Article

An Integrated Analytical Approach Reveals Trichome Acylsugar Metabolite Diversity in the Wild Tomato *Solanum Pennellii*

Daniel B. Lybrand [1], Thilani M. Anthony [1], A. Daniel Jones [1] and Robert L. Last [1,2,*]

[1] Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI, USA; dblybrand@ucdavis.edu (D.B.L); thilani@chemistry.msu.edu (T.M.A); jonesar4@msu.edu (A.D.J)
[2] Department of Plant Biology, Michigan State University, East Lansing, MI, USA
* Correspondence: lastr@msu.edu

Abstract: Acylsugars constitute an abundant class of pest- and pathogen-protective Solanaceae family plant specialized metabolites produced in secretory glandular trichomes. *Solanum pennellii* produces copious triacylated sucrose and glucose esters, and the core biosynthetic pathway producing these compounds was previously characterized. We performed untargeted metabolomic analysis of *S. pennellii* surface metabolites from accessions spanning the species range, which indicated geographic trends in acylsugar profile and revealed two compound classes previously undescribed from this species, tetraacylglucoses and flavonoid aglycones. A combination of ultrahigh performance liquid chromatography high resolution mass spectrometry (UHPLC-HR-MS) and NMR spectroscopy identified variations in number, length, and branching pattern of acyl chains, and the proportion of sugar cores in acylsugars among accessions. The new dimensions of acylsugar variation revealed by this analysis further indicate variation in the biosynthetic and degradative pathways responsible for acylsugar accumulation. These findings provide a starting point for deeper investigation of acylsugar biosynthesis, an understanding of which can be exploited through crop breeding or metabolic engineering strategies to improve endogenous defenses of crop plants.

Keywords: acylsugar; wild tomato; *Solanum pennellii*; secretory glandular trichome; specialized metabolism; intraspecific variation; metabolomics

1. Introduction

Plants produce thousands of lineage-specific compounds termed specialized metabolites [1]. Unlike the highly conserved core pathways common to nearly all plants, specialized metabolic pathways evolve rapidly, leading to tremendous structural and functional diversity (*e.g.*, terpenoids, flavonoids) [2–4]. These pathways and their products provide a chemical palette that mediates interactions between plants and biotic or abiotic stressors in their environments. Many of these specialized metabolites accumulate in specialized structures, including epidermal secretory glandular trichomes (SGTs) [5], which act as a first line of defense against herbivores and pathogens [6–9].

Plants of the genus *Solanum*, which includes tomato, potato, and eggplant, possess multiple types of SGTs that produce a diverse array of specialized metabolites [10,11] including acylsugars, which accumulate to up to 20% leaf dry mass in the wild tomato *Solanum pennellii* [12]. *S. pennellii* acylsugars consist of sucrose or glucose cores esterified with fatty acid acyl groups of variable length and branching pattern (Fig. 1). These compounds defend *S. pennellii* and other *Solanum* species from insect pests including silverleaf whitefly (*Bemisia tabaci*), western flower thrips (*Frankliniella occidentalis*), and army beetworm (*Spodoptera exigua*) [13–15]. The antioviposition and antiherbivory properties of *S. pennellii* acylsugars prompted efforts to breed cultivated tomato (*Solanum lycopersicum*) varieties with *S. pennellii*-like acylsugar profiles [13,15,16]. Such efforts are aided by

 doi:10.20944/preprints202008.0702.v1

knowledge of the genetic loci underlying acylsugar biosynthesis [15–18]. The *S. pennellii* core acylsugar biosynthetic pathway consists of three BAHD-family acylsugar acyltransferases (ASATs) that sequentially transfer acyl groups from coenzyme A (CoA) donors to a sucrose acceptor, yielding triacylsucroses (Fig. 1) [19]. As shown in Figure 1, enzymes involved in acyl CoA biosynthesis [*e.g.*, acyl CoA synthetase (ACS), enoyl CoA hydratase (ECH), and isopropylmalate synthase3 (IPMS3)] affect the structures of triacylsucroses by modulating the available pool of acyl CoA donors [20,21], while enzymes that selectively cleave acyl chains from intact acylsugars [acylsugar acylhydrolases (ASHs)] or hydrolyze triacylsucroses to form triacylglucoses [acylsucrose fructofuranosidase1 (ASFF1)] influence steady-state acylsugar profiles and facilitate rapid acylsugar turnover [22–25].

Previous studies of acylsugar metabolism in *S. pennellii* focused on either the chemical substructures of acylsugars (*i.e.*, sugar cores and acyl chains; [21,26,27]) or the enzymes that synthesize and degrade acylsugars (ACS, ASATs, ASFF1, ASHs, ECH, IPMS3) [19–23]. This work revealed that *S. pennellii* accumulates a mixture of acylglucoses and acylsucroses [26,28], collectively containing acyl chains with at least 13 unique structures [21,26,27]. Although most studies focused on the southern Peruvian *S. pennellii* LA0716 [29] acylsugars, [19,22,23,27,30–33], Shapiro and co-workers quantified abundance of acylsugar substructures from 19 accessions of *S. pennellii* distributed across the range of the species [26], while Ning and co-workers analyzed acylsugar acyl chains in 14 *S. pennellii* accessions to determine the genetic basis for differential accumulation of 3-methylbutanoate and 2-methylpropanoate acyl chains in northern and southern regions of Peru [21]. Knowledge of relative acylsugar substructure abundances among related species or populations provides insight into the biosynthesis of these compounds, which facilitates use of acylsugars in crop defense and illuminates evolution of specialized metabolic pathways [15,21].

A complete understanding of acylsugar biosynthesis and evolution requires knowledge of specific acylsugar structures as revealed by metabolomic approaches, including untargeted liquid chromatography mass spectrometry (LC-MS) of intact molecules or structural resolution by NMR spectroscopy. For example, an early report on *S. pennellii* acylsugar metabolism in which Burke and co-workers partially characterized acylglucoses in *S. pennellii* LA0716 by NMR [28] provided information on the number of acyl chains and established esterification on the 2-, 3-, and 4-positions, which later facilitated discovery of the three ASATs constituting the core acylsugar biosynthetic pathway [19,34]. A combination of untargeted metabolomic analysis of acylsugars in *Solanum habrochaites* and *Petunia axillaris* [35,36] using LC-MS and NMR spectroscopy facilitated elucidation of core acylsugar pathways in these species [19,37].

Structural information about *in planta* intact acylsugars is essential for discovering and characterizing enzymes in the acylsugar biosynthetic pathway of a single species and for comparing pathways between species. To create a more complete picture of acylsugar diversity in *S. pennellii*, we combined untargeted ultrahigh performance liquid chromatography-high resolution mass spectrometry (UHPLC-HR-MS) and NMR spectroscopy to characterize the SGT metabolome of 16 *S. pennellii* accessions, revealing variation in levels of 43 specialized metabolites including 39 acylsugars. We initially annotated all metabolites based on mass spectra and subsequently purified and resolved structures of selected acylsugars by NMR. Multivariate statistical analyses of these profiling data recognized specific compounds that distinguish various *S. pennellii* accessions from one another. Our analyses confirmed previous reports showing that acyl chain complement drives acylsugar variation between *S. pennellii* accessions [21,26], and revealed a positive correlation between expression of the *ASFF1* gene that facilitates acylsucrose hydrolysis [22] and acylglucose accumulation. We also observed tetraacylglucoses and methyl flavonoids, two classes of compounds previously undescribed in *S. pennellii* SGTs.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

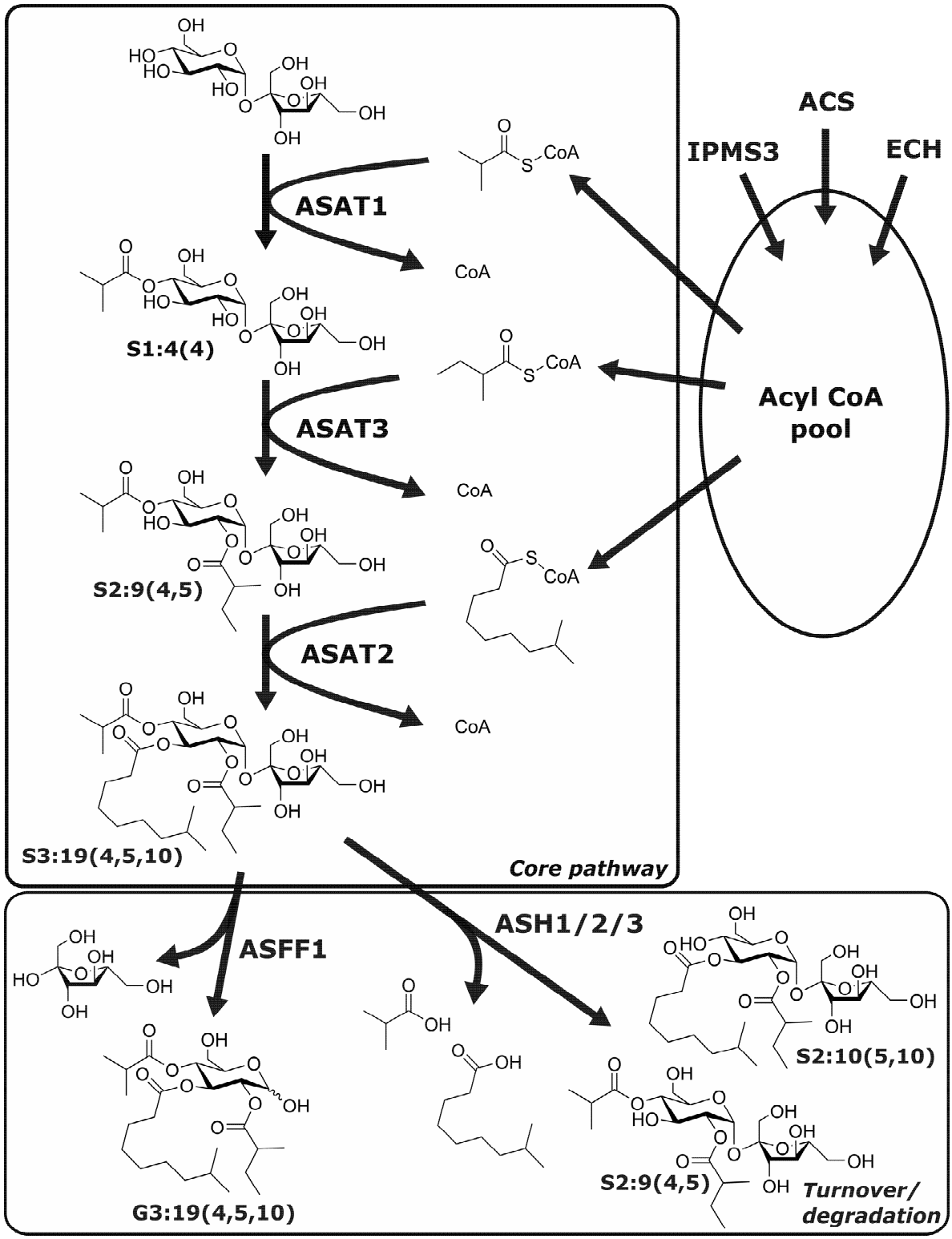


Figure 1. The acylsugar biosynthetic pathway in *S. pennellii*. Enzymes including IPMS3, ACS, and ECH contribute to production of the acyl CoA pool. ASATs constitute the core acylsugar pathway and transfer acyl chains from acyl CoA molecules to a sucrose core. ASFF1 and ASHs catalyze acylsugar turnover or degradation by hydrolyzing the fructose moiety of the sugar core and acyl chains, respectively. Acylsugar nomenclature is as follows: the first letter indicates the sugar core ("S" for sucrose, "G" for glucose); the number before the colon indicates the number of acyl chains; the number after the colon indicates the sum of carbons in all acyl chains; the numbers in parentheses indicate the number of carbons in individual acyl chains. ACS – acyl CoA synthetase; ASAT – acylsucrose acyltransferase; ASFF – acylsucrose fructofuranosidase; ASH – acylsugar hydrolase; CoA – coenzyme A; ECH – enoyl CoA hydratase; IPMS – isopropyl malate synthase.

 doi:10.20944/preprints202008.0702.v1

2. Results

2.1 Experimental Design

Previous studies indicated that wild tomato species including *Solanum pennellii* exhibit intraspecific variation in the amounts and types of acylsugars produced [21,26,38]. To identify geographic trends in acylsugar quantity and quality in *S. pennellii*, we extracted compounds from the surfaces of leaflets from 16 accessions spanning the 1500-km geographic range of the species (Fig. 2). Six biological replicates of each accession were sampled to capture intra-accession metabolic diversity. We included eight accessions from the northern portion of Peru (north range) and eight from the southern portion (south range) and classified two clusters of accessions within the south range by region including the southernmost Atico group and the Pisco group. A group of accessions from the Nazca region, described as *S. pennellii var. puberulum*, are trichome-deficient and exhibit minimal accumulation of acylsugars and transcripts of genes associated with acylsugar metabolism [26,39]; our pilot experiments confirmed the absence of detectable acylsugars in this group and these accessions were excluded from this study. All extracts were analyzed by UHPLC-HR-MS using positive-mode electrospray ionization.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

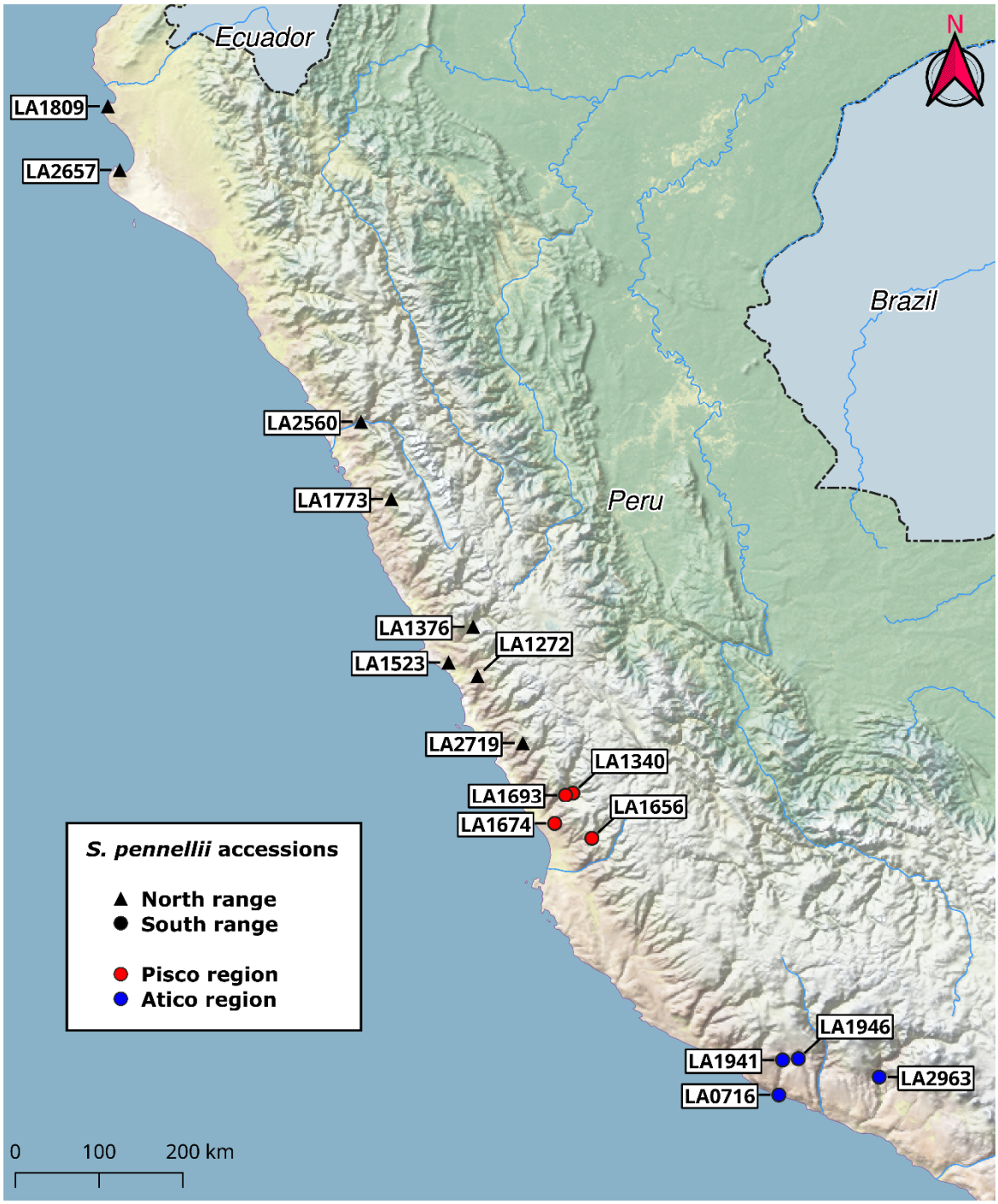


Figure 2. Locations of *S. pennellii* accessions used in this study across the geographic range of the species in Peru. Accessions classified as belonging to the north range are denoted with black triangles, those classified as belonging to the south range with circles. South range accessions are further classified by region (red for Pisco, blue for Atico). This map was created using the open source QGIS software (http://qgis.org) and GPS coordinates for accession locations provided by the C.M. Rick Tomato Genetics Resource Center (TGRC; http://tgrc.ucdavis.edu).

 doi:10.20944/preprints202008.0702.v1

126 *2.2 Untargeted metabolomics reveals acylsugars and flavonoids in trichomes*

127 Automated feature extraction and deconvolution of compound ions detected by
128 UHPLC-HR-MS analysis of leaf dip extracts followed by filtering to remove low-quality features
129 resulted in detection of 54 metabolic features. Based on annotation of collision induced dissociation
130 spectra and comparisons to previously characterized trichome-localized metabolites in *Solanum* spp.
131 [23,28,40] we categorized all 54 metabolic features as putative acylsugars or flavonoid aglycones. All
132 annotated acylsugars possessed either a six-carbon monosaccharide core or a 12-carbon disaccharide
133 core based on analysis of neutral losses from pseudomolecular ions and *m*/*z* of product ions.
134 Acylglucoses sharing acylation patterns are resolved as distinct α and β anomers by
135 reverse-phase chromatography but some acylglucose β anomers co-elute with α anomers of
136 later-eluting acylglucose isomers, precluding direct determination of the number of acylglucoses
137 present in a sample from the number of acylsugar metabolic features detected. Examination of
138 chromatograms and associated spectra for all features categorized as acylglucoses revealed that 32
139 metabolic features with distinct retention times identified as acylglucoses collectively represent α
140 and β anomers of 21 acylglucoses containing unique acyl chain complements. This consolidation
141 reduced the 54 metabolic features assigned *in silico* to 43 metabolic features. We categorized these
142 features as 18 triacylsucroses, 19 triacylglucoses, two tetraacylglucoses, and four flavonoids (Tables
143 1 and 2).
144

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

145 **Table 1.** Annotations of acylsugars in *S. pennellii*. RT = retention time (min); m/z_{acc} = accurate
146 $[M+NH_4]^+$ mass measured; m/z_{ex} = exact mass calculated from formula; Δm (ppm) = parts per million
147 error between m/z_{ex} and m/z_{acc}; fragment m/z = ions used for acyl chain determinations.

Name	RT	Formula	m/z_{acc}	m/z_{ex}	Δm (ppm)	Fragment m/z
Triacylsucroses						
S3:12(4,4,4)	2.21	$C_{24}H_{40}O_{14}$	570.2778	570.2756	3.9	373.1872, 285.1326, 197.0809, 127.0395
S3:13(4,4,5)	2.34	$C_{25}H_{42}O_{14}$	584.2926	584.2913	2.2	387.2010, 299.1507, 197.0809, 127.0395
S3:14(4,5,5)	2.59	$C_{26}H_{44}O_{14}$	598.3075	598.3069	1.0	401.2178, 313.1668, 211.0951, 127.0396
S3:15(5,5,5)	3.00	$C_{27}H_{46}O_{14}$	612.3229	612.3226	0.5	415.2348, 313.1661, 211.0974, 127.0395
S3:16(5,5,6)	3.44	$C_{28}H_{48}O_{14}$	626.3392	626.3382	1.6	429.2489, 327.1810, 211.0946, 127.0373
S3:16(4,4,8)	3.67	$C_{28}H_{48}O_{14}$	626.3387	626.3382	0.8	429.2489, 285.1360, 197.0809, 127.0395
S3:17(4,5,8)	4.23	$C_{29}H_{50}O_{14}$	640.3543	640.3539	0.6	443.2710, 299.1541, 211.0974, 127.0395
S3:17(4,4,9)	4.47	$C_{29}H_{50}O_{14}$	640.3536	640.3539	-0.5	443.2646, 285.1341, 197.0773, 127.0396
S3:18(4,4,10)-1	5.45	$C_{30}H_{52}O_{14}$	654.3699	654.3695	0.6	457.2864, 285.1360, 197.0837, 127.0395
S3:18(4,4,10)-2	5.71	$C_{30}H_{52}O_{14}$	654.3699	654.3695	0.6	457.2864, 285.1360, 197.0837, 127.0395
S3:19(4,5,10)-1	6.24	$C_{31}H_{54}O_{14}$	668.3856	668.3852	0.6	471.3013, 299.1507, 211.0974, 127.0395
S3:19(4,5,10)-2	6.52	$C_{31}H_{54}O_{14}$	668.3855	668.3852	0.5	471.3013, 299.1507, 211.1003, 127.0395
S3:20(5,5,10)	7.57	$C_{32}H_{56}O_{14}$	682.4011	682.4008	0.4	485.3103, 313.1661, 211.0974, 127.0395
S3:20(4,4,12)	8.20	$C_{32}H_{56}O_{14}$	682.4009	682.4008	0.2	485.3146, 285.1360, 197.0837, 127.0395
S3:21(5,5,11)	8.56	$C_{33}H_{58}O_{14}$	696.4166	696.4165	0.1	499.3259, 313.1661, 211.0974, 127.0395
S3:21(4,5,12)	9.14	$C_{33}H_{58}O_{14}$	696.4161	696.4165	-0.6	499.3259, 299.1472, 211.0974, 127.0395
S3:22(5,5,12)	10.26	$C_{34}H_{60}O_{14}$	710.4319	710.4321	-0.3	513.3442, 313.1661, 211.0974, 127.0395

148

149

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

150 **Table 1.** (cont'd)

Name	RT	Formula	m/z_{acc}	m/z_{ex}	Δm (ppm)	Fragment m/z
Triacylsucroses (cont'd)						
S3:23(5,6,12)	11.24	$C_{35}H_{62}O_{14}$	724.4471	724.4478	-1.0	527.3616, 327.1810, 211.0946, 127.0373
Triacylglucoses						
G3:12(4,4,4)	2.76; 2.84	$C_{18}H_{30}O_9$	408.2235	408.2228	1.7	373.1872, 285.1326, 197.0809, 127.0395
G3:13(4,4,5)	3.12; 3.24	$C_{19}H_{32}O_9$	422.2392	422.2385	1.7	387.2014, 299.1501, 197.0801, 127.0396
G3:14(4,5,5)	3.70; 3.83	$C_{20}H_{34}O_9$	436.2547	436.2541	1.4	401.2178, 299.1466, 211.0951, 127.0374
G3:15(5,5,5)	4.42; 4.58	$C_{21}H_{36}O_9$	450.2705	450.2698	1.6	415.2308, 313.1626, 211.0974, 127.0395
G3:16(5,5,6)	5.23; 5.41	$C_{22}H_{38}O_9$	464.2859	464.2854	1.1	429.2489, 327.1810, 211.0946, 127.0395
G3:16(4,4,8)-1	5.56; 5.80	$C_{22}H_{38}O_9$	464.2861	464.2854	1.5	429.2529, 285.1360, 197.0809, 127.0395
G3:16(4,4,8)-2	5.80; 6.04	$C_{22}H_{38}O_9$	464.2857	464.2854	0.7	429.2529, 285.1360, 197.0809, 127.0395
G3:17(4,5,8)-1	6.46; 6.72	$C_{23}H_{40}O_9$	478.3011	478.3011	0.0	443.2628, 299.1472, 211.0946, 127.0395
G3:17(4,5,8)-2	6.71; 6.99	$C_{23}H_{40}O_9$	478.3008	478.3011	-0.6	443.2628, 299.1472, 197.0781, 127.0373
G3:18(4,4,10)-1	8.03; 8.33	$C_{24}H_{42}O_9$	492.3168	492.3167	0.2	457.2779, 285.1326, 197.0809, 127.0395
G3:18(4,4,10)-2	8.33; 8.64	$C_{24}H_{42}O_9$	492.3170	492.3167	0.6	457.2779, 285.1326, 197.0809, 127.0395
G3:19(4,5,10)-1	9.05; 9.34	$C_{25}H_{44}O_9$	506.3328	506.3324	0.8	471.2970, 299.1472, 211.0974, 127.0395
G3:19(4,5,10)-2	9.34; 9.66	$C_{25}H_{44}O_9$	506.3328	506.3324	0.8	471.2970, 299.1507, 211.0946, 127.0395
G3:20(5,5,10)	10.47; 10.72	$C_{26}H_{46}O_9$	520.3486	520.3480	1.2	485.3146, 313.1661, 211.0974, 127.0395
G3:20(4,4,12)	11.10; 11.42	$C_{26}H_{46}O_9$	520.3483	520.3480	0.6	485.3103, 285.1326, 197.0809, 127.0395
G3:21(5,5,11)	11.47; 11.75	$C_{27}H_{48}O_9$	534.3637	534.3637	0.0	499.3215, 313.1626, 211.0974, 127.0373
G3:21(4,5,12)	12.10; 12.40	$C_{27}H_{48}O_9$	534.3636	534.3637	-0.2	499.3290, 299.1507, 211.0974, 127.0395

151

152

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

Table 1. (cont'd)

Name	RT	Formula	m/z_{acc}	m/z_{ex}	Δm (ppm)	Fragment m/z
Triacylglucoses (cont'd)						
G3:22(5,5,12)	13.10; 13.36	$C_{28}H_{50}O_9$	548.3794	548.3793	0.2	513.3442, 313.1661, 211.0974, 127.0395
G3:23(5,6,12)	14.02; 14.29	$C_{29}H_{52}O_9$	562.3938	562.3950	-2.1	527.3471, 327.1848, 211.0960, 127.0372
Tetraacylglucoses						
G4:14(2,4,4,4)	3.54; 3.79	$C_{20}H_{32}O_{10}$	450.2343	450.2334	2.0	415.1946, 327.1417, 239.0891, 197.0809, 127.0373
G4:15(2,4,4,5)	4.10; 4.45	$C_{21}H_{34}O_{10}$	464.2502	464.2491	2.4	429.2162, 341.1562, 239.0922, 197.0837, 127.0395

Table 2. Annotations of flavonoids in *S. pennellii*. RT = retention time (min); m/z_{acc} = accurate [M+H]+ mass measured; m/z_{ex} = exact mass calculated from formula; Δm (ppm) = parts per million error between m/z_{ex} and m/z_{acc}; core = putative flavonol core based on molecular formula; # Me = number of methyl groups based on molecular formula and mass spectrum (Appendix Fig. S2).

Name	RT	Formula	m/z_{acc}	m/z_{ex}	Δm (ppm)	Core	# Me
Flavonoid A	3.04	$C_{17}H_{14}O_6$	315.0869	315.0863	1.9	kaempferol	2
Flavonoid C	3.17	$C_{18}H_{16}O_7$	345.0980	345.0969	3.2	quercetin	3
Flavonoid D	4.00	$C_{19}H_{18}O_7$	359.1137	359.1125	3.3	quercetin	4
Flavonoid B	4.90	$C_{18}H_{16}O_6$	329.1025	329.1020	1.5	kaempferol	3

2.3 Acylsugar core composition varies across the S. pennellii *geographic range*

Absolute quantification of total acylsucroses and acylglucoses in 16 accessions of *S. pennellii* revealed variation in total acylsugar accumulation (from 133 µmol/g dry weight (DW) in LA2657 to 340 µmol/g DW in LA2560) and relative abundance of acylglucoses and acylsucroses (from 42% acylglucoses in LA2963 to 95% acylglucoses in LA0716) (Table 3; Fig. S1A,B). While we found no discernable geographic trends in total acylsugar accumulation (Table 3; Fig. S1A), higher relative abundance of acylglucoses was observed in southern accessions compared with northern accessions. In the northern extent of the range, acylglucose composition varied from 56% (LA2657) to 70% (LA2719), while in the southern span, values ranged from 77% (LA1693) to 95% (LA0716) acylglucose. The south accession LA2963 is a notable exception to this trend, showing a lower acylglucose composition (42%) than any other accession (Table 3; Fig. S1B).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

Table 3. Acylsugar accumulation and percent acylglucose in accessions of *S. pennellii* as determined by UHPLC-MS-MS. Values are presented as mean ± SD (n = 6). Results of analysis of variance and Tukey's mean-separation test are indicated as letters. Accessions that do not have at least one letter in common are significantly different from one another ($p < 0.05$). The range and region of each accession within Peru is also indicated.

Accession	Total acylsugars (μmol/g DW)	Tukey's MST	% acylglucose	Tukey's MST	Range	Region
LA1809	136 ± 27	B	69 ± 4	CDE	North	
LA2657	133 ± 28	B	56 ± 6	EF	North	
LA2560	340 ± 64	A	65 ± 6	CDE	North	
LA1773	237 ± 98	AB	66 ± 2	CDE	North	
LA1376	261 ± 105	AB	70 ± 7	CDE	North	
LA1523	158 ± 68	B	65 ± 8	CDE	North	
LA1272	163 ± 109	B	58 ± 4	DEF	North	
LA2719	218 ± 44	AB	70 ± 3	BCDE	North	
LA1340	166 ± 39	B	80 ± 10	ABC	South	Pisco
LA1693	193 ± 88	AB	77 ± 7	ABCD	South	Pisco
LA1674	248 ± 83	AB	90 ± 4	A	South	Pisco
LA1656	269 ± 54	AB	90 ± 4	AB	South	Pisco
LA1946	257 ± 94	AB	82 ± 20	ABC	South	Atico
LA1941	244 ± 66	AB	95 ± 2	A	South	Atico
LA2963	183 ± 38	B	42 ± 6	F	South	Atico
LA0716	238 ± 75	AB	95 ± 2	A	South	Atico

2.4 Variable acyl chain and sugar composition yield acylsugar diversity

Annotation of the 39 acylsugars present in our dataset revealed 26 unique molecular formulas, including multiple structural isomers (Table 1). As no alternative sugar cores other than glucose and sucrose have been reported from *S. pennellii*, this isomerism is likely driven by variation in acyl chains or positions of specific acylations. Six pairs of structural isomers had indistinguishable mass spectra (Fig. 3A; Table 1). This suggests two possible non-mutually exclusive types of acylsugar structural isomerism: acylsugars with similar complements of acyl chains but differing in acyl chain positions (positional isomers), and acylsugars bearing acyl chains with identical chemical formulas but different branching patterns (acyl chain structural isomers). The latter hypothesis is supported by previous reports of unbranched, *iso*-branched, and *anteiso*-branched acyl chains in *S. pennellii* acylsugars [18,21,26]. Additional structural isomers differ in the number of carbons present in individual constituent acyl chains (Fig. 3B; Table 1), while maintaining constant the total number of acyl carbons. The presence of two tetraacylglucoses in the dataset, G4:14(2,4,4,4) and G4:15(2,4,4,5), also indicates variation in the number of acylsugar acylations. In contrast, neither mono- nor di-acylated sugars were observed, although these are intermediates in tri- and tetraacylated sugar biosynthesis [19].

All but one of the annotated triacylsucroses [S3:17(4,4,9)] in our dataset show patterns of acyl group neutral losses in their mass spectra that mirror those observed in at least one triacylglucose (Table 1). We hypothesized that pairs of acylsucroses and acylglucoses with similar neutral mass losses possessed identical acyl chain complements, consistent with the current model of *S. pennellii* acylsugar biosynthesis in which cleavage of acylsucrose glycosidic bonds by ASFF1 removes the β-fructofuranose rings to yield acylglucoses [22]. These observations indicate that variation in the identity of acyl chains, number of acyl chains, and identity of sugar core all contribute to the acylsugar diversity in *S. pennellii*. While the presence of multiple acylsugar structural isomers with identical mass spectra implies isomeric acyl chains, and the similarity in neutral losses between

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

acylsucroses and acylglucoses suggests identical chain elemental composition but not necessarily topology, the mass spectrometry techniques applied could not establish key structural features, leading us to resolve their structures using NMR .

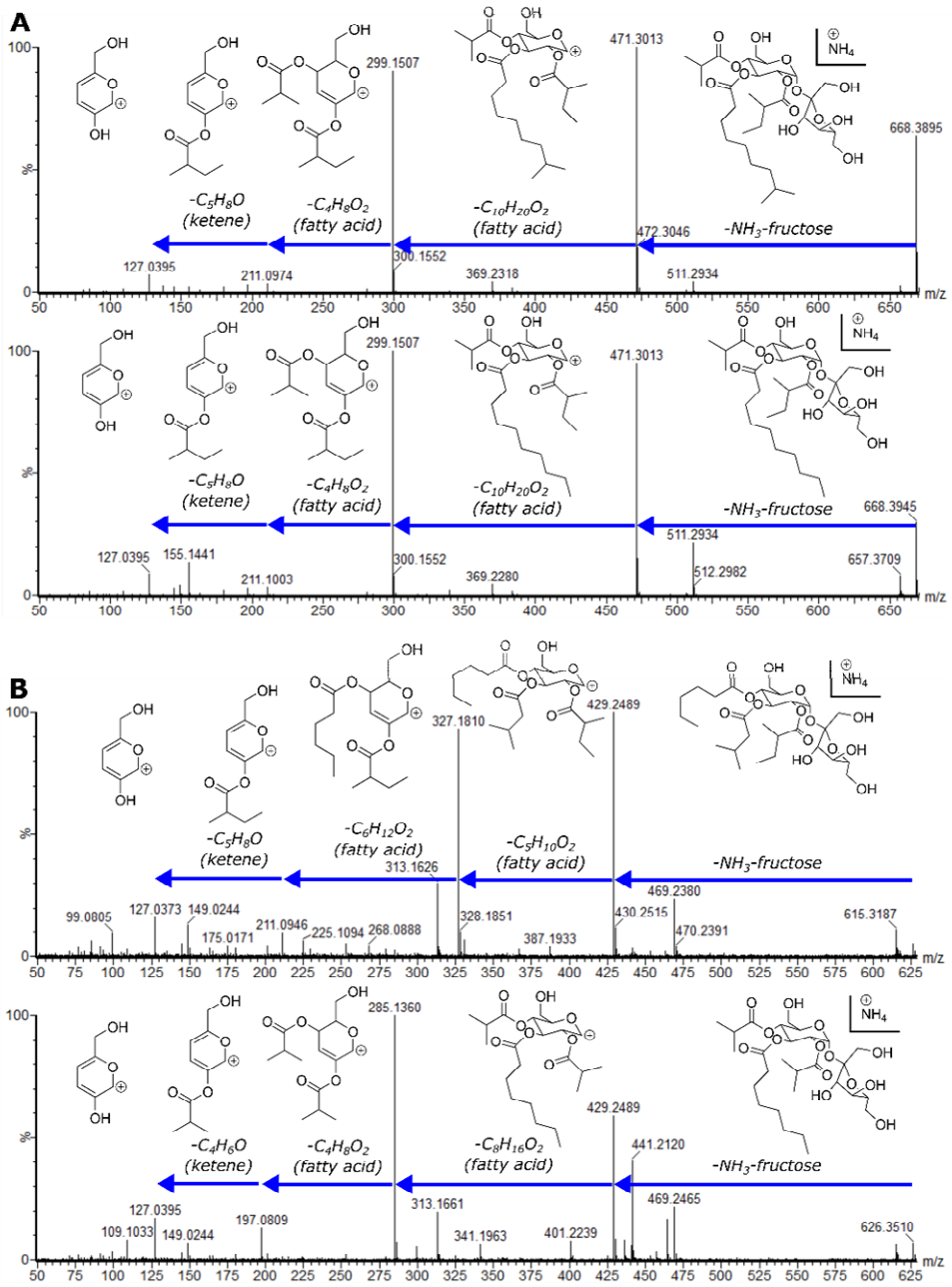


Figure 3. CID mass spectra of acylsugar structural isomers. (A) Mass spectra of S3:19(4,5,10)-1 (top) and S3:19(4,5,10)-2 (bottom). Structures of both compounds were resolved by NMR (Fig. 4). (B) Mass spectra of S3:16(5,5,6) (top) and S3:16(4,4,8) (bottom). Structures of these compounds were not resolved by NMR, and the specific branching patterns and positions of acyl chains are unknown.

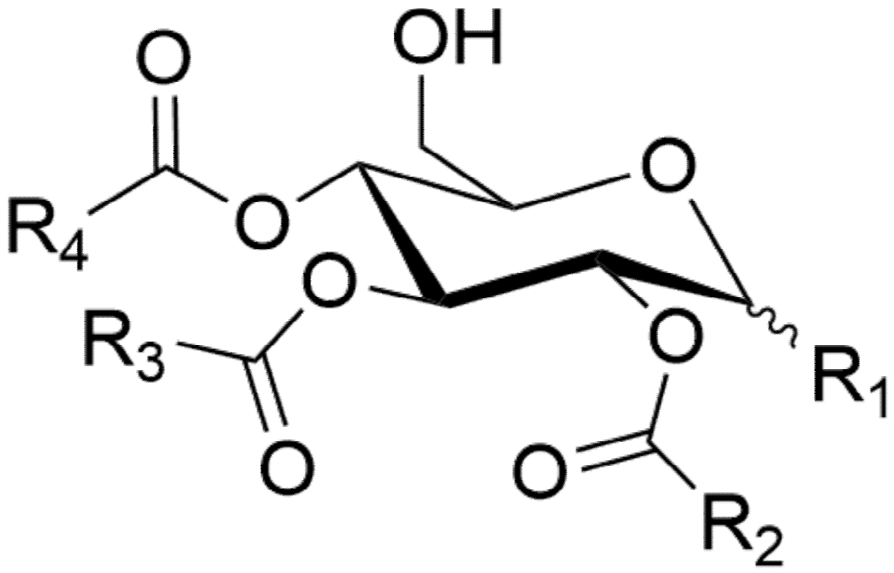
2.5 NMR spectroscopy resolves structural relationships between acylsugars

We selected 10 acylsugars for purification and structural resolution by NMR, including five triacylsucroses (S3:12(4,4,4), S3:18(4,4,10)-1, S3:18(4,4,10)-2, S3:19(4,5,10)-1, and S3:19(4,5,10)-2) and five triacylglucoses (G3:12(4,4,4), G3:18(4,4,10)-1, G3:18(4,4,10)-2, G3:19(4,5,10)-1, and G3:19(4,5,10)-2) (Table 1; Fig. 4; See Tables S6- S15, Figs. S3-62 for NMR chemical shifts and spectra). NMR spectroscopy confirmed that all examined disaccharide-containing acylsugars possess a sucrose core while all monosaccharide acylsugars are based on glucose, consistent with previous

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

analyses of *S. pennellii* acylsugars [23,26,28]. NMR analysis further revealed that all are acylated at the 2-, 3-, and 4- hydroxyls of the pyranose ring, also consistent with previous reports [23,28]. The structures of two compounds, G3:12(4,4,4) and S3:19(4,5,10)-1, matched two previously published acylsugar structures [23,28].

We tested the hypotheses that acylsugar isomers with indistinguishable mass spectra possess either identical complements of acyl chains attached to different positions of the sugar core or isomeric acyl chains with different branching patterns. The structures of four pairs of isomers were compared, including two pairs each of acylsucrose and acylglucose isomers (S3:18(4,4,10)-1/2, S3:19(4,5,10)-1/2, G3:18(4,4,10)-1/2, G3:19(4,5,10)-1/2). In each case, both isomers had identical configurations of acyl chains at the 2- and 4- positions. However, for all four isomeric pairs, we observed an *iso*-branched 10-carbon acyl chain (R3 = $(Me)_2CH(CH_2)_6$) in the earlier-eluting isomer and an unbranched 10-carbon acyl chain (R3 = $Me(CH_2)_8$) in the later-eluting isomer (Fig. 4). This demonstrates that acylsugar diversity is influenced by differences in acyl chain branching patterns as well as variation in the molecular formulas of constituent acyl chains. We also compared the structures of acylsucroses and acylglucoses with similar neutral loss patterns. The acylation pattern of each of the five purified acylsucroses was identical to that of its analogous purified acylglucose (*e.g.*, S3:12(4,4,4) and G3:12(4,4,4); S3:19(4,5,10)-1 and G3:19(4,5,10)-1; Fig. 4). This is consistent with the hypothesis that these acylsucroses are biosynthetic precursors of the analogous acylglucoses.



S3:12(4,4,4) R_1 = Fru; R_2 = R_3 = R_4 = Me_2CH

G3:12(4,4,4) R_1 = OH; R_2 = R_3 = R_4 = Me_2CH

S3:18(4,4,10)-1 R_1 = Fru; R_2 = R_4 = Me_2CH; R_3 = $Me_2CH(CH_2)_6$

G3:18(4,4,10)-1 R_1 = OH; R_2 = R_4 = Me_2CH; R_3 = $Me_2CH(CH_2)_6$

S3:18(4,4,10)-2 R_1 = Fru; R_2 = R_4 = Me_2CH; R_3 = $Me(CH_2)_8$

G3:18(4,4,10)-2 R_1 = OH; R_2 = R_4 = Me_2CH; R_3 = $Me(CH_2)_8$

S3:19(4,5,10)-1 R_1 = Fru; R_2 = $MeCH_2MeCH$; R_3 = $Me_2CH(CH_2)_6$; R_4 = Me_2CH

G3:19(4,5,10)-1 R_1 = OH; R_2 = $MeCH_2MeCH$; R_3 = $Me_2CH(CH_2)_6$; R_4 = Me_2CH

S3:19(4,5,10)-2 R_1 = Fru; R_2 = $MeCH_2MeCH$; R_3 = $Me(CH_2)_8$; R_4 = Me_2CH

G3:19(4,5,10)-2 R_1 = OH; R_2 = $MeCH_2MeCH$; R_3 = $Me(CH_2)_8$; R4 = Me_2CH

Figure 4. NMR-resolved structures of acylsugars purified from *S. pennellii*. For acylsucroses, the R_1 group is observed only in the α configuration. For each acylglucose, two distinct anomers are observed with group R_1 in either the α or β configuration. Fru = β-fructofuranose.

2.6 Flavonoids vary by core and degree of methylation

Our dataset contained four methylated flavonoid aglycones. Methyl flavonoid molecular formulas were consistent with di-, tri-, and tetramethylated derivatives of tetra- and pentahydroxylated flavonols (Table 2), resembling the methylated myricetins observed in *S. habrochaites* and *S. lycopersicum* [40–42]. Two kaempferol-like tetrahydroxylated flavonoids were observed possessing two and three methylations (denoted as flavonoids A and B), while two quercetin-like pentahydroxylated flavonoids were observed possessing three and four methylations (flavonoids C and D). As *S. lycopersicum* accumulates glycosylated derivatives of the flavonols kaempferol and quercetin (tetra- and pentahydroxylated, respectively) in type VI trichomes [6,42], we hypothesized that the methylated flavonoids observed in *S. pennellii* leaf dips were also kaempferol- and quercetin-derived. While analysis of flavonoid mass spectra indicated molecular formulas and presence of methyl groups, few low-mass fragment ions were present in the spectra to aid in further structural assignment previously demonstrated with myricetin derivatives (Fig. S2) [43]. Nevertheless, our results indicated flavonoid diversity in terms of both flavonol core and degree of methylation.

2.7 Multivariate analysis implicates short branched acyl chains in north-south acylsugar variation

We used the full dataset representing 43 specialized metabolites in 16 accessions of *S. pennellii* to identify metabolite-based differences between accessions across the geographic range. Due to overlapping retention times observed with some acylglucose anomers and the resulting difficulty in assigning accurate abundances to individual acylglucoses, we used the original dataset containing 54 metabolite features obtained prior to spectral interpretation instead of the dataset containing the 43 unique metabolites. Unsupervised principal component analysis (PCA) of metabolites signal abundances of all accessions revealed clear separation of accessions in the north range from those in the south range with the exception of two outliers (Fig. 5). These samples both represent individuals of south range accession LA1946 that cluster with north range samples; we hypothesize that this is due to seed contamination or sample tracking error.

An orthogonal partial least squares/projection to latent structures discriminant analysis (OPLS-DA) model successfully classified 100% of north range samples and 94% of south range samples (Table 4), indicating that metabolite features ranked by the model were good predictors of geographic origin (Table S16). The abundance of each metabolite feature was correlated with either the north or south range samples. Structural characteristics of the five metabolite features showing the strongest quantitative correlation with either sample class were compared. Three acylglucoses [G3:15(5,5,5), G3:16(5,5,6), G3:21(5,5,11)] and two acylsucroses [S3:16(5,5,6), S3:21(5,5,11)] showed the strongest correlation with north range accessions, while four acylglucoses (G3:12(4,4,4), G3:13(4,4,5), G3:18(4,4,10)-2, G3:19(4,5,10)-2) and one acylsucrose (S3:18(4,4,10)-2), showed the strongest correlation with south range accessions. Four acylsugars enriched in the south range (G3:12(4,4,4), G3:18(4,4,10)-2, G3:19(4,5,10)-2, S3:18(4,4,10)-2) were structurally characterized by NMR in this study (Fig. 4), while a fifth [G3:14(4,5,5)] was annotated in previous work [23]. All four-carbon acyl chains in these acylsugars are 2-methylpropanaote, while only one of the five-carbon chains in the G3:14(4,5,5) compound is 3-methylbutanoate; the other five-carbon acyl chains in G3:14(4,5,5) and G3:19(4,5,10)-2 are 2-methylbutanoate. While we cannot definitively identify the branching pattern of five-carbon acyl chains in the metabolites associated with the north range, our findings agree with previously observed trends in *S. pennellii* favoring accumulation of four-carbon 2-methylpropanoate chains in southern accessions and five-carbon 3-methylbutanoate chains in northern accessions, with five-carbon 2-methylbutanoate chains abundant across the range [21,26].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

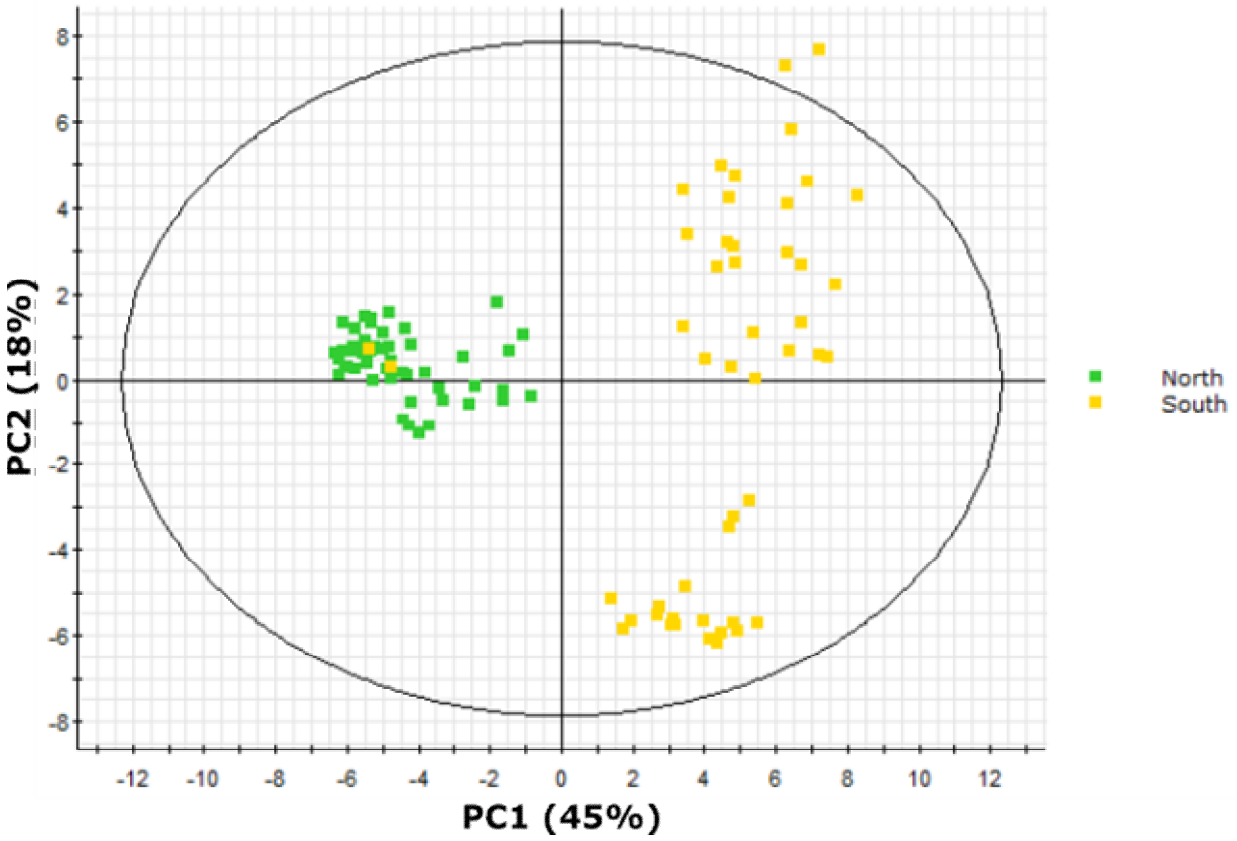


Figure 5. PCA scores plot of samples from 16 *S. pennellii* accessions from across Peru separated by abundances of 54 metabolite features identified in trichome extracts by UHPLC-HR-MS. Samples from the North range are indicated in green, while samples from the South range are indicated in yellow (see Fig. 2 for details on geographic range). Principal component 1 (PC1) accounted for approximately 45% of the variance in the dataset and drove strong separation between north and south accessions, while PC2 accounted for 18% of the variance and associated primarily with variation within the south range accessions.

Table 4. OPLS-DA model performance. The table indicates the percentage of test samples that each model classified correctly, incorrectly, or was unable to classify.

	n	% Correct	% Incorrect	% Unknown
Full range (N = 96)				
North	48	100	0	0
South	48	94	4	2
South range (N = 48)				
Pisco	24	67	4	29
Atico	24	77	4	19
Atico region (N = 24)				
LA0716/LA1941/LA1946	18	97	0	3
LA2963	6	100	0	0

2.8 Variation in medium-length acyl chains drives variation within the south range

As our PCA also indicated substantial intragroup variation in south range samples (Fig. 5), we performed additional multivariate analyses to distinguish profiles within south range plant extracts (Fig. 6). These accessions form two distinct geographic clusters from the Pisco or Atico regions (Fig.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

2). An OPLS-DA model of Pisco and Atico samples successfully classified 67% of Pisco region samples and 77% of Atico region samples but misclassified or was unable to classify 28% of all samples (Table 4), indicating that this model performed poorly when compared to our north/south range OPLS-DA model. However, the model still recognized metabolites that had strong quantitative correlation with either the Pisco or Atico region samples (Table S17). The top five compounds demonstrating strong correlation with Pisco region samples comprised four acylglucoses [G3:16(4,4,8)-1, G3:16(4,4,8)-2, G3:17(4,5,8)-1, and G3:17(4,5,8)-2] and one acylsucrose [S3:17(4,4,9)], while the top five metabolites correlating with the Atico region samples consisted of two acylglucoses [G3:20(4,4,12), G3:21(4,5,12)], two acylsucroses [S3:20(4,4,12), S3:21(4,5,12)], and one flavonoid (flavonoid A). The medium-length acyl chains (defined here as possessing more than five carbons) in correlative features show a sharp distinction between the two regions with four of five acylsugars more abundant in the Pisco region accessions bearing an eight-carbon acyl chain and all four acylsugars that are more abundant in the Atico region containing a 12-carbon acyl chain, while four- and five-carbon acyl chains have a similar distribution between the two regions. This suggests medium-length acyl chain variation as the key driver in separation of accessions from the Pisco and Atico region.

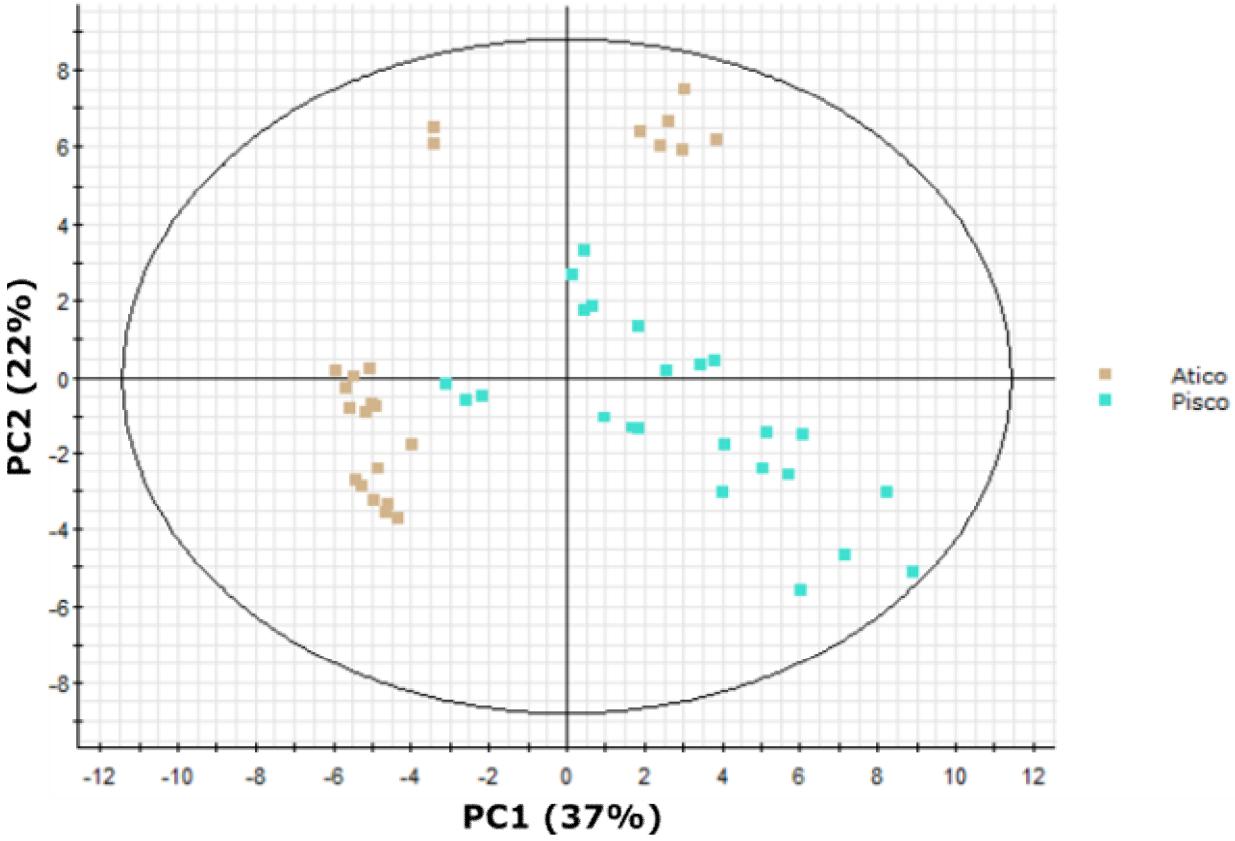


Figure 6. PCA scores plot of samples from eight *S. pennellii* accessions from the southern portion of the range of the species in Peru separated by abundances of 54 metabolite features identified in trichome extracts by UHPLC-HR-MS. Samples from the Atico region are indicated in tan, while samples from the Pisco region are indicated in light blue (see Fig. 2 for details on regions). Separation is observed between Atico and Pisco samples. However, Atico region samples exhibit bimodal clustering. PC1 accounted for 47% of variance and described most of the variation between accession LA2963 samples and other Atico region accessions, while PC2 accounted for 35% of variance and described primarily variation within the main Atico cluster.

2.9 LA2963 segregates from other Atico region accessions due to high acylsucrose content

 doi:10.20944/preprints202008.0702.v1

The bimodal clustering of Atico region samples revealed by PCA (Fig. 6) indicates chemical diversity among accessions that are < 150 km apart, in contrast to previously reported trends in acylsugar diversity observed between accessions separated by > 1000 km (Fig. 5; Table S.12) [21,26]. We explored this diversity using multivariate analysis. PCA of the four Atico region accessions (LA0716, LA1941, LA1946, and LA2963) showed two major clusters (Fig. 7). One cluster contained all biological replicates of accessions LA0716 and LA1941 along with four samples of LA1946 [both LA1946 samples outside this cluster represent outliers that clustered with north range accessions in our north/south PCA (Fig. 5)]. The other major cluster contained all samples of accession LA2963.

An OPLS-DA model discriminating between LA2963 samples and all other Atico region accessions correctly classified 97% of samples from the main Atico cluster and 100% of LA2963 samples (Table 4), indicating that metabolite features ranked by the model were good predictors of sample group (Table S18). Compounds that were more abundant in the main Atico cluster included four acylglucoses [G3:18(4,4,10)-1, G3:18(4,4,10)-2, G3:19(4,5,10)-1, G3:20(4,4,12)] and one flavonoid (flavonoid A) while molecules correlated with the anomalous accession LA2963 included five acylsucroses [S3:18(4,4,10)-1, S3:19(4,5,10)-1, S3:19(4,5,10)-2, S3:20(4,4,12), S3:21(4,5,12)]. This indicates relative acylsucrose and acylglucose abundance as the key driver of separation between LA2963 samples and other Atico region samples, and was consistent with our initial analysis of *S. pennellii* sugar core abundance, which indicated accession LA2963 as an outlier among southern accessions that possessed low acylglucose content (Table 3; Fig. S1B).

Two of the acylsucroses correlated with accession LA2963 (S3:18(4,4,10)-1, S3:19(4,5,10)-1) have structures consistent with precursors of two acylglucoses correlated with the main Atico cluster (G3:18(4,4,10)-1, G3:19(4,5,10)-1; described above; Fig. 4), while a third compound correlated with LA2963 [S3:20(4,4,12)] has a fragmentation pattern consistent with a possible precursor of another Atico cluster-correlated acylglucose [G3:20(4,4,12)] (Table 1). The ASFF1 enzyme hydrolyzes acylsucroses yielding acylglucoses in *S. pennellii* LA0716 [22]. We hypothesized that low ASFF1 activity in plants of accession LA2963 relative to other Atico region accessions contributed to the low accumulation of acylglucoses in this accession and corresponding high accumulation of acylsucroses. To test this hypothesis, we investigated whether there is a correlation between the relative accumulation of acylglucoses and expression of the *ASFF1* gene by saponification of acylsugar extracts and UHPLC-MS-MS sugar core quantification, and relative quantification of *ASFF1* transcript abundance by RT-qPCR in paired leaflets from three biological replicates of *S. pennellii* LA0716 and LA2963 (Fig. 8). Acylglucoses constituted 94% of acylsugars in LA0716 but only 38% of acylsugars in LA2963 (Fig. 8A), consistent with our previous sugar core quantification results (Table 3; Fig. S1B), while *ASFF1* transcripts were 2.9-fold more abundant in LA0716 than in LA2963 (Fig. 8B). Linear regression analysis indicated a positive correlation between *ASFF1* transcript abundance and percentage of acylsugars accumulating as acylglucoses (R^2 = 0.84; Fig. 8C). This correlation supports the role of the *ASFF1* gene in determining acylsugar core composition in *S. pennellii* [22] and further suggests a role for transcriptional regulation of *ASFF1* in intraspecific sugar core variation.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

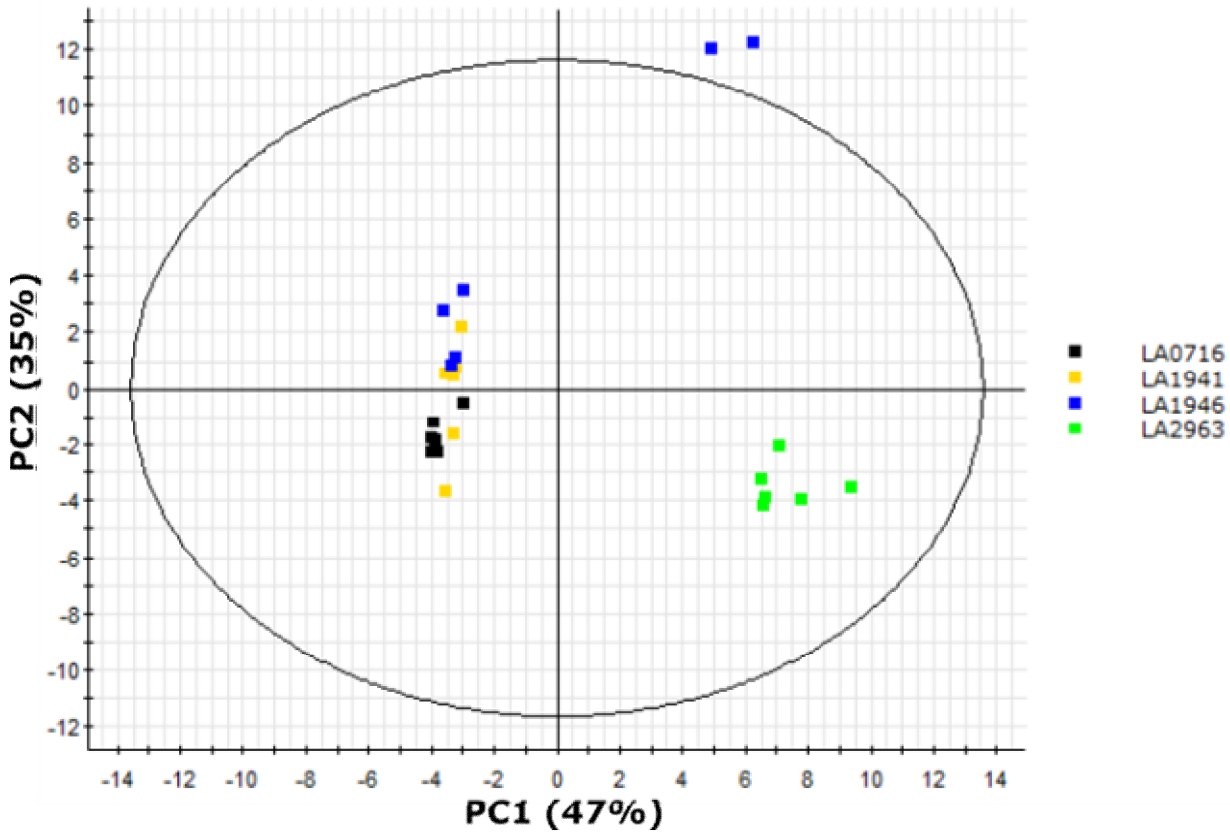


Figure 7. PCA scores plot of samples from four *S. pennellii* accessions in the Atico region of Peru separated by abundances of 54 metabolite features identified in trichome extracts by UHPLC-HR-MS. Samples from accession LA0716 are indicated in black, samples from accession LA1941 in yellow, samples from accession LA1946 in blue, and samples from accession LA2963 in green (see Fig. 2 for details of the Atico region). PC1 accounted for 47% of variance and described most of the variation between accession LA2963 samples and other Atico region accessions, while PC2 accounted for 35% of variance and described primarily variation within the main Atico cluster.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

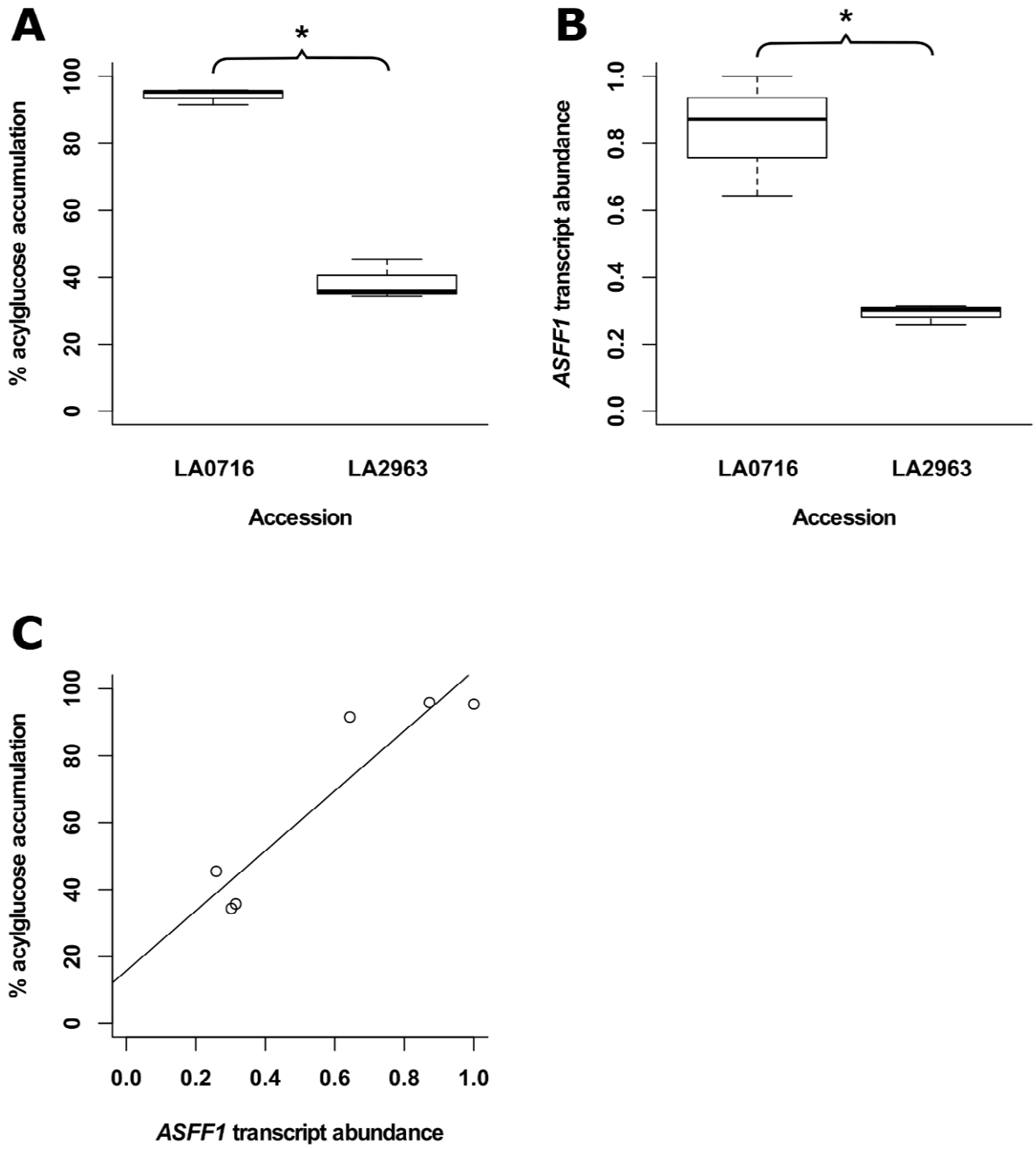


Figure 8. Analysis of acylglucose accumulation and *ASFF1* transcript abundance in paired leaflets of *S. pennellii* accessions LA0716 and LA2963. (A) Percentage of total acylsugars accumulating as acylglucoses. (B) Relative abundance of *ASFF1* transcripts. (C) Linear regression of *ASFF1* transcript abundance and percentage of percentage of acylsugars accumulating as acylglucoses ($R^2 = 0.84$). "*" indicates $p < 0.05$ (analysis of variance); $n = 3$ for both accessions.

3. Discussion

To capitalize on the protective properties of acylsugars, plant breeders are creating tomato lines with altered acylsugar profiles and increased insect resistance [15–18,44,45]. This process is facilitated by knowledge of acylsugar protective properties [13] and the genetic basis for acylsugar biosynthesis and diversity [15–18]. Characterization of the acylsugars found in *S. pennellii* is essential for elucidating and evaluating protective benefits of specific compounds and of pathways involved

in acylsugar biosynthesis and degradation. Our analysis of metabolites extracted from the surface of *S. pennellii* leaflets annotated a total of 43 specialized metabolites consisting of 18 acylsucroses, 21 acylglucoses, and four flavonoids. UHPLC-MS analysis alone indicated the presence of two tetraacylglucoses (Table 1), a type of acylsugar previously unknown in *S. pennellii*, as well as four methyl flavonoids (Table 2), a class of compounds known from the related tomatoes *S. lycopersicum* and *S. habrochaites* but previously unknown in this species [6,40–43]. A combination of UHPLC-MS and NMR spectroscopy indicated both acyl chain length and branching pattern as mechanisms of acylsugar isomerism and confirmed that acylglucose structures are consistent with acylsucrose hydrolysis products (Table 1; Fig. 4, Fig. 3). Multivariate analysis of our UHPLC-MS dataset provided additional confirmation of the differential accumulation of short branched acyl chains in acylsugars from northern and southern *S. pennellii* accessions (Fig. 5; Table S16) and revealed geographic variation between smaller sub-regions within the range of the species (Fig. 6; Fig. 7). Accessions from the Pisco and Atico regions were distinguished by enrichment of eight-carbon acyl chains in the former and 12-carbon acyl chains in the latter (Table S17). Within the Atico region, the acylsugar profile of accession LA2963 differs from that of nearby accessions primarily due to low acylglucose abundance compared to other Atico region accessions (Fig. 8A). The new dimensions of acylsugar variation discovered in this work demonstrate that aspects of acylsugar biosynthesis and degradation within and beyond the core pathway await characterization.

Our findings indicate that additional acyltransferase activities involved in *S. pennellii* acylsugar biosynthesis remain to be identified. We annotated two previously unreported tetraacylglucoses, both containing acetyl groups (Table 1). While tetraacylated sugars with acetyl groups are common in *S. habrochaites* and *S. lycopersicum*, they are absent from published analyses of *S. pennellii* acylsugars [21,26,35]. Thus far, three ASATs involved in acylsugar biosynthesis, each performing a single acylation step, have been identified in *S. pennellii* [19]. The presence of tetraacylglucoses in this species requires a fourth acylation step, which could be performed by one of the previously described acyltransferases from the *S. pennellii* acylsugar pathway (*i.e.,* ASAT1/2/3) or by an acyltransferase not previously implicated in acylsugar biosynthesis. In *S. lycopersicum*, acetylation of triacylsucroses is performed by ASAT4 [46]. The *S. pennellii* ASAT4 locus is therefore worth investigating as a candidate acylsugar acetyltransferase in this species. Explicit searches using extracted ion chromatograms of anticipated ion masses revealed no tetraacylsucroses. These compounds may accumulate at levels below the detection threshold. Alternatively, tetraacylglucoses may not be hydrolysis products of tetraacylsucroses, but rather derived via direct acylation of triacylglucoses. Further characterization of acyltransferases is necessary to determine the origins of tetraacylated sugars in *S. pennellii*.

Our analysis also revealed previously unreported intraspecific differences in acylsugar acyl chain accumulation. Prior studies confirmed eight- and 12-carbon acyl chains in *S. pennellii* acylsugars [21,26], and we identified differential accumulation of acylsugars containing these acyl chains between accessions from the Pisco and Atico regions (Table S17). Differences in abundance of eight- and 12-carbon acyl chain-containing acylsugars may reflect differences in acyl CoA availability, ASAT-catalyzed incorporation of acyl chains into acylsugars, or acylsugar turnover. Interspecific variation at genetic loci encoding enoyl CoA hydratase (ECH) and acyl CoA synthetase (ACS) enzymes leads to a high proportion of 10-carbon relative to 12-carbon acyl chains in *S. pennellii* LA0716 and a high proportion of 12-carbon relative to 10-carbon acyl chains in *S. lycopersicum* M82 [20]. Differences in substrate specificity of these enzymes among *S. pennellii* accessions could lead to variation in medium-length acyl CoA pools and subsequent incorporation of medium-length acyl chains into acylsugars. Alternatively, variation in ASAT affinity for acyl CoAs among accessions may explain differences in acyl chain incorporation even if similar acyl CoA pools are present across accessions. Finally, differential accumulation of eight- and 12-carbon acyl chain-containing acylsugars may reflect differences in acylsugar turnover rates between *S. pennellii* accessions. The ASH carboxylesterase enzymes facilitate acylsugar degradation in *S. lycopersicum* and *S. pennellii* primarily by removing acyl chains from the 3-position of acylsucroses and acylglucoses [23]. NMR spectra of acylsugars in *S. pennellii* consistently show medium-length R3 chains while groups R2 and R4 are exclusively short four- or five-carbon acyl chains (Fig. 4) [23], suggesting that eight- and

12-carbon acyl chains could be removed by ASHs. Mass spectra indicated the presence of eight- and 12-carbon acyl chains in our dataset but the corresponding acylsugars were not selected for purification and structural resolution by NMR. As both straight and branched 8- and 12-carbon acyl chains have been observed in *S. pennellii* acylsugars [21,26,27], further structural characterization of these acylsugars by NMR is warranted for deeper exploration of their biosynthetic origins.

In addition to uncovering variation in acylsugar acyl chains, we identified variation in sugar core proportion within the Atico region. The proportion of acylglucoses in the southern accession LA2963 (42%) was less than half of that observed for nearby accessions from the Atico region (82-95%) (Table 3; Fig. S1). We tested the hypothesis that the proportion of acylsugars accumulating as acylglucoses in *S. pennellii* could be associated with levels of the ASFF1 enzyme, which hydrolyzes acylsucroses to acylglucoses in accession LA0716 [22]. A combination of UHPLC-MS-MS and RT-qPCR demonstrated that the percentage of acylsugars accumulating as acylglucoses correlated with abundance of *ASFF1* transcripts in two accessions from the Atico region, LA0716 and LA2963 (Fig. 8C). Combined with the observation that knockout of the *ASFF1* gene in *S. pennellii* LA0716 abolishes acylglucose accumulation [22], our findings suggest that low levels of ASFF1 expression lead to a low proportion of acylglucoses in *S. pennellii* LA2963. Additional work is needed to dissect the mechanism leading to differences in transcript accumulation in these southern accessions.

Our current understanding of acylsugar biosynthesis in *Solanum* was achieved primarily through interspecific comparison of acylsugar phenotypes and analysis of variation in the underlying genetic loci [19–22,46]. The intraspecific variations in *S. pennellii* acylsugar phenotype reported here provide a basis for further pathway analyses. The dimensions of acylsugar variation within *S. pennellii* are potentially linked to all known components of acylsugar metabolism including enzymes in auxiliary pathways that generate acylsugar precursors (*e.g.*, IMPS3, ECH, ACS), activities of the core acylsugar biosynthetic pathway (*i.e.*, ASATs), and enzymes that degrade or remodel acylsugars (*e.g.*, ASHs and ASFF1). The presence of tetraacylglucoses indicates undiscovered core pathway acyltransferase activity in the form of new ASATs or broader substrate specificity of existing ASATs. Differential accumulation of eight-carbon and 12-carbon acyl chain-containing acylsugars among *S. pennellii* accessions may reflect variation in biosynthesis of medium-length acyl CoA precursors to acylsugars by enzymes like ACS and ECH, variation in ASAT affinity for medium-length acyl CoAs in the core acylsugar pathway, or variation in ASH affinity for medium-length acyl chain-containing acylsugars during acylsugar degradation and turnover. The correlation between relative acylsugar core abundance and *ASFF1* expression indicates a role for gene regulation in affecting acylsugar composition. Further investigation of acylsugar structures, biochemical characterization of enzymes in the pathway, and an understanding of genetic regulatory networks governing pathway expression will facilitate efforts to improve the endogenous defenses of Solanaceae crops with a variety of techniques ranging from marker-assisted selection to CRISPR/Cas9-mediated gene editing and synthetic biology approaches.

4. Materials and Methods

4.1 Plant material

Seeds of all *S. pennellii* accessions were obtained from the C.M. Rick Tomato Genetics Resource Center (TGRC; University of California, Davis, CA). Seeds were treated with 2.6% sodium hypochlorite for 30 min and subjected to three 5-min rinses in de-ionized water before sowing on moist Whatman grade 1 filter paper (Sigma-Aldrich, St. Louis, MO) in Petri dishes. Seeds were kept in the dark at room temperature and transplanted upon germination. Additional details of plant growth are in Table S1.

4.2 Acylsugar extraction

Single leaflets from the youngest fully expanded leaves of individual *S. pennellii* plants at 16 weeks post-germination were harvested and placed into pre-washed 10 x 75 mm borosilicate glass test tubes. Leaflets were collected from six individual plants of each *S. pennellii* accession, with an

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 August 2020 doi:10.20944/preprints202008.0702.v1

empty test tube included as a process blank. To each tube, 1 mL of a 3:3:2 mixture of acetonitrile/isopropanol/water containing 0.1% formic acid and 0.25 μM telmisartan internal standard was added. Tubes were vortexed for 30 s and solvent decanted into 2-mL glass autosampler vials. Equal volumes of each extract (excluding the process blank) were combined to create a pooled quality control (QC) sample. Vials were sealed with polytetrafluoroethylene (PTFE)-lined caps and stored at -20°C for later processing.

4.3 Metabolomic analysis by UHPLC-MS

Aliquots of *S. pennellii* acylsugar extracts, process blank, and QC sample were diluted 100-fold in 1:1 methanol/water containing 0.1% formic acid in new 2-mL autosampler vials. Five aliquots of the diluted process blank and QC samples were prepared and analyzed. Analyte samples were injected in a randomized order while process blank and QC samples were injected at regular intervals. Samples were subjected to UHPLC-MS analysis using an Acquity I-class pump coupled to a G2-XS QToF mass spectrometer (Waters Corporation, Milford, MA). Separations were performed by reverse phase (C18) chromatography using a 20 min gradient and mass spectra were acquired in continuum format from 2 to 18 min using quasi-simultaneous acquisition of low- and high-energy spectra (MS^E). Additional details of the UHPLC-MS method are in Table S2.

4.4 Untargeted metabolomics data processing

For untargeted metabolomic analysis, data were initially processed using Progenesis QI v2.4 software (Nonlinear Dynamics Ltd., Newcastle, UK). Leucine enkephalin lockmass correction (*m/z* 556.2766) was applied during run importation and all runs were aligned to retention times of a bulk pool run automatically selected by the software. Peak picking was carried out on features eluting between 2.15 and 14.5 min using an automatic sensitivity level of 5 (most sensitive) without restriction on minimum chromatographic peak width. This resulted in detection of 2361 compound ions. Spectral deconvolution was carried out considering the following possible adduct ions: M+H-H_2O, M+H, M+NH_4, M+Na, M+K, M+C_2H_8N, 2M+H, 2M+NH_4, 2M+Na, 2M+K, 2M+C_2H_8N. After deconvolution, 1559 compound ions remained.

To remove features from the dataset introduced by solvents, glassware, or instrumentation, several filters were applied to the 1559 compound ions remaining after deconvolution. Compounds with the highest mean abundance in process blank samples, maximum abundance less than 0.5% of the most abundant compound in the dataset, or a coefficient of variation > 20% across QC samples were excluded from the dataset. This reduced the total number of metabolic features to 54.

Further analysis of compound signals extracted by Progenesis QI software was executed using EZinfo v3.0.2 software (Umetrics, Umeå, Sweden). For principal component analysis (PCA), data were subjected to logarithmic transformation and scaled to unit variance ("autoscaled"). For orthogonal partial least squares/projection to latent structures discriminant analysis (OPLS-DA), no data transformation was applied, and Pareto scaling was implemented. Generation of OPLS-DA models was carried out as follows: for each model, the relevant data files were divided into three subsets, each subset containing data files representing two of six biological replicates from each accession considered by the model. The first data subset contained data files representing the first pair of biological replicates from each relevant accession in the randomized injection list, while the second and third subsets contained the second and third pairs of biological replicates, respectively. The three data subsets, each representing one-third of the relevant data, were used as training sets to generate three independent OPLS-DA models. Each model was then used to classify the remaining two-thirds of the data not used in generation of the model, representing four of six biological replicates from each accession considered by the model. All OPLS-DA model statistics reported represent averages of the three independent models.

For all metabolic features extracted with Progenesis QI and used in downstream analyses with EZinfo, spectra were interpreted using MassLynx v4.2 software (Waters Corporation). Accurate masses of all features in all raw data files were obtained by applying the Continuous Lockmass Correction feature of the Accurate Mass Measure module. All precursor ions (annotated as either

 doi:10.20944/preprints202008.0702.v1

$[M+NH_4]^+$ or $[M+H]^+$ adducts) were selected from the low-energy function while all fragment ions were assigned based on the high-energy function. Observed *m/z* values for precursor and product ions as well as neutral loss masses were compared to theoretical values generated using ChemDraw v19.0 software (PerkinElmer, Inc., Waltham, MA). For acylsugars, molecular formulas were determined by comparing accurate *m/z* values of $[M+NH_4]^+$ pseudomolecular ions to theoretical *m/z* values of hypothetical acylsugar $[M+NH_4]^+$ adducts. The molecular formulas of all acyl chain components from individual acylsugars were inferred by a similar process using ketene and fatty acid neutral losses from pseudomolecular precursor ions observed in the high-energy function. While acylium product ions representing acyl chains appear in many spectra, their occurrence is inconsistent across compounds, especially in those of low abundance. Therefore, all acyl chain assignments were made using the neutral loss data, which could be unambiguously interpreted for all spectra. For flavonoids, molecular formulas were determined by comparing accurate *m/z* values of $[M+H]^+$ pseudomolecular ions to theoretical *m/z* values of hypothetical flavonoid $[M+H]^+$ adducts.

4.5 Acylsugar quantification

Acylsugars were quantified from untargeted UHPLC-MS data by integration of extracted ion chromatogram peaks using the QuanLynx module of MassLynx software (Waters Corporation). All acylsucroses and acylglucoses detected in the metabolomics dataset were quantified using a standard curve of two each of purified acylsucroses and acylglucoses [S3:12(4,4,4), S3:18(4,4,10)-1, G3:12(4,4,4), and G3:18(4,4,10)-1] of authenticated concentrations at 0.3125, 0.625, 1.25, 2.5, and 5.0 µM. Acylsugars containing fewer than 18 carbons in all acyl chains were quantified using the G3:12(4,4,4) or S3:12(4,4,4) response factor while acylsugars containing 18 or more carbons were quantified using the response factor of G3:18(4,4,10)-1 or S3:18(4,4,10)-1. All quantifications were performed using extracted ion chromatograms of the *m/z* value for the relevant $M+NH_4^+$ adduct using a mass window of *m/z* 0.05. When multiple acylsugar isomers (including anomers) were present, all acylsugars of a given molecular formula were quantified using a single extracted ion chromatogram. The retention time window was adjusted for each compound based on the number of isomers and retention time differences between isomers. Telmisartan was used as an internal reference for all quantifications.

For quantification of total acylsugar cores (*i.e.*, sucrose and glucose), acylsugar extracts or purified acylsugars were saponified and sugar core quantified using UHPLC-MS-MS. For each acylsugar analyte or standard, a 20-µL aliquot was evaporated to dryness in a 1.7-mL microfuge tube using a vacuum centrifuge and dissolved in 200 µL of a 1:1 methanol/3 M aqueous ammonia solution. The saponification reactions were incubated at room temperature for 48 hours at which point solvent was removed by vacuum centrifuge at room temperature. The dried residue was dissolved in 200 µL of 10 mM ammonium bicarbonate (pH 8.0) in 90% acetonitrile containing 0.5 µM $^{13}C_{12}$-sucrose and 0.5 µM $^{13}C_6$-glucose as internal standards and transferred to a 2-mL glass vial. Samples were subjected to UHPLC-MS-MS analysis. Levels of sucrose and glucose were quantified using a standard curve of the corresponding sugar at final concentrations of 3.13, 6.25, 12.5, 25, and 50 µM. Details of the UHPLC-MS-MS method are in Table S3.

4.6 RNA extraction, cDNA synthesis, and qPCR

Relative *ASFF1* transcript levels were measured using a published method [22]. Briefly, single leaflets from the youngest fully-expanded leaf of three biological replicates each of 12-week-old *S. pennellii* LA0716 and LA2963 plants were harvested and powdered under liquid nitrogen prior to RNA extraction using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany), cDNA synthesis using SuperScript III reverse transcriptase (Invitrogen, Carlsbad, CA), and qRT-PCR analysis using SYBR Green PCR Master Mix on a QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems, Warrington, UK). RT_ASFF_F and RT_ASFF_R primers were used to detect *ASFF1* transcript; RT_EF-1a_F/R, RT_actin_F/R, and RT_ ubiquitin_F/R primers were used to detect transcripts of the *EF-1α*, *actin*, and *ubiquitin* genes, respectively (Table S4). Relative levels of *ASFF1* transcript were

determined using the ΔΔCt method [47] and nor3zed to the geometric mean of *EF-1α*, *actin*, and *ubiquitin* transcript levels.

4.7 Acylsugar purification

Purifications were performed using a Waters 2795 Separations Module (Waters Corporation) and an Acclaim 120 C18 HPLC column (4.6 x 150 mm, 5 μm; ThermoFisher Scientific, Waltham, MA) with a column oven temperature of 30°C and flow rate of 2 mL/min. For acylsucrose purification, the mobile phase consisted of water (solvent A) and acetonitrile (solvent B). For acylglucose purification, methanol was used as solvent B. Fractions were collected using a 2211 Superrac fraction collector (LKB Bromma, Stockholm, Sweden).

For purification of acylsucroses, acylsugars were extracted from mature plants of the *S. pennellii* LA0716 *asff1-1* mutant [22], which exclusively accumulates acylsucroses. Surface metabolites from ~75 g leaflets were extracted in 500 mL methanol containing 0.1% formic acid. This extract was dried under vacuum with a rotary evaporator and the resulting residue dissolved in ~3 mL acetonitrile containing 0.1% formic acid. Quantification of this solution by UPLC-MS-MS indicated a concentration of ~150 mM total acylsucroses. This extract was diluted 14-fold in 70% acetonitrile containing 0.1% formic acid. Acylsucroses were purified by pooling fractions from 10 injections of 50 μL each. A linear gradient of 45% B at 0 min, 60% B at 30 min, 100% B at 30.01 min held until 35 min, and 45% B at 35.01 min held until 40 min was used. Fractions were collected at 10-s intervals into tubes containing 300 μL 0.1% formic acid in water. The S3:12(4,4,4) compound eluted between 1 and 2 min; the S3:18(4,4,10)-1 compound eluted between 12 and 14 min; the S3:18(4,4,10)-2 compound eluted between 14 and 16 min; the S3:19(4,5,10)-1 compound eluted between 17 and 19 min; and the S3:19(4,5,10)-2 compound eluted between 19 and 21 min.

For purification of acylglucoses, acylsugars were extracted from mature plants of *S. pennellii* LA0716, which accumulates > 90% acylglucoses [26]. Surface metabolites were extracted from ~75 g leaflets as described for acylsucrose purification. UHPLC-MS-MS quantification of the resulting solution indicated ~500 mM acylglucose concentration. This extract was diluted 20-fold in 1:1 methanol/water containing 0.1% formic acid. Acylglucoses were purified by pooling fractions after 20 injections of 50 μL each. A linear gradient of 5% B at 0-1 min, 60% B at 2 min, 100% B at 32 min held until 35 min, and 5% B at 36 min held until 40 min was used. The G3:12(4,4,4) compound eluted between 6 and 7 min; the G3:18(4,4,10)-1 compound eluted between 17 and 18 min; the G3:18(4,4,10)-2 compound eluted between 18 and 19 min; the G3:19(4,5,10)-1 compound eluted between 20 and 21 min; and the G3:19(4,5,10)-2 compound eluted between 21 and 22 min.

Purity of acylsugar fractions was verified by UHPLC-MS using an LC-20AD HPLC (Shimadzu, Kyoto, Japan) coupled to a G2-XS QToF mass spectrometer (Waters Corporation). Separations were performed using an Ascentis Express C18 HPLC column (2.1 x 100 mm, 2.7 μm; Supelco, Bellefonte, PA). The mobile phases consisted of 100 mM ammonium formate, pH 3.4 (solvent A) and 100 mM ammonium formate, pH 3.4, in 90% methanol (solvent B). Five-microliter aliquots were injected onto the column and eluted with a linear gradient of 5% B at 0-1 min, 60% B at 1.01 min, 100% B at 8 min, and 5% B at 8.01-10 min. The solvent flow rate was 0.4 mL/min and the column temperature was 40°C. Analyses were performed using positive-ion mode electrospray ionization and sensitivity mode analyzer parameters. Source parameters were: capillary voltage at 3.00 kV, sampling cone voltage at 40 V, source offset at 80 V, source temperature at 100°C, desolvation temperature at 350°C, cone gas flow at 50.0 L/hour, and desolvation gas flow at 600.0 L/hour. Quasi-simultaneous mass spectrum acquisition at low and high collision energy conditions (MS^E) was performed over an *m/z* range of 50 to 1500 with a scan time of 0.5 s. Adduct ions were obtained using a collision potential of 6.0 V; fragment ions were obtained using a collision potential ramp of 15 to 40 V. Spectra were acquired in centroid format.

Pure acylsugar fractions were pooled and solvent removed using a vacuum centrifuge. Residues were reconstituted in 1 mL 3:3:2 acetonitrile/isopropanol/water with 0.1% formic acid, transferred to 2-mL glass autosampler vials, sealed with PTFE-lined caps, and stored at -20°C. Aliquots of purified acylsugars were quantified using the saponification method described above.

4.8 NMR spectroscopy

NMR spectra (^{1}H, gCOSY, gHSQC, gHMBC, and ^{1}H-^{1}H *J*-resolved spectra were collected at the Max T. Rogers NMR Facility at Michigan State University using a DDR 500 MHz NMR spectrometer (Agilent, Santa Clara, CA) equipped with a 7600AS 96-sample autosampler running VnmrJ v3.2A software. ^{13}C spectra were collected on the same instrument at 125 MHz. All spectra were referenced to non-deuterated chloroform solvent signals (δH = 7.26 (s) and δC = 77.2 (t) ppm). Additional details of the NMR data collection methods are in Table S5.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1: Supplementary file 1 (supp1.docx) containing Tables S1-18 and Figures S1-62, and Supplementary file 2 (supp2.xlsx) containing the complete processed dataset for the findings presented.

Author Contributions: Conceptualization, D.B.L and R.L.L.; methodology, D.B.L. and A.D.J.; formal analysis, D.B.L., T.M.A., and A.D.J.; investigation, D.B.L and T.M.A; writing—original draft preparation, D.B.L; writing—review and editing, D.B.L, T.M.A, A.D.J., and R.L.L; supervision, A.D.J. and R.L.L.; funding acquisition, A.D.J. and R.L.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Foundation, grant number IOS-1546617 to A.D.J. and R.L.L. D.B.L. was also supported by the National Institutes of Health predoctoral training grant T32-GM110523.

Acknowledgments: The authors thank Philip Engelgau, Hannah Parks, and Dr. Rachel E. Kerwin for helpful comments on the manuscript. The authors also thank Dr. Anthony L. Schilmiller at the Michigan State University Mass Spectrometry and Metabolomics Core (MSMC) facility for technical assistance with data collection. Mass spectra were collected in the laboratory of A.D.J. and at the MSMC facility. All NMR spectra were collected at the Max T. Rogers NMR facility at Michigan State University.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Pichersky, E.; Lewinsohn, E. Convergent Evolution in Plant Specialized Metabolism. *Annu. Rev. Plant Biol.* **2011**, *62*, 549–566, doi:10.1146/annurev-arplant-042110-103814.
2. Pichersky, E.; Raguso, R.A. Why do plants produce so many terpenoid compounds? *New Phytol.* **2018**, *220*, 692–702, doi:10.1111/nph.14178.
3. Tahara, S. A Journey of Twenty-Five Years through the Ecological Biochemistry of Flavonoids. *Biosci. Biotechnol. Biochem.* **2007**, *71*, 1387–1404, doi:10.1271/bbb.70028.
4. Milo, R.; Last, R.L. Achieving Diversity in the Face of Constraints: Lessons from Metabolism. *Science* **2012**, *336*, 1663–1667, doi:10.1126/science.1217665.
5. Liu, Y.; Jing, S.-X.; Luo, S.-H.; Li, S.-H. Non-volatile natural products in plant glandular trichomes: chemistry, biological activities and biosynthesis. *Nat. Prod. Rep.* **2019**, *36*, 626–665, doi:10.1039/c8np00077h.
6. Kang, J.-H.; Liu, G.; Shi, F.; Jones, A.D.; Beaudry, R.M.; Howe, G.A. The Tomato *odorless-2* Mutant Is Defective in Trichome-Based Production of Diverse Specialized Metabolites and Broad-Spectrum Resistance to Insect Herbivores. *Plant Physiol.* **2010**, *154*, 262–272, doi:10.1104/pp.110.160192.
7. Luu, V.T.; Weinhold, A.; Ullah, C.; Dressel, S.; Schoettner, M.; Gase, K.; Gaquerel, E.; Xu, S.; Baldwin, I.T. *O*-Acyl Sugars Protect a Wild Tobacco from Both Native Fungal Pathogens and a Specialist Herbivore. *Plant Physiol.* **2017**, *174*, 370–386, doi:10.1104/pp.16.01904.
8. Weinhold, A.; Baldwin, I.T. Trichome-derived O-acyl sugars are a first meal for caterpillars that tags them for predation. *Proc. Natl. Acad. Sci.* **2011**, *108*, 7855–7859, doi:10.1073/pnas.1101306108.

9. Yang, T.; Stoopen, G.; Wiegers, G.; Mao, J.; Wang, C.; Dicke, M.; Jongsma, M.A. Pyrethrins Protect Pyrethrum Leaves Against Attack by Western Flower Thrips, *Frankliniella occidentalis*. *J. Chem. Ecol.* **2012**, *38*, 370–377, doi:10.1007/s10886-012-0097-7.
10. McDowell, E.T.; Kapteyn, J.; Schmidt, A.; Li, C.; Kang, J.-H.; Descour, A.; Shi, F.; Larson, M.; Schilmiller, A.; An, L.; et al. Comparative Functional Genomic Analysis of *Solanum* Glandular Trichome Types. *Plant Physiol.* **2011**, *155*, 524–539, doi:10.1104/pp.110.167114.
11. Nakashima, T.; Wada, H.; Morita, S.; Erra-Balsells, R.; Hiraoka, K.; Nonami, H. Single-Cell Metabolite Profiling of Stalk and Glandular Cells of Intact Trichomes with Internal Electrode Capillary Pressure Probe Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2016**, *88*, 3049–3057, doi:10.1021/acs.analchem.5b03366.
12. Fobes, J.F.; Mudd, J.B.; Marsden, M.P.F. Epicuticular Lipid Accumulation on the Leaves of *Lycopersicon pennellii* (Corr.) D'Arcy and *Lycopersicon esculentum* Mill. *Plant Physiol.* **1985**, *77*, 567–570, doi:10.1104/pp.77.3.567.
13. Leckie, B.M.; D'Ambrosio, D.A.; Chappell, T.M.; Halitschke, R.; De Jong, D.M.; Kessler, A.; Kennedy, G.G.; Mutschler, M.A. Differential and Synergistic Functionality of Acylsugars in Suppressing Oviposition by Insect Herbivores. *PLOS ONE* **2016**, *11*, e0153345, doi:10.1371/journal.pone.0153345.
14. Mirnezhad, M.; Romero-González, R.R.; Leiss, K.A.; Choi, Y.H.; Verpoorte, R.; Klinkhamer, P.G.L. Metabolomic analysis of host plant resistance to thrips in wild and cultivated tomatoes. *Phytochem. Anal.* **2010**, *21*, 110–117, doi:10.1002/pca.1182.
15. Smeda, J.R.; Schilmiller, A.L.; Anderson, T.; Ben-Mahmoud, S.; Ullman, D.E.; Chappell, T.M.; Kessler, A.; Mutschler, M.A. Combination of Acylglucose QTL reveals additive and epistatic genetic interactions and impacts insect oviposition and virus infection. *Mol. Breed.* **2018**, *38*, 3, doi:10.1007/s11032-017-0756-z.
16. Leckie, B.M.; De Jong, D.M.; Mutschler, M.A. Quantitative trait loci increasing acylsugars in tomato breeding lines and their impacts on silverleaf whiteflies. *Mol. Breed.* **2012**, *30*, 1621–1634, doi:10.1007/s11032-012-9746-3.
17. Leckie, B.M.; De Jong, D.M.; Mutschler, M.A. Quantitative trait loci regulating sugar moiety of acylsugars in tomato. *Mol. Breed.* **2013**, *31*, 957–970, doi:10.1007/s11032-013-9849-5.
18. Leckie, B.M.; Halitschke, R.; De Jong, D.M.; Smeda, J.R.; Kessler, A.; Mutschler, M.A. Quantitative trait loci regulating the fatty acid profile of acylsugars in tomato. *Mol. Breed.* **2014**, *34*, 1201–1213, doi:10.1007/s11032-014-0110-7.
19. Fan, P.; Miller, A.M.; Liu, X.; Jones, A.D.; Last, R.L. Evolution of a flipped pathway creates metabolic innovation in tomato trichomes through BAHD enzyme promiscuity. *Nat. Commun.* **2017**, *8*, 2080, doi:10.1038/s41467-017-02045-7.
20. 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.; et al. Evolution of a plant gene cluster in Solanaceae and emergence of metabolic diversity. *eLife* **2020**, *9*, e56717, doi:10.7554/eLife.56717.
21. Ning, J.; Moghe, G.; Leong, B.; Kim, J.; Ofner, I.; Wang, Z.; Adams, C.; Jones, A., Daniel; Zamir, D.; Last, R., L. A feedback insensitive isopropylmalate synthase affects acylsugar composition in cultivated and wild tomato. *Plant Physiol.* **2015**, *169*, 1821–1835, doi:10.1104/pp.15.00474.
22. Leong, B.J.; Lybrand, D.B.; Lou, Y.-R.; Fan, P.; Schilmiller, A.L.; Last, R.L. Evolution of metabolic novelty: A trichome-expressed invertase creates specialized metabolic diversity in wild tomato. *Sci. Adv.* **2019**, *5*, eaaw3754, doi:10.1126/sciadv.aaw3754.

23. Schilmiller, A.L.; Gilgallon, K.; Ghosh, B.; Jones, A.D.; Last, R.L. Acylsugar Acylhydrolases: Carboxylesterase-Catalyzed Hydrolysis of Acylsugars in Tomato Trichomes. *Plant Physiol.* **2016**, *170*, 1331–1344, doi:10.1104/pp.15.01348.
24. Wang, Z.; Jones, A.D. Profiling of Stable Isotope Enrichment in Specialized Metabolites Using Liquid Chromatography and Multiplexed Nonselective Collision-Induced Dissociation. *Anal. Chem.* **2014**, *86*, 10600–10607, doi:10.1021/ac502205y.
25. Wang, Z. Dynamics, distribution and development of specialized metabolism in glandular trichome of tomato and its wild relatives. Ph.D., Michigan State University: United States -- Michigan, 2015.
26. Shapiro, J.A.; Steffens, J.C.; Mutschler, M.A. Acylsugars of the wild tomato *Lycopersicon pennellii* in relation to geographic distribution of the species. *Biochem. Syst. Ecol.* **1994**, *22*, 545–561, doi:10.1016/0305-1978(94)90067-1.
27. Slocombe, S.P.; Schauvinhold, I.; McQuinn, R.P.; Besser, K.; Welsby, N.A.; Harper, A.; Aziz, N.; Li, Y.; Larson, T.R.; Giovannoni, J.; et al. Transcriptomic and Reverse Genetic Analyses of Branched-Chain Fatty Acid and Acyl Sugar Production in *Solanum pennellii* and *Nicotiana benthamiana*. *Plant Physiol.* **2008**, *148*, 1830–1846, doi:10.1104/pp.108.129510.
28. Burke, B.A.; Goldsby, G.; Mudd, B.J. Polar epicuticular lipids of *Lycopersicon pennellii*. *Phytochemistry* **1987**, *26*, 2567–2571, doi:10.1016/S0031-9422(00)83879-0.
29. Correll, D.S. *The potato and its wild relatives.*; Contributions from Texas Research Foundation; Texas Research Foundation: Renner, Texas, 1962;
30. Ghangas, G.S.; Steffens, J.C. UDPglucose: fatty acid transglucosylation and transacylation in triacylglucose biosynthesis. *Proc. Natl. Acad. Sci.* **1993**, *90*, 9911–9915, doi:10.1073/pnas.90.21.9911.
31. Kuai, J.P.; Ghangas, G.S.; Steffens, J.C. Regulation of Triacylglucose Fatty Acid Composition (Uridine Diphosphate Glucose:Fatty Acid Glucosyltransferases with Overlapping Chain-Length Specificity). *Plant Physiol.* **1997**, *115*, 1581–1587, doi:10.1104/pp.115.4.1581.
32. Li, Alice.X.; Steffens, J.C. An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. *Proc. Natl. Acad. Sci.* **2000**, *97*, 6902–6907, doi:10.1073/pnas.110154197.
33. Li, A.X.; Eannetta, N.; Ghangas, G.S.; Steffens, J.C. Glucose Polyester Biosynthesis. Purification and Characterization of a Glucose Acyltransferase. *Plant Physiol.* **1999**, *121*, 453–460, doi:10.1104/pp.121.2.453.
34. Fan, P.; Miller, A.M.; Schilmiller, A.L.; Liu, X.; Ofner, I.; Jones, A.D.; Zamir, D.; Last, R.L. In vitro reconstruction and analysis of evolutionary variation of the tomato acylsucrose metabolic network. *Proc. Natl. Acad. Sci.* **2016**, *113*, E239–E248, doi:10.1073/pnas.1517930113.
35. Ghosh, B.; Westbrook, T.C.; Jones, A.D. Comparative structural profiling of trichome specialized metabolites in tomato (*Solanum lycopersicum*) and *S. habrochaites*: acylsugar profiles revealed by UHPLC/MS and NMR. *Metabolomics* **2014**, *10*, 496–507, doi:10.1007/s11306-013-0585-y.
36. Liu, X.; Enright, M.; Barry, C.S.; Jones, A.D. Profiling, isolation and structure elucidation of specialized acylsucrose metabolites accumulating in trichomes of *Petunia* species. *Metabolomics* **2017**, *13*, 85, doi:10.1007/s11306-017-1224-9.
37. Nadakuduti, S.S.; Uebler, J.B.; Liu, X.; Jones, A.D.; Barry, C.S. Characterization of Trichome-Expressed BAHD Acyltransferases in *Petunia axillaris* Reveals Distinct Acylsugar Assembly Mechanisms within the Solanaceae. *Plant Physiol.* **2017**, *175*, 36–50, doi:10.1104/pp.17.00538.
38. Kim, J.; Kang, K.; Gonzales-Vigil, E.; Shi, F.; Jones, A.D.; Barry, C.S.; Last, R.L. Striking Natural Diversity in Glandular Trichome Acylsugar Composition Is Shaped by Variation at the Acyltransferase2 Locus in the Wild Tomato *Solanum habrochaites*. *Plant Physiol.* **2012**, *160*, 1854–1870, doi:10.1104/pp.112.204735.

39. Mandal, S.; Ji, W.; McKnight, T.D. Candidate Gene Networks for Acylsugar Metabolism and Plant Defense in Wild Tomato *Solanum pennellii*. *Plant Cell* **2020**, *32*, 81–99, doi:10.1105/tpc.19.00552.
40. Schmidt, A.; Li, C.; Shi, F.; Jones, A.D.; Pichersky, E. Polymethylated Myricetin in Trichomes of the Wild Tomato Species *Solanum habrochaites* and Characterization of Trichome-Specific 3′/5′- and 7/4′-Myricetin O-Methyltransferases. *Plant Physiol.* **2011**, *155*, 1999–2009, doi:10.1104/pp.110.169961.
41. Schmidt, A.; Li, C.; Jones, A.D.; Pichersky, E. Characterization of a flavonol 3-O-methyltransferase in the trichomes of the wild tomato species *Solanum habrochaites*. *Planta* **2012**, *236*, 839–849, doi:10.1007/s00425-012-1676-0.
42. Kim, J.; Matsuba, Y.; Ning, J.; Schilmiller, A.L.; Hammar, D.; Jones, A.D.; Pichersky, E.; Last, R.L. Analysis of Natural and Induced Variation in Tomato Glandular Trichome Flavonoids Identifies a Gene Not Present in the Reference Genome. *Plant Cell* **2014**, *26*, 3272–3285, doi:10.1105/tpc.114.129460.
43. Li, C.; Schmidt, A.; Pichersky, E.; Shi, F.; Jones, A.D. Identification of methylated flavonoid regioisomeric metabolites using enzymatic semisynthesis and liquid chromatography-tandem mass spectrometry. *Metabolomics* **2013**, *9*, 92–101, doi:10.1007/s11306-012-0451-3.
44. Maluf, W.R.; Maciel, G.M.; Gomes, L.A.A.; Cardoso, M. das G.; Gonçalves, L.D.; Silva, E.C. da; Knapp, M. Broad-Spectrum Arthropod Resistance in Hybrids between High- and Low-Acylsugar Tomato Lines. *Crop Sci.* **2010**, *50*, 439–450, doi:10.2135/cropsci2009.01.0045.
45. Alba, J.M.; Montserrat, M.; Fernández-Muñoz, R. Resistance to the two-spotted spider mite (*Tetranychus urticae*) by acylsucroses of wild tomato (*Solanum pimpinellifolium*) trichomes studied in a recombinant inbred line population. *Exp. Appl. Acarol.* **2009**, *47*, 35–47, doi:10.1007/s10493-008-9192-4.
46. Schilmiller, A.L.; Charbonneau, A.L.; Last, R.L. Identification of a BAHD acetyltransferase that produces protective acyl sugars in tomato trichomes. *Proc. Natl. Acad. Sci.* **2012**, *109*, 16377–16382, doi:10.1073/pnas.1207906109.
47. Pfaffl, M.W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **2001**, *29*, 45e–445, doi:10.1093/nar/29.9.e45.