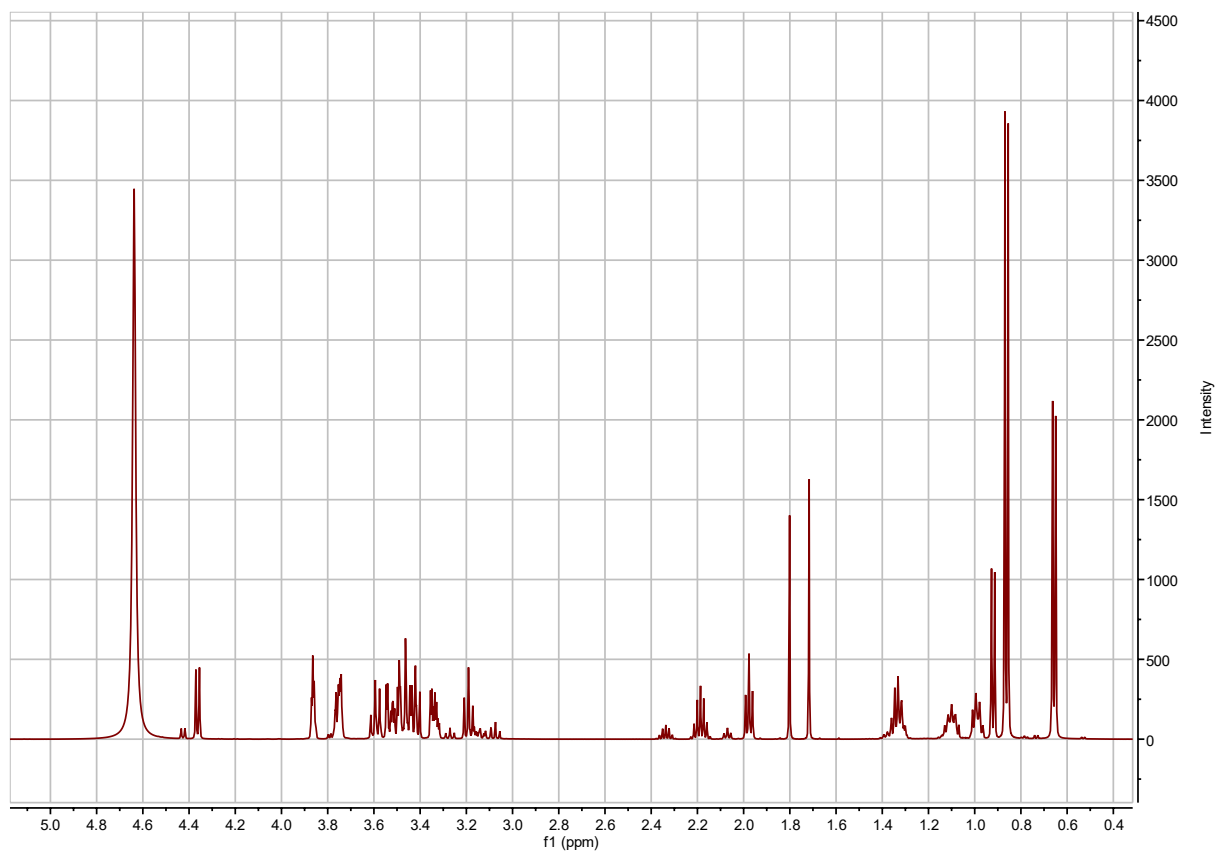
Figure S2.74. 4-*O*-arabinopyranosyl *myo*-inositol derived from saponified AI4:18(2,4,4,8) ^{1}H NMR.
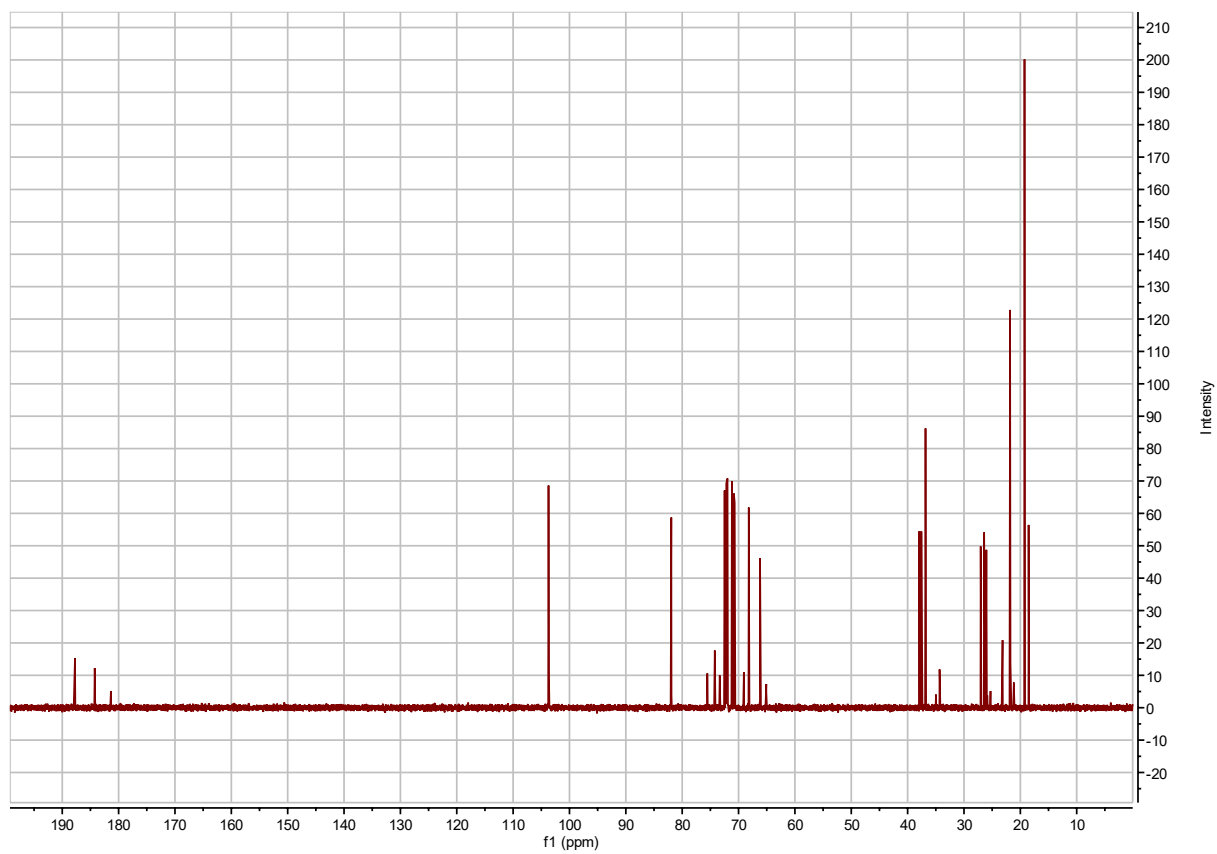
Figure S2.75. 4-*O*-arabinopyranosyl *myo*-inositol derived from saponified AI4:18(2,4,4,8) ^{13}C NMR.
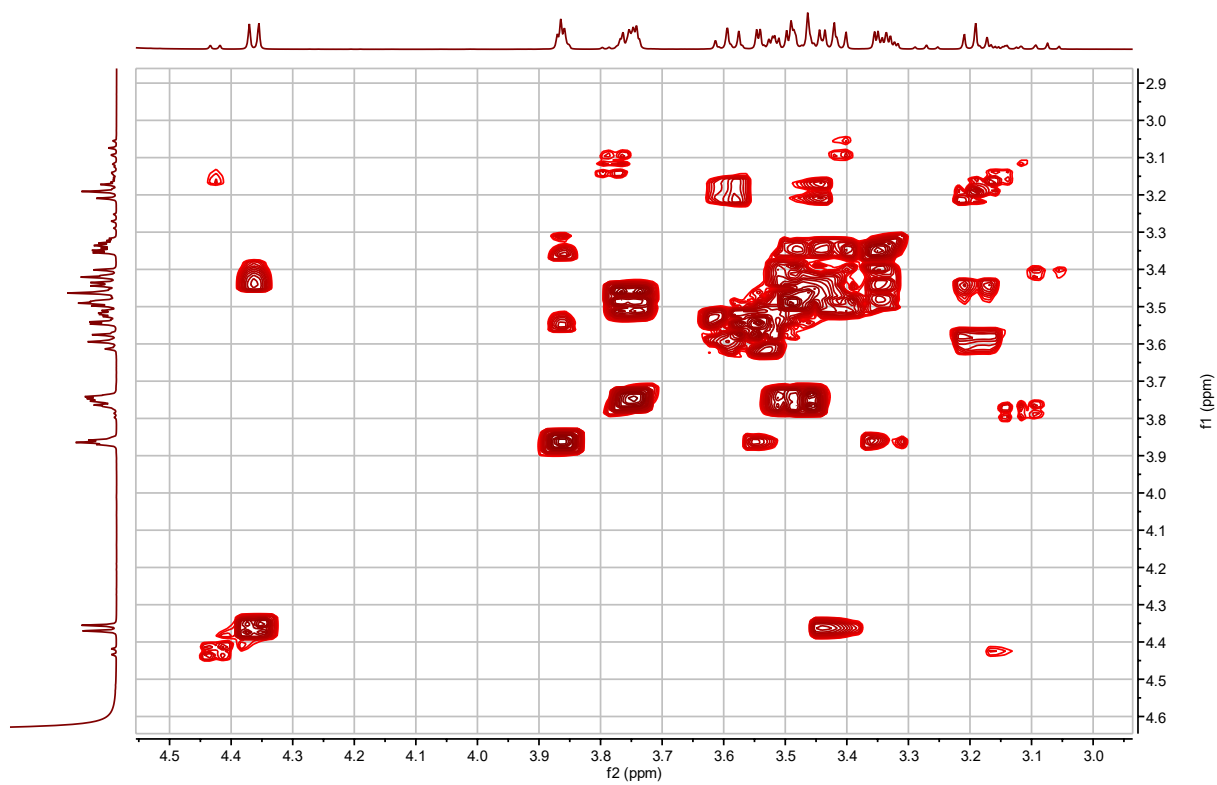
Figure S2.76. 4-*O*-arabinopyranosyl *myo*-inositol derived from saponified AI4:18(2,4,4,8) [1]**H-**[1]**H COSY.**
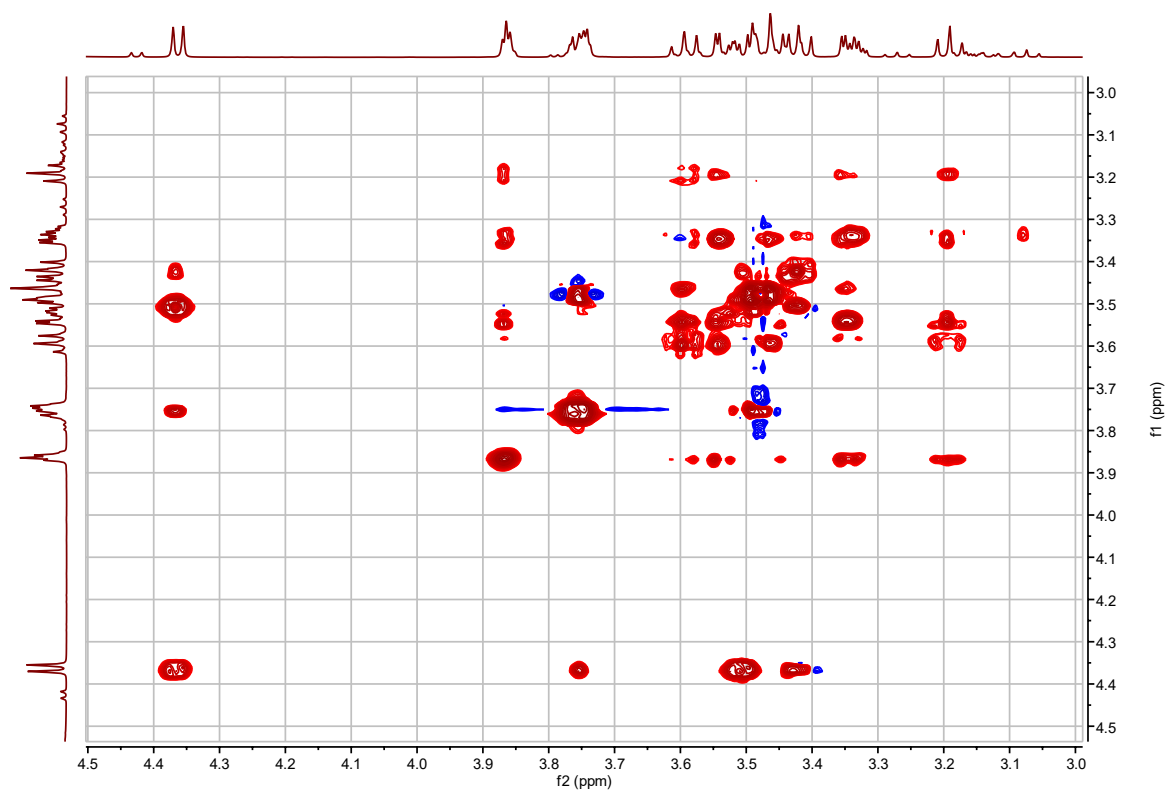
**Figure S2.77. 4-*O*-arabinopyranosyl *myo*-inositol derived from saponified AI4:18(2,4,4,8)
¹H-¹H TOSCY.**
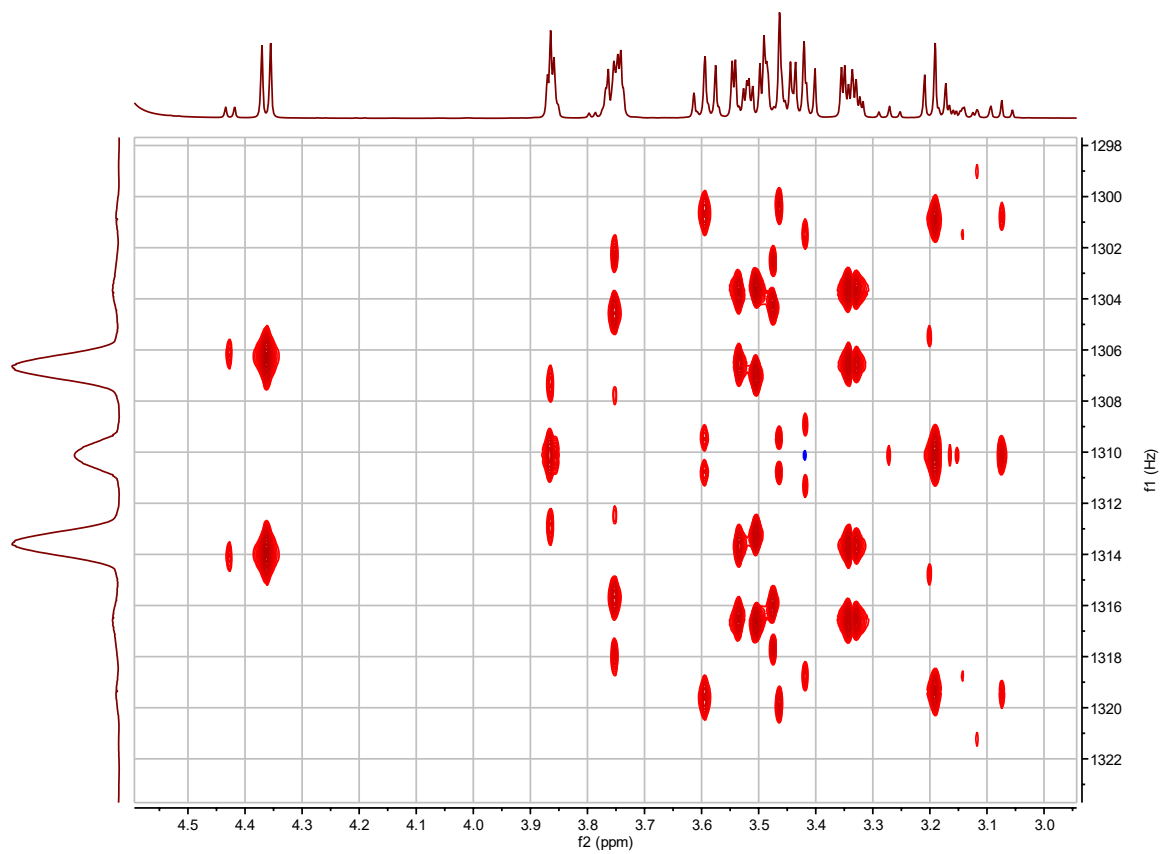
Figure S2.78. 4-*O*-arabinopyranosyl *myo*-inositol derived from saponified AI4:18(2,4,4,8) *J*-resolved.
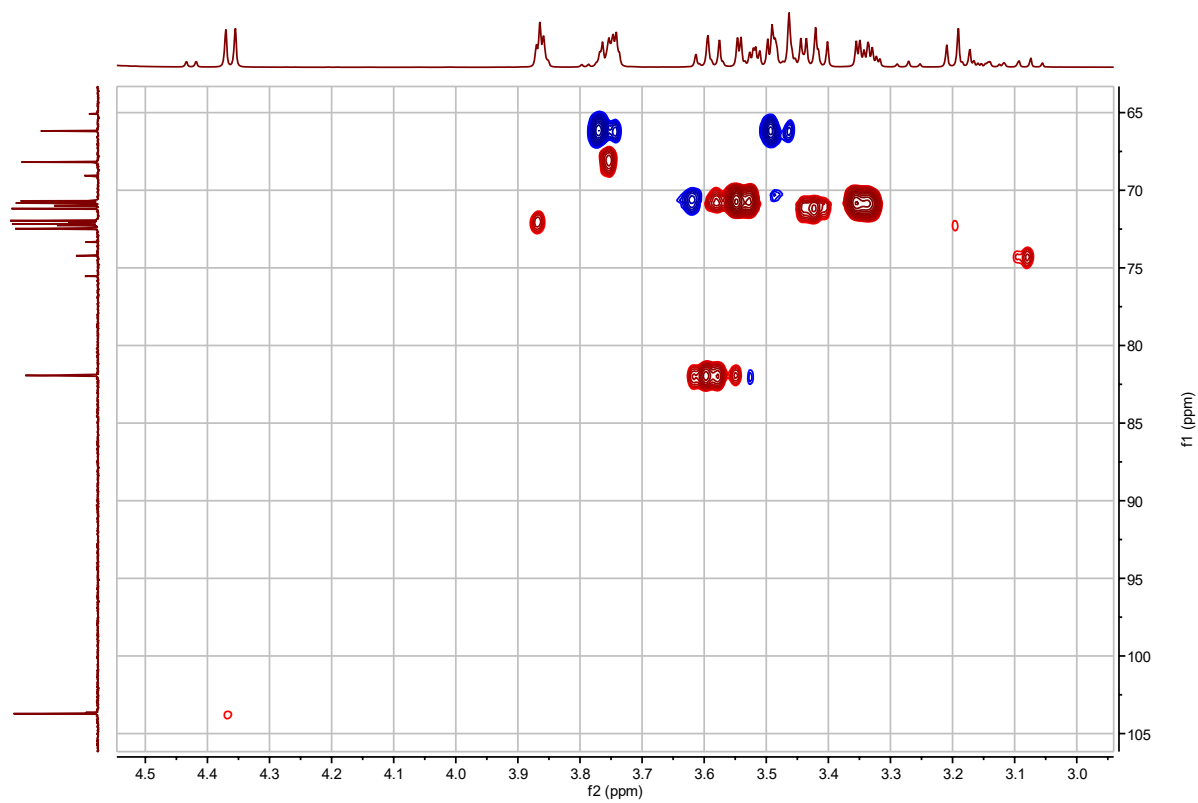
**Figure S2.79. 4-*O*-arabinopyranosyl *myo*-inositol derived from saponified AI4:18(2,4,4,8)
¹H-¹³C HSQC.**

Figure S2.80. 4-*O*-arabinopyranosyl *myo*-inositol derived from saponified AI4:18(2,4,4,8) ^{1}H-^{13}C H2BC.
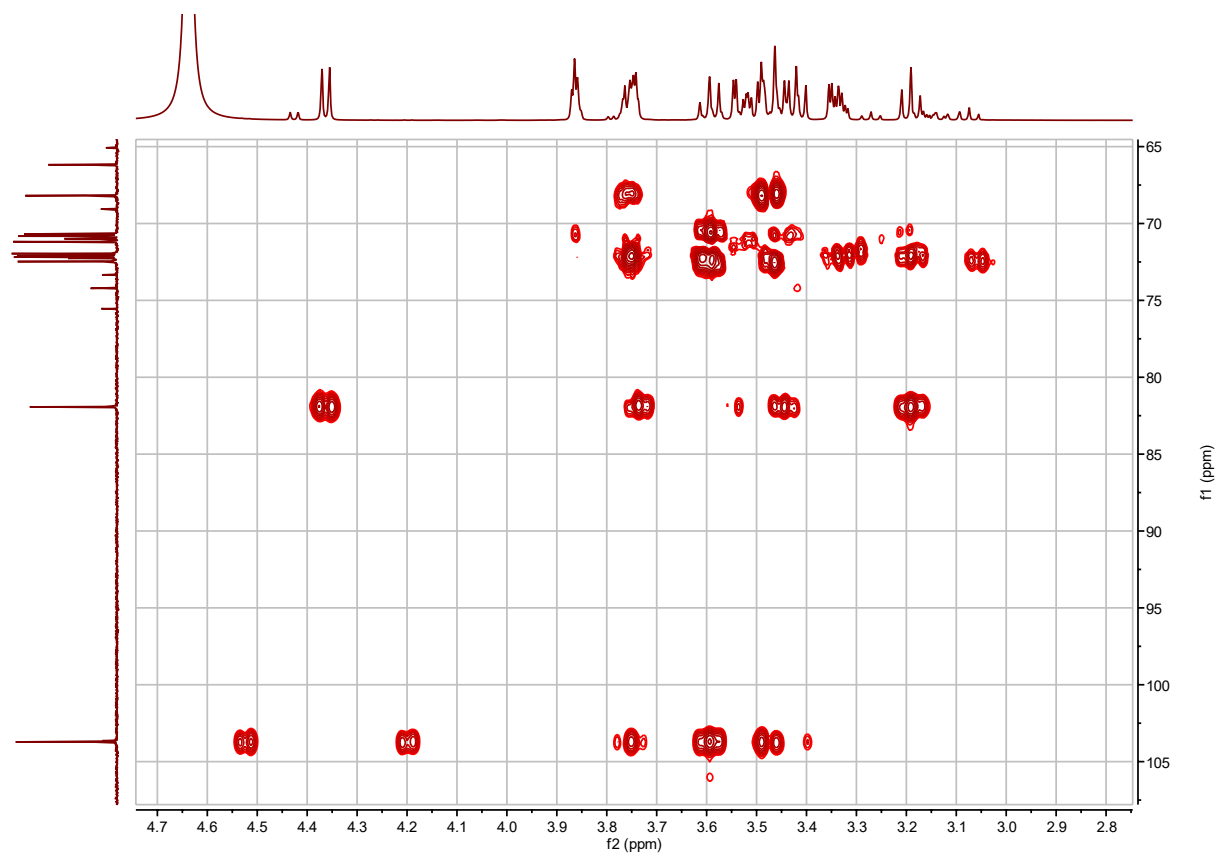
Figure S2.81. 4-*O*-arabinopyranosyl *myo*-inositol derived from saponified AI4:18(2,4,4,8) ¹H-¹³C HMBC.
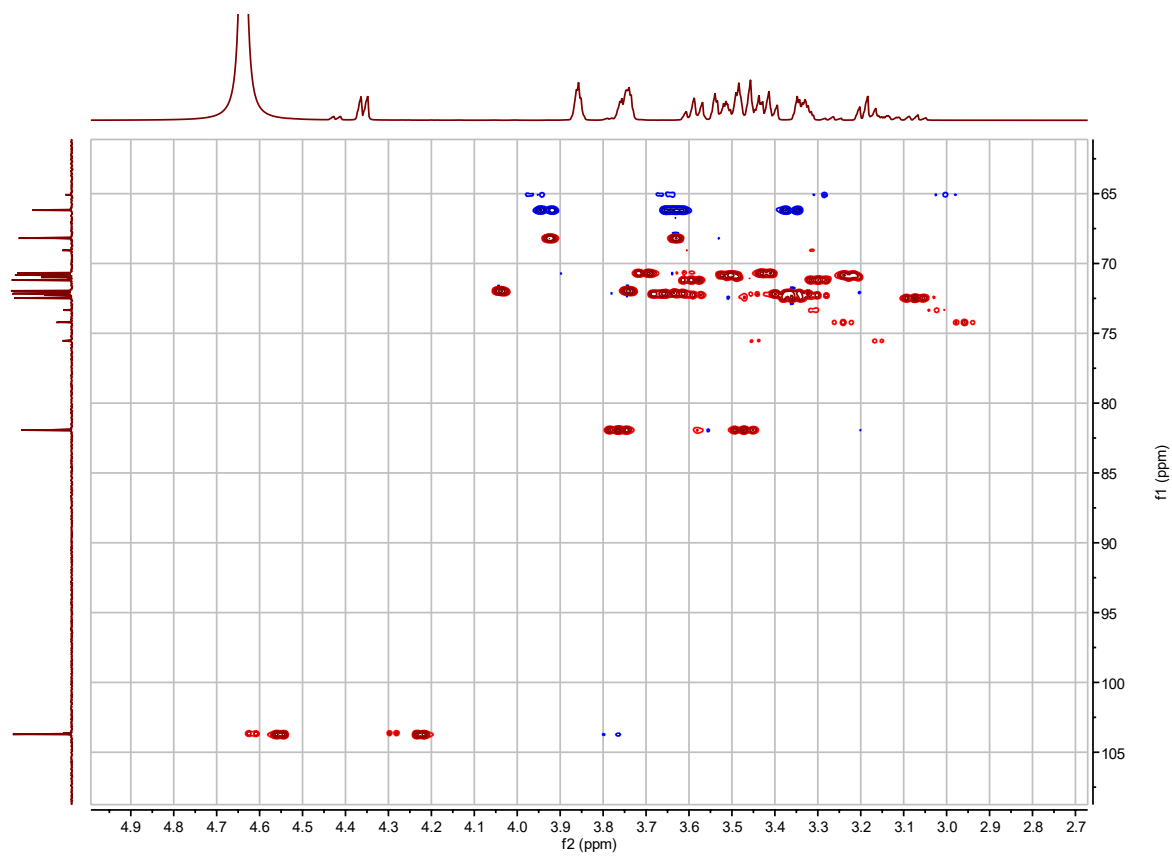
Figure S2.82. 4-*O*-arabinopyranosyl *myo*-inositol derived from saponified AI4:18(2,4,4,8) ¹H-¹³C coupled HSQC.

Table S2.47. NMR metadata for the Agilent DDR2 500 MHz instruments.

Facility supervisor:	Dr. Daniel Holmes
Analyst	Paul D. Fiesel
Instrument location	MSU Max T. Rogers NMR Facility
Facility instrument title	Ahriman and Ormuzd
Manufacturer	Agilent
Field frequency lock	Acetonitrile-d_3; D_2O
Additional solute	None
Solvent	CD_3CN: 600 μL; D_2O: 600 μL
Chemical shift standard	CH_3CN-d_3 (δ_H = 1.94 and δ_C = 118.70 ppm); H_2O-d_2 (δ_H = 4.36 ppm)
Concentration standard	None
Instrument	Agilent DDR2 500 MHz with 7600AS 96 autosamplers
Geographic location of instrument	42.7288, -84.4745
Magnet	499.91 MHz
Probe	OneNMR Probe with Protune accessory for hands-off tuning
Acquisition software	VnmrJ 4.2A
Sample details	Kontes NMR tube, 8 in, Temperature @ 298K, no spinning

Table S2.48. NMR metadata for the Varian Inova 600 MHz instrument.

Facility supervisor:	Dr. Daniel Holmes
Analyst	Paul D. Fiesel
Instrument location	MSU Max T. Rogers NMR Facility
Facility instrument title	Sobek
Manufacturer	Varian
Field frequency lock	Acetonitrile-d_3
Additional solute	None
Solvent	CD_3CN: 600 μL
Chemical shift standard	CH_3CN-d_3 (δ_H = 1.94 and δ_C = 118.70 ppm)
Concentration standard	None
Instrument	Varian Inova 600 MHz
Geographic location of instrument	42.7288, -84.4745
Magnet	599.77 MHz
Probe	Nalorac 5 mm PFG switchable probe pretuned for 1H, 13C
Acquisition software	VnmrJ 4.2A
Sample details	Kontes NMR tube, 8 in, Temperature @ 298K, no spinning

Table S2.49. NMR metadata for the Bruker Avance NEO 600 MHz instrument.

Facility supervisor:	Dr. Daniel Holmes
Analyst	Paul D. Fiesel
Instrument location	MSU Max T. Rogers NMR Facility
Manufacturer	Bruker
Field frequency lock	Acetonitrile-d_3
Additional solute	None
Solvent	CD_3CN: 600 μL
Chemical shift standard	CH_3CN-d_3 (δ_H = 1.94 and δ_C = 118.70 ppm)
Concentration standard	None
Instrument	Bruker Avance NEO 600 MHz NMR with shielded magnet
Geographic location of instrument	42.7288, -84.4745
Magnet	600.32 MHz
Probe	5 mm nitrogen cryogenic HCN Prodigy probe
Acquisition software	TopSpin 4.1.1
Sample details	Kontes NMR tube, 8 in, Temperature @ 298K, no spinning

Table S2.50. NMR metadata for the Bruker Avance NEO 800 MHz instrument.

Facility supervisor:	Dr. Daniel Holmes
Analyst	Paul D. Fiesel
Instrument location	MSU Max T. Rogers NMR Facility
Manufacturer	Bruker
Field frequency lock	Acetonitrile-d$_3$
Additional solute	None
Solvent	CD$_3$CN: 600 μL
Chemical shift standard	CH$_3$CN-d$_3$ (δ_H = 1.94 and δ_C = 118.70 ppm)
Concentration standard	None
Instrument	Bruker Avance NEO 800 MHz NMR with shielded magnet
Geographic location of instrument	42.7288, -84.4745
Magnet	800.33 MHz
Probe	5 mm helium cryogenic HCN probe
Acquisition software	TopSpin 4.1.1
Sample details	Kontes NMR tube, 8 in, Temperature @ 298K, no spinning
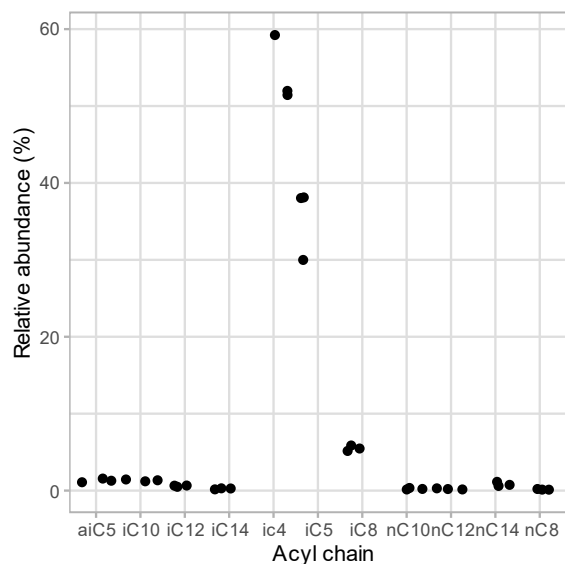
Figure S2.83. Acyl chain composition of *S. melongena* acylsugars. Relative abundance of acyl chains shown between three *S. melongena* leaf surface extracts. Straight acyl chains were identified with authentic reference standards and iso-branched chains were identified with NIST mass spectral library searches. Hydroxylated acyl chains were not included in this analysis and only detected acyl chains are included.

CHAPTER 3:

ACYLINOSITOL BIOSYNTHESIS WITHIN *SOLANUM* GLANDULAR TRICHOMES

Abstract

Plants synthesize a diverse array of lineage- and tissue-specific compounds called specialized metabolites. These exhibit diverse bioactivities involved in plant defense and communication as well as human medicine and foods. Acylsugars, one class of specialized metabolites, are produced within glandular trichomes of Solanaceae (nightshade) family plants and act as anti-insect and anti-fungal molecules. The vast acylsugar structural diversity observed serves as an excellent model for investigating plant metabolic evolution. Acylinositols, one type of acylsugar, are produced in multiple major *Solanum* genus clades and exhibit enormous structural diversity and unique structures with inositol core and chain glycosylations and hydroxyacyl chains. The observation that acylinositols differ in core structure and acylchains from well-characterized acylsucroses, suggested that their biosynthetic pathways are distinct. Considering this, understanding acylinositol biosynthesis can provide insights into their evolution and enables study of their biological roles. In this study, a brinjal eggplant, *Solanum melongena*, tissue-specific transcriptome was utilized to uncover an *in vitro* pathway capable of synthesizing a triacylinositol with chromatographic and mass spectral characteristics identical to a plant-produced triacylinositol. This pathway knowledge was transferred to study trichome acylinositol metabolism in *Solanum quitoense*, a South American fruit crop, identifying an enzyme capable of catalyzing the third inositol acylation. Utilizing a previously developed transcriptome and transient gene silencing protocol, I provided evidence that *Solanum nigrum*, a species residing in a different *Solanum* major clade than *S. melongena* and *S. quitoense*, utilizes a similar acylinositol biosynthetic pathway. This study highlights the usefulness of comparative biochemistry to uncover evolutionary mechanisms underlying metabolic novelty.

Introduction

Acylsugars exhibit enormous chemical diversity as documented in Chapter 2, and this diversity presumably impacts their biological activities against insects, fungi, and microbes. In fact, multiple reports demonstrate that acylsugar structural differences can influence multiple mechanisms of pest deterrence. For example, swapping sugar cores and acyl chains had differential impacts against the mortality of the arthropod species *Cacopsylla pyricola*, *Manduca sexta*, *Manduca nicotianae*, and *Tetranychus urticae* (Puterka et al., 2003). Additionally, the relationship between acylsugar structure and biological activity is not always consistent between different arthropod species (Leckie et al., 2016; Puterka et al., 2003). This acylsugar structure-function relationship is still underdeveloped, and tools such as near isogenic plant lines, differing only in the introduced transgenic genes, with varied acylsugar components can help clarify this association. Understanding the biosynthesis of unusual acylsugar traits can enable development of these tools.

Acylinositols were identified throughout multiple *Solanum* major clades including Clade II, DulMo, VANAns, and Potato clades (Chapter 2; Kerwin et al., unpublished) and these metabolites exhibit multiple atypical traits such as hydroxylated medium length acyl chains and glycosylation. From the phylogenetic distribution of acylinositols, we can infer that this trait arose near the *Solanum* crown node, ~14 mya, however, the lack of resolution between the major *Solanum* clades limits our ability to infer the number of evolutionary origins (Gagnon et al., 2022). Uncovering acylinositol biosynthetic enzymes and pathways from members within different major clades would provide insight into acylinositol evolutionary history and enable testing of their biological function.

Inositol acetyltransferases were recently identified to create tetraacylated compounds in the brinjal eggplant, *Solanum melongena*, and the South American fruit crop lulo, *Solanum quitoense* (Leong et al., 2020). *S. melongena* SmASAT3-L1 catalyzes the R4 acetylation of AI3:16(i4,i4,i8), while *S. quitoense* SqTAIAT acetylates I3:22(2,10,10) and I3:24(2,10,12) at R6. Both enzymes are likely orthologous or paralogous to ASATs in the *S. lycopersicum* acylsucrose pathway, specifically SlASAT3 and SlASAT4, suggesting the acylinositol pathway shares biosynthetic enzymes with acylsucrose biosynthesis.

In contrast, the first steps of trichome acylinositol biosynthesis remain unclear. *S. quitoense* gene knockdown experiments implicated the genes *SqASAT1H*, *SqASAT3H*, and *SqASAT4H* to be involved in producing acylinositols (Leong et al., 2022). *In vitro* enzyme assays suggest the pathway begins with SqASAT1H, an SlASAT1 outparalog (Koonin, 2005), which acylates *myo*-inositol with a C10 or C12 medium acyl chain. SqASAT4H and SqASAT3H, enzymes homologous to SlASAT4 and SlASAT3 can further acylate the resultant monoacylinositol with an acetyl chain and another medium acyl chain, respectively. However, the resulting enzymatically-produced triacylinositol did not coelute with plant-produced triacylinositols by LC-MS suggesting one or more pathway components are missing.

Here I propose an *in vitro* triacylinositol biosynthetic pathway for trichomes of the brinjal eggplant and identify a new enzyme as putatively involved in *S. quitoense* acylinositol biosynthesis. Utilizing the recently generated eggplant tissue-specific transcriptome (Chapter 2), we tested *S. melongena* candidate acyltransferases for forward acylating activity starting with *myo*-inositol and reverse activity with purified triacylinositols. Multiple one pot enzyme combinations of SmASAT1-L, SmASAT3, and SmASAT3-L7 synthesized I2:8(i4,i4). A fourth enzyme, SmASAT3-INOSITOL (SmASAT3-I), acylated I2:8(i4,i4) with medium length acyl

chains. However, the *in vitro* triacylinositol products coeluted by LC-MS with plant-produced

triacylinositol only under conditions that promote pH-induced acyl chain rearrangement. I

transferred this pathway knowledge to *S. quitoense* and identified SqASAT3-I, a trichome-

expressed homolog of SmASAT3-I. VIGS and reverse enzyme assays supported the hypothesis

that SqAST3-I performs the third acylation in *S. quitoense* acylinositol pathway. I then tested if

the more distantly related species *S. nigrum* shared a similar acylinositol pathway to *S.*

melongena and *S. quitoense* by reducing expression of SnASAT1-L and SnASAT3 mRNA with

VIGS. Indeed, knockdown of SnASAT1-L and SnASAT3 expression decreased acylinositol

accumulation suggesting that the three species share a similar pathway and acylinositol

biosynthesis originated in their shared common ancestor.

Results

S. melongena in vitro acylinositol pathway discovery

      I tested the seven *S. melongena* candidate enzymes identified in Chapter 2 for activity

acylating *myo*-inositol with both short and medium length acyl-CoAs (iC4, nC8, nC10, nC12,

nC14), chain lengths related to chains found in *S. melongena* acylsugars. We hypothesized that

the SlASAT1 homolog SMEL_07g013870, SmASAT1-LIKE (SmASAT1-L), would catalyze

this first acylation as it shares 92% amino acid identity to SqASAT1H which acylates *myo*-

inositol with nC10 and nC12-CoAs (Leong et al., 2022). Indeed, SmASAT1-L was a candidate

enzyme that showed activity acylating *myo*-inositol with the medium acyl chains nC8-, nC10-,

nC12-, and nC14-CoAs (Figure S3.5). I also tested the seven candidate enzymes for activity with

the short chain iC4-CoA and *myo*-inositol and found that both SmASAT1-L and the ortholog of

SlASAT3, SMEL_12g015770 (SmASAT3), were active (Figure S3.4), suggesting that the

pathway could begin with either short or medium chain acylation. Testing the hypothesis that the

pathway begins with medium chain acylation, addition of either the SmASAT3 paralog

SmASAT3-LIKE7 (SmASAT3-L7; SMEL_07g013880) or another SmASAT3 paralog

SmASAT3-L1 to 'one-pot' assays with SmASAT1-L produced I2:12(i4,n10) (Figure S3.6). The

I2:12 products created by the ASAT3 paralogs did not coelute suggesting different positions are

acylated (Figure S3.6). After heat-inactivation, the addition of SmASAT3 and iC4-CoA to the

above one-pot assays, now a two-step assay, produced I3:18(i4,i4,n10) (Figure S3.7). However,

this triacylinositol did not coelute with any plant-produced I3:18(i4,i4,10) product (Figure S3.7),

indicating different positions are acylated. Testing the hypothesis that the pathway begins with

iC4 acylation, we found that multiple two enzyme combinations in one-pot assays could produce

the diacylinositol I2:8(i4,i4), with the combinations of SmASAT1-L+SmASAT3, SmASAT1-

L+SmASAT3-L7, and SmASAT3+SmASAT3-L7 exhibiting the highest activity (Figure 3.1).

However, further addition of candidate enzymes and medium chain CoAs to one-pot and

multistep assays failed to add a medium acyl chain to create a triacylinositol.
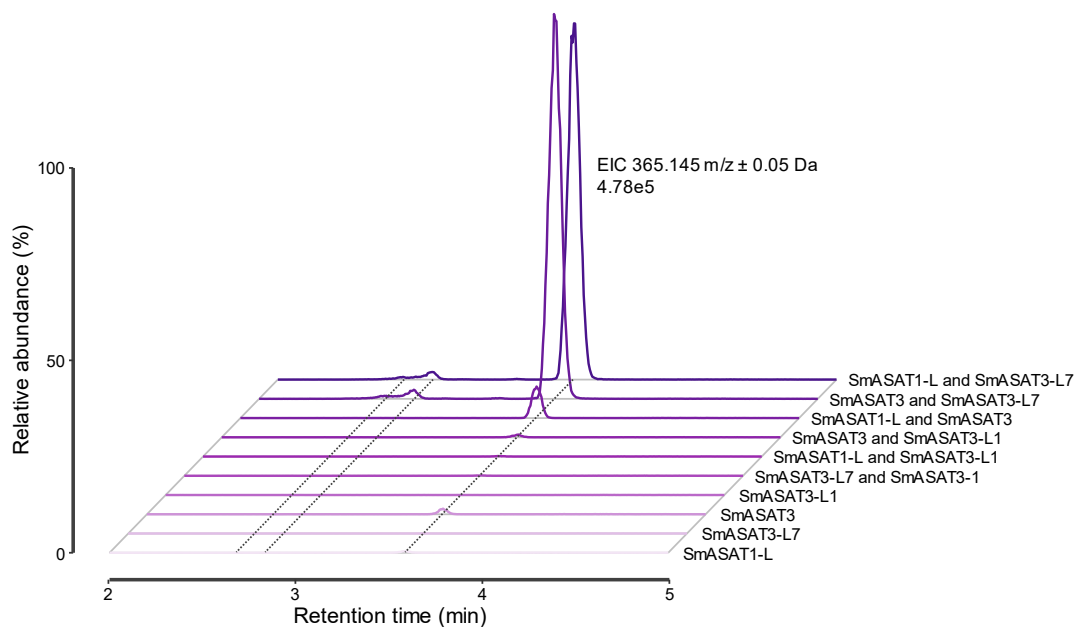
Figure 3.1. Multiple two enzyme combinations of SmASATs produce the diacylinositol I2:8(i4,i4). Products were formed in one-pot assays supplied with *myo*-inositol and iC4-CoA buffered at pH 8.0 and were analyzed with the 7 min I2:8 LC inlet method. Extracted ion

Figure 3.1. (cont'd)
chromatograms display the formate adduct of I2:8, 365.145 *m/z*.

        Taking advantage of the ability of BAHD acyltransferases to act in the reverse direction

(Leong et al., 2020; Lou et al., 2021; Schenck et al., 2022), I sought an acyltransferase capable of

removing an acyl chain from purified triacylinositols and found evidence that a BAHD distantly

related to characterized ASATs catalyzes a third acylation. These reactions are useful because

observed activity confirms an enzyme is acting upon a 'correct' position, information not

gleaned in the forward assays. I tested the candidate enzymes for reverse activity against the *in*

vivo products I3:18(i4,i4,n10) and I3:22(i4,i4,i14-OH) and found SMEL_06g025230 catalyzed

the removal of the medium acyl chain from all three triacylinositols. In contrast,

SMEL_08g018390 removed the iC10 chain from I3:18(i4,i4,i10), producing I2:8(i4,i4) at a

~1000-fold lower signal intensity than the de-acylated product formed by SMEL_06g025230

(Figure 3.2). Phylogenetic analysis revealed that SMEL_06g025230 resides within a clade of 16

S. melongena BAHDs with no characterized ASATs (Chapter 2). With no close homology for

naming this enzyme, we named this enzyme SmASAT3-INOSITOL (SmASAT3-I) based on the

hypothesis that it performs third acylation in the pathway and that it acts upon inositol esters.
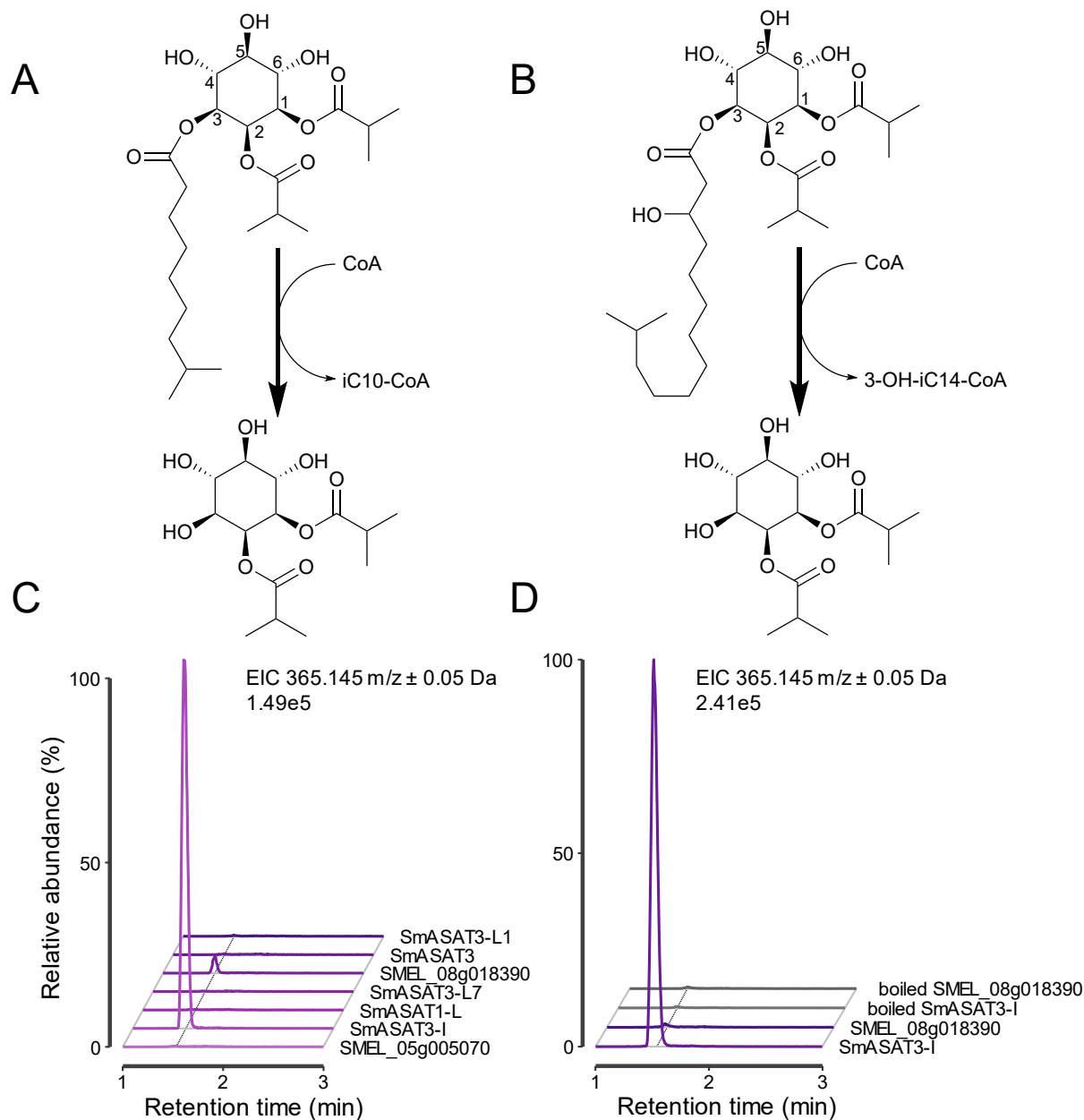
Figure 3.2. SmASAT3-I catalyzes the removal of medium acyl chains from purified *S. melongena* triacylinositols. (A) Reaction scheme for the removal of the iC10 chain from I3:18(i4,i4,i10) observed in (C). (B) Reaction scheme for the removal of the 3-OH-iC14 chain from I3:18(i4,i4,3-OH-i14) observed in (D). (C) LC-MS analysis of reverse *in vitro* assays supplied with I3:18(i4,i4,i10). Extracted ion chromatograms (EIC) display the formate adduct of I2:8(i4,i4), 365.145 *m/z*. (D) LC-MS analysis of reverse in vitro assays supplied with I3:22(i4,i4,3-OH-i14). In this set of assays, only the two enzymes with activity in panel C were used. Extracted ion chromatograms (EIC) display the formate adduct of I2:8(i4,i4), 365.145 *m/z*.

Considering that SmASAT3-I did not demonstrate measurable activity acylating

I2:8(i4,i4) generated by SmASAT3-L1 and SmASAT3-L7, we tested for acyl chain position

differences between the forward- and reverse-produced I2:8 based upon LC-MS coelution

(Figure 3.3A). Indeed, the two products do not coelute, consistent with the hypothesis that

different inositol positions are acylated considering identical acyl chains are present. This result

is reminiscent of *S. quitoense* acylinositol enzymes that produce triacylinositols acylated at

positions different than plant-produced triacylinositols (Leong et al., 2022).

A possible role for intramolecular rearrangement in acylinositol biosynthesis

  The positional differences led us to consider two hypotheses: first, an untested enzyme is

involved and second, that acyl chain rearrangement produces I2:8 acylated at the 'correct'

position, which is subsequently acted upon by SmASAT3-I. Because intramolecular

nonenzymatic acyl chain rearrangement was described previously, and occurs at higher rates

with increasing pH values from 6-8 and greater (Fan et al., 2015; Leong et al., 2022; Lou et al.,

2021), I investigated this hypothesis by testing if the forward assay I2:8 rearranges at pH 8.0 and

if so, whether SmASAT3-I can act upon one of the rearranged isomer(s). As hypothesized, one

pot forward assays with SmASAT1-L and SmASAT3 produced one I2:8(i4,i4) isomer at pH 6.0

and multiple isomers at pH 8.0, with one pH 8.0 isomer coeluting with I2:8 generated from the

reverse reaction of SmASAT3-I against I3:22(i4,i4,3-OH-i14) (Figure 3.3A). SmASAT3-I and

nC14-CoA were then added to the diacylinositol-producing forward assays after a heat-enzyme

inactivation step. One and two peaks corresponding to I3:22(i4,i4,n14) were detected in these

assays conducted at pH 6.0 and 8.0, respectively (Figure 3.3B), and the peak exclusive to the pH

8.0 reaction coeluted with plant produced I3:22(i4,i4,n14) (Figure 3.3B and S3.3). This result is

consistent with the hypothesis that acyl chain rearrangement is involved in acylinositol

biosynthesis.
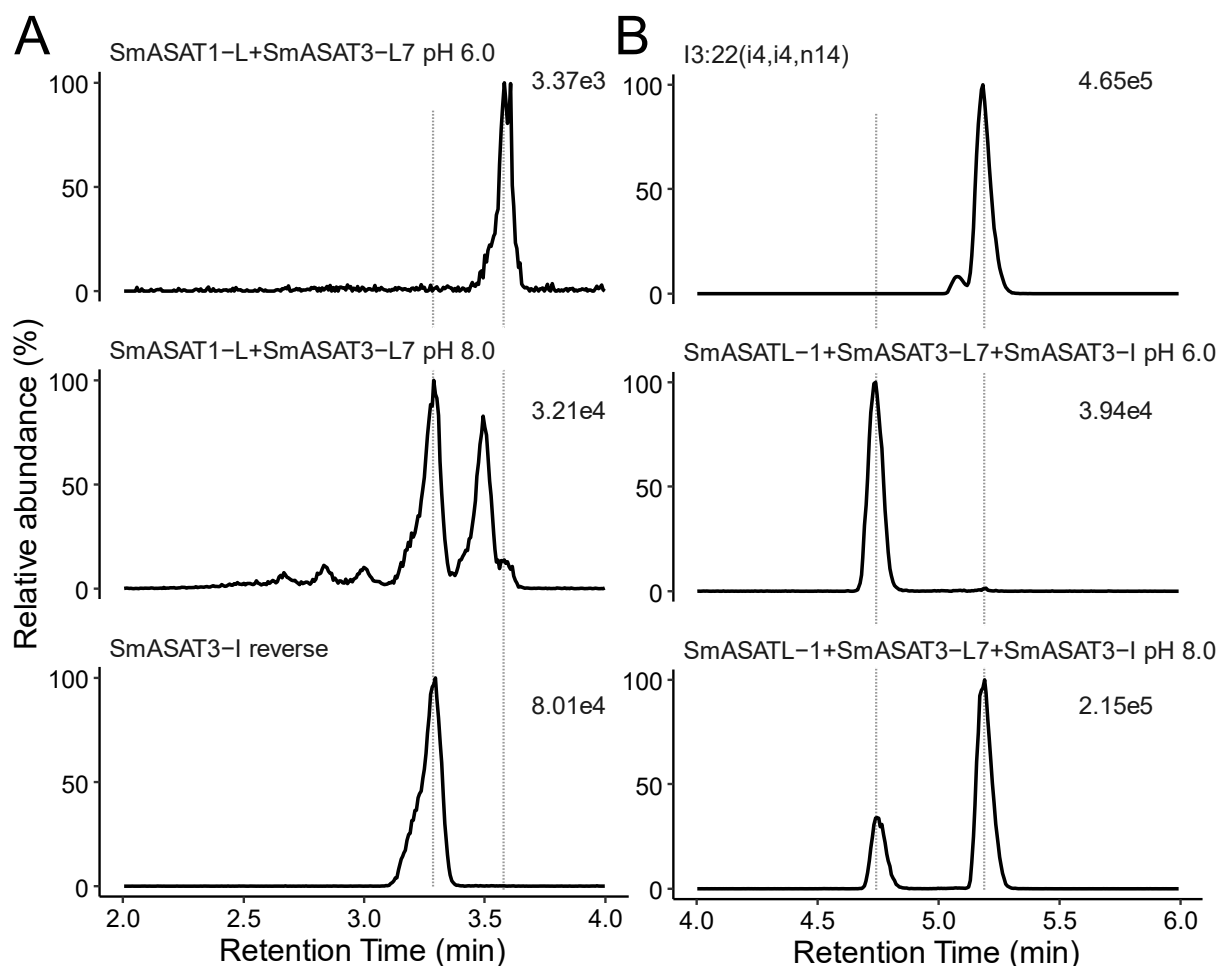
Figure 3.3. Analysis of di- and triacylinositol *in vitro* assay products at pH 6.0 or pH 8.0.
(A) LC-MS analysis of I2:8(i4,i4) *in vitro* assay products created in one-pot reactions with
SmASAT1-L and SmASAT3-L7 at pH 6.0 or pH 8.0. SmASAT3-I reverse assay product
generated from I3:22(i4,i4,3-OH-i14). Forward assay products generated with SmASAT1-L and
SmASAT3-L7 at pH 6.0 (top) and pH 8.0 (middle) with a heat inactivation following the first
reaction. (B) LC-MS analysis of I3:22(i4,i4,n14) *in vitro* assay products generated in a two-step
assay with SmASAT1-L, SmASAT3-L7, and SmASAT3-I. The first step included SmASAT1-L,
SmASAT3-L7, and iC4-CoA generating I2:8(i4,i4). After a heat inactivation step, SmASAT3-I
and nC14-CoA were added generating I3:22(i4,i4,n14). Reactions were conducted at pH 6.0
(middle) or pH 8.0 (bottom) and their products were compared to purified I3:22(i4,i4,n14) (top).

Evidence for *S. quitoense* ASAT3-I homolog in acylinositol biosynthesis

The identification of an eggplant acyltransferase not orthologous or paralogous to

previously characterized ASATs suggested a possible explanation for the incomplete *S.*

quitoense in vitro acylinositol pathway and led us to hypothesize the involvement of a

SmASAT3-I homolog in synthesis of *S. quitoense* acylinositols. A tBLASTn search of

SmASAT3-I against the *S. quitoense* transcriptome (Moghe et al., 2017) identified the

homologous transcript c25595_g1 (SqASAT3-INOSITOL, SqASAT3-I). This transcript was

>340-fold enriched in trichome compared with shaved petiole, and highly expressed in glandular

trichomes at 12,000 average reads (Figure 3.4A) (Moghe et al., 2017). I tested the hypothesis that

SqASAT3-I is involved in acylinositol biosynthesis with *in vitro* 'reverse' assays. Consistent

with this hypothesis, the heterologously expressed SqASAT3-I gene product catalyzed the

removal of a nC10 acyl chain from I3:22(2,10,10) to create I2:12(2,10) (Figure 3.5). Similarly,

SqASAT3-I gene expression knockdown, validated by qPCR (Figure S3.1), resulted in a

statistically significant decrease in I3:22(2,10,10) and increase in I2:12(2,10) abundance (Figure

3.4C-F). These results support the hypothesis that SqASAT3-I is the third acylating enzyme in *S.*

quitoense acylinositol biosynthesis.
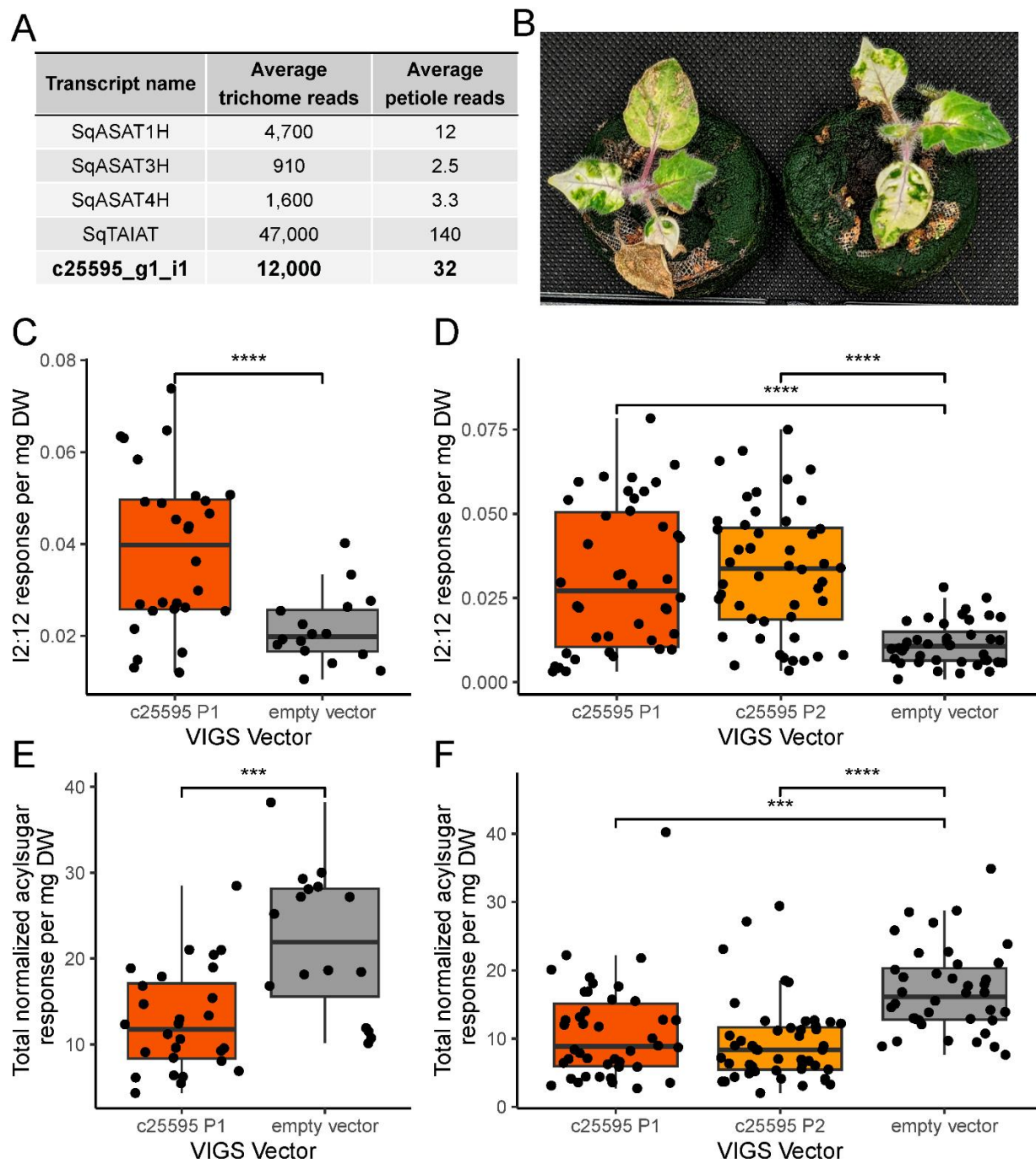
Figure 3.4. VIGS of ASAT3-I in *S. quitoense*. (A) The transcript c25595_g1 (SqASAT3-I) is highly expressed in glandular trichomes, similar to previously characterized *S. quitoense* ASATs (Leong et al., 2022, 2020). Expression data, rounded to two significant figures are derived from Moghe et al. (2017) and describe expression in isolated glandular trichomes and trichome-less petioles. (B) Plants with silenced phytoene desaturase gene expression were used as visual markers for when tissue from the experimental plants should be collected. (C-F) Acylsugar analysis of c25595_g1-targeted and empty vector plants. (C-D) Comparison of I2:12(2,10) LC-MS response between c25595_g1 targeted plants and empty vector control plants with (C) and

Figure 3.4. (cont'd)

(D) representing data from two independent experiments. (E-F) Comparison of total acylsugar LC-MS response between c25595_g1 targeted plants and empty vector control plants with (E) and (F) representing data from two independent experiments. Acylsugar peak area in (C-F) were normalized to the internal standard telmisartan and leaf tissue dry weight (DW). ***$P < 0.001$; ****$P < 0.0001$. Statistical comparisons were conducted with Welch's two-sample t test. For experiment one (C, E), c25595_P1 targeted leaf samples, n = 28; empty vector leaf samples n = 16. For experiment two (D, F), c25595_P1 targeted leaf samples, $n = 42$; c25595_P2 targeted leaf samples, $n = 48$; empty vector leaf samples, $n = 40$.
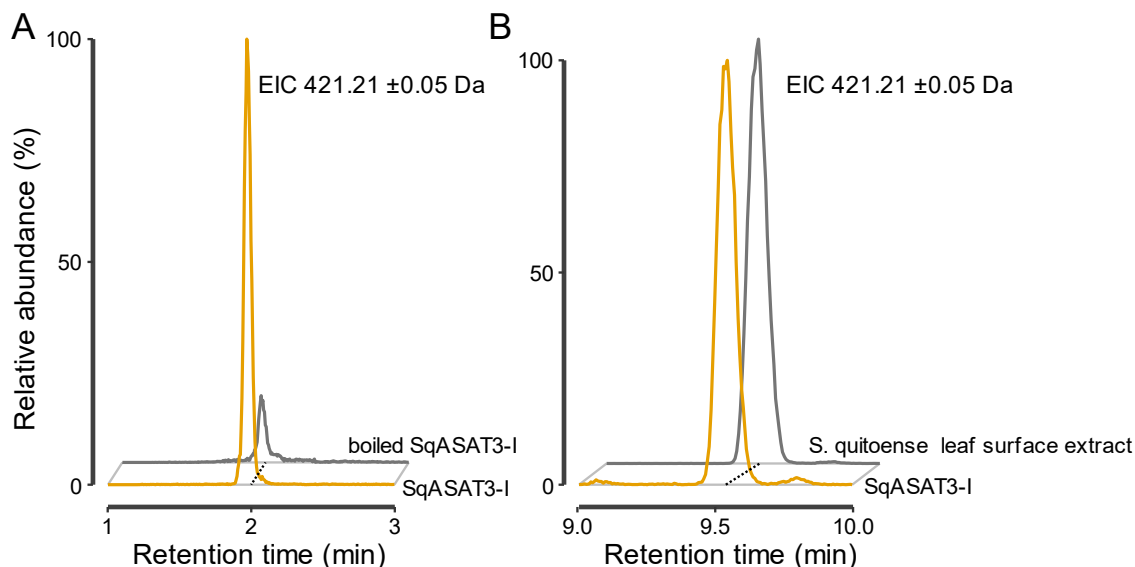
Figure 3.5. SqASAT3-I acts in the reverse direction to remove a C10 acyl chain from I3:22(2,10,10). Both extracted ion chromatograms display the formate adduct of I2:12(2,10), 421.21 *m/z*. (A) LC-MS analysis of SqASAT3-I reverse reaction products when supplied with free CoA and I3:22(2,10,10). (B) The I2:12(2,10) SqASAT3-I reverse reaction product in (A) coelutes with the plant-produced I2:12(2,10).

Knockdown of ASAT1-L and ASAT3 homologs in *S. nigrum* supports a shared acylinositol pathway

I tested if the *S. nigrum* acylinositol pathway is similar to that in *S. quitoense* and *S. melongena* by knocking down gene expression of *S. nigrum* ASAT1-L and ASAT3 homologs using an established and highly effective VIGS method (Lou et al., 2021). The *S. nigrum* homologs of SmASAT1-L and SmASAT3 were identified with BLAST searches yielding the transcripts c64578_g1, SnASAT1-L, and c71009_g1, SnASAT3, both of which were highly expressed and enriched in glandular trichome tissue compared to shaved petioles (Table 3.1).

SnASAT1-L and SnASAT3 gene expression knockdown, validated by qPCR (Figure S3.2),

resulted in a statistically significant decrease in the ratio of acylinositols to acylglucoses (Figure

3.6). Noting that *S. nigrum* resides in DulMo, a different major clade than the Clade II members

S. melongena and *S. quitoense* (Gagnon et al., 2022), these collective results support a

acylinositol pathway shared between the *Solanum* major clades of DulMo and Clade II with a

common evolutionary origin.

Table 3.1. Tissue-specific expression of characterized *S. nigrum* ASATs, and ASAT1-L and ASAT3. Expression data are derived from Moghe et al., (2017). The genes SnASAT1, SnASAT2, and SnAGAT1 were previously characterized (Lou et al., 2021).

Transcript name	Average trichome reads	Average shaved petiole reads
SnASAT1	2382	11
SnASAT2	1416	17
SnAGAT1	3408	31
SnASAT1-L	**1397**	**11**
SnASAT3	**1815**	**61**
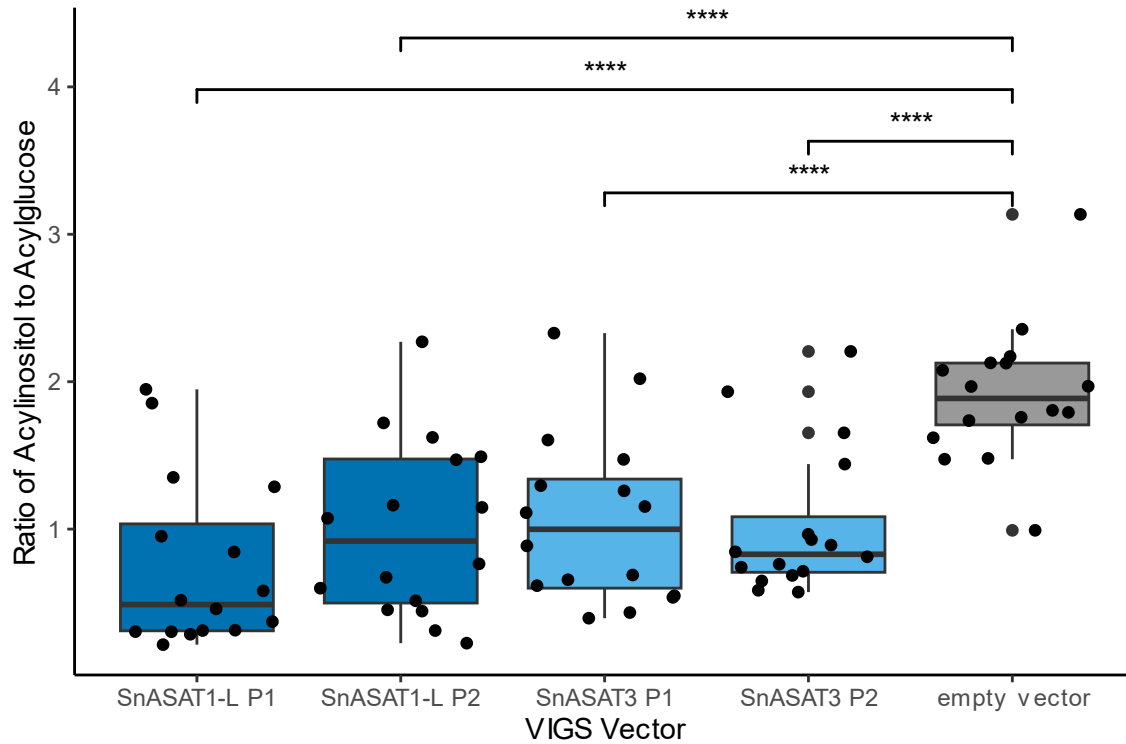
Figure 3.6. VIGS of ASAT1-L and ASAT3 in *S. nigrum*. The ratio of acylinositol and acylglucose responses were compared between ASAT targeted plants, SnASAT1-L and

Figure 3.6. (cont'd)
SnASAT3, and empty vector control plants. To calculate the acylinositol to acylglucose ratio, acylsugar peak area was normalized to the internal standard telmisartan. Values for acylinositols (I3:16, I3:17, I3:18) and acylglucoses (G2:12, G2:13, G2:14, G2:15, G3:14, G3:15, G3:16, G3:17) were then summed. ****$P < 0.0001$. Statistical comparisons were conducted with Welch's two-sample t test. Samples from 16 plants were collected for each VIGS vector.

Discussion

Aided by extensive available genetic resources and remarkable small molecule diversity, the Solanaceae family is a model system for studying specialized metabolism evolution. Harnessing the Solanaceae family helped uncover evolutionary mechanisms and biosynthetic pathways underlying terpene, steroidal glycoalkaloid, tropane alkaloid, and acylsugar diversity (Fiesel et al., 2022). In acylsugar biosynthesis, gene loss, duplication, and neofunctionalization were demonstrated to drive acylsugar phenotypic differences (Fan et al., 2017; Schilmiller et al., 2015). Application of this system uncovered the putative *S. melongena*, *S. quitoense,* and *S. nigrum* acylinositol pathways and identified familiar mechanisms underlying acylsugar pathway evolution.

Here, I describe evidence for an *in vitro* triacylinositol pathway using *S. melongena* enzymes, which produces a triacylinositol that coelutes with a plant-produced triacylinositol. I transferred this putative pathway to two non-model species, taking advantage of available transcriptomic and gene silencing resources. In *S. quitoense*, I identified a previously undescribed enzyme, SqASAT3-I, which can perform the third acylation *in vitro*. In *S. nigrum*, I obtained genetic evidence that SnASAT1-L and SnASAT3 are involved in acylinositol biosynthesis suggesting existence of a biosynthetic pathway shared between *S. melongena*, *S. quitoense*, and *S. nigrum*.

Non-enzymatic rearrangement may play a role in acylinositol biosynthesis

Non-enzymatic acyl chain rearrangement is required for the outlined *S. melongena in vitro* pathway, and I hypothesize that this rearrangement occurs *in planta* based on three lines of reasoning. First, of the seven *S. melongena* BAHDs with the highest trichome gene expression, no enzyme produced detectable levels of I1:4 from the reverse assay-produced I2:8(i4,i4). Additionally, acyl chain rearrangement occurs at pH values >7 which fall within reported plant cytosolic pH values (Moseyko and Feldman, 2001; Shen et al., 2013), although, it is important to note that glandular trichomes are atypical plant cells and could have an atypical acidic cytosolic pH which limits acyl chain rearrangement. While unusual, non-enzymatic reactions including intramolecular rearrangement were described in other natural product pathways such as in the synthesis of the bacterially-derived cladoniamides D-G and siderophore compounds (Bouthillette et al., 2022; Du et al., 2014; Wuest et al., 2009). Further work characterizing the site of acylsugar biosynthesis and the site's pH may provide evidence supporting this rearrangement hypothesis.

SqASAT3-I catalyzes the third acylation in the proposed *S. quitoense* acylinositol pathway

Two lines of evidence support the hypothesis that SqASAT3-I catalyzes the medium chain acylation of diacylinositols. VIGS-mediated SqASAT3-I gene expression reduction led to decreased acylsugar accumulation and an increase in the accumulation of the diacylinositol I2:12(2,10). Additionally, heterologously expressed SqASAT3-I protein catalyzed the reverse reaction to remove a C10 chain from purified I3:22(2,10,10) to produce I2:12(2,10). These data are consistent with the hypothesis that SqASAT3-I is the third acylating enzyme in *S. quitoense* acylinositol biosynthesis.

Acylinositol biosynthesis is distinct from acylsucrose biosynthesis and evolved through multiple evolutionary mechanisms

My results and the previous *S. quitoense* acylinositol investigations (Fiesel et al., 2023; Leong et al., 2022, 2020) suggest acylinositols arose through multiple evolutionary changes creating a distinctly different pathway from characterized acylsucrose pathways. Our data are consistent with the hypothesis that gene duplication generated additional ASAT1 copies, resulting in the duplicate ASAT1-L exhibiting divergent inositol acylating activity from ASAT1 (Leong et al., 2022; Lou et al., 2021). The relaxed selection upon ASAT1-L may have enabled the enzyme's acyl acceptor specificity to shift to another common sugar, *myo*-inositol. This story of duplication leading to altered acylsugar phenotypes is common. For example, the wild tomato *Solanum pennellii* LA0716 'flipped' acylsucrose pathway evolved in part because of a duplicated ASAT3 copy (Fan et al., 2017). This ASAT3 duplicate exhibits altered acyl acceptor specificity to acylate monoacylsucroses as well as the 2-position on the pyranose sucrose ring rather than the 3'-position on the furanose sucrose ring. These changes as well as SpASAT2 protein sequence changes resulted in production of acylsucroses acylated at different positions than the cultivated tomato acylsucroses.

In our model, acylinositol biosynthesis includes a co-opted Clade III BAHD, ASAT3-I, which is phylogenetically distinct from other characterized ASATs. Previous phylogenetic analysis revealed that ASAT3-I resides within a clade consisting of 16 *S. melongena* BAHDs and CaPun1, capsaicin synthase, which is the only characterized gene in this BAHD subgroup (Chapter 2; Moghe et al., 2023). CaPun1 is predicted to catalyze a nitrogen acylation unusual for Clade III BAHDs (Moghe et al., 2023; Stewart Jr et al., 2005). ASAT3-I also exhibits atypical acylating activity but with an OH-C14 acyl chain (Figure 3.2), suggesting that the co-option of

ASAT3-I was involved in the evolution of acylsugar hydroxyacyl chains. Analysis of enzymes related to ASAT3-I and CaPun1 may identify further atypical enzymatic activities underlying plant specialized metabolites.

The proposed pathways provide a starting point for investigating the enormous interspecific acylinositol structural differences identified in Chapter 2. In fact, analysis of only two species – *S. melongena* and *S. quitoense* – led me to document significant divergences in their proposed biosynthetic pathways. For example, unlike the proposed *S. melongena* pathway, the proposed *S. quitoense* pathway begins with a medium chain acylation and produces acylsugars with only C2, C10, and C12 acyl chains. Acylinositol structural differences with unidentified biochemical mechanisms include the contrasting acylinositol disaccharide core structures between *S. melongena* and *S. quitoense* and acylation position differences between *S. melongena* and *S. nigrum*. Investigating these differences equipped with the transcriptomic resources and gene knockdown/knockout techniques for these three species will better our understanding of how acylsugars evolve (Chapter 2) (Hurney, 2018; Lou et al., 2021).

Synthetic biology application of acylinositol pathway enzymes

The acylinositol biosynthetic enzymes discovered here and previously described (Leong et al., 2022, 2020) will be useful synthetic biology tools in understanding the how different sugar cores impact acylsugar biological activities as well as generating new acylsugars not observed in plants. Our current understanding of how different acylsugar traits impact plant defense can be expanded with tools created through synthetic biology. For example, transgenic production of near-isogenic lines with varied acylsugar structures will accelerate the study of the relationship between acylsugar structures and their protective functions. Additionally, acylsugar biosynthetic enzymes can be mixed and matched to produce new acylsugars not observed in plants potentially

producing compounds with varied bioactivities against different plant pests. This approach has been validated *in vitro* and its utilization *in planta* may increase plant resilience against pests (Schenck et al., 2022).

Methods

<u>Gene cloning, heterologous protein expression and purification</u>

All *S. melongena* candidate genes were cloned into pET28b as described in Chapter 2. SqASAT3-I was cloned into pET28b(+) (MilliporeSigma, Burlington, MA, USA) for protein expression. The SqASAT3-I coding sequence was amplified from *S. quitoense* cDNA with primers listed in Table S3.1 and Q5 2X Hotstart master mix (New England Biolabs, Ipswich, MA, USA), and the PCR amplicon was purified by agarose gel separation and extraction with the Monarch DNA Gel Extraction Kit (New England Biolabs). Using 2X Gibson Assembly Master Mix (New England Biolabs), the c25595_g1 PCR amplicon was inserted into a doubly digested BamHI/XhoI pET28b(+) through Gibson assembly according to the manufacturer's instructions. The construct was transformed into BL21 Rosetta(DE3) cells (MilliporeSigma) and verified by colony PCR and Sanger sequencing using T7 promoter and terminator primers (Table S3.1). Michigan State University Research Technology Support Facility Genomics Core (East Lansing, MI, USA) performed the Sanger sequencing. The remaining *S. melongena* and *S. quitoense* candidate genes were cloned into pET28b as described previously (Chapter 2; Leong et al., 2022, 2020).

Candidate enzymes were expressed as previously described (Chapter 2). In brief, 50 mL LB cultures with 1% glucose (w/v) were inoculated from transformation colonies and incubated overnight at 37ºC and 225 rpm. Secondary 1 L cultures were inoculated with 15 mL of the overnight culture and were incubated at 37ºC and 225 rpm. After an OD600 of 0.5 was reached,

the cultures were chilled on ice for 25 min and isopropylthio-β-galactoside was added to a final

concentration of 50 µM, except for SmASAT3-L7 which had a final concentration of 25 µM.

Cultures were then incubated for 16 hours at 16⁰C and 180 rpm. Cells were harvested by

centrifugation at 4,000 rpm for 10 minutes at 4⁰C.

S. melongena BAHDs were purified as described in Chapter 2. *S. quitoense* BAHDs were

purified following the protocol described for SaASAT3-L1 in Chapter 2.

Enzyme assays

Enzyme assays were conducted similar to those in Chapter 2 and in Leong et al., (2022).

All assays were buffered in 100 mM sodium phosphate buffer at pH 6.0 or 8.0 as noted.

For forward assays, acyl-CoAs (MilliporeSigma) were supplied at a final concentration of

0.1 mM each, *myo*-inositol was added to a final concentration of 10 mM. For reverse assays, free

CoA (MilliporeSigma, Burlington, MA, USA) was added to a final concentration of 1 mM.

Purified acylsugar substrates were dried down using a vacuum centrifuge and redissolved in

ethanol:water:formic acid (1:1:0.001). One microliter of the prepared acylsugars were used as

acyl acceptors. Six microliters of each enzyme were added to a final volume of 60 µL. For

negative controls, 6 µL of enzyme that was heat inactivated at 95°C for 10 minutes was

substituted in place of untreated enzyme. In general, assays were incubated at 30°C for 30

minutes, unless noted otherwise. For two-step assays, the first step substrates were incubated at

30°C for 30 minutes, then heat-inactivated at 65C for 10 minutes. Then, the second step reagents

were added and incubated at 30°C for 30 minutes. After incubation, 120 µL of

acetonitrile:isopropanol:formic acid (1:1:0.001) with 1.5 µM telmisartan (MilliporeSigma) stop

solution was added. Reactions were then spun at 17,000 x g for 10 minutes to remove precipitate.

Supernatant was placed in autosampler vials and analyzed by LC-MS.

<u>VIGS analysis</u>

Constructs for VIGS targeting c25595_g1, SnASAT1-L, and SnASAT3 were assembled by adaptation of a previously published method (Dong et al., 2007). pTRV2-LIC was linearized through PstI digestion and subsequent purification from a 1% agarose gel using the New England Biolabs Monarch gel purification kit. The c25595_g1 VIGS target sequences were PCR amplified from *S. quitoense* cDNA using primers with adaptors for ligation (Table S3.1). The SnASAT1-L and SnASAT3 target sequences were PCR amplified from *S. nigrum* leaf cDNA using primers with adaptors for ligation (Table S3.1). In separate 5 µL reactions, The PCR amplicons and linearized vector were incubated with 1X NEB 2.1 reaction buffer, 5 mM dNTP (dATP for PCR product; dTTP for linearized pTRV2-LIC), and T4 polymerase (New England Biolabs) for 30 min at 22ºC followed by 20 min at 70ºC. One µL and 2 µL of the pTRV2-LIC and PCR amplicon reactions, respectively, were mixed together and incubated for 2 min at 65ºC followed by 10 min at 22ºC. The constructs were then transformed into chemically competent cells, Top10 (ThermoFisher Scientific, Waltham, MA, USA) following manufacturer's protocol.

S. quitoense seedlings were inoculated following a modified protocol (Leong et al., 2022; Velásquez et al., 2009). *S. quitoense* seeds were sterilized with 10% bleach (v/v) for 10 minutes while being rocked at 24 rpm with a GyroMini nutating mixer (Labnet, Edison, NJ, USA), and subsequently rinsed with distilled, sterile water 5-6 times. Seeds were soaked in sterile 1000 ppm Giberellin A3 (GoldBio, St. Louis, MO, USA) overnight and then transferred to Whatman filter paper (MilliporeSigma). Seeds were stored in the dark at room temperature for 7 days and then transferred to peat pots (Jiffy, Zwijndrecht, Netherlands). Peat pots were incubated at 25°C, 16/8-h day/night light cycle, and ~70 µmol m^{-2}s^{-2} photosynthetic photon flux density with cool white fluorescent bulbs. Peat pots were covered with a humidity dome, 22 x 11 x 3 inches

(Growers Supply Company) for the first five days. When seedlings had one true leaf and were about an inch tall, inoculation cultures were prepared following Velásquez et al. (2009) with modifications described by Leong et al. (2022). Two independent VIGS experiments were conducted with the c25595_P1 construct used in the first experiment and both c25595_P1 and c25595_P2 constructs used in the second experiment.

Gene expression of SnASAT1-L and SnASAT3 in *S. nigrum* was silenced following a previously developed protocol (Lou et al., 2021) which adapted a published vacuum infiltration protocol (Hartl et al., 2008).

Acylsugar abundances in VIGS leaf surface extracts were quantified by LC-MS with the QuanLynx function in MassLynx v4.1 (Waters Corporation, Milford, MA, USA) as previously described (Lou et al., 2021; Lybrand et al., 2020). In short, QuanLynx generated extracted ion chromatograms for the formate adduct of each acylsugar with a mass window of *m/z* of 0.05. Peak areas were integrated and normalized to the peak area of the telmisartan internal standard to create acylsugar response values. For the *S. quitoense* VIGS data, the acylsugar response values were then normalized to the dry weight of extracted tissue for each sample.

Acylsugar extractions

Acylsugars were extracted from VIGS plants as previously described (Leong et al., 2019; Lou and Leong, 2019). Briefly, half of a leaf was placed into a 1.5 mL tube (Dot Scientific, Inc., Burton, MI, USA) containing 1 mL extraction solvent (3:3:2 acetonitrile:isopropanol:water, 0.1% formic acid, 1 µM telmisartan (internal standard) (MilliporeSigma)). The extractions were rocked at 24 rpm for 2 min with a GyroMini nutating mixer (Labnet). Solvent was extracted and placed in a 2 mL glass autosampler vial (Restek, Bellefonte, PA, USA) and sealed with a 9 mm cap with a PTFE/silicone septum (J.G. Finneran, Vineland, NJ, USA).

VIGS qPCR analysis

qPCR analysis was conducted following a modified protocol (Leong et al., 2022, 2020; Lou et al., 2021). RNA was extracted from VIGS leaf samples with the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) with on-column DNase digestion (Qiagen) following the manufacturer's instructions. RNA was quantified with a Nanodrop 2000c instrument (ThermoFisher Scientific). cDNA was synthesized from 1 µg of RNA and SuperScript II Reverse Transcriptase for *S. nigrum* samples and SuperScript III Reverse Transcriptase for *S. quitoense* samples. *S. quitoense* cDNA was diluted 100-fold, and *S. nigrum* cDNA was then diluted 40-fold. *S. quitoense* qPCRs used SYBR Power Green PCR Master Mix (ThermoFisher Scientific) while *S. nigrum* qPCRs used SYBR Green PCR Master Mix (ThermoFisher Scientific). RT_c25595_1_F/R, RT_SnASAT1-L_1_F/R, RT_SnASAT3_1/2_F/R, RT_Sq_ACTIN_1_F/R, RT_Sq_EF1A_1_F/R, RT_Sn_ACTIN1_F/R, and RT_Sn_ACTIN3_F/R primers were used to detect c25595_g1, SnASAT1-L, SnASAT3, SqACTIN, SqEF1α, SnACTIN1, and SnACTIN3 transcripts, respectively, at a final concentration of 200 nM (Table S3.1). Primers had amplification efficiencies within 90-110%. Michigan State University Research Technology Support Facility Genomics Core (East Lansing, MI, USA) carried out the qPCRs using a QuantStudio 7 Flex Real-Time PCR System (Applied Bio-systems). The following temperature cycling conditions were used: 50℃ for 2 min, 95℃ for 10 min, and 40 cycles of 95℃ for 14 s and 60℃ for 1 min. Relative expression values for c25595_g1, SnASAT1-L, and SnASAT3 were calculated with the ΔΔCT method and normalized to the geometric mean of housekeeping gene transcript levels, ACTIN and EF1α for *S. quitoense* and ACTIN1 and ACTIN3 for *S. nigrum*, and the mean expression values from empty vector plants. Three to four technical replicates were used for all qPCRs.

<u>LC-MS analysis</u>

In vitro and *in planta* acylsugars were analyzed by LC-MS with the methods described below. For all analyses, 10 µL of sample were injected into a column kept at 40ºC, using binary solvent gradients with a flow rate of 0.3 mL/min, unless otherwise noted.

VIGS plant extracts were analyzed with a 7-min LC gradient using a LC-20ADvp ternary pump (Shimadzu, Kyoto, Japan) coupled to a Waters Xevo G2-XS QToF mass spectrometer (Waters Corporation) and equipped with an Ascentis Express C18 HPLC column (10 cm x 2.1 mm, 2.7 µm; Supelco). The gradient with a flow rate of 0.4 mL/min was as follows: 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, held at 5% B until 7 min. Ions were acquired from *m/z* 50 to 1200 with a scan time of 0.1 s and three acquisition functions with different collision potentials (0, 25, 60 V). Lock mass calibration referenced to the leucine enkephalin [M+H]⁻ ion was applied during data acquisition. The ESI⁻ parameters were as follows: capillary voltage, 2 kV; sampling cone voltage, 60 V; source temperature, 100°C; desolvation temperature 350°C; cone gas flow, 50 L/Hr; desolvation gas flow, 600 L/Hr.

Enzyme assay products were analyzed with a Waters Acquity UPLC coupled to a Waters Xevo G2-XS QToF mass spectrometer (Waters Corporation) equipped with electrospray ionization (ESI) source. All enzyme assays used binary solvent gradients with 100% acetonitrile as solvent A and 0.1% formic acid in water as solvent B. Enzyme assay products were analyzed with a general 7-min gradient described in Chapter 2 and Table S2.29, unless noted otherwise. The I2:8(i4,i4) enzyme assays shown in Figure 3.1 and 3.6 used a modified 7-min gradient as follows: 2% B at 0 min, 30% B at 3 min, 99% B at 4 min, held at 99% B until 5 min, 2% B at 5.01 min, held at 2% B until 7 min. The I1:4(i4) enzyme assays shown in Figure S3.4 were

analyzed with a 9-min gradient using an Acquity BEH Amide column (10 cm x 2.1 mm, 130 Å, 1.7 μm; Waters) which was as follows: 95% B at 0 min 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.

For coelution analysis between enzymatically- and plant-produced I3:18(4,4,10), a 14-min linear gradient was used with an Acquity UPLC BEH C18 column (10 cm x 2.1 mm, 130 Å, 1.7 μm; Waters), kept at 40°C, on the same instrument used for enzyme assay analysis. The binary solvent, linear gradient was as follows: 5% B at 0 min, 40% B at 2 min, 90% B at 10 min, 100% B at 10.01 min, held at 100% B until 12 min, 5% B at 12.01 min, held at 5% B until 14 min.

For coelution analysis between enzymatically- and plant-produced I3:22(i4,i4,n14), a 14-min linear gradient was used with an Acquity UPLC BEH C18 column (10 cm x 2.1 mm, 130 Å, 1.7 μm; Waters), kept at 40°C, on the same instrument used for enzyme assay analysis. The binary solvent, linear gradient was as follows: 5% B at 0 min, 60% B at 2 min, 70% B at 8 min, 100% at 10 min, held at 100% B until 12 min, 5% B at 12.01 min, held at 5% B until 14 min.

For all enzyme assays, the following ESI$^-$ parameters were used: capillary voltage, 2 kV; sampling cone voltage, 35 V; source temperature, 100°C; desolvation temperature 350°C; cone gas flow, 50 L/Hr; desolvation gas flow, 600 L/Hr. Ions were acquired from m/z 50 to 1200 with a scan time of 0.1 s and three acquisition functions with different collision potentials (0, 25, 60 V). Lock mass calibration referenced to the leucine enkephalin [M+H]$^-$ ion was applied during data acquisition.

Acknowledgements

pET28b. I acknowledge the Michigan State University RTSF Mass Spectrometry and Metabolomics

Core Facilities for LC-MS analysis support.

REFERENCES

Bouthillette, L.M., Aniebok, V., Colosimo, D.A., Brumley, D., MacMillan, J.B., 2022. Nonenzymatic reactions in natural product formation. Chem. Rev. 122, 14815–14841. https://doi.org/10.1021/acs.chemrev.2c00306

Dong, Y., Burch-Smith, T.M., Liu, Y., Mamillapalli, P., Dinesh-Kumar, S.P., 2007. A ligation-independent cloning tobacco rattle virus vector for high-throughput virus-induced gene silencing identifies roles for NbMADS4-1 and -2 in floral development. Plant Physiol. 145, 1161–1170. https://doi.org/10.1104/pp.107.107391

Du, Y.-L., Williams, D.E., Patrick, B.O., Andersen, R.J., Ryan, K.S., 2014. Reconstruction of cladoniamide biosynthesis reveals nonenzymatic routes to bisindole diversity. ACS Chem. Biol. 9, 2748–2754. https://doi.org/10.1021/cb500728h

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. https://doi.org/10.1038/s41467-017-02045-7

Fan, P., Miller, A.M., Schilmiller, A.L., Liu, X., Ofner, I., Jones, A.D., Zamir, D., Last, R.L., 2015. In vitro reconstruction and analysis of evolutionary variation of the tomato acylsucrose metabolic network. Proc. Natl. Acad. Sci. U.S.A. 113, E239-248. https://doi.org/10.1073/pnas.1517930113

Fiesel, P.D., Kerwin, R.E., Jones, A.D., Last, R.L., 2023. Trading acyls and swapping sugars: metabolic innovations in Solanum trichomes. bioRxiv https://doi.org/10.1101/2023.06.05.542877

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. https://doi.org/10.1039/D2NP00003B

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

Hartl, M., Merker, H., Schmidt, D.D., Baldwin, I.T., 2008. Optimized virus-induced gene silencing in *Solanum nigrum* reveals the defensive function of leucine aminopeptidase against herbivores and the shortcomings of empty vector controls. New Phytol. 179, 356–365. https://doi.org/10.1111/j.1469-8137.2008.02479.x

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

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

Leckie, B.M., D'Ambrosio, D.A., Chappell, T.M., Halitschke, R., De Jong, D.M., Kessler, A., Kennedy, G.G., Mutschler, M.A., 2016. Differential and synergistic functionality of acylsugars in suppressing oviposition by insect herbivores. PLoS ONE 11, e0153345. https://doi.org/10.1371/journal.pone.0153345

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

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

Leong, B.J., Lybrand, D.B., Lou, Y.-R., Fan, P., Schilmiller, A.L., Last, R.L., 2019. Evolution of metabolic novelty: a trichome-expressed invertase creates specialized metabolic diversity in wild tomato. Sci. Adv. 5, eaaw3754. https://doi.org/10.1126/sciadv.aaw3754

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

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

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

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

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

Moseyko, N., Feldman, L.J., 2001. Expression of pH-sensitive green fluorescent protein in *Arabidopsis thaliana*. Plant Cell Environ. 24, 557–563. https://doi.org/10.1046/j.1365-3040.2001.00703.x

Puterka, G.J., Farone, W., Palmer, T., Barrington, A., 2003. Structure-function relationships affecting the insecticidal and miticidal activity of sugar esters. J. Econ. Entomol. 96, 636–644. https://doi.org/10.1093/jee/96.3.636

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

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

Shen, J., Zeng, Y., Zhuang, X., Sun, L., Yao, X., Pimpl, P., Jiang, L., 2013. Organelle pH in the Arabidopsis endomembrane system. Mol. Plant 6, 1419–1437. https://doi.org/10.1093/mp/sst079

Stewart Jr, C., Kang, B.-C., Liu, K., Mazourek, M., Moore, S.L., Yoo, E.Y., Kim, B.-D., Paran, I., Jahn, M.M., 2005. The Pun1 gene for pungency in pepper encodes a putative acyltransferase. Plant J. 42, 675–688. https://doi.org/10.1111/j.1365-313X.2005.02410.x

Velásquez, A.C., Chakravarthy, S., Martin, G.B., 2009. Virus-induced gene silencing (VIGS) in *Nicotiana benthamiana* and tomato. JoVE J. Vis. Exp. e1292. https://doi.org/10.3791/1292

Wuest, W.M., Sattely, E.S., Walsh, C.T., 2009. Three siderophores from one bacterial enzymatic assembly line. J. Am. Chem. Soc. 131, 5056–5057. https://doi.org/10.1021/ja900815w
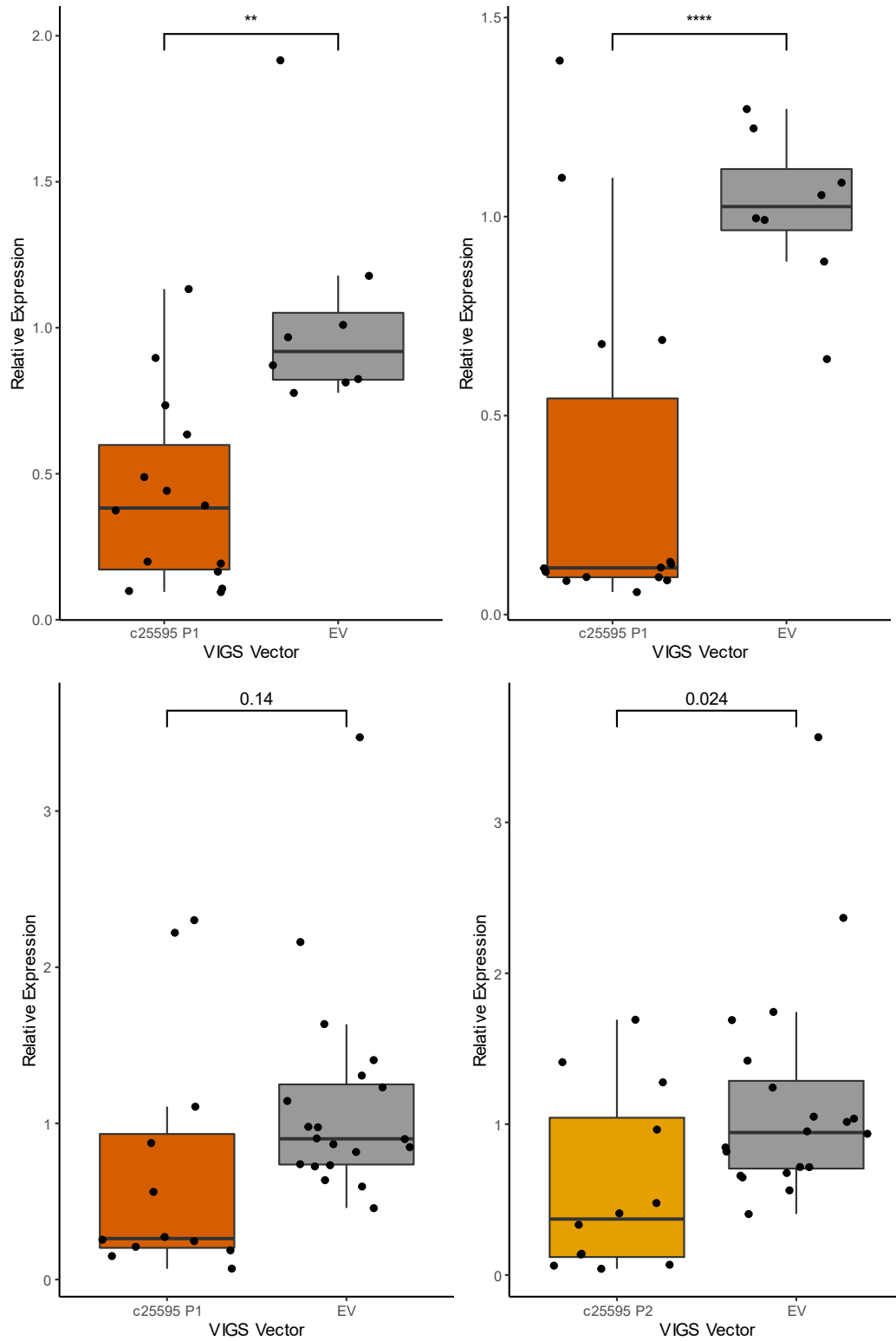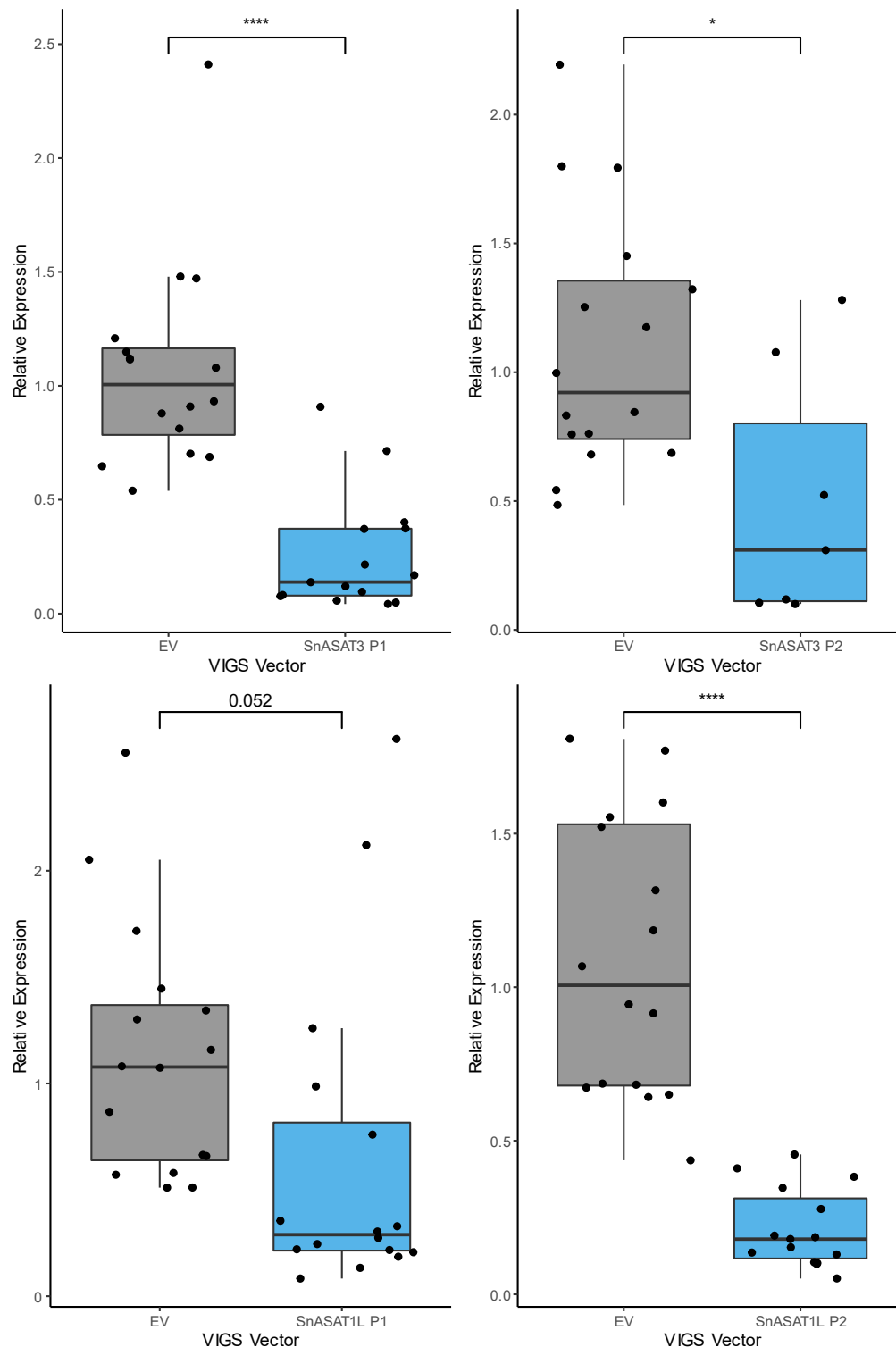
Figure S3.1. Relative transcript abundance of SqASAT3-I in VIGS plants. (A) SqASAT3-I relative transcript abundance from leaf three (left) and leaf four (right). c25595 P1, *n* = 16; empty

Figure S3.1. (cont'd)

vector, $n = 8$. (B) SqASAT3-I relative transcript abundance from plants with two different gene fragments targeting SqASAT3-I, c25595 P1 (left) and c25595 P2 (right). c25595 P1, $n = 12$; c25595 P2, $n = 12$; empty vector, $n = 20$. **$P < 0.01$, ****$P < 0.0001$. Statistical comparisons were conducted with Welch's two-sample t test. 3-4 technical replicates were used for each sample.

Figure S3.2. Relative transcript abundance of SnASAT1-L and SnASAT3 in VIGS plants.
SnASAT1-L was targeted in plants with two different constructs, SnASAT1-L P1 (top left) and
SnASAT1-L P2 (top right). SnASAT3 was targeted in plants with two different constructs,
SnASAT3 P1 (bottom left) and SnASAT3 P2 (bottom right). Top left: SnASAT1-L P1, $n = 16$;
empty vector, $n = 15$. Top right: SnASAT1-L P2, $n = 7$; empty vector, $n = 16$. Bottom left:

Figure S3.2. (cont'd)
SnASAT3 P1, $n = 16$; empty vector, $n = 16$. Bottom right: SnASAT3 P2, $n = 14$; empty vector, $n = 16$. $*P < 0.05$, $****P < 0.0001$. Statistical comparisons were conducted with Welch's two-sample t test. 3-4 technical replicates were used for each sample.
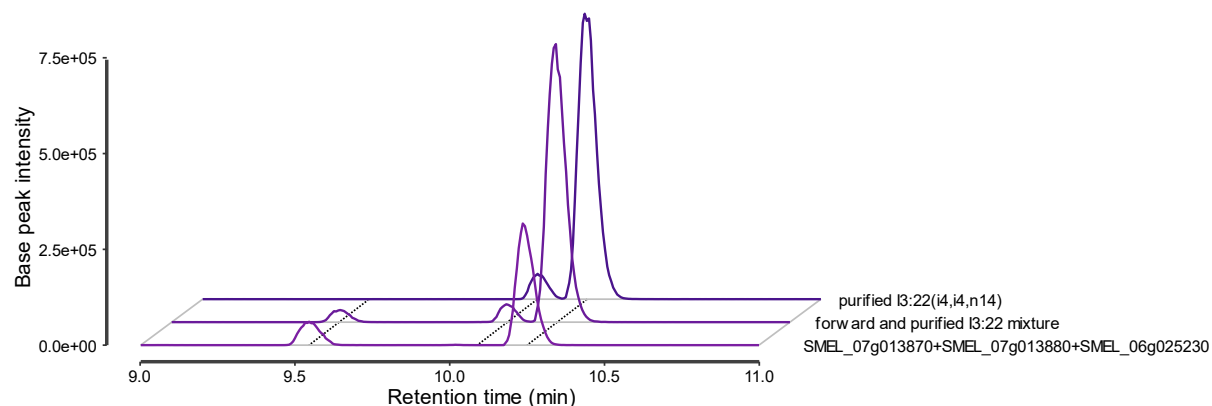
Figure S3.3. Coelution analysis of I3:22(i4,i4,n14) *in vitro* **assay product.** Extracted ion chromatograms display the formate adduct of I3:22, 575.35 *m/z*. The enzymatically-produced I3:22 was synthesized in a two-step enzyme reaction with SmASAT1-L, SmASAT3-L7, *myo*-inositol, and iC4-CoA in the first step. After a heat-inactivation, SmASAT3-I and nC14-CoA were added for the second step incubation.
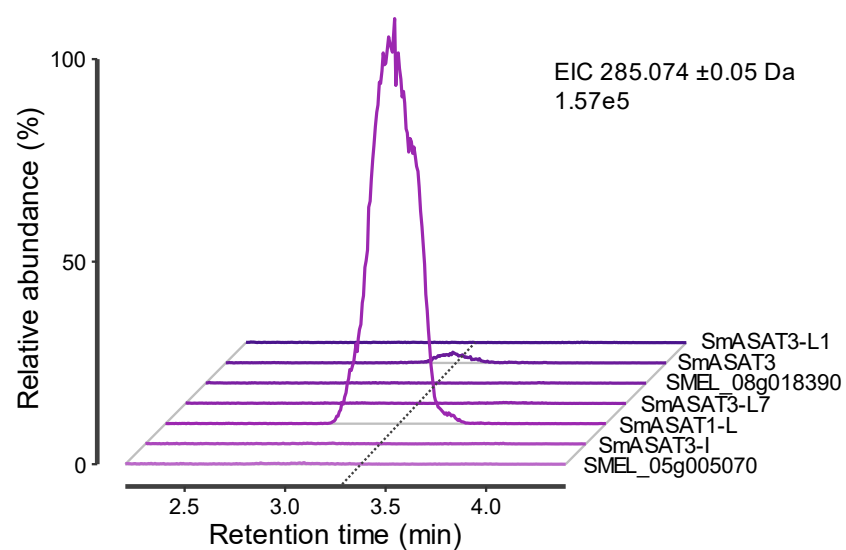
Figure S3.4. SmASAT1-L and SmASAT3 acylate myo-inositol with iC4-CoA. Extracted ion chromatograms display the chloride adduct of I1:4(i4), 285.074 *m/z*. Products were analyzed with the 9 min BEH amide LC method. Assays were conducted at pH 8.0.
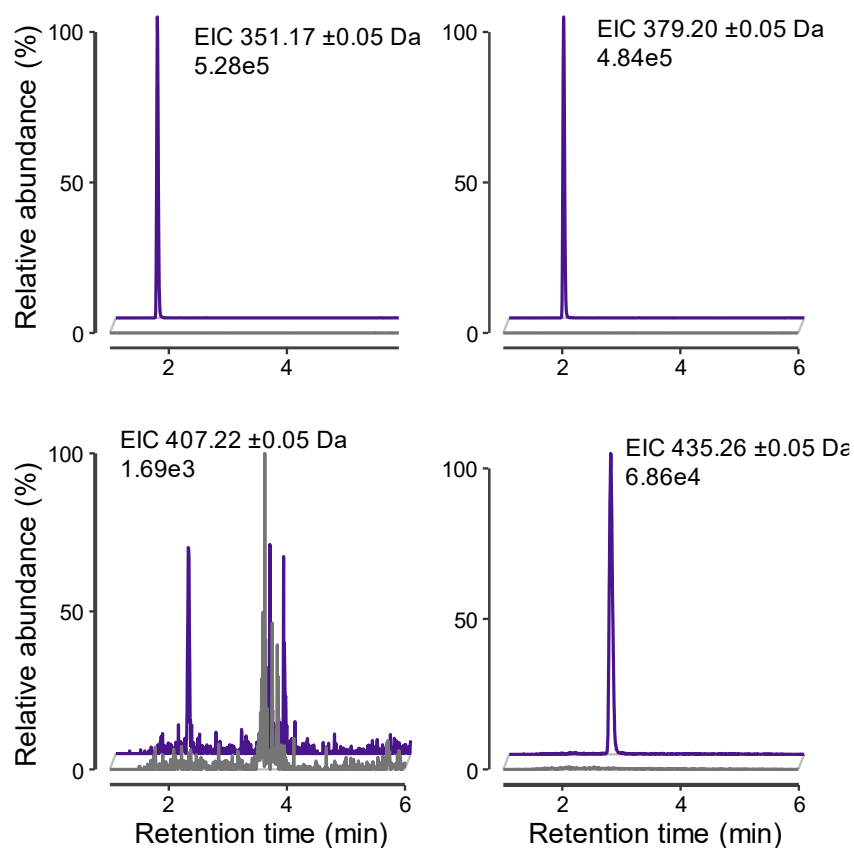
Figure S3.5. SmASAT1-L acylates *myo*-inositol with medium length acyl chains. nC8- (top left), nC10- (top right), nC12- (bottom left), nC14-CoAs (bottom left) were supplied to SmASAT1-L (purple traces) or boiled SmASAT1-L (gray traces) and *myo*-inositol. The top left chromatogram displays the formate adduct for I1:8(nC8), 351.17 *m/z*. The top right chromatogram displays the formate adduct for I1:10(nC10), 379.20 *m/z*. The bottom left chromatogram displays the formate adduct for I1:12(nC12), 407.22 *m/z*. The bottom right chromatogram displays the formate adduct for I1:14(nC14), 435.26 *m/z*. All reactions were conducted at pH 6.0.
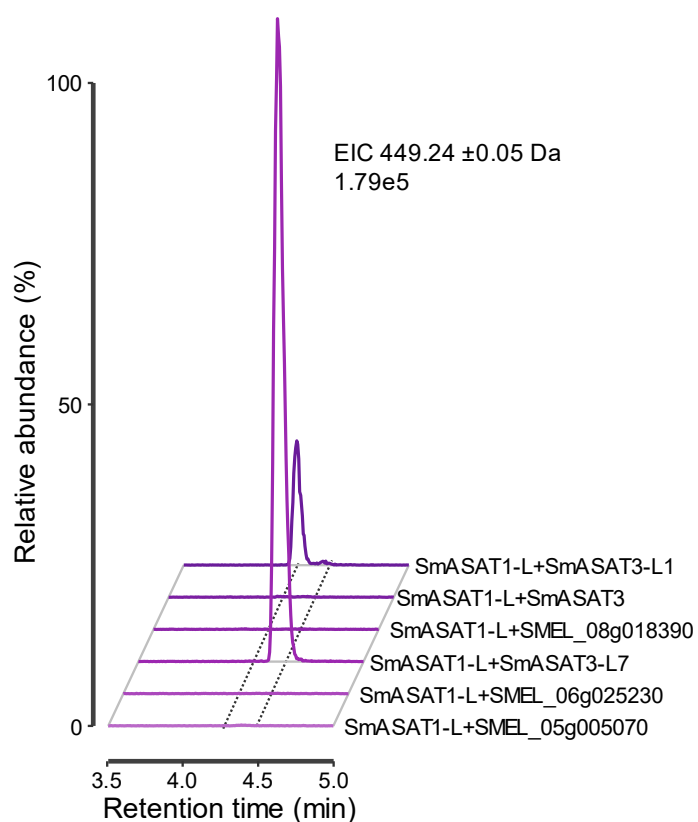
Figure S3.6. I2:14(i4,n10) formed in one-pot assays with the enzyme combinations of SmASAT1-L+SmASAT3-L7 and SmASAT1-L+SmASAT3-L1. The forward assays were conducted at pH 6.0 and were analyzed with the general 7-min enzyme assay LC gradient. Extracted ion chromatograms display the formate adduct of I2:14(i4,n10), 449.24 *m/z*.
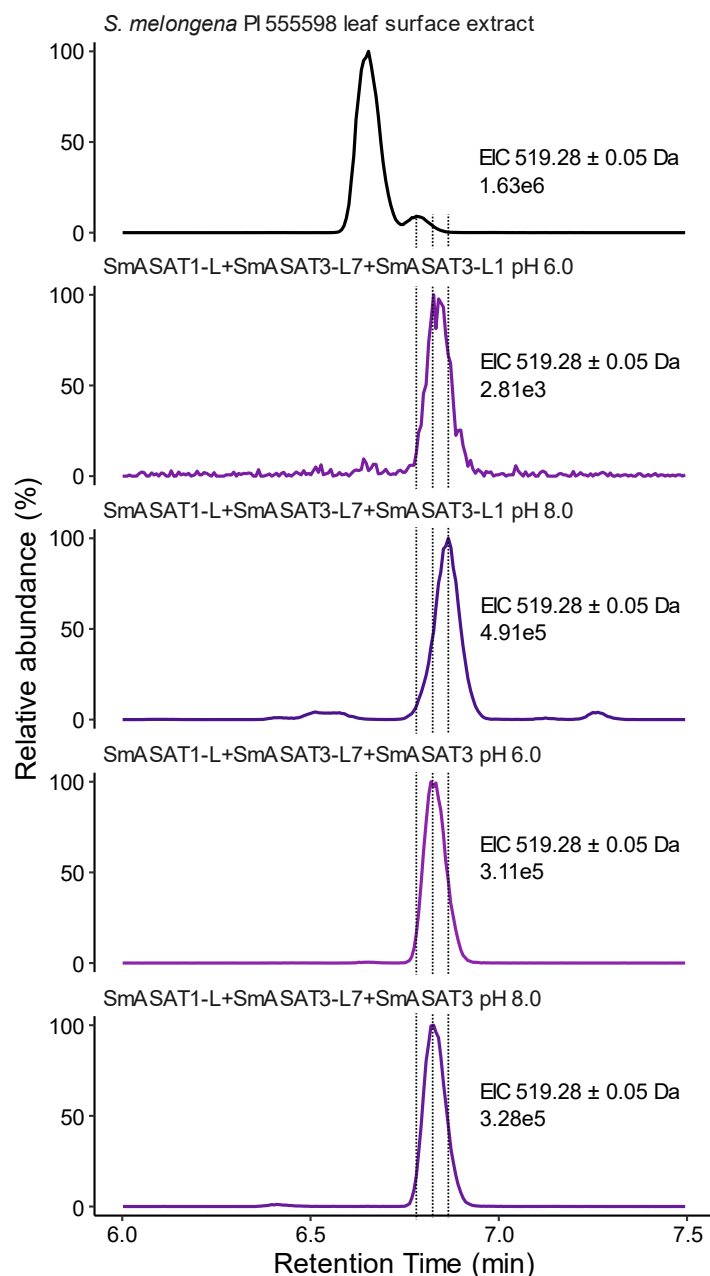
Figure S3.7. In two-step assays, SmASAT3 and SmASAT3-L1 acylate I2:14(i4,n10) to produce I3:18(i4,i4,n10) that does not coelute with any plant-produced I3:18. The first step included SmASAT1-L, SmASAT3-L7, iC4-CoA, and iC10-CoA generating I2:14(i4,n10). After a heat inactivation step, iC4-CoA and SmASAT3 or SmASAT3-L1 were added generating I3:18(i4,i4,n10). Reactions were conducted at pH 6.0 or pH 8.0 and their products were compared to I3:18 from a *S. melongena* plant extract (top). The largest peak in the plant extract is the NMR characterized I3:18(i4,i4,i10) and the smaller peak is likely I3:18(i4,i4,n10). This annotation of the smaller peak is supported by two pieces of evidence: acylinositols with straight medium acyl chains tend to elute after their isomers with iso-branched medium acyl chains (Chapter 2), and GC-MS analysis of acyl chain composition detected nC10 acyl chains (Chapter 2). Extracted ion chromatograms display the formate adduct of I3:18, 519.28 *m/z*.

Table S3.1. Oligonucleotides used in this study.

Sequence name	Sequence	Description
c25595_g1_F	agcatgactggtggacagcaaatgggtcggATGTCATGTGTATATAAAGCTAGCTTCTTT	Sequence used for cloning into pET28b with BamHI/XhoI site
c25595_g1_F	gccggatctcagtggtggtggtggtggtgcAAATAATGTTCACCTTAGTAACGTTTTGAA	Sequence used for cloning into pET28b with BamHI/XhoI site
T7_promoter	TAATACGACTCACTATAGGG	Primer for Sanger sequencing and colony PCR of pET28b plasmids
T7_terminator	GCTAGTTATTGCTCAGCGG	Primer for Sanger sequencing and colony PCR of pET28b plasmids
RT_c25595_1_F	GCCAAGGGTCCATGTAACAAG	For qPCR analysis of c25595 expression
RT_c25595_1_R	TGAAGAAGTTGCTCGTCACG	For qPCR analysis of c25595 expression
RT_SnASAT1-L_1_F	CGGATCCCGTAGAGATCAAA	For qPCR analysis of SnASAT1-L expression
RT_SnASAT1-L_1_R	CATCCAAAGTCCATGTCACG	For qPCR analysis of SnASAT1-L expression
RT_SnASAT3_1_F	GCCTTCTTCTACCCCAAACC	For qPCR analysis of SnASAT3 expression
RT_SnASAT3_1_R	TTCGGACATTGAACGATTGA	For qPCR analysis of SnASAT3 expression
RT_SnASAT3_2_F	AATGAGTGCGTGCATGTCAC	For qPCR analysis of SnASAT3 expression
RT_SnASAT3_2_R	CTGTAGCGGCCCAATCATTAATG	For qPCR analysis of SnASAT3 expression
RT_Sq_ACTIN_1_F	GTGTGATGGTTGGCATGGGG	For qPCR analysis of SqActin expression
RT_Sq_ACTIN_1_R	GGTGTTCCTCAGGGGCAACA	For qPCR analysis of SqActin expression
RT_Sq_EF1A_1_F	CTCAAGGCTGAGCGTGAACG	For qPCR analysis of SqEF1alpha expression
RT_Sq_EF1A_1_R	GCACAGTCAGCCTGAGAGGT	For qPCR analysis of SqEF1alpha expression
RT_Sn_ACTIN1_F	ACAATTGGTGCTGAGCGTTT	For qPCR analysis of SnActin1 expression
RT_Sn_ACTIN1_R	TTTCAGGTGGAGCAACAACC	For qPCR analysis of SnActin1 expression
RT_Sn_ACTIN3_F	TCACAGAGCGTGGTTACTCG	For qPCR analysis of SnActin3 expression
RT_Sn_ACTIN3_R	CTGCTTCCATTCCGATCATT	For qPCR analysis of SnActin3 expression
Sq_c25595_g1_i1_VIGS_01_F	CGACGACAAGACCCTCATGGGCTGGAAGTCGAAGA	Sequence used for cloning into pTRV2
Sq_c25595_g1_i1_VIGS_01_R	GAGGAGAAGAGCCCTGCTAATGCAGCCCAATCCCT	Sequence used for cloning into pTRV2

Table S3.1. (cont'd)

Sq_c25595_g1_i1_VIGS_02_F	CGACGACAAGACCCTTTCCTTTGGACCATCGCCAA	Sequence used for cloning into pTRV2
Sq_c25595_g1_i1_VIGS_02_R	GAGGAGAAGAGCCCTTTTGGGGGTAGCTGAGGTGA	Sequence used for cloning into pTRV2

CHAPTER 4:

CONCLUSIONS AND FUTURE DIRECTIONS

The Solanaceae family continues to be a rich resource for understanding metabolic evolution. In this dissertation, an enormous amount of acylsugar diversity was revealed within the *Solanum* genus making use of available germplasm. Utilizing this diversity helped uncover pieces of the acylinositol biosynthetic pathway, revealing familiar evolutionary mechanisms of gene duplication and loss leading to metabolic diversity. This research greatly expands our knowledge of plant diversity and provides exciting opportunities for uncovering further metabolic evolution.

The acylsugar chemical analyses presented in Chapter 2 expand and redefine our understanding of acylsugar structural diversity. Acylsucroses were thought to be the predominant acylsugar type in contrast to the sporadically observed acylglucoses and rarely observed acylinositols, as described in Chapter 1 (Fiesel et al., 2022). Contrasting with this paradigm, I found acylinositols were widespread with the megadiverse *Solanum* genus, being detected in multiple major *Solanum* clades representing ~1100 of the c. 2700 Solanaceae species. In fact, I hypothesized that the largest *Solanum* clade, Clade II, lost acylsucrose biosynthesized as supported by a lack of detectable acylsucroses in this clade. Considering that acylsugars have been studied for decades, how did these widespread acylinositols go overlooked? A few reasons play a role in this. First, prior work in the Solanaceae has been biased by the tractable acylsucrose-producer cultivated tomato and other readily available ornamental plants such as *Nicotiana*, *Salpiglossis*, and *Petunia* species. In addition, the most cultivated acylinositol-producer, *S. melongena*, hardly produces glandular trichomes (Figure S2.12). This broad study of *Solanum* acylsugars not only showed that acylinositols are widespread but also uncovered 18 acyl chains and two new sugar cores not previously observed. With 33 acyl chains (Chapter 2, Lou et al., 2021; Leong et al., 2020; Herrera-Salgado et al., 2005) and six inositol acylation

positions, there are 1.3 billion theoretically possible acylinositol structures outnumbering the previously estimated 820 million theoretically possible acylsucrose structures (Fan et al., 2019). This acylinositol estimate does not even account for glycosylated acylinositols and glycohydroxyacylhexoses, demonstrating that this study greatly expands our knowledge of acylsugar structural diversity.

Plants synthesize a tremendous amount of chemical diversity, yet characterizing this remains technically challenging and time consuming. This was exemplified by the analysis of *Solanum* acylsugars in Chapter 2 in which more than a hundred acylsugars could be identified in a single species. Because acylsugar isomers were not always resolved into distinct peaks, our reported acylsugar numbers are almost certainly an underestimate of the true diversity. Incorporating orthogonal techniques such as ion mobility can increase compound separation without compromising on acquisition time. Along with chromatographic resolution limitations, compound identification continues to be a major bottleneck in plant metabolomic studies (Chaleckis et al., 2019; da Silva et al., 2015). In the work described in Chapter 2, acylsugars were identified manually, a time-intensive method. Other software methods exist but exhibit significant limitations. For example, MS/MS database queries can accurately identify compounds, but many databases lack plant metabolites. Other methods such as the machine learning based CANOPUS bin compounds into classes (Dührkop et al., 2021), but do not provide specific compound annotations. The scientific community must address these data collection and analysis limitations in order to enhance our ability to characterize plant chemical diversity.

Our understanding of plant chemistry is also limited by available germplasm. The *Solanum* acylsugar survey covered many of the minor clades composing Clade II, VANAns, and DulMo clades, yet many clades were unrepresented (e.g. Micracantha, Crinitum, Erythrotrichum,

and Thomasiifolium) as restricted by obtainable germplasm. Expanding access to diverse germplasm is key to understanding plant diversity, however navigating the complex legal and ethical constraints of importing plant germplasm remains difficult (Buck and Hamilton, 2011; Prathapan et al., 2018; Secretariat of the Convention on Biological Diversity, 2011). While this issue is beyond the scope of this dissertation, as a scientific community we should work collaboratively with international research groups to characterize and understand plant diversity (Pearce et al., 2020).

With the acylsugar characterizations in Chapter 2 and available genetic resources, pieces of the acylinositol pathway were discovered revealing familiar evolutionary themes. Gene duplication was repeatedly observed as an important driver in acylsugar structure evolution likely due to relaxed selection on duplicated genes leading to altered enzymatic activities (Carretero-Paulet and Fares, 2012; Moghe et al., 2014; Ohno, 2013; Panchy et al., 2016). This was observed within the tomato clade in which ASAT3 duplicated and neofunctionalized to acylate a different sucrose hydroxyl (Fan et al., 2017). Similarly, a duplicated ASAT1, ASAT1-L, with *myo*-inositol acylating activity led to the evolution of acylinositols within the *Solanum* genus. In addition, gene loss impacted acylsugar phenotypes both in the tomato clade, e.g. ASAT4 and IPMS (Kim et al., 2012; Ning et al., 2015), and the eggplant clade, ASAT3-L1 (Chapter 3). More recently, we observed the co-option of "new" acyltransferases, i.e acyltransferases not orthologous or paralogous to characterized ASATs, with unique acylating activities. Two of these "new" acyltransferases are SnAGAT1 which represents the first reported ASAT to acylate glucose (Lou et al., 2021) and SmASAT3-I which exhibits unique activity with hydroxyacyl chains (Chapter 4).

The investigation of 31 out of more than 1200 *Solanum* species reported a tremendous amount of acylsugar diversity, indicating we are just scratching the surface of plant metabolic diversity. Expanding this analysis to different tissue types, specialized metabolite classes, and additional *Solanum* clades promise to uncover further chemical diversity and provide insights into plant metabolic evolution. One exciting prospect lies in addressing the evolution, diversity, and phylogenetic distribution of the recently identified root acylsugars through a phylogenetically-guided metabolomic analysis (Korenblum et al., 2020; Kerwin et al., unpublished).

Commensurate with our expanded understanding of *Solanum* acylsugar diversity described in Chapter 2 comes many exciting opportunities to uncover how these compounds are produced and evolved. Exploring acylinositol glycoside structural variation and differences in acyl chain structural and positional differences between *S. melongena* and *S. nigrum*, as documented in Chapter 2 and previous reports (Hurney, 2018; Lou et al., 2021), are both exciting avenues for future work. Additionally, acylsugar-like compounds were reported to be produced outside of the Solanaceae family (Liu et al., 2019; Moghe et al., 2023), and their unknown biosynthetic pathways introduce many research opportunities. Of the non-Solanaceae acylsugars, investigation of the Asteraceae *Inula ssp.* produced acylinositols would provide further insights into acylinositol evolution and biosynthesis (Sun et al., 2021; Wang et al., 2013; Wu et al., 2015; Zou et al., 2008). This dissertation advanced our understanding of acylsugar diversity, biosynthesis, and evolution, and I expect this work will serve as a foundation for other researchers.

REFERENCES

Buck, M., Hamilton, C., 2011. The Nagoya Protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilization to the convention on biological diversity. Review of European Community & International Environmental Law 20, 47–61. https://doi.org/10.1111/j.1467-9388.2011.00703.x

Carretero-Paulet, L., Fares, M.A., 2012. Evolutionary dynamics and functional specialization of plant paralogs formed by whole and small-scale genome duplications. Molecular Biology and Evolution 29, 3541–3551. https://doi.org/10.1093/molbev/mss162

Chaleckis, R., Meister, I., Zhang, P., Wheelock, C.E., 2019. Challenges, progress and promises of metabolite annotation for LC–MS-based metabolomics. Current Opinion in Biotechnology, Analytical Biotechnology 55, 44–50. https://doi.org/10.1016/j.copbio.2018.07.010

da Silva, R.R., Dorrestein, P.C., Quinn, R.A., 2015. Illuminating the dark matter in metabolomics. Proceedings of the National Academy of Sciences 112, 12549–12550. https://doi.org/10.1073/pnas.1516878112

Dührkop, K., Nothias, L.-F., Fleischauer, M., Reher, R., Ludwig, M., Hoffmann, M.A., Petras, D., Gerwick, W.H., Rousu, J., Dorrestein, P.C., Böcker, S., 2021. Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. Nat Biotechnol 39, 462–471. https://doi.org/10.1038/s41587-020-0740-8

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. https://doi.org/10.1016/j.pbi.2019.03.005

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. https://doi.org/10.1038/s41467-017-02045-7

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

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

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

Korenblum, E., Dong, Y., Szymanski, J., Panda, S., Jozwiak, A., Massalha, H., Meir, S., Rogachev, I., Aharoni, A., 2020. Rhizosphere microbiome mediates systemic root

metabolite exudation by root-to-root signaling. P. N. A. S. U.S.A. 117, 3874–3883. https://doi.org/10.1073/pnas.1912130117

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

Liu, Y., Jing, S.-X., Luo, S.-H., Li, S.-H., 2019. Non-volatile natural products in plant glandular trichomes: chemistry, biological activities and biosynthesis. Nat. Prod. Rep. 36, 626–665. https://doi.org/10.1039/C8NP00077H

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

Moghe, G., Irfan, M., Sarmah, B., 2023. Dangerous sugars: Structural diversity and functional significance of acylsugar-like defense compounds in flowering plants. Curr. Opin. Plant Biol. 73, 102348. https://doi.org/10.1016/j.pbi.2023.102348

Moghe, G.D., Hufnagel, D.E., Tang, H., Xiao, Y., Dworkin, I., Town, C.D., Conner, J.K., Shiu, S.-H., 2014. Consequences of whole-genome triplication as revealed by comparative genomic analyses of the wild radish *Raphanus raphanistrum* and three other Brassicaceae species. Plant Cell 26, 1925–1937. https://doi.org/10.1105/tpc.114.124297

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

Ohno, S., 2013. Evolution by gene duplication. Springer Science & Business Media.

Panchy, N., Lehti-Shiu, M., Shiu, S.-H., 2016. Evolution of gene duplication in plants. Plant Physiol. 171, 2294–2316. https://doi.org/10.1104/pp.16.00523

Pearce, T.R., Antonelli, A., Brearley, F.Q., Couch, C., Campostrini Forzza, R., Gonçalves, S.C., Magassouba, S., Morim, M.P., Mueller, G.M., Nic Lughadha, E., Obreza, M., Sharrock, S., Simmonds, M.S.J., Tambam, B.B., Utteridge, T.M.A., Breman, E., 2020. International collaboration between collections-based institutes for halting biodiversity loss and unlocking the useful properties of plants and fungi. PLANTS, PEOPLE, PLANET 2, 515–534. https://doi.org/10.1002/ppp3.10149

Prathapan, K.D., Pethiyagoda, R., Bawa, K.S., Raven, P.H., Rajan, P.D., 172 CO-SIGNATORIES FROM 35 COUNTRIES, 2018. When the cure kills—CBD limits biodiversity research. Science 360, 1405–1406. https://doi.org/10.1126/science.aat9844

Secretariat of the Convention on Biological Diversity, 2011. Nagoya Protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their

utilization to the convention on biological diversity : text and annex. (Report). Secretariat of the Convention on Biological Diversity. https://doi.org/10.25607/OBP-789

Sun, C.-P., Jia, Z.-L., Huo, X.-K., Tian, X.-G., Feng, L., Wang, C., Zhang, B.-J., Zhao, W.-Y., Ma, X.-C., 2021. Medicinal Inula species: phytochemistry, biosynthesis, and bioactivities. Am. J. Chin. Med. 49, 315–358. https://doi.org/10.1142/S0192415X21500166

Wang, C., Zhang, X., Wei, P., Cheng, X., Ren, J., Yan, S., Zhang, W., Jin, H., 2013. Chemical constituents from Inula wissmanniana and their anti-inflammatory activities. Arch. Pharm. Res. 36, 1516–1524. https://doi.org/10.1007/s12272-013-0143-1

Wu, J., Tang, C., Yao, S., Zhang, L., Ke, C., Feng, L., Lin, G., Ye, Y., 2015. Anti-inflammatory inositol derivatives from the whole plant of *Inula cappa*. J. Nat. Prod. 78, 2332–2338. https://doi.org/10.1021/acs.jnatprod.5b00135

Zou, Z.-M., Xie, H.-G., Zhang, H.-W., Xu, L.-Z., 2008. Inositol angelates from the whole herb of *Inula cappa*. Fitoterapia 79, 393–394. https://doi.org/10.1016/j.fitote.2007.11.031