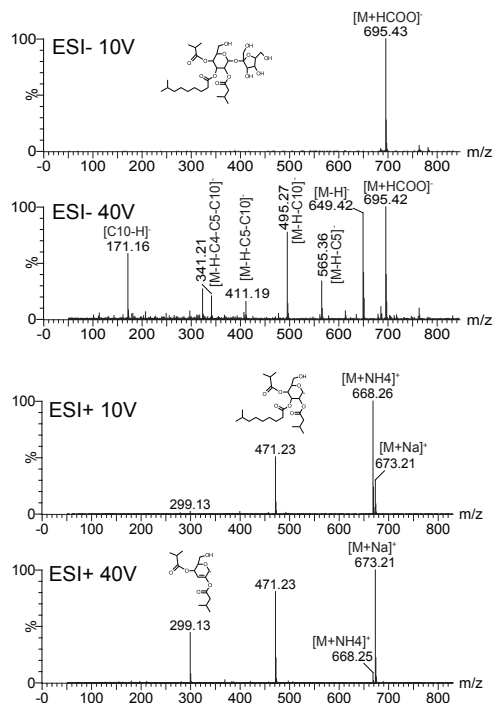
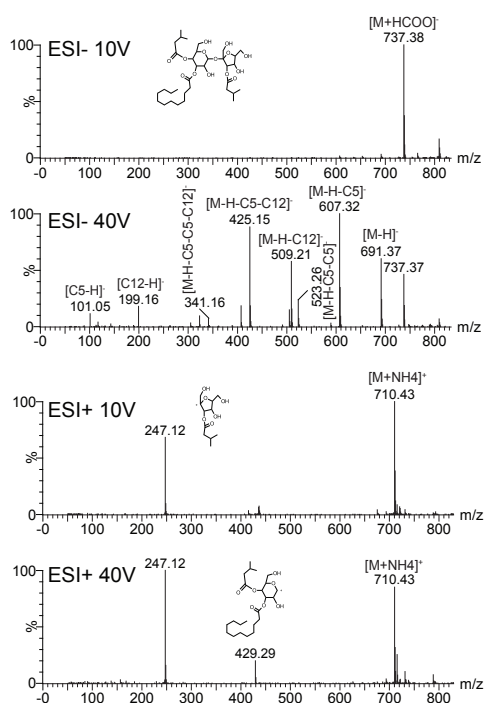


M82 S3:22 (5,5,12)

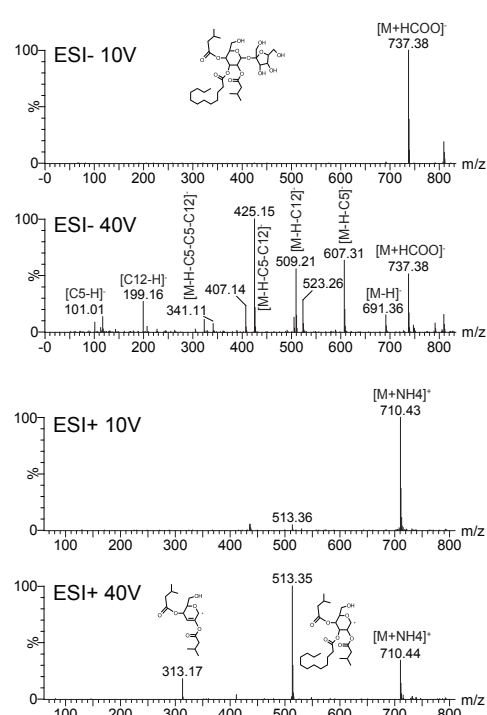
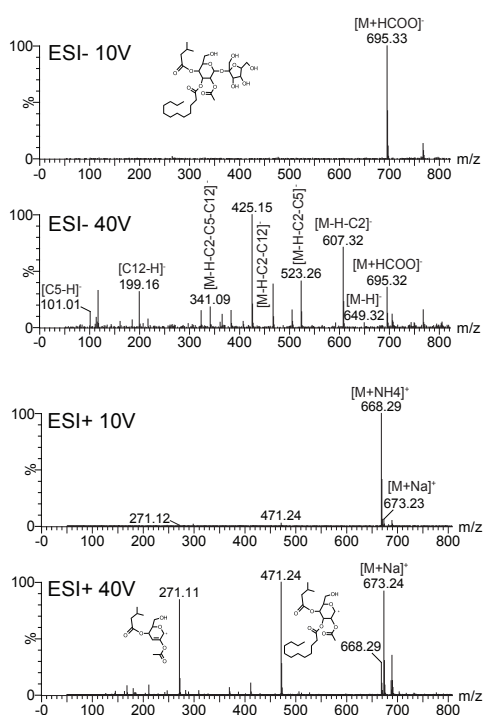
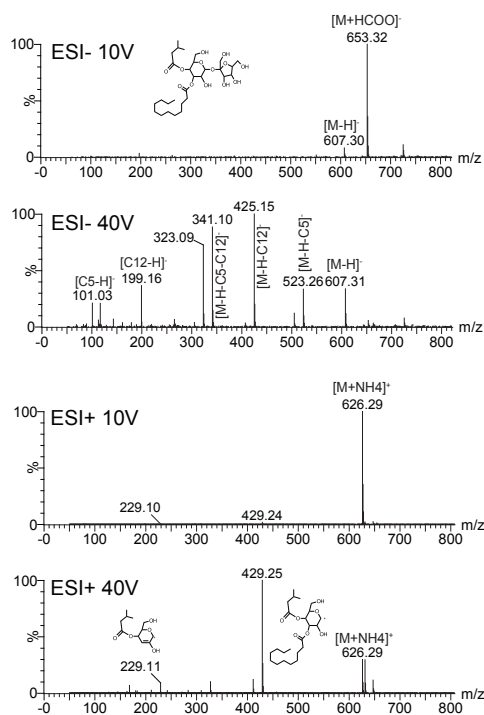
LA0716 S3:19 (4,5,10)



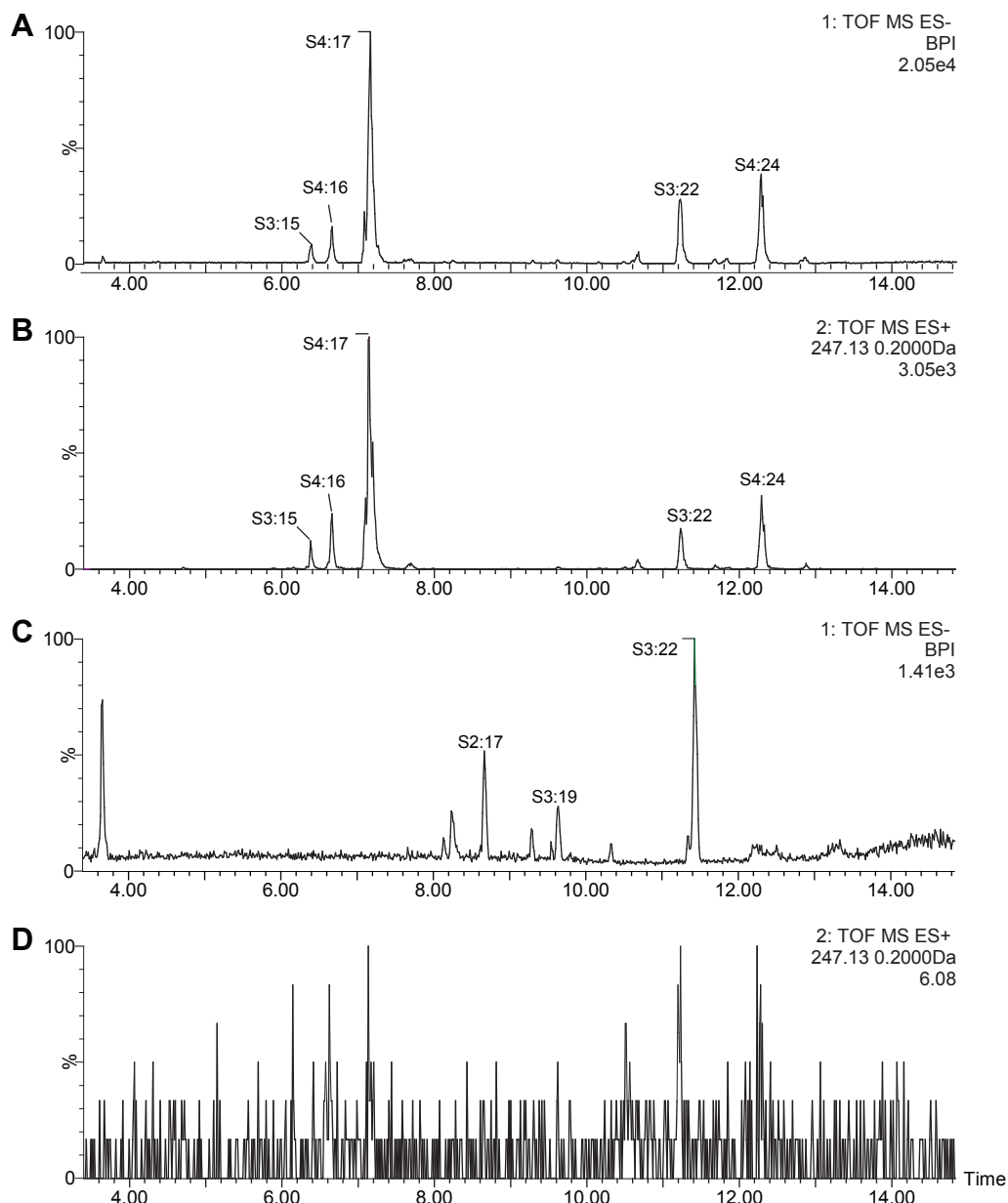
IL11-3 S2:17 (5,12)

IL11-3 S3:19 (2,5,12)

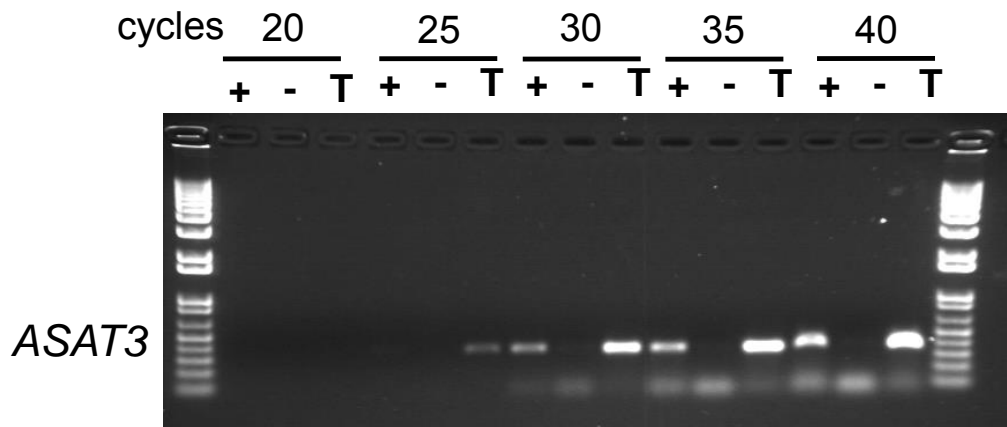
IL11-3 S3:22 (5,5,12)



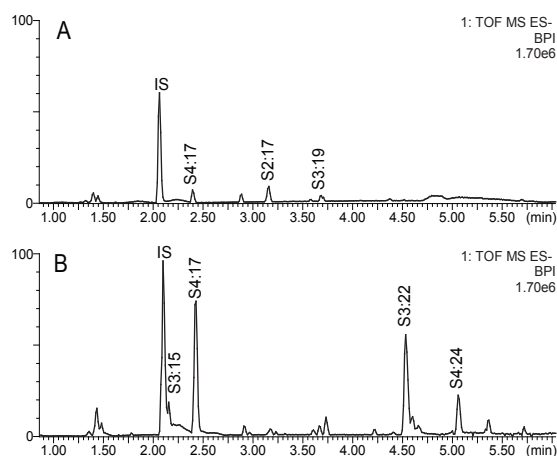
Supplemental Figure S1. Mass spectra of abundant acylsucroses from M82, LA0716 and IL11-3. Negative and positive ionization mode mass spectra are shown for acylsucroses from each genotype. For each ionization mode, spectra are shown using a low aperture 1 voltage (10 V) and a higher voltage (40 V) to induce fragmentation. Fragmentation in negative mode reveals the length of the acyl chains present due to loss of acyl groups as ketenes, and fragmentation in positive mode causes glycosidic bond cleavage to reveal which sugar ring the acyl chains are present on (Ghosh et al., 2013). S2:17 (IL11-3) and S3:22 (M82) structures are verified by NMR, other structures are inferred based on fragmentation and comparison to other known structures.



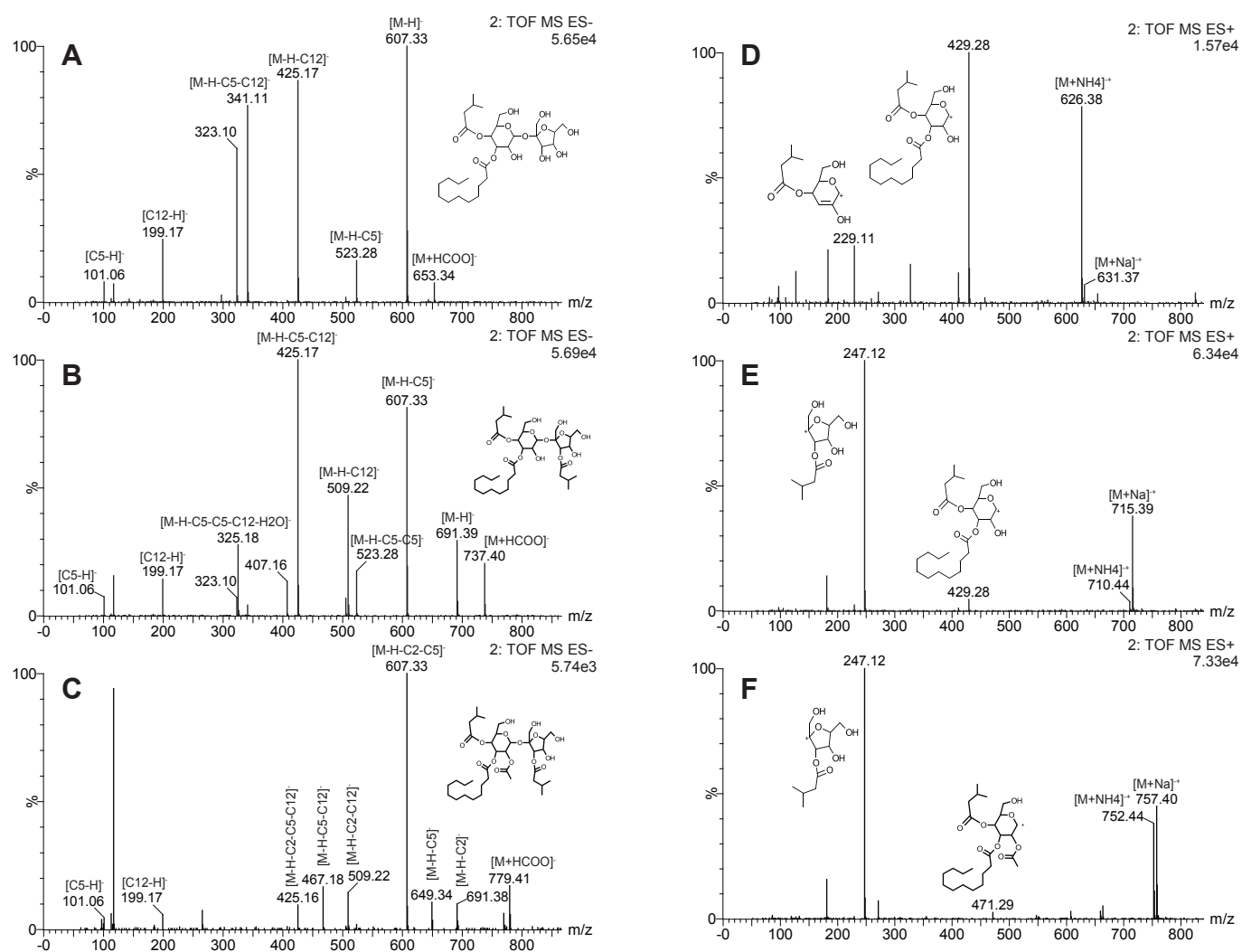
Supplemental Figure S2. IL11-3 acylsugars lack the R_3 acyl group normally seen on M82 acylsucroses. Acylsucroses from M82 extracts show a characteristic fragment of m/z 247.13 in positive ion mode MS that corresponds to the furanose ring with a single five-carbon acyl chain. In contrast, acylsucroses in IL11-3 lack this m/z 247.13 fragment after fragmentation under identical conditions. A, base peak intensity (BPI) chromatogram for M82 acylsugars ionized in negative mode; B, extracted ion chromatogram (XIC) for m/z 247.13 from M82 acylsugars ionized in positive mode; C, BPI chromatogram for IL11-3 acylsugars ionized in negative mode; D, XIC for m/z 247.13 from IL11-3 acylsugars ionized in positive mode showing absence of this characteristic fragment for C5 acylation of the furanose ring.



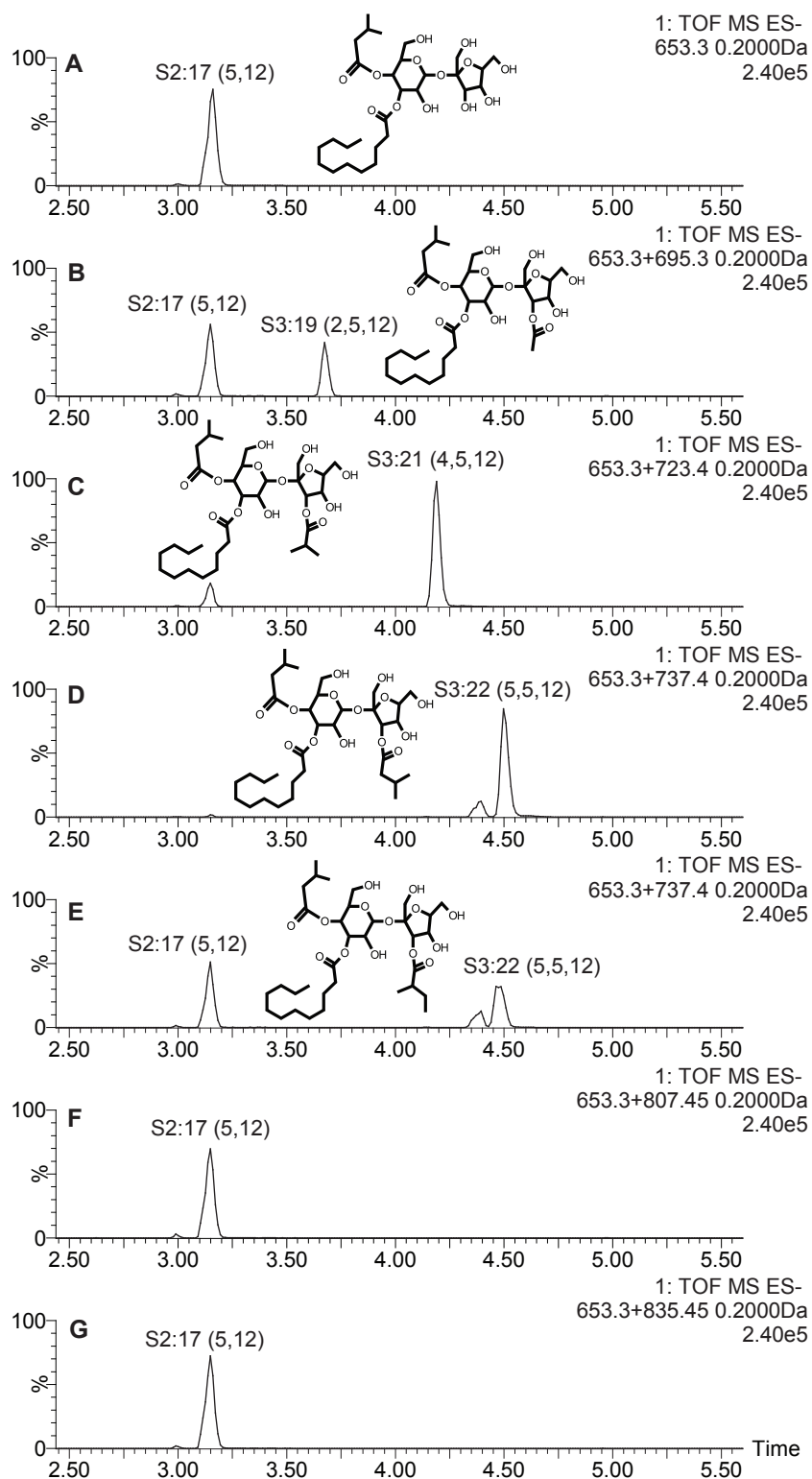
Supplemental Figure S3. RT-PCR analysis of *Sl-ASAT3* expression in M82 plants. RNA from stem and petiole tissue (+), trichomes isolated from stems and petioles (T), and stem and petioles after removal of the trichomes (-) was used for RT-PCR analysis of *Sl-ASAT3* transcript levels. Aliquots were taken after the indicated number of cycles. Expression was detected in the intact stem and petiole tissue as well as in the isolated trichomes. Very little or no expression was detected in the underlying stem and petiole tissue after removal of the trichomes. The gel image is a representative image from two biological replicates.



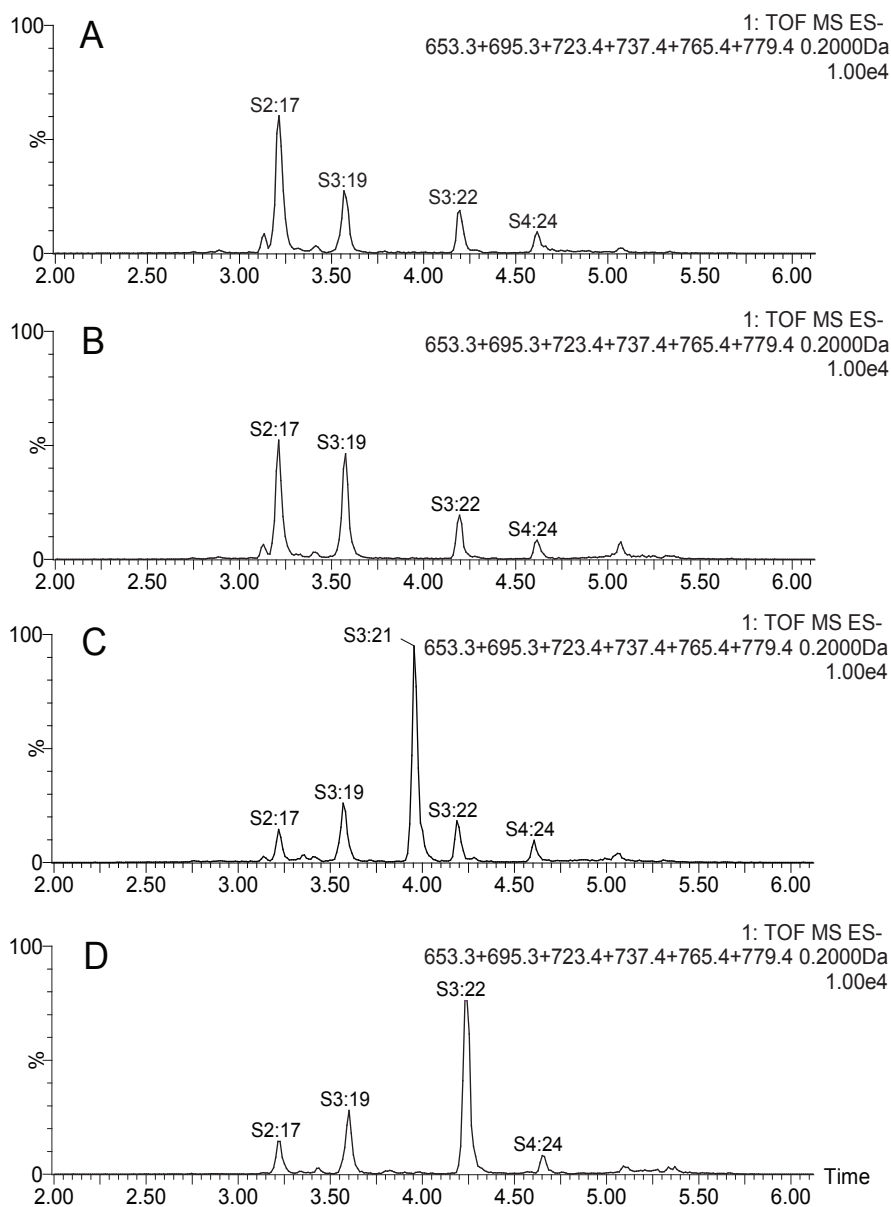
Supplemental Figure S4. Base peak intensity chromatograms for SI-ASAT3 transgenic lines. A, T_1 plant number 10-3 from RNAi of SI-ASAT3 in M82 plants. B, T_0 line number 25 from expression of SI-ASAT3 in IL11-3 plants. S3:15, S4:17, S3:22 and S4:24 all have an acyl chain on the furanose ring. Both chromatograms are representative of acylsugar profiles from transgenic plants with strong phenotypes.



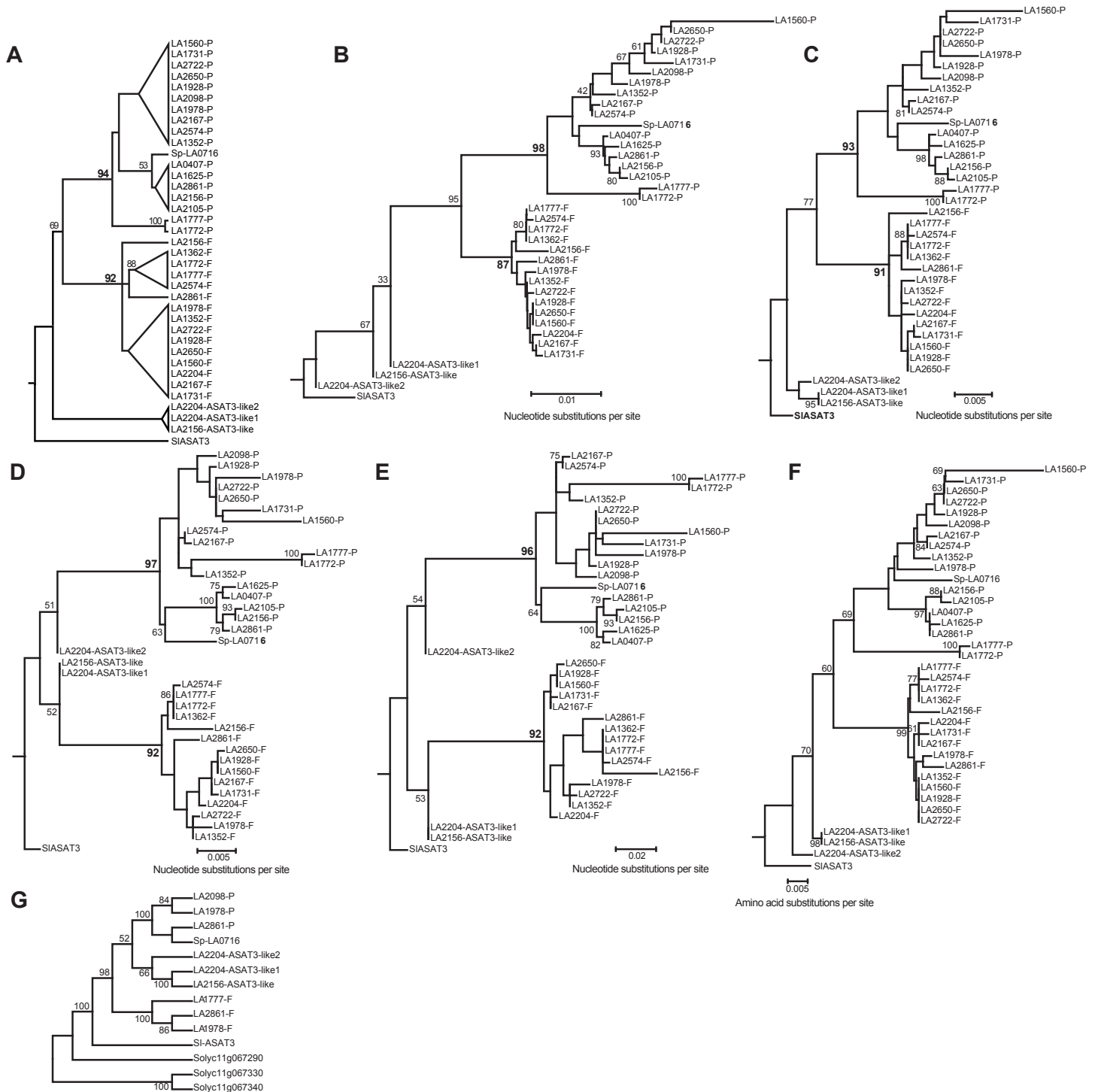
Supplemental Figure S5. Mass spectra of SI-ASAT3 assay products when using S2:17 as the acceptor substrate. Collision energy was set at 25 eV to induce fragmentation. Negative mode MS are shown in A-C and positive mode MS are shown in D-F. (A,D) S2:17 (5,12) substrate from IL11-3; (B,E) S3:22 (5,5,12) product from incubation of SI-ASAT3 with S2:17 and iC5-CoA; (C,F) S4:24 (2,5,5,12) product from SI-ASAT3 + S2:17 + iC5-CoA followed by addition of SI-ASAT4 + C2-CoA. Structures of products are inferred based on mass spectra and co-chromatography with known structures.



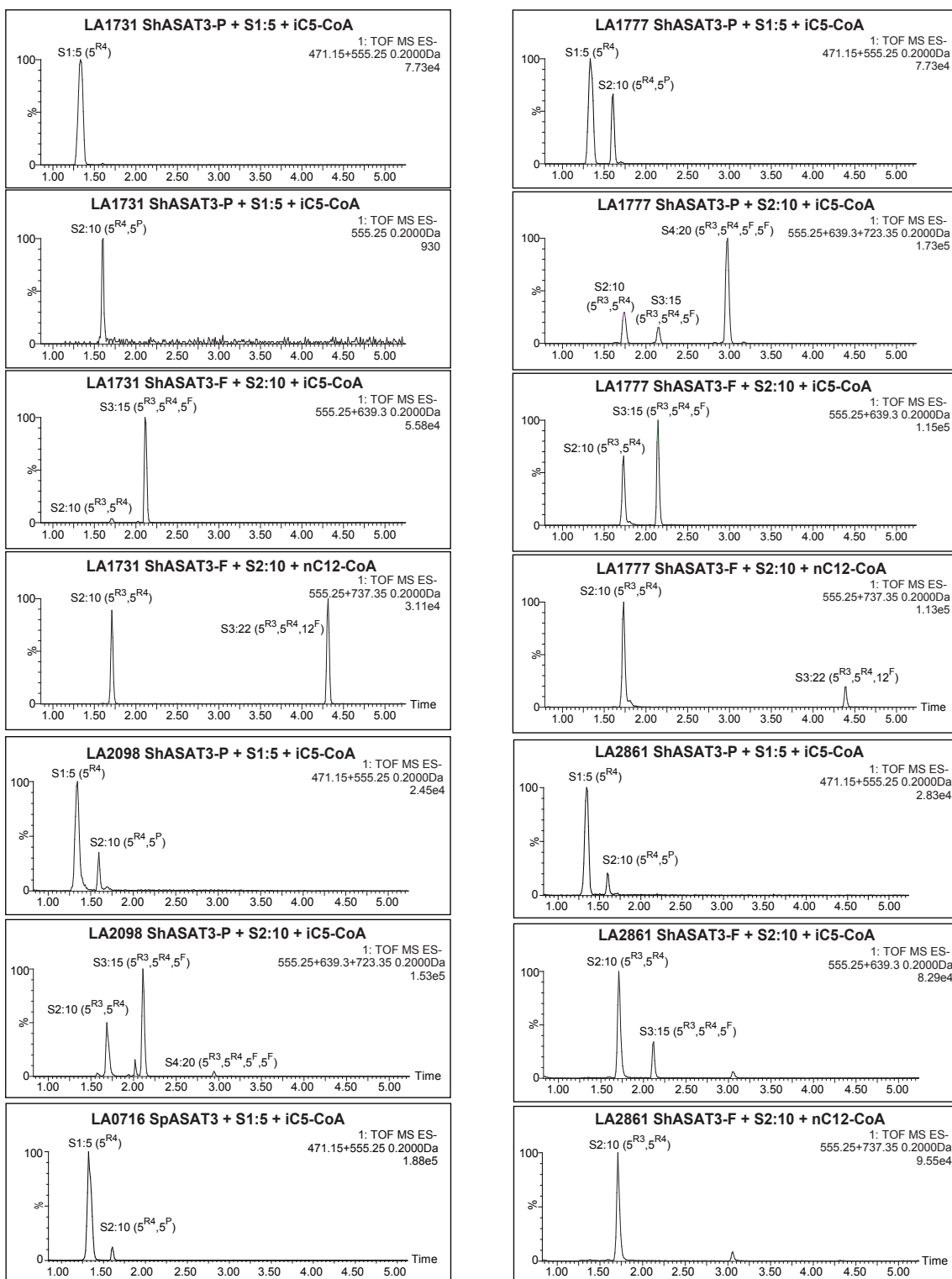
Supplemental Figure S6. SI-ASAT3 assays with S2:17 (5,12) and a variety of available acyl-CoA donors. (A) S2:17 substrate from IL11-3; (B) SI-ASAT3 + S2:17 + C2-CoA; (C) SI-ASAT3 + S2:17 + iC4-CoA; (D) SI-ASAT3 + S2:17 + iC5-CoA; (E) SI-ASAT3 + S2:17 + aiC5-CoA; (F) SI-ASAT3 + S2:17 + iC10-CoA; (G) SI-ASAT3 + S2:17 + nC12-CoA. Structures of products shown are inferred.



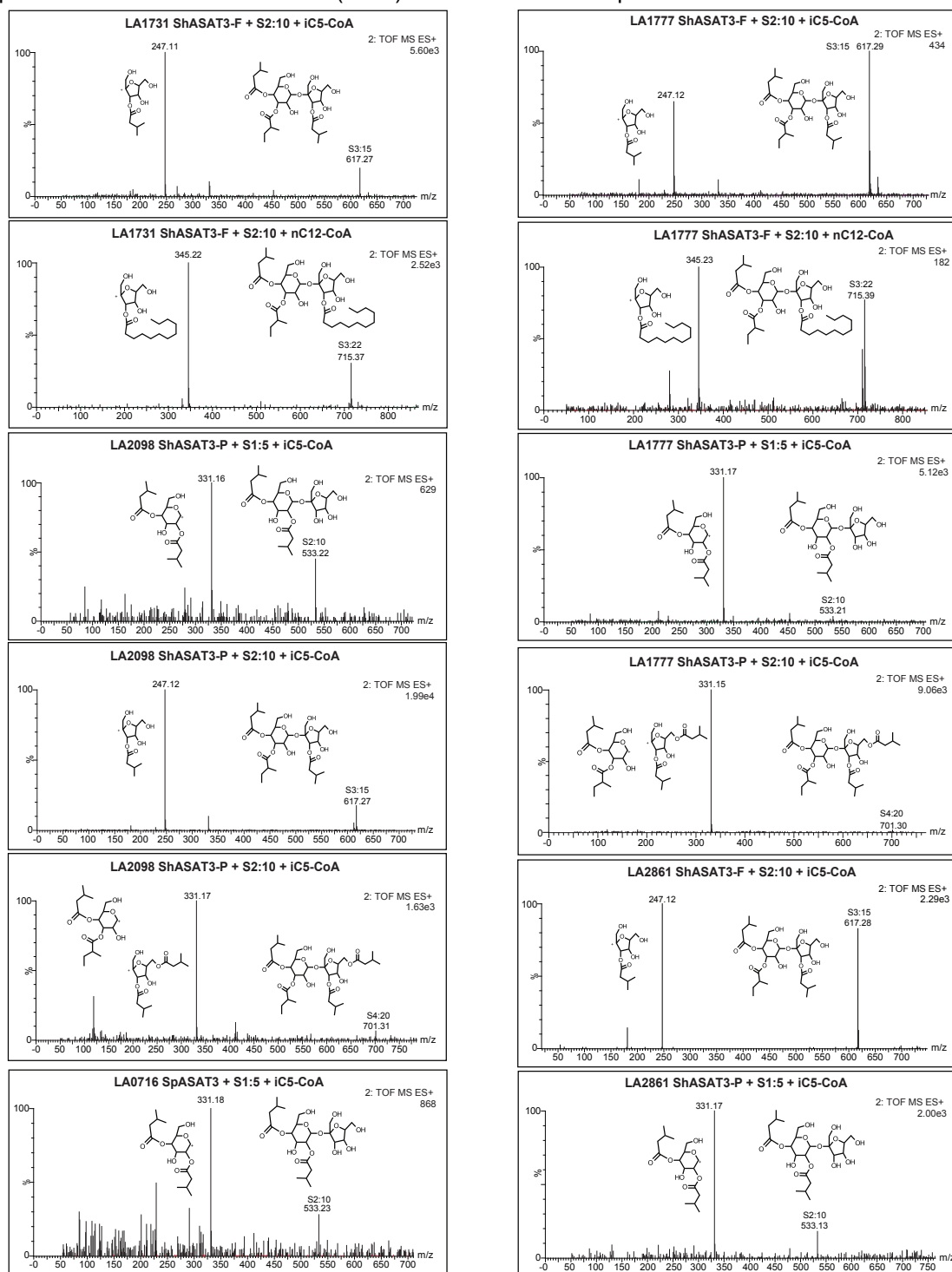
Supplemental Figure S7. SI-ASAT3 activity using acylsugar acceptors collected from RNAi plants. This mixture included both S2:17 (5,12) and S3:19 (2,5,12) and allowed for testing SI-ASAT3 preference for acceptors with and without acetylation. Acylsucroses were partially purified from T_1 plants from RNAi line #10 and used in assays with recombinant SI-ASAT3 enzyme and selected acyl-CoA donors. Shown are extracted ion chromatograms for the acylsugar substrate mixture (A), and the reaction products after incubation with C2-CoA (B), iC4-CoA (C), and iC5-CoA (D).



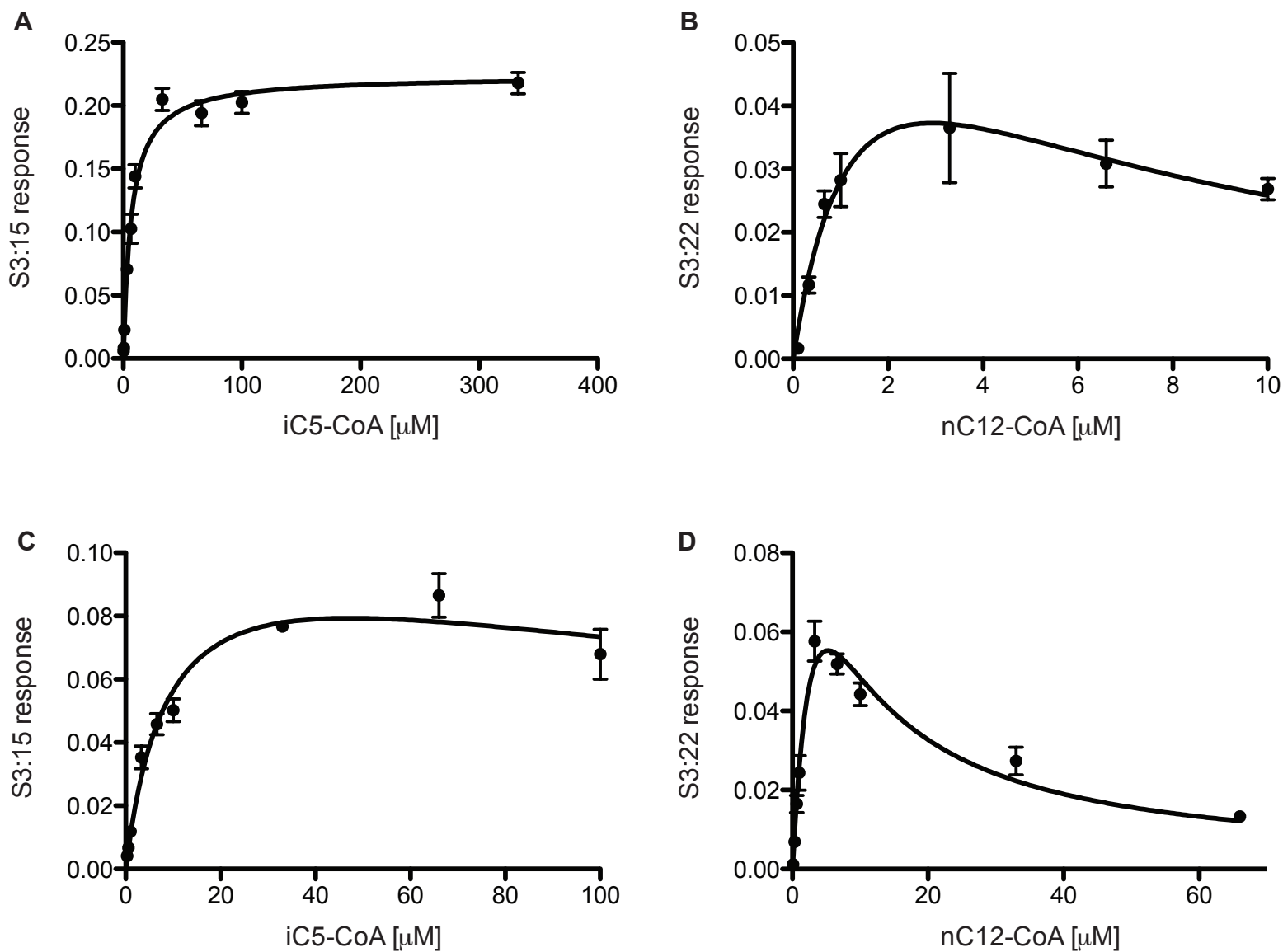
Supplemental Figure S8. Phylogenetic relationships between different ASAT3 sequences. (A): Condensed version of the phylogenetic tree shown in Figure 7B. All branches with < 70 bootstrap support, except the Sp-LA0716 branch, were condensed. (B-E): Phylogenetic trees obtained using different strategies. In all scenarios, bootstrap values for branches with support < 60 are not shown, except for the nodes involving the ASAT3-like sequences. Bootstrap values were calculated using 1000 replicates. In all representations, SI-ASAT3 was used as an outgroup, and the outgroup branch has been compressed for representation purposes. The four strategies used are as follows. (B): Neighbor-Joining (NJ) tree obtained in MEGA6 using the T92 model with gamma distributed rates modeled using gamma = 0.5. (C): Maximum likelihood (ML) tree obtained in MEGA6 using the TN93 model with uniform site rates. (D): ML tree obtained in RAxML using the GTRGAMMA model with self-estimated gamma. (E): ML tree obtained in RAxML using the GTRGAMMA model with invariant sites used for estimating gamma. (F): Rooted NJ tree of protein sequences using JTT model and uniform rates. (G): An unrooted ML bootstrap consensus tree obtained using the same parameters used for generating Figure 7B, using randomly chosen Sh-ASAT3-P and -F sequences, Sp-ASAT3, SI-ASAT3 and using additional *S. lycopersicum* BAHD sequences located nearby SI-ASAT3 on chromosome 11. Note that SI-ASAT3 is the correct ortholog of the Sh-ASAT3 sequences. The sequence alignment used to generate trees A-E is in Supplemental Dataset 2. The sequence alignments used to generate tree F is in Supplemental Dataset 3 and tree G is in Supplemental Dataset 4.



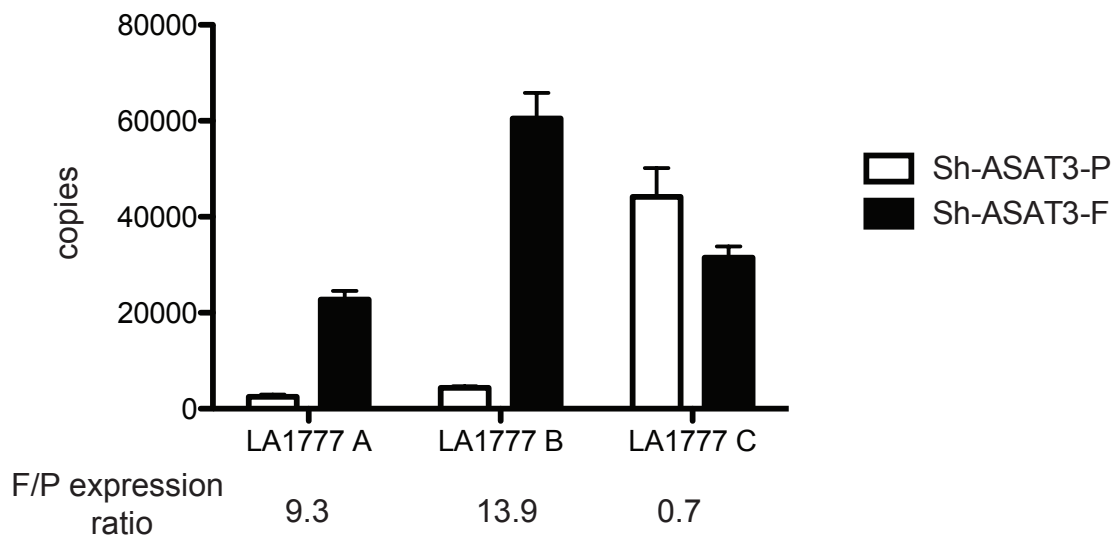
Supplemental Figure S9. Extracted ion chromatograms from analysis of Sh-ASAT3 enzyme assay products. Samples were analyzed using negative ion mode MS. The m/z for unreacted substrate and detected products are shown for each assay performed using Sp-ASAT3 and the various Sh-ASAT3 enzymes. The enzyme, acceptor substrate, and acyl-CoA substrate for each reaction is listed above the XIC.



Supplemental Figure S10. Positive ion mode mass spectra for Sh-ASAT3 enzyme assay products. Mass spectra obtained in positive ion mode MS (collision energy = 25 eV) were used to determine which ring (pyranose or furanose) the acyl chain was added to in the reaction. In general, the positive charge is more stabilized when retained by the furanose ring fragment after cleavage of the glycosidic bond. Therefore, for triacylsucroses, the m/z 247.1 ion indicates a single C5 chain on the furanose ring and m/z 345.2 indicates a C12 on the furanose ring. For the products shown here, the m/z 331.16 fragments indicate two C5's on the same ring. Since the structure of the S1:5 substrate is known to have the acyl chain on the pyranose ring, the S2:10 product having a m/z 331.16 fragment is inferred to have the second C5 also on the pyranose ring. In the case of the S4:20 (5,5,5,5) product obtained using the S2:10 substrate (that has both chains on the pyranose ring), the m/z 331.15 fragment corresponds to two C5's on both the furanose ring and the pyranose ring. Actual positions of the added acyl chains on either the pyranose or furanose ring cannot be determined by MS and is inferred.



Supplemental Figure S11. Determination of apparent K_m and K_i values for acyl-CoA substrates. The S2:10 acceptor substrate was held at saturating concentration and acyl-CoA concentrations were varied. Product peak area divided by the internal standard peak area (response) was plotted for each concentration of acyl-CoA and assays were done in triplicate. A,B LA1777 Sh-ASAT3-F; C,D LA1731 Sh-ASAT3-F. Non-linear regression using Michaelis-Menten model (A) or substrate inhibition model (B-D) in the GraphPad Prism5 software package was used to determine apparent K_m and K_i values and standard errors.



Supplemental Figure S12. Analysis of Sh-ASAT3-P and -F expression in trichomes of LA1777 plants. Transcript levels were determined by qRT-PCR for three LA1777 individuals and the relative levels of Sh-ASAT3-P and -F were compared to the acylsugar phenotype. Two LA1777 plants (LA1777 A and B) that produced acylsucroses having a single acyl chain on the furanose exhibited a higher expression level of Sh-ASAT3-F compared to Sh-ASAT3-P. One LA1777 plant that produced some acylsucroses having two short acyl chains on the furanose ring showed a higher level of Sh-ASAT3-P expression relative to Sh-ASAT3-F (LA1777 C). Acylsugar phenotype data for the three LA1777 plants is presented in Supplemental Dataset 1.

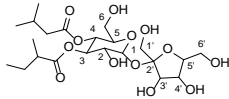
Supplemental Table S1. NMR chemical shifts of S2:17 (5,12) from IL11-3

S2:17 (5,12)		
Purified from <i>S. lycopersicum</i> IL11-3		
HRMS: (ESI) m/z calculated for $C_{30}H_{55}O_{15}$ ([M+HCOO ⁻]): 653.3390, found 653.3343		
Material recovered: ~ 300 μ g		
NMR solvent: CD ₃ CN		
Carbon # (group)	¹ H (ppm)	¹³ C (ppm)
1(CH)	5.45 (d, $J = 3.8$ Hz)	92.6 ^b
2(CH)	3.70 (m) ^a	70.9 ^b
3(CH)	5.29 (t, $J = 10.0$ Hz)	73.5 ^b
3-O-		
-1(CO)		174.0 ^c
-2(CH ₂)	2.24 ^a	34.8 ^b
-3(CH ₂)	1.53 ^a	25.5 ^b
-4-10(CH ₂ -CH ₂ -CH ₂ -CH ₂ -CH ₂ - CH ₂ -CH ₂ -)	1.24-1.28 (br. s) ^a	29.7- 32.5(7 carbon resonances)
-11(CH ₂)	1.28 ^a	23.4 ^b
-12(CH ₃)	0.87 (t, $J = 7.2$ Hz)	14.4 ^b
4(CH)	4.88 (t, $J = 9.9$ Hz)	69.2 ^b
4-O-		
-1(CO)		172.9 ^c
-2(CH ₂)	2.11 ^{a,b}	43.6 ^b
-3(CH)	1.98 ^a	25.9 ^c
-4(CH ₃) x 2	0.90 ^a	22.4 ^b
5(CH)	4.04 ^a	72.3 ^b
6(CH ₂)	3.45, 3.55 ^{a,b}	61.4 ^b
1'(CH ₂)	3.50, 3.55 ^b	65.0 ^b
2'(C)	-	105.0 ^c
3'(CH)	3.95 (d, $J = 7.3$ Hz)	80.3 ^b
4'(CH)	4.00 ^a	75.1 ^b
5'(CH)	3.70 ^a	83.6 ^b
6'(CH ₂)	3.60, 3.70 ^{a,b}	61.9 ^b
^a Determined by COSY. ^b Determined by HSQC. ^c Determined by HMBC		

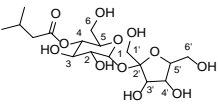
Supplemental Table S2. Genes in the IL11-3 mapping interval. List of the predicted genes in the *S. lycopersicum* Heinz 1706 genome interval that contains the locus controlling acylsucrose furanose ring acylation. Percentages of reads for each gene out of the total reads are shown for trichome RNAseq data from M82 and LA0716.

ITAG2.3 Gene Model	ITAG2.3 annotation	SGN unigene	M82 SGT cDNA RNAseq	LA0716 SGT cDNA RNAseq
Solyc11g067260	Hydrolase alpha/beta fold family protein (AHRD V1 ***- D7LIQ3_ARALY); contains Interpro domain(s) IPR000073 Alpha/beta hydrolase fold-1	SGN-U577685	-	0.0008%
Solyc11g067270	Acyltransferase-like protein (AHRD V1 **-- Q589X8_TOBAC); contains Interpro domain(s) IPR003480 Transferase	-	0.0030%	0.6700%
Solyc11g067280	MYB-CC type transfactor (AHRD V1 **-- Q3LHL3_SOLTU); contains Interpro domain(s) IPR006447 Myb-like DNA-binding region, SHAQKYF class	SGN-U576104	0.0004%	0.0038%
Solyc11g067290	Acyltransferase-like protein (AHRD V1 ***- Q589X7_TOBAC); contains Interpro domain(s) IPR003480 Transferase	-	-	-
Solyc11g067300	Lipid A export ATP-binding/permease protein msbA (AHRD V1 *--- MSBA_CHRVO); contains Interpro domain(s) IPR003439 ABC transporter-like	-	-	-
Solyc11g067310	Uncharacterized ABC transporter ATP-binding protein/permease C9B6.09c (AHRD V1 *--- YNT9_SCHPO); contains Interpro domain(s) IPR003439 ABC transporter-like	-	-	-
Solyc11g067320	Sulfotransferase 5a (AHRD V1 ***- C5H9P5_BRARP); contains Interpro domain(s) IPR000863 Sulfotransferase	-	-	-
Solyc11g067330	Acyltransferase (Fragment) (AHRD V1 **-- C1JZ69_DATST); contains Interpro domain(s) IPR003480	-	-	-
Solyc11g067340	Acyltransferase (Fragment) (AHRD V1 **-- C1JZ69_DATST); contains Interpro domain(s) IPR003480	SGN-U564111	-	-
Solyc11g067350	Acyltransferase (Fragment) (AHRD V1 **-- C1JZ70_SOLPN); contains Interpro domain(s) IPR003480	-	-	-
Solyc11g068350	F-box family protein (AHRD V1 ***- D7KV05_ARALY); contains Interpro domain(s) IPR013101 Leucine-rich repeat2	-	-	-
Solyc11g068360	Cc-nbs-lrr, resistance protein	SGN-U592772	-	-
Solyc11g068370	BZIP transcription factor (AHRD V1 ***- Q93XM6_ARATH); contains Interpro domain(s) IPR011616 bZIP transcription factor, bZIP-1	SGN-U597450	-	-
Solyc11g068380	Unknown Protein (AHRD V1)	-	-	-
Solyc11g068390	Unknown Protein (AHRD V1)	-	-	-
Solyc11g068400	Ubiquinol-cytochrome C reductase (AHRD V1 ***- D6BQNO_9ROSI); contains Interpro domain(s) IPR008027	SGN-U570518	-	-
Solyc11g068410	Unknown Protein (AHRD V1)	-	-	-
Solyc11g068420	Ribosomal protein (AHRD V1 ***- Q6ER67_ORYSJ); contains Interpro domain(s) IPR002143 Ribosomal protein L1	SGN-U581188	0.0050%	0.0180%
Solyc11g068430	Ferredoxin (AHRD V1 ***- Q9AUE1_IMPBA); contains Interpro domain(s) IPR010241 Ferredoxin [2Fe-2S]	SGN-U569608	0.1500%	0.0140%
Solyc11g068440	Glucan endo-1 3-beta-glucosidase 7 (AHRD V1 ***- B6T478_MAIZE); contains Interpro domain(s) IPR000490 Glycoside hydrolase, family 17	SGN-U585972	0.0058%	0.0210%

Supplemental Table S3. NMR chemical shifts of S2:10 (i5,ai5) produced *in vitro*

S2:10 (5,5)		
		
HRMS: (ESI) m/z calculated for $C_{22}H_{38}NO_{16}$ ($[M+NO_3^-]$): 572.2196, found: 572.2227		
Material recovered: < 500 μg		
NMR solvent: CD₃CN		
Carbon # (group)	¹ H (ppm)	¹³ C (ppm)
1(CH)	5.45 (d, $J = 3.9$ Hz)	92.5 ^b
2(CH)	3.70 (m) ^a	70.7 ^b
3(CH)	5.33 (t, $J = 9.8$ Hz) ^a	73.2 ^b
3-O-		
-1(CO)		177.2 ^c
-2(CH)	2.33 ^a	41.6 ^b
-2'(CH ₃)	1.05 (d, $J = 7.0$ Hz)	16.7 ^b
-3(CH ₂)	1.43 (m), 1.61 (m)	27.2 ^b
-4(CH ₃)	0.87 (t, $J = 7.4$ Hz)	11.7 ^b
4(CH)	4.91 (t, $J = 9.9$ Hz)	69.0 ^b
4-O-		
-1(CO)		173.0 ^c
-2(CH ₂)	2.09-2.19 (m)	43.5 ^b
-3(CH)	1.98 ^a	26.0 ^c
-4(CH ₃) x 2	0.90 (d, $J = 6.7$ Hz), 0.90 (d, $J = 6.7$ Hz)	22.4 ^b
5(CH)	4.09 ^a	71.8 ^b
6(CH ₂)	3.45, 3.57 ^{a,b}	61.3 ^c
1'(CH ₂)	3.51, 3.56 ^{b,c}	64.6 ^c
2'(C)	-	104.9 ^c
3'(CH)	3.95 (d, $J = 8.0$ Hz)	80.0 ^b
4'(CH)	4.00 (t, $J = 7.9$ Hz)	74.8 ^b
5'(CH)	3.70 ^a	83.4 ^b
6'(CH ₂)	3.61, 3.69 ^a	61.9 ^b
^a Determined by COSY. ^b Determined by HSQC. ^c Determined by HMBC		

Supplemental Table S4. NMR chemical shifts of S1:5 (i5) produced *in vitro*

S1:5		
		
HRMS: (ESI) m/z calculated for $C_{18}H_{31}O_{14}$ ([$M+HCOO^-$): 471.1719, found 471.1772		
Material recovered: < 500 μ g		
NMR solvent: D_2O		
Carbon # (group)	1H (ppm)	^{13}C (ppm)
1(CH)	5.44 (d, $J = 3.8$ Hz)	92.1 ^b
2(CH)	3.63 (m) ^a	71.0 ^b
3(CH)	4.02 (m) ^a	74.0 ^b
4(CH)	4.90 (t, $J = 9.8$ Hz)	70.5 ^b
4-O-		
-1(CO)		175.3 ^c
-2(CH ₂)	2.33 (d, $J = 7.3$ Hz)	43.0 ^b
-3(CH)	2.06 (m)	25.4 ^c
-4(CH ₃) x 2	0.94 (d, $J = 6.7$ Hz)	21.6 ^b
5(CH)	3.92 ^a	70.6 ^b
6(CH ₂)	3.54, 3.62 ^a	62.4 ^b
1'(CH ₂)	3.80 ^b	62.4 ^b
2'(C)	-	103.7 ^c
3'(CH)	4.21 (d, $J = 8.9$ Hz)	76.4 ^b
4'(CH)	4.00 (t, $J = 7.9$ Hz)	74.1 ^b
5'(CH)	3.88 ^a	81.4 ^b
6'(CH ₂)	3.67 ^a	61.3 ^b
^a Determined by COSY. ^b Determined by HSQC. ^c Determined by HMBC		

Supplemental Table S5. Primer sequences for PCR-based markers for genetic mapping and ASAT3 amplification.

Markers were either available from the Sol genomics website (solgenomics.net) or created by us based on polymorphic sites between M82 and LA0716 sequences. Restriction enzymes for CAPS (cleaved amplified polymorphic sequence) markers are shown with the expected sizes for each genotype.

Marker	Forward Primer	Reverse Primer	SL2.30 genomic location	Enzyme	M82 (bp)	LA0716 (bp)
TG400	TCCAAATCCACCACCTATCC	AGCATTGCTCCCTGCTAAAG	48602852-48603255	Hinfl	404	300+104
C2_At3g44880	ACAAGATTTCGTCGAAATTCTC	ACCACATCCATTAATGACCATCC	49286882-49287233	Dral	255+97	357
72113	CTTGCTAGAATCCACACTCACATAACC	CAACAGGTCGATGAAAATGATACAAGA	49595187-49595635	none	448	396
233B07	GAGGGTGTTCATTTCGTTTCAGTTTG	CGCACCCGTTGGCATCTTTG	49786194-49786582	ApaLI	256+131	387
139J14	ATTAAGATTGCGGAAAGACG	GATCGTTTTTCGCCTTGGGTTTGA	49889898-49890139	Scal	198+46	242
36B01T7	CGGCCATGTCGAAAGTAGAAAG	TCGAACCGGAGAGCAAAAATG	49977351-49977685	BsaHI	136+162+58	355
U600389F1R2	GCCAACGGGGAACAAAGCAGAC	TAAGACGGGAGGCAGTATTGAT	50068571-50068804	Rsal	192+63	255
M8190	CTTTCTCCCTTTACCGTCTGA	TTGGTGGGCGTAAATTCATAGCA	50096563-50096835	Rsal	273	38+46+189
AT6820F2R1	AATCAGGCGCGTTTAGACCAA	GCTCCCTAACAAAGACGACTCAACAACCT	50156957-50157474	Ndel	518	~320+190
38413F1R1	GGGAGACCCGAATGAGTTTG	GGAGTTCATCAGGGTAGTAGAGAC	50255627-50255833	AluI	207	127+80
C2_At2g27730	TGGAGAAAGAGAAGCTGGAGAAGC	TCCTTTGACATTAGGTACCAACCC	50296574-50297481	HaeIII	802+106	578+217+106
303G16-3	CCACCCAACCAACTTAGTACAAT	ATGGGCGAAATGCTGAGACAA	50402330-50403216	XmnI	811+24	541+271+24
C2_At3g54470	TCCTGACTTTGGTTCTAAGCTTAGATCG	TCAAATATTAAGAAGTTGTGCTTGTCTGC	50521808-50522472	Rsal	446+219	700
cLEC24C3	AGATCGGCAAATGATCCAAG	ACTTGTGGCGAAAAATGAGG	50619743-50621023	EcoRI	959+322	1285
C2_At5g60600	TTGCTTCAAGTTGCAGAATGCC	ACCAGGCAAGTGTGACGTCTTCTCTC	51067682-51067895	Rsal	124+90	214
T302	TGGCTCATCCTGAAGCTGATAGCGC	AGTGTACATCCTTGCCATTGACT	51877416-51878207	none	792	~950

RT-PCR Primers

SI-ASAT3-RT	TACGCCACCAATTCCACTGA	CATTTAGCCCATCTCCACTTTGTTC
Sh-ASAT3-P	CACATAAATTTCTCTGCATTCTCCATA	GCCCAATCATTATGAATTTAGAGAGAG
Sh-ASAT3-F	TTATGGATCACATAAATCTTCTCTACAT	GCCCAATCATTATAAATTTAGAGACAC

Primers for cloning ASAT3 for expression in *E. coli*

SI-ASAT3, Sp-ASAT3, and Sh-ASAT3-P	GCTAGCATGGCATCATCAACAATTAT	CTCGAGTTATTTGGTTGATTCAACAACCTGG
Sh-ASAT3-F	GCTAGCATGGCATCATCAAAAATGATATCT	CTCGAGTTATTTGGTTGATTCAACAACCTGG