Evolution of metabolic novelty: A trichome-expressed invertase creates specialized metabolic diversity in wild tomato

Bryan J. Leong¹*, Daniel B. Lybrand²*, Yann-Ru Lou², Pengxiang Fan², Anthony L. Schilmiller³, Robert L. Last¹,2†

Plants produce a myriad of taxonomically restricted specialized metabolites. This diversity—and our ability to correlate genotype with phenotype—makes the evolution of these ecologically and medicinally important compounds interesting and experimentally tractable. Trichomes of tomato and other nightshade family plants produce structurally diverse protective compounds termed acylsugars. While cultivated tomato (Solanum lycopersicum) strictly accumulates acylsuccoses, the South American wild relative Solanum penellii produces copious amounts of acylglucoses. Genetic, transgenic, and biochemical dissection of the S. penellii acylglucose biosynthetic pathway identified a trichome gland cell–expressed invertase-like enzyme that hydrolyzes acylsuccoses (Sopen03g040490). This enzyme acts on the pyranose ring–acylated acysuccoses found in the wild tomato but not on the furanose ring–decorated acysuccoses of cultivated tomato. These results show that modification of the core acylsucrose biosynthetic pathway leading to loss of furanose ring acylation set the stage for co-option of a general metabolic enzyme to produce a new class of protective compounds.

INTRODUCTION

Plants synthesize hundreds of thousands of structurally diverse and lineage-, tissue-, or cell type–specific specialized metabolites (1). These compounds act as antiherbivory, antimicrobial, or allelopathic compounds (2–4). The enzymes that produce these specialized metabolites arise from a variety of proteins through gene duplication and neo- or subfunctionalization. Examples include diversification of existing specialized metabolic enzymes (5) and co-option of core metabolic activities (6). In the latter mechanism, gene duplication can remove stabilizing selective pressure on one of the paralogs associated with core metabolism, enabling evolution of the novel enzymatic activity without disrupting growth and development. This leads to new and modified pathways that produce structurally and functionally diverse specialized metabolites.

Glandular trichome–synthesized acylated sugars (“acylsugars”) are structurally diverse specialized metabolites found throughout the Solanaceae (7–13). These compounds have documented roles in direct and indirect protection against herbivores and microbes (14, 15), as well as allelopathic properties (15, 16). Their low toxicity to vertebrates generates interest in generating plant breeding strategies for deploying acylsugars in crop protection (17, 18). These metabolites consist of a sugar core—typically sucrose—with aliphatic chains of variable length, structure, and number attached by ester linkages. Acylsugars were reported from genera across the Solanaceae family, including Datura, Nicotiana, Petunia, Physalis, Salpiglossis, and Solanum with single species producing at least three dozen chromographically distinct acylsugars (10, 13, 16, 19–22).

In recent years, several evolutionarily related enzymes were implicated in the core acylsucrose biosynthetic pathways in species across the family, including the cultivated tomato Solanum lycopersicum, Petunia axillaris, and Salpiglossis sinuata (7, 9–12, 23). These biosynthetic pathways consist of trichome–expressed BAHD-family (BEAT, AHCT, HCBT, and DAT) acysugar acyltransferases (ASATs) (7, 9, 23), which sequentially transfer acyl groups from acyl–coenzyme A (acyl-CoA) substrates to specific hydroxyl groups of sucrose (7, 9, 23).

The cultivated tomato biosynthetic network is well characterized, with four ASATs—SIASAT1 to SIASAT4—catalyzing consecutive reactions to produce tri- and tetra-acylated sucrosses. SIASAT1 acts first by transferring an acyl chain to the R4 hydroxyl of the pyranose ring of sucrose, and SIASAT2 transfers an acyl chain to the R3 position of the monoacylated sucrose (23). Next, SIASAT3 acylates the diacylated sucroses at the furanose ring R3′ position (9). SIASAT4 completes the pathway by transferring an acetyl group to the pyranose ring R2 position of a triacylsucrose (7, 22). Enzyme promiscuity and the presence of an array of acyl-CoAs result in the production of a diverse group of acylsucroses in S. lycopersicum (11, 24).

The metabolic diversity in acylsugars is even greater in the broader Solanum genus. The wild relative of tomato, Solanum penellii LA0716, is a prime example, producing a mixture of abundant acylsucroses that are distinct from those in S. lycopersicum. While S. lycopersicum accumulates acylsucroses with two or three acylations on the pyranose ring and a single acylation at the furanose ring R3′ position (termed F-type acylsucroses), S. penellii accumulates distinct P-type triacylsucroses acylated only on the pyranose R2, R3, and R4 positions (11). P-type acylsucroses are synthesized by S. penellii orthologs of the S. lycopersicum ASAT1, ASAT2, and ASAT3 enzymes (Fig. 1A). The different acylation pattern observed in S. penellii results from altered substrate specificity and acylation position of SpASAT2 and SpASAT3 relative to their S. lycopersicum counterparts (11).

S. penellii LA0716 has other acylsugar characteristics that differentiate it from cultivated tomato (Fig. 1B). First, it produces copious
characterized *S. pennellii* acylsucrose biosynthetic enzymes (9, 11), no complete acylglucose metabolic pathway has yet been described. This is despite the fact that acylglucoses were also characterized in several additional Solanaceae species (16, 20). A previously proposed partial *S. pennellii* pathway invoked two glucosyltransferases capable of creating 1-O-acetyl-D-glucose from uridine diphosphate (UDP)–glucose and free fatty acids of differing structures (27). This mechanism proposed a second step in which a serine carboxypeptidase-like (SCPL) acyltransferase catalyzed disproportionation of two 1-O-isobutyryl-D-glucose molecules to yield one 1,2-O-diisobutyryl-D-glucose (28, 29). However, this pathway is unlikely to function in vivo as the 1,2,4-O-triacylglycerols obtained in vitro differ from the 2,3,4-O-triacylglycerols observed in *S. pennellii* both in the number (two instead of three) and in the position of acyl chains: *S. pennellii* acylglucoses bear chains at the R2, R3, and R4 positions rather than at the R1 position (19).

In contrast to the unsubstantiated published biosynthetic pathway, compelling qualitative trait locus (QTL) and biochemical results implicate multiple genetic loci in acylglucose accumulation in *S. pennellii* LA0716. The combination of three *S. pennellii* regions on chromosomes 3, 4, and 11 causes *S. lycopersicum* breeding line CU071026 to accumulate acylsugars comprising up to 89% acylglucoses (30). The presence of QTLs on both chromosomes 3 and 11 yields detectable acylglucoses, while addition of the chromosome 4 locus leads to elevated accumulation. Notably, chromosome 4 and 11 QTLs respectively include the *SpASAT2* and *SpASAT3* genes responsible for accumulation of P-type acylsugars in *S. pennellii* (9, 23). The agreement of QTL and biochemical data is consistent with the hypothesis that *SpASAT2* and *SpASAT3* produce P-type acylsaccharides that are substrates for a chromosome 3 factor that then synthesize triacylglycerols.

We report the characterization of the plant specialized metabolic invertease–like enzyme acylacrylc sugar fructofuranosidase I (*SpASFF1*; Sopen03g040490), a chromosome 3 β-fructofuranosidase capable of cleaving the glycosidic bond of P-type acylsugars. Genetic and transgenic plant approaches demonstrate that *S. pennellii* LA0716 acylglucose production requires *SpASFF1*. This work also documents a three-gene epistatic interaction between the *SpASAT2*, *SpASAT3*, and *SpASFF1* loci that conditions high-level acylglucose accumulation. While yeast invertase and other variants involved in core metabolism have been studied since the 19th century (31–33), this work documents a new type of role for β-fructofuranosidase–type enzymes in specialized metabolism. These results extend our understanding of evolutionary mechanisms leading to trichome specialized metabolic diversity by demonstrating how neofunctionalization led to co-option of invertase from general metabolism into a cell type–specific specialized metabolic network.

**RESULTS**

*An S. pennellii chromosome 3 locus is necessary for acylglucose production from P-type acylsugars*

Published QTL mapping studies indicate that introgression of *S. pennellii* LA0716 loci on chromosomes 3, 4, and 11 leads to accumulation of acylglucoses in a cultivated tomato *S. lycopersicum* background (30, 34). Three introgression lines harboring individual acylglucose QTLs in the *S. lycopersicum* background were screened, but none of the single introgressions in lines IL3-5, IL4-1, or IL11-3 (35) yielded detectable leaf acylglucoses (fig. S1). These observations are consistent with the hypothesis that multiple *S. pennellii* loci are

Fig. 1. Three *S. pennellii* LA0716 regions condition acylglucose accumulation. (A) Examples of nuclear magnetic resonance (NMR)–resolved *S. lycopersicum* and *S. pennellii* acylsugar structures. Acylsugars from *S. lycopersicum* are composed of sucrose acylated on both the pyranose and furanose rings (“F-type”). *S. pennellii* acylsugars are a mixture of sucrose-based (“P-type”) and glucose-based compounds with acylation exclusively on the pyranose ring. (B) Acylsugar electrospray ionization (ESI)–mode liquid chromatography–mass spectrometry (LC-MS) analysis of *S. lycopersicum* M82 and *S. pennellii* LA0716 acylglucoses and acylsucroses. (C) Left: Representation of *S. pennellii* chromosomal introgressions in BIL6521 × BIL6180 progeny that contain quantitative trait loci (QTLs) affecting acylglucose biosynthesis (30). The black portions of the chromosomes correspond to *S. pennellii* introgressions, while the white portions correspond to the chromosomal regions in the BIL62 background. Right: ESI–mode liquid chromatography–mass spectrometry (LC-MS) analysis of BIL6180 compared with the BIL6180 × BIL6521 F2 progeny reveals acylglucose accumulation in the hybrid, but not in BIL6180. All ESI–mode acylsugars were identified as formate adducts. Chr., chromosome.

amounts of acylsugars that render the plant sticky, representing up to ~20% of leaf dry weight (19, 25). Second, the vast majority of *S. pennellii* LA0716 acylsugars are glucose molecules with three acyl chains (termed “acylglycerols”) (Fig. 1, A and B), while only 7 to 16% of total acylsugars are acylsucroses (26). In contrast to the well-
needed for *S. lycopersicum* acylglucose accumulation. There are low but detectable levels of acylglucoses (87% of total acylsugars; fig. S2) in backcross inbred line BIL6521 (36), which contains *S. pennellii* LA0716 introgressions from chromosomes 1, 3, and 11. This BIL accumulates four acylglucoses (table S1), with the major one, G3:22 (5,5,12) (figs. S3 and S4), resembling the pyranose ring of the P-type acylsucrose S3:22 (5,5,12)-P detected at low levels in trichomes of the single chromosome 11 introgression line, IL11-3 (9). BIL6521 accumulates a P-type acylsucrose, S3:22 (figs. S3 and S4). These results are consistent with the hypothesis that the chromosome 3 region is necessary for acylglucose production but only when P-type acylsucroses are produced. Note that, in our nomenclature, “S” and “G” refer to a sucrose or glucose core, respectively, and 3:22 (5,5,12) indicates that there are three ester-linked acyl chains of 5, 5, and 12 carbons, for a total of 22 chain carbons (9). When nuclear magnetic resonance (NMR)–derived structural information is available, superscripts indicate acyl chain positions with R representing the pyranose ring and R’ representing the furanose ring (Fig. 1A).

Because the acylsugar levels in BIL6521, which lack a chromosome 4 introgression, were much lower than most other lines, we tested the impact of adding a chromosome 4 introgression carrying the SpASAT2 locus. A cross was made between BIL6521 and BIL6180, a recombinant line harboring introgressions on chromosomes 4, 5, and 11, which includes both the SpASAT2 and SpASAT3 loci (Fig. 1C). BIL6180 was previously found to produce only P-type acylsucroses as a result of the chromosome 4 and 11 introgressions (Fig. 1C); however, it accumulated significantly higher overall levels of acylsucroses compared to BIL6521 (fig. S2) and other short chain–containing P-type acylsucroses not present in BIL6521. If all P-type acylsucroses are substrates for an *S. pennellii* LA0716 factor on chromosome 3, we predict that both of the corresponding acylglucoses G3:15 (5,5,5) and G3:22 (5,5,12) would accumulate in a line harboring the chromosome 3, 4, and 11 introgressions. The F2 progeny of BIL6521 × BIL6180, genotyped as heterozygous for the *S. pennellii* chromosome 3 and 4 introgressions and homozygous for the *S. pennellii* chromosome 11 region, produced these two predicted acylglucoses (Fig. 1C and figs. S3 and S4). These findings, in combination with the published QTL results, indicate that the *S. pennellii* chromosome 3 introgression is necessary for acylglucose biosynthesis and suggests that P-type acylsucroses are acylglucose biosynthetic precursors.

**The chromosome 3 locus encodes a glandular trichome–expressed β-fructofuranosidase**

We sought candidate glycoside hydrolase (GH) genes in the 1.7-Mb QTL AG3.2, the acylglucose-associated region from *S. pennellii* LA0716 previously mapped to the bottom of chromosome 3 (Fig. 2A) (30). Three of the 238 genes in this region of the *S. lycopersicum* Heinz 1706 genome assembly SL2.50 annotation (37) are predicted as encoding GHs (members of the GH32, GH35, and GH47 families; table S2). We focused on the GH32 family Soyent03g040490 gene because all previously characterized members of the family have β-fructofuranosidase or fructosyltransferase activity (38). As acylsucroses and β-fructofuranosides, we hypothesized that the GH32 enzyme cleaves the glycosidic bond of P-type acylsucroses to generate acylglucoses. On the basis of the full results of this study, we designate this gene *ACYLSUCROSE FRUCTOFURANOSIDASE 1* (ASFF1).

*S. lycopersicum* acylsugars accumulate in type I/IV glandular trichome tip cells (39), and trichome tip cell–specific gene expression is a hallmark of all characterized acylsugar biosynthetic genes (e.g., ASAT1/2/3/4, IPMS3) (7, 9, 23, 24). We used a reporter gene approach to ask whether SpASFF1 exhibits trichome-specific expression. The 1.8-kb region immediately upstream of the SpASFF1 open reading frame (ORF) in the *S. pennellii* LA0716 genome drove expression of a green fluorescent protein–β-glucuronidase (GFP-GUS) fusion protein in *S. lycopersicum* M82 plants. A GFP signal in transformed plants was observed in the tip cells of type I/IV trichomes but not in the trichome stalk cells or underlying stem epidermis (Fig. 2B). This result is consistent with a role of SpASFF1 enzyme in type I/IV trichome metabolism.

We cross-validated the trichome-enriched expression pattern of ASFF1 in *S. pennellii* LA0716 using quantitative reverse transcription PCR (qRT-PCR). ASFF1 transcript levels were 3.7-fold higher in trichomes of *S. pennellii* LA0716 stems than in underlying shaved stem tissue (*P* < 0.001, Welch two-sample t test) (Fig. 2C). The observed enrichment of transcripts in trichome samples is similar to analysis of previously identified acylsugar biosynthetic genes from tomato, petunia, and tobacco (7, 12, 15, 23, 24). Together, transcript enrichment in trichomes and restriction of gene expression to trichome tip cells support the hypothesis that SpASFF1 acts in acylsugar biosynthesis.

**Acylglucoses accumulate in *S. pennellii* LA0716 but not in *S. lycopersicum* M82** (8, 19). However, ASFF1 is predicted to encode a full ORF in both the *S. lycopersicum* Heinz 1706 and the *S. pennellii* LA0716 genomes (37, 40). In contrast to enrichment of ASFF1 transcripts in trichomes of *S. pennellii* LA0716, ASFF1 transcripts are half as abundant in *S. lycopersicum* M82 trichomes as in underlying stem tissue (*P* < 0.05, Welch two-sample t test) (Fig. 2C). Together, the tissue- and species-level specificity of ASFF1 expression is consistent with a role for the gene in acylglucose biosynthesis.

**Gene editing reveals that SpASFF1 is necessary for *S. pennellii* LA0716 acylglucose accumulation**

We used CRISPR-Cas9–mediated gene editing in *S. pennellii* LA0716 to test whether SpASFF1 is necessary for acylglucose accumulation. Two small guide RNAs (sgRNAs) targeting the third SpASFF1 locus (Fig. 1C). However, three homoygous T1 mutants were obtained with different site-specific mutations, each of which is predicted to cause complete loss of function through translational frameshifts and premature protein termination. Two of them (spasff1-1 and spasff1-1-2), which carry 228–base pair (bp) and 276-bp indels, respectively, are derived from segregation of one heteroallelic T0 plant. The third mutant (spasff1-2) with a 1-bp insertion is the descendant of a homozygous T0 mutant.

Results from a liquid chromatography–mass spectrometry (LC-MS) analysis of leaf surface metabolites from these lines were consistent with the hypothesis that SpASFF1 is necessary for acylglucose biosynthesis. All spasff1 lines failed to accumulate detectable acylglucoses (Fig. 3A and fig. S5). Three homozygous T1 mutants were obtained with different site-specific mutations, each of which is predicted to cause complete loss of function through translational frameshifts and premature protein termination. Two of them (spasff1-1 and spasff1-1-2), which carry 228–base pair (bp) and 276-bp indels, respectively, are derived from segregation of one heteroallelic T0 plant. The third mutant (spasff1-2) with a 1-bp insertion is the descendant of a homozygous T0 mutant.

Results from a liquid chromatography–mass spectrometry (LC-MS) analysis of leaf surface metabolites from these lines were consistent with the hypothesis that SpASFF1 is necessary for acylglucose biosynthesis. All spasff1 lines failed to accumulate detectable acylglucoses (Fig. 3A and fig. S5). Three homozygous T1 mutants were obtained with different site-specific mutations, each of which is predicted to cause complete loss of function through translational frameshifts and premature protein termination. Two of them (spasff1-1 and spasff1-1-2), which carry 228–base pair (bp) and 276-bp indels, respectively, are derived from segregation of one heteroallelic T0 plant. The third mutant (spasff1-2) with a 1-bp insertion is the descendant of a homozygous T0 mutant.

**SpASFF1 converts pyranose ring–acylated P-type acylsucroses to acylglucoses both in vivo and in vitro**

The results described above strongly suggest that SpASFF1 converts pyranose ring–acylated P-type acylsucroses to acylglucoses. In addition, IL3-5 does not accumulate detectable acylglucoses despite having *S. pennellii* ASFF1, suggesting that F-type acylsucroses are not substrates for SpASFF1 (Fig. 4A and fig. S7). We took a transgenic
approach to ask whether SpASFF1 alone is sufficient to confer acylglucose accumulation in a P-type acylsucrose-accumulating background. We transformed the P-type acylsucrose-producing *S. lycopersicum* double-introgression BIL6180 (Fig. 4B) with a transferred DNA (T-DNA) containing the ORF of SpASFF1 and the 1.8 kb immediately upstream of its start codon (Fig. 4C). In addition to the P-type acylsucroses in the parental BIL6180, the SpASFF1 transgenics accumulated major hexose acylsugars with MS characteristics consistent with G3:15 (5,5,5) and G3:22 (5,5,12) (Fig. 4C and fig. S8). The acyl chain composition of these acylglucoses matches the S3:15 (5 R2,5 R3,5 R4) and S3:22 (5 R2,5 R4,12 R3) P-type acylsucroses detected in BIL6180 (fig. S4). Acylglucoses in the transgenic lines are also identical to those seen in BIL6521 × BIL6180 based on LC retention time and MS fragmentation. This confirms that SpASFF1 converts *S. pennellii* P-type acylsucroses produced by SpASAT2 and SpASAT3 to acylglucoses. Together, these in vivo results indicate that the presence of SpASFF1 is sufficient to yield acylglucoses in vivo when P-type acylsucroses are present, but not in plants accumulating only F-type acylsucroses.

In vitro assays supported the hypothesis that SpASFF1 accepts P-type but not F-type acylsucroses as substrates. Initial attempts to express SpASFF1 fusion proteins in *Escherichia coli* did not produce soluble protein. For this reason, recombinant His-tagged SpASFF1 was expressed using the *Nicotiana benthamiana* transient expression system (41). The enzyme was tested with both P-type and F-type acylsucrose substrates purified from *S. pennellii* asff1 and *S. lycopersicum* M82, respectively. Consistent with in vivo observations, SpASFF1 demonstrated hydrolytic activity with purified P-type S3:19 (4 R4,5 R2,10 R3) (42), yielding a compound with a mass/charge ratio \((m/z)\) consistent with a G3:19 (4,5,10) structure (Fig. 5A and fig. S9). In contrast, SpASFF1 demonstrated no hydrolytic activity with F-type S3:22 (5 R4,5 R3′,12 R3) (8) (Fig. 5B), suggesting that the presence of an acyl chain on the sucrose furanose ring prevents enzymatic hydrolysis. We further observed that SpASFF1 activity was undetectable with unmodified sucrose, while a commercially available yeast invertase hydrolyzed sucrose but not S3:19 (fig. S10). This SpASFF1 in vitro substrate specificity corroborates the in vivo results showing that acylglucoses only accumulate in lines containing P-type acylsucroses.

**DISCUSSION**

The results described above show that *S. pennellii* LA0716 synthesizes acylglucoses from P-type acylsucroses via the action of a previously uncharacterized trichome invertase (Fig. 5A), a homolog of
the most venerable enzyme in the history of biochemistry, yeast invertase. The canonical GH32 enzyme was first characterized in the 1840s through studies of “optical inversion” of cane sugar (sucrose) into a mixture of glucose and fructose. The enzyme was assayed two decades later, and its study by Maud Leonora Menten and Leonor Michaelis led to the theory of enzyme kinetics early in the 20th century (31). Since that time, other general metabolic activities have been identified for diverse GH32 β-fructofuranosidases, including plant glycan biosynthesis, cell wall modification, and hormone metabolism (43, 44).

Our results stand in contrast to the previously proposed direct synthesis of acylglucoses from UDP-glucose and free fatty acids.
Three GH genes were identified in the third acylglucose-linked Solanum pennellii to acylglucoses. Acylsucroses (Fig. 1A), we hypothesized acylsucrose and acylglucose biosynthesis (30). In a region not implicated in acylglucose accumulation in QTL mapping studies (10, in a region not implicated in acylglucose accumulation in QTL mapping studies (30). In contrast, QTLs linked to acylglucose accumulation in S. pennellii on chromosomes 4 and 11 include SpASAT2 and SpASAT3, suggesting a connection between acylsucrose and acylglucose biosynthesis (23, 30).

As the acylsucroses and acylglucoses in S. pennellii differ only by the presence or absence of a furanose ring (Fig. 1A), we hypothesized that a GH converts the S. pennellii acylsucroses (42) to acylglucoses. Three GH genes were identified in the third acylglucose-linked QTL on chromosome 3. These genes represent members of GH families 32, 35, and 47 (table S2). Most characterized plant GH35 enzymes act as β-galactosidases, while GH47 enzymes function as α-mannosidases in posttranslational protein modification (47, 48). Thus, these were not compelling candidates for cleavage of acylated sucrose substrates. Conversely, GH32 enzymes act on a variety of β-fructofuranoside in plants, including sucrose and fructans (38, 43). Our results indicate that SpASFF1 is a “derived” β-fructofuranosidase, with an active site that can accommodate pyranose, but not furanose-acylated sucrose esters. Understanding the structural features that allow SpASFF1 to hydrolyze P-type acylsucroses could inform engineering of novel specialized metabolites in plants and microbes.

We identified the GH32 SpASFF1 β-fructofuranosidase as being necessary and sufficient for conversion of P-type acylsucroses into acylglucoses. The most direct evidence is that ablation of the SpASFF1 gene using CRISPR-Cas9 gene editing led to acylsucrose-accumulating wild tomato S. pennellii LA0716 mutants with undetectable acylglucoses, showing that the enzyme is necessary for production of acylglucoses (Fig. 3). Multiple lines of genetic and biochemical evidence support the hypothesis that SpASFF1 uses P-type acylsucrose substrates. For example, no acylglucoses were detected in the F-type acylsucrose-producing introgression line IL3–5, despite the presence of SpASFF1 in the introgressed region (Fig. 4A and fig. S1). In contrast, transgenic trichome expression of the SpASFF1 invertase in the P-type acylsucrose-producing SpASAT2 and SpASAT3 double introgression line S. lycopersicum BIL6180 resulted in acylglucose accumulation (Fig. 4). Our in vitro assay results support the in vivo evidence that P-type acylsucroses are SpASFF1 substrates. In vitro assays with recombinant SpASFF1 demonstrated conversion of the purified P-type S3:19 (4R,5R,10R) to the cognate acylglucose G3:19 (4,5,10) (Fig. 5A). In contrast, the enzyme was inactive against F-type S3:22 (5R,4R,12R) (Fig. 5B) and did not hydrolyze uncylated sucrose (fig. S10).

Together, these data indicate that S. pennellii acylglucose metabolism results from evolution of a three-gene epistatic system, where the innovation of P-type acylsucrose synthesis by modification of the core BAHD acyltransferases potentiated evolution of SpASFF1 to produce acylglucoses. Our results reveal that a member of the GH32 β-fructofuranosidase enzyme family acquired expression in the trichome glandular tip cell (Fig. 2) and the ability to cleave acylated sucrose (Fig. 5), which led to an increase in the diversity of Solanaceae trichome specialized metabolites. This is a remarkable evolutionary innovation, where a member of an enzyme family long recognized as important in general metabolism was co-opted into specialized metabolism by the “blind watchmaker” of evolution.

Acylsugar accumulation is widespread throughout the Solanaceae with occurrences in genera as distantly related as Salpiglossis and Solanum, sharing a last common ancestor >30 million years (Ma) ago (8, 10, 49). While acylsugars show wide structural variation in the number and length of acyl chains throughout the family, sucrose is the most prominent sugar core. Acylsucroses not only accumulate in genera whose lineages diverged <20 Ma ago, such as Solanum and Physalis (8, 21), but also accumulate in species representing earlier diverging lineages, including Salpiglossis and Petunia (10, 12). In addition, acyl chains are present on the furanose ring in at least some members of each of these genera, suggesting that accumulation of F-type acylsucroses evolved long ago.

Although apparently limited in distribution relative to acylsucroses, acylglucoses occur in diverse genera including Solanum, Datura,
and Nicotiana (16, 19, 20). Despite the fact that acylglucose accumulation is common to species in both Solanum and Nicotiana, which diverged approximately 24 Ma ago, the differences in SpASFF1 gene expression and SpASAT substrate specificity that facilitated acylglucose accumulation in S. pennellii arose in the ~7 Ma since divergence from the last ancestor in common with S. lycopersicum (11, 49, 50). This supports independent evolutionary origins of acylglucoses in distinct lineages. In the Solanum genus, P-type acylsucroses are a prerequisite for acylglucose accumulation. The predominance of F-type acylsucroses within the Solanaceae may explain the relative rarity of acylglucoses in the family. However, characterization of the ASAT enzymes responsible for acylsucrose biosynthesis in Soliglossis, Petunia, and Solanum demonstrates multiple changes in enzyme substrate specificity throughout the evolutionary history of the acylsucrose pathway (10–12). Plasticity of the acylsugar pathway may have given rise to P-type acylsucroses multiple times throughout evolutionary history. If so, this would provide independent opportunities for co-option of acylglucose-producing GHs into the acylsugar pathway. Are the enzymes responsible for hydrolyzing acylsucroses to yield acylglucoses restricted to the GH32 family or have other enzyme families evolved in different acylglucose-accumulating lineages? Whether and to what extent multiple origins of acylglucose biosynthesis share common features remains to be explored.

Over the past decade, discovery of pathways and enzymes of plant specialized metabolism has improved at an increasing rate. Before this, the taxonomic restriction of specialized metabolism biased
Fig. 5. SpASFF1 cleaves a P-type S3:19 acylsucrose but not an F-type S3:22 acylsucrose. (A) LC-MS analysis of in vitro enzyme assay products indicates that SpASFF1 hydrolyzed P-type S3:19 (4R,5R,10R) acylsucrose, yielding two compounds with m/z 533.3. This m/z is consistent with an acylglucose product with a G3:19 (4,5,10) configuration; the two peaks represent the α and β anomers of the acylglucose. (B) LC-MS analysis of in vitro assays with F-type S3:22 (5R,5R,12R) acylsucrose indicates no hydrolysis products with SpASFF1. Acylglucose structure is inferred from collision-induced dissociation-mediated fragmentation (fig. S9). All ESI+ mode acylsugars were identified as formate adducts.

Deep analysis toward pathways found in model organisms: for instance, glucosinolates in Arabidopsis, cyclic hydroxamic acids in maize and other well-studied grasses, and isoflavonoids in Medicago and soybean. Marked improvements in sensitivity and selectivity of MS- and NMR-based analytical chemistry helped broaden the scope of well-studied metabolic networks (S1, S2). In parallel, development of species-agnostic DNA sequencing and functional genomics screening tools (notably, virus-induced gene silencing and genetic transformation) permitted rigorous correlation of in vitro activities and in vivo phenotypes. The rapid advancement of gene editing techniques using CRISPR-Cas on agriculturally important and undomesticated species markedly expands the specialized metabolism functional genomics toolkit. These methods allow not only direct tests of in vivo function but also elimination of the T-DNA by simple genetic crossing. For example, the spsff1 mutant lines can be used to understand the impacts of acylsucrases versus acylglucoses on the fitness of S. pennellii both in the greenhouse and in the field. These studies could lead to crops with engineered natural pesticides, broaden our understanding of the roles of specialized metabolites in mediating environmental interactions, and inform our understanding of the mechanisms underpinning specialized metabolic evolution.

Materials and methods

Plant material

Seeds of S. lycopersicum M82 were obtained from the C.M. Rick Tomato Genetics Resource Center (TGRC; University of California, Davis, CA); seeds of IL3-5, BIL6180, and BIL6521 were obtained from D. Zamir (Hebrew University of Jerusalem, Rehovot, Israel) (36); seeds of S. pennellii LA0716 were provided by M. Mutschler (Cornell University, Ithaca, NY). Seeds were treated with half-strength bleach for 30 min and rinsed three times in deionized water for 5 min before sowing on moist filter paper in petri dishes. Upon germination, seedlings were transferred to soil. Young plants were grown in 9-cm pots in a peat-based propagation mix (Sun Gro Horticulture, Agawam, MA). S. lycopersicum and introgression lines were watered four times weekly with deionized water and supplemented once weekly with half-strength Hoagland’s solution; S. pennellii was watered once weekly with deionized water and supplemented once weekly with half-strength Hoagland’s solution. Plants used for analysis were grown in a growth chamber under a 16-hour photoperiod [190 μmol m⁻² s⁻¹ photosynthetic photon flux density (PPFD)] with 28°C day and 22°C night temperatures set to 50% relative humidity. BIL lines used for crosses were grown in a soil mix consisting of four parts SUREMIX (Michigan Grower Products Inc., Galesburg, MI) to one part sand in a greenhouse with a daytime maximum temperature of 30°C and a nighttime minimum temperature of 16°C; sunlight was supplemented with high-pressure sodium bulbs on a 16-hour light/8-hour dark cycle. For seed production, S. pennellii asf1 T₀ plants were grown in soil containing one part Canadian sphagnum (Mossor Lee Co., Millston, WI), one part coarse sand (Quikrete, Atlanta, GA), one part white pumice (Everwood Farm, Brooks, OR), and one part redwood bark (Sequioa Bark Sales, Reedley, CA) supplemented with 1.8 kg of crushed oyster shells (Down to Earth Distributors Inc., Eugene, OR), 1.8 kg of hydrated lime (Bonide Products Inc., Oriskany, NY), and 0.6 kg of triple super phosphate (T & N Inc., Foristell, MO) per cubic meter.

Acylsugar analysis

The acylsugar extraction interactive protocol is available in Protocols.io at http://dx.doi.org/10.17504/protocols.io.xj2fkqe. Leaf surface acylsugars were extracted from single leaflets with 1 ml of a mixture of isopropanol (J.T.Baker, Phillipsburg, NJ)/acetonitrile (Sigma-Aldrich, St. Louis, MO)/water (3:3:2) with 0.1% formic acid and 1 μM telmisartan (Sigma-Aldrich, St. Louis, MO) as a high-performance liquid chromatography (HPLC) standard. The leaf tissue was gently agitated on a rocker in this extraction solvent for 2 min. The extraction solvent was collected and stored in 2-mL LC-MS vials at −80°C.

LC-MS samples (both enzyme assays and plant samples) were run on a Waters ACQUITY UPLC coupled to a Waters Xevo G2-XS QToF mass spectrometer. Ten microliters of the acylsugar extracts were injected into an Ascentis Express C18 HPLC column (10 cm by 2.1 mm, 2.7 μm) (Sigma-Aldrich, St. Louis, MO), which was maintained at 40°C. The LC-MS methods used the following solvents: 10 mM ammonium formate (pH 2.8) as solvent A and 100% acetonitrile as solvent B. Compounds were eluted using one of two gradients.

A 7-min linear elution gradient consisted of 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, and held at 5% B until 7 min. A 21-min linear elution gradient consisted of 5% B at 0 min, 60% B at 3 min, 100% B at 15 min, held at 100% B until 18 min, 5% B at 18.01 min, and held at 5% B until 21 min.
The MS settings were as follows for negative ion-mode electrospray ionization (ESI–): capillary voltage, 2.00 kV; source temperature, 100°C; desolvation temperature, 350°C; desolvation nitrogen gas flow rate, 600 liters/hour; cone voltage, 35 V; and mass range, m/z 50 to 1000 (with spectra accumulated at 0.1 s per function). Three separate acquisition functions were set up to test different collision energies (0, 15, and 35 V).

The MS settings for positive ion-mode ESI were as follows: capillary voltage, 3.00 kV; source temperature, 100°C; desolvation temperature, 350°C; desolvation nitrogen gas flow rate, 600 liters/hour; cone voltage, 35 V; and mass range, m/z 50 to 1000 (with spectra accumulated at 0.1 s per function). Three separate acquisition functions were set up to test different collision energies (0, 15, and 45 V). Lock mass correction was performed using leucine enkephalin as the reference compound for data acquired in both negative and positive ion mode.

Acylsugar quantification
To accurately quantify total acylsugars, samples were saponified before LC-MS analysis and sugar cores were quantified with authentic isotopically labeled standards. A leaflet was immersed in 2 ml of dichloromethane (VWR International, Radnor, PA) and 500 µl of water with 30-s vortexing. After phase separation, 1 ml of the dichloromethane layer was removed to a borosilicate glass vial and evaporated to dryness under flowing air. Dried samples were dissolved in 1 ml of acetonitrile and the reaction was incubated in a sealed 1.5-ml microcentrifuge tube (Sigma-Aldrich, St. Louis, MO). Glucose and sucrose standard solutions were processed using the same protocol for acylsugars, as described above.

qPCR analysis
Tissue of 10-week-old S. pennellii LA0716 and S. lycopersicum M82 were harvested as follows: stems were flash-frozen in liquid nitrogen, and trichomes were shaved into 1.5-ml microcentrifuge tubes with a clean razor blade. Trichomes and denuded stems were kept in liquid nitrogen and ground with plastic micropestles in 1.5-ml microcentrifuge tubes. RNA was extracted from ground trichomes and stems (six biological replicates for each species and tissue type) using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. For each sample, 250 ng of RNA was quantified using a NanoDrop 2000c (Thermo Fisher Scientific, Waltham, MA) to synthesize cDNA using SuperScript III Reverse Transcriptase (Invitrogen, Carlsbad, CA). qRT-PCR was carried out using SYBR Green PCR Master Mix on a QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems, Warrington, UK) using the following cycling conditions: 48°C for 30 min, 95°C for 10 min, 40 cycles of 95°C for 15 s and 60°C for 1 min, followed by melt curve analysis. RT_ASF1_F and RT_ASF1_R primers were used to detect the 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-1a, actin, and ubiquitin genes, respectively (table S3).
biological replicate, relative levels of the ASFF1 transcript were determined using the ΔΔCt method (53) and normalized to the geometric mean of EF-1α, actin, and ubiquitin transcript levels.

**Genotyping of progeny of BIL6521 × BIL6180**

DNA from the progeny of the cross between BIL6521 and BIL6180 was extracted from leaf material that were archived on FTA (Flinders Technology Associates) PlantSaver cards (GE Healthcare, Uppsala, Sweden) and purified according to the manufacturer’s specifications. Extracted DNA from the FTA cards were used for PCR amplification with GoTaq Green Mastermix to genotype the sample using 04g011460_Marker_Indel-F/R and ASFF_Chr3_Indel_002_F/R (table S3).

**DNA construct assembly**

All Sanger DNA sequencing confirmation in this study was performed with the indicated sequencing primers at the Research Technology Support Facility Genomics Core, Michigan State University, East Lansing, MI.

For proSpASFF1:SpASFF1 ORF (pK7WG), a 1.8-kb region of the upstream region and ORF of SpASFF1 was split into four amplicons using four sets of primers: ASFF_001_F/R, ASFF_002_F/R, ASFF_003_F/R, and ASFF_004_F/R (table S3). The first and fourth amplicon contained adapters for assembly into pENTR-D-TOPO that has been digested with NotI/AscI, respectively. The construct was assembled using New England Biolabs (NEB) Gibson Assembly according to manufacturer specifications (NEB, Ipswich, MA). The construct was verified by Sanger sequencing using M13 Forward, T7 promoter primers, and cloning primers. The insert was sub-cloned into pK7WG (54) using the LR Clonase II Enzyme mix (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer’s instructions. Presence of the insert was determined by colony PCR using ASFF_001F/R. Completed vectors were transformed into Agrobacterium strain AGL0. Leaf material from recovered plants were archived on FTA PlantSaver cards (GE Healthcare, Uppsala, Sweden) and genotyped by PCR amplification with GoTaq Green Mastermix and pK7WG-Kan-F/R primers (table S3).

For proSpASFF1-GFP/GUS (pKGWFS7), a 1.8-kb region of the upstream region of ASFF1 was amplified from S. pennellii LA0716 genomic DNA using ASFF_promoter_F1/R1 primers (table S3). pENTR-D-TOPO was digested with NotI/AscI to linearize the vector and create overhangs compatible for Gibson Assembly. The amplicon also contained adapters for insertion into pENTR-D-TOPO digested with NotI/AscI. Constructs were Sanger-sequenced using M13F/R primers in addition to the ASFF_promoter_F1/R1 primers. LR Clonase II Enzyme mix was used to subclone the fragment into pKGWFS7 (54). The construct was transformed into AGL0 for plant transformation using the described protocol.

The CRISPR-ASFF1 vector was constructed as follows. CRISPR sgRNAs were designed using the site finder toolset in Geneious v10 (www.geneious.com). Two target sequences located on the exon were selected for their high on-target activity scores, based on a published algorithm (55), and low off-target scores against published S. pennellii genome database (56). Each sgRNA was obtained as a gBlock synthesized in vitro by IDT (Integrated DNA technologies) (www.idtdna.com) (table S3) and subsequently assembled with pICH47742:2x3SS-5'UTR-hCas9(STOP)-NOST [Addgene plasmid no. 49771, provided by S. Kamoun (Sainsbury Lab, Norwich, UK)] (57), pICH41780 (Addgene plasmid no. 48019) and pAGM4723 (Addgene plasmid no. 48015, both gifts from S. Marillonnet) (58), and pICSL11024 [Addgene plasmid no. 51144, a gift from J. D. Jones (Sainsbury Lab, Norwich, UK)] using Golden Gate Assembly. In short, the restriction-ligation reactions (20 µl) were set up by mixing 15 ng of synthesized sgRNAs with 1.5 µl of T4 ligase buffer (NEB), 320 U of T4 DNA ligase (NEB), 1.5 µl of bovine serum albumin (BSA; 0.1 mg/ml; NEB), 8 U of BpiI (Thermo Fisher Scientific), and 100 to 200 ng of the intact plasmids. The reactions were incubated at 37°C for 30 s, followed by 26 cycles (37°C, 3 min; 16°C, 4 min), and then incubated at 50°C for 5 min and 65°C for 5 min. The ligated products were directly used to transform E. coli competent cells. Positive clones were chosen based on colony PCR and sequenced at the Michigan State University Research Technology Support Facility (MSU RTSF) using the pAGM4723_SeqF1, pAGM4723_SeqR1, pICSL11024_SeqF1, pICH47742CAS9_SeqF2, pICH47742_SeqR1, pICLH41780_SeqR1, and ASFF_SeqR primers (table S3). The construct was transformed into S. pennellii LA0716 using the plant transformation protocol described below. Leaf material from recovered plants were archived on FTA PlantSaver cards (GE Healthcare, Uppsala, Sweden) and genotyped by PCR amplification with ASFF_F/R, followed by Sanger sequencing with ASFF_SeqR (table S3).

For spasff1 line transcript analysis, RNA was extracted from spasff1-1/1/1-1-2 lines using RNeasy Plant Mini Kit according to the kit specifications (Qiagen, Venlo, The Netherlands). RNA was quantified using a NanoDrop 2000c (Thermo Fisher Scientific, Waltham, MA). One microgram of RNA was used for cDNA synthesis using Superscript II Reverse Transcriptase according to the manufacturer’s specifications. The ASFF1_transcript_amb_01F/R primers (final concentration, 0.5 µM) were used to amplify the region within the ASFF1 CDS (coding sequence), which was cloned into pMINI-T 2.0 (NEB, Ipswich, MA). T7 and SP6 promoter primers were used for Sanger sequence confirmation of the inserts, and ClustalW was used for alignment of the transcripts (www.ebi.ac.uk/Tools/msa/clustalo/).

**Competent cell preparation and transformation of constructs into Agrobacterium**

A single colony of AGL0 or LBA4404 Agrobacterium was inoculated into two 5-ml cultures of YEP media [10 g of yeast extract, 10 g of Bacto peptone, and 5 g of NaCl per liter (pH 7)] with rifampicin (50 µg/ml). Cultures were incubated in borosilicate glass test tubes (18 mm by 150 mm) with foam plugs overnight at 30°C, shaken at 200 rpm. One hundred ninety milliliters of LB was inoculated with 10 ml of the overnight cultures in an autoclaved 500-ml Erlenmeyer flask. Cultures were grown in a shaking incubator (200 rpm) at 30°C to OD_600 (optical density at 600 nm) = 1.0, incubated in ice for 10 min, and centrifuged at 4°C in 50-ml conical tubes at 3200g for 5 min. Pellets were resuspended in 1 ml of sterile 20 mM CaCl_2 and 100 -µl aliquots; dispensed into sterile, prechilled 1.7-ml microcentrifuge tubes; snap-frozen using liquid nitrogen; and stored at −80°C.

For Agrobacterium transformation, 1 µg of construct DNA purified using an Omega EZNA Plasmid DNA Mini Kit 1 (Omega Bio-Tek, Norcross, GA) was added to the frozen Agrobacterium aliquots on ice. Cells were thawed in a 37°C water bath for 5 min, mixed well by flicking, and snap-frozen in liquid nitrogen. Cells were thawed, and 1 ml of YEP was added to the tube. The transformations were incubated at 28°C and shaken at 200 rpm for 4 hours. Cells were centrifuged at 17,000g for 30 s, the supernatant was decanted, and the cell pellet was resuspended in 100 µl of fresh YEP. The cell pellet was resuspended, and the entire suspension was plated onto an LB plate with spectinomycin (100 µg/ml).
Presence of the insert containing vector was verified by colony PCR. Colonies were collected with a pipette tip and resuspended in 20 µl of sterile water. Two microliters of the cell suspension was added to a PCR tube with a reaction with a final volume of 25 µl. GoTaq Green Mastermix (2×) was used for the colony PCR according to the manufacturer’s specifications (Promega, Madison, WI). Primers (final concentration, 0.4 µM) pertaining to the insert were used for amplification.

**Plant transformation**

In all cases, petri plates containing plant tissue were sealed with a single layer of micropore paper tape (3M, Maplewood, MN). Transformation of *S. lycopersicum* and *S. pennellii* LA0716 was performed using AGL0 using a modification of published protocols (59, 60). Fifty to sixty seeds were incubated in 40% bleach and agitated for 5 min. Seeds were rinsed six times, each with 40 ml of sterile double-distilled H2O with 5 min of rocking and decanting of wash solution. A flame-sterilized spatula was used to distribute the seeds onto the surface of 1/2× MSO medium (Murashige and Skoog media with no sucrose) (59) in a Phytatray II (Sigma-Aldrich, St. Louis, MO). Containers were incubated at 25°C on a 16-hour light/8-hour dark cycle with a light intensity of 70 µmol m⁻² s⁻¹ PPFD.

At day 8 for *S. lycopersicum* or day 11 for *S. pennellii* LA0716, the seedlings were removed from the 1/2× MSO medium jar. The hypocotyl and radicle were excised and discarded. The cotyledon explant was placed on a sterile petri dish. One to two millimeters was placed on the surface of a sterile D1 media plate (59) on which was placed on a sterile petri dish and, using a scalpel, calluses were cut and then placed onto fresh 2Z media plates with the same selection. Subsequently, explants were transferred to new 2Z plates every 2 weeks. Throughout the process, dying tissue was removed, and growing tissue was placed on the media. Five to eight weeks after cocultivation, shoots were harvested from the explants and placed into Phytatray II (Sigma-Aldrich, St. Louis, MO) containing 100 ml of Murashige and Skoog salts, 3% sucrose, and Nitsch Vitamins, 0.8% bactoagar, pH 6.0 (MSSV) media (Murashige and Skoog salts, 3% sucrose, and Nitsch Vitamins, 0.8% bactoagar, pH 6.0) (60) supplemented with timentin (100 µg/ml), kanamycin (50 µg/ml), and indole-3 butyric acid (1 µg/ml). MSSV-containing Phytatray were incubated under the same growth conditions (16-hour light/8-hour dark cycle at 70 µmol m⁻² s⁻¹ PPFD). Shoots were monitored for leaf and root production, and shoots with roots and leaves were placed into pots containing Redi-Earth soil. Flats were covered with a plastic dome under the same growth conditions. Domes were removed from flats after 3 to 4 days.

**Transient expression and purification of SpASFF1 protein**

The ASFF1 CDS was amplified from *S. pennellii* LA0716 trichome cDNA using ASEFF_F and ASEFF_R primers (table S3) and cloned into the pGEM backbone using the pGEM-T Easy cloning kit (Promega, Madison, WI). The ASFF1 CDS was subsequently reamplified with the (pEQ-H-T)-ASFF-His_F and (pEQ-H-T)-ASFF-His_R primers (table S3) to add adapters for Gibson Assembly. The resulting PCR product was transferred to pEQ-H-T vector (41) previously digested with NruI–HF and SmaI restriction enzymes (NEB, Ipswich, MA) using 2× Gibson Assembly master mix (NEB, Ipswich, MA) according to the manufacturer’s instructions to create an expression clone coding for the full-length protein with a C-terminal 6x His tag (ASFF1-HT-pEAQ). The completed vector was subsequently transformed into LBA4404 cells, as described above. For transient expression, an Agrobacterium tumefaciens LBA4404 strain carrying the ASFF1-HT-pEAQ construct was streaked onto LB agar containing rifampicin (50 µg/ml) and kanamycin (50 µg/ml) and incubated for 3 days at 28°C. Single colonies were used to inoculate 250-ml Erlenmeyer flasks containing 50 ml of YEP medium with rifampicin (50 µg/ml) and kanamycin (50 µg/ml); cultures were incubated at 28°C and shaken at 300 rpm overnight. Cultures were harvested by centrifugation at 800g at 20°C for 20 min. The supernatant was discarded, and the resulting loose pellet was resuspended in 50 ml of buffer A [10 mM 2-ethanesulfonic acid (MES; Sigma-Aldrich, St. Louis, MO) at pH 5.6 and 10 mM MgCl2]. This cell suspension was centrifuged at 800g at 20°C for 20 min, and the resulting pellet was resuspended to a final OD600 = 1.0 with buffer A. A 200 mM solution of acetylsyringone (Sigma-Aldrich, St. Louis, MO) dissolved in DMSO was added to the suspension at a final concentration of 200 µM, and the suspension was incubated at room temperature with gentle rocking for 4 hours. This suspension was infiltrated into fully expanded leaves of 6-week-old *N. benthamiana* plants using a needleless 1-ml tuberculin syringe. Plants were grown under a 16-hour photoperiod (70 µmol m⁻² s⁻¹ PPFD) and at a constant 22°C set to 70% relative humidity. At 8 days after infiltration, 28 g of infiltrated leaves was harvested, deveined, and flash-frozen in liquid nitrogen. Tissue was powdered under liquid nitrogen with a mortar and pestle and added to 140 ml of ice-cold buffer B [25 mM 3-[1-(2-hydroxyethyl)piperazin-1-yl]propane-1-sulfonic acid (EPPS) at pH 8.0, 1.5 M NaCl, 1 mM EDTA with 2 mM dithiothreitol (DTT), 1 mM benzamide, 0.1 mM phenylmethylsulfonyl fluoride, and 1 mM PMSF].
10 μM trans-epoxysuccinyl-1-leucylamido(4-guanidinobutane (E-64), and 5% (w/v) polyvinylpolypyrrolidone (PVPP); all reagents were obtained from Sigma-Aldrich (St. Louis, MO) except DTT, which was obtained from Roche Diagnostics (Risch-Rotkreuz, Switzerland). The mixture was stirred for 4 hours at 4°C, filtered through six layers of Miracloth, and centrifuged at 27,000g at 4°C for 30 min. The supernatant was decanted and passed through a 0.22-μm polyethersulfone filter (EMD Millipore, Billerica, MA) before being loaded onto a HisTrap HP 1-ml affinity column and eluted using a gradient of 10 to 500 mM imidazole in buffer B using an ÄKTA start FPLC module (GE Healthcare, Uppsala, Sweden). Fractions were analyzed by SDS–polyacrylamide gel electrophoresis, and the presence of ASFF1-HT was confirmed by immunoblot using the BG–His-1 monoclonal antibody (Roche, Mannheim, Germany) to detect His-tagged proteins. Purified ASFF1-HT was subsequently transferred to 100 mM sodium acetate (pH 4.5) with 50% glycerol using a 10DG desalting column (Bio-Rad, Hercules, CA). Protein was quantified against a standard curve of BSA (Thermo Fisher Scientific, Waltham, MA) using a modified Bradford reagent (Bio-Rad, Hercules, CA) according to the manufacturer’s instructions.

Enzyme assays

For activity assays, 100 ng of ASFF1-HT or 1 μg of Saccharomyces cerevisiae invertase (catalog no. 14504, Grade VII, Sigma-Aldrich, St. Louis, MO) and 0.1 nmol F- or P-type acylscurose or 10 nmol sucrose (Sigma-Aldrich, St. Louis, MO) were added to 30 μl of 50 mM sodium acetate (pH 4.5) in 250-μl thin-wall PCR tubes. Reactions were incubated for 1 hour at 30°C and stopped by the addition of 60 μl of 1:1 acetonitrile/isopropanol containing 1.5 μM tleimisartan as internal standard and centrifuged for 10 min at 16,000g to remove precipitated protein. The supernatant was transferred to 2-ml autosampler vials with 250-μl glass inserts and analyzed by LC-MS as described above.

Statistical analysis

All statistical analyses were performed using the “stats” R package (R Core Team, 2017). One-way analysis of variance (ANOVA) was executed on acylscugar data using the “aov” command. Between- and within-group variances were determined using the sum-of-squares values obtained from ANOVA; these values were subsequently used to determine the power of the ANOVA using the “power.anova.test” function. Analysis by Tukey’s post hoc mean-separation test was executed using the “TukeyHSD” command, with the results of one-way ANOVA as input. Welch two-sample t tests were executed on transcript abundance data using the “t.test” command. The power of these analyses was determined using the “power.t.test” function.
27. J. P. Kuai, C. S. Ghangas, J. C. Steffens, Regulation of triacylglycerol fatty acid composition in wild tomato. *Plant Physiol.* **169**, 1821–1835 (2015).

26. J. A. Shapiro, J. C. Steffens, M. A. Mutschler, Acylsugars of the wild tomato *Lycopersicon pennelli* in relation to geographic distribution of the species. *Biochem. Syst. Ecol.* **22**, 545–561 (1994).

25. J. P. Kuai, C. S. Ghangas, J. C. Steffens, Regulation of triacylglycerol fatty acid composition (uridine diphosphate glucose: Fatty acid acyltransferase with overlapping chain-length specificity). *Plant Physiol.* **115**, 1581–1587 (1997).

24. J. Ning, G. D. Moghe, B. Leong, J. Kim, I. Ofner, Z. Wang, C. Adams, A. D. Jones, D. Zamir, R. L. Last, A feedback-insensitive isopropylmalate synthase affects acylsugar composition in cultivated and wild tomato. *Plant Physiol.* **169**, 1821–1835 (2015).

23. J. F. Obreja, J. B. Mudd, M. P. F. Marsden, Epicuticular lipid accumulation on the leaves of *Lycopersicon pennelli* (Corr.) D'arcy and *Lycopersicon esculentum* Mill. *Plant Physiol.* **77**, 567–570 (1985).

22. J. A. Shapiro, J. C. Steffens, M. A. Mutschler, Acylsugars of the wild tomato *Lycopersicon pennelli* and metabolic network. *Plant Physiol.* **169**, 1821–1835 (2015).

21. J. P. Kuai, C. S. Ghangas, J. C. Steffens, Regulation of triacylglycerol fatty acid composition in cultivated and wild tomato. *Plant Physiol.* **169**, 1821–1835 (2015).

20. J. A. Shapiro, J. C. Steffens, M. A. Mutschler, Acylsugars of the wild tomato *Lycopersicon pennelli* in relation to geographic distribution of the species. *Biochem. Syst. Ecol.* **22**, 545–561 (1994).

19. J. P. Kuai, C. S. Ghangas, J. C. Steffens, Regulation of triacylglycerol fatty acid composition (uridine diphosphate glucose: Fatty acid acyltransferase with overlapping chain-length specificity). *Plant Physiol.* **115**, 1581–1587 (1997).

18. A. X. Li, N. Eannetta, G. S. Ghangas, J. C. Steffens, Glucose polyester biosynthesis. Purification and characterization of a glucose acyltransferase. *Plant Physiol.* **121**, 453–460 (1999).

17. A. X. Li, J. C. Steffens, An acyltransferase catalyzing the formation of diacylglycerol is a serine carboxypeptidase-like protein. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6902–6907 (2000).

16. B. M. Leckie, D. M. De Jong, M. A. Mutschler, Quantitative trait loci regulating sugar moieties of acylsugars in tomato. *Mol. Breed.* **31**, 957–970 (2013).

15. L. Michaelis, M. L. Menten, K. A. Johnson, R. S. Goody, The original Michaelis constant: Translation of the 1913 Michaelis-Menten paper. *Biochemistry* **50**, 8264–8269 (2011).

14. M. A. Sainz-Polo, M. Ramírez-Escudero, A. Lafraya, B. González, J. Marín-Navarro, J. Polaina, J. Sanz-Aparicio, Three-dimensional structure of Saccharomyces inverte: Role of a non-catalytic domain in oligomerization and substrate specificity. *J. Biol. Chem.* **288**, 9755–9766 (2013).

13. H. Wan, L. Wu, Y. Yang, G. Zhou, Y.-L. Ruan, Evolution of sucrose metabolism: The dichotomy of invertases and beyond. *Trends Plant Sci.* **23**, 163–177 (2018).

12. J. R. Smeda, A. L. Schilmiller, T. Anderson, S. Ben-Mahmoud, D. E. Ullman, T. M. Chappell, A. Kessler, M. A. Mutschler, Combination of acylglucose QTL reveals additive and epistatic genetic interactions and impacts insect oviposition and virus infection. *Mol. Breed.* **38**, 3 (2018).

11. Y. Eshed, D. Zamir, An introgression line population of *Lycopersicon pennelli* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* **141**, 1147–1162 (1995).

10. I. Ofner, J. Lashbrooke, T. Pleban, A. Aharoni, D. Zamir, Solanum pennelli backcross inbred lines (BILs) link small genomic bins with tomato traits. *Plant J.* **87**, 151–160 (2016).

9. Tomato Genome Consortium, The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485**, 635–641 (2012).

8. W. Van den Ende, W. Lammens, A. Van Laere, L. Schroeven, K. Le Roy, Donor and acceptor substrate specificity among plant glycoside hydrolase family 32 enzymes. *FEBS Lett.* **276**, 5788–5790 (2009).

7. T. Nakashima, H. Wada, S. Morita, R. Erra-Balsells, K. Hiraoka, N. Nonami, Single-cell metabolite profiling of stalk and glandular cells of intact trichomes with internal electrode capillary probe pressure electrophoresis ionization mass spectrometry. *Anal. Chem.* **88**, 3049–3057 (2016).

6. A. Bolger, F. Scosza, M. E. Bolger, C. Lanz, F. Maumus, T. Tohge, H. Quesneville, S. Alseek, I. Särensen, G. Lichtenstein, E. A. Fich, M. Conte, H. Keller, K. Schneeburger, R. Schwacke, I. Ofner, J. Vrebalov, Y. Xu, S. Osorio, S. A. Altlos, E. Schijlen, J. M. Jiménez-Gómez, M. Ryngajllo, C. Wilen, R. Orchard, H. W. Virgin, J. Listgarten, D. E. Root, Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).

5. P. H. Dui, A. Scott, J. A. Weinstein, F. A. Ran, S. Konermann, V. Agarwala, Y. Li, E. J. Fine, X. Wu, O. Shalem, T. J. Craddock, L. A. Marraffini, G. Bao, F. Zhang, DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).

4. K. Belah, A. Chaparro-Garcia, S. Kamoun, V. Nekrasov, Plant genome editing made easy: Targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* **9**, 3 (2013).

3. E. Weber, C. Engler, R. Gruetzer, S. Werner, S. Marillonnet, A modular cloning system for standardized assembly of multigene constructs. *PLOS ONE* **6**, e16765 (2011).

2. S. McCormick, in *Plant Tissue Culture Manual*, K. Lindsey, Ed. (Springer Netherlands, 1997), vol. 13, pp. 311–319.

1. J. J. Fillatti, J. Kiser, R. Rose, L. Comai, Effortive transfer of a glycoside tolerance gene into tomato using a binary Agrobacterium tumefaciens vector. *Nat. Biotechnol.* **11**, 726–730 (1987).

---

**Acknowledgments:** We thank M. Mutschler, D. Zamir, and the C. M. Rick TGRG (University of California, Davis, CA) for providing germplasm essential to this study. Plant transformation was performed by K. Imre. We are also grateful to R. Chetletat and X. Qin for advice about S. pennelli transformation and D. Jones, L. Chen, and the MSU RTSF Mass Spectrometry and Metabolomics facility for analytical chemistry advice. We thank C. Schenck for editorial suggestions and E. Pichersky and members of the MSU Solanaceae Specialized Metabolism Evolution group for advice. Funding: This work was funded by National Science Foundation grant IOS-PPGPB-1546617 (to R.L.L.) and National Institute of General Sciences of the National Institutes of Health graduate training grant no. T32-GM110523 (to B.J.L. and D.B.L.). Author contributions: Design: B.J.L., D.B.L., Y.-R.L., P.F., A.L.S., and R.L.L. Experiments and data analysis: B.J.L., D.B.L., Y.-R.L., P.F., A.L.S., and R.L.L. Writing: B.J.L., D.B.L., Y.-R.L., P.F., A.L.S., and R.L.L. Competing interests: The authors declare that they have no competing interests. Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors. Sequence data for *SpASFF1* are available in GenBank: *SpASFF1* (MK297330). The following materials require a material transfer agreement: *peAQ-HT*, pK7WG, pKGFW57, pICH47742:2x35S-5′UTR-hCaS9(STOP)-NOST, pCIH41780, pAGM4723, and pCLSI11024. Requests for biological materials or data should be submitted to R.L.L. at last@msu.edu.

Submitted 20 December 2018
Accepted 6 March 2019
Published 24 April 2019

10.1126/sciadv.aaw3754

**Citation:** B. J. Leong, D. B. Lybrand, Y.-R. Lou, P. Fan, A. L. Schilmiller, R. L. Last, Evolution of metabolic novelty: A trichome-expressed invertase creates specialized metabolic diversity in wild tomato. *Sci. Adv.* **5**, eaaaw3754 (2019).