The Genetics of Whitefly Resistance in Tomato

Dissertation

zur Erlangung des akademischen Grades

Doctor rerum naturalium (Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät I

Biowissenschaften

der Martin-Luther-Universität Halle-Wittenberg,

vorgelegt von

Frau Micha Gracianna Devi, S.P., M.Sc.

geboren am 02.07.1991 in Bogor, Indonesia

Erstgutachter: Prof. Dr. Alain Tissier

Zweitgutachter: Prof. Dr. Jörg Degenhardt

Drittgutachter: Dr. P.M. Petra Bleeker

Tag der Verteidigung: 18.03.2022

With knowledge comes power.

With wisdom comes judgement.

With mandate comes authority.

2021

Table of Contents

Abstract

Tomato plants (*Solanum lycopersicum*) have adapted strategies by inducing or constitutively activating their defense mechanisms to mitigate a wide range of environmental factors for optimum growth and development. Whiteflies (*Bemisia tabaci*) have serious economic consequences as they impact tomato yield via feeding on leaves and vectoring plant viruses. Breeding by introgression of chromosome fragments from resistant wild species, such as *Solanum habrochaites*, provides a potential solution to develop whitefly resistant varieties. In the present work, the response of two populations of *Solanum habrochaites* sp. *glabratum* x *S. lycopersicum* against whitefly infection was evaluated using phenotypic and genotypic observation of tomato leaves as well as biochemical analyses using untargeted metabolomics by liquid and gas chromatography mass spectrometry. This information led to the identification of two QTL regions in chromosome 5 and 11, which positively correlated with oviposition and colocalized with a newly reported candidate susceptibility factor involved in polyamine biosynthesis. QTL11 includes a region with *Acylsugar Acyl Transferase 3 (ASAT3)*, a gene that is associated with increased adult whitefly mortality, increased density of type IV trichomes, and higher abundance of acylsugars. However, an isoform of *ASAT3* found in *S habrochaites* sp. *glabratum* was hypothesized to be regiospecific in the transfer of acyl-groups to a different position than the ones that have previously been characterized in *S. lycopersicum*. Furthermore, with reference to acylsugar biosynthesis, two genes coding for *acyl-CoA synthetases* which are specific to short and medium chained fatty acid were characterized via enzyme assay and virus induced gene silencing. This study shows that integration of different approaches and techniques including metabolomics, genomics, and transcriptomics is required to unravel the complexity of tomato resistance/susceptibilty traits to whitefly, the here described results may help breeders to select biomarkers for finetuning the metabolic composition of tomatoes.

Keywords: whitefly, tomato, LC/GC-MS, polyamines, acylsugar, QTL analysis, metabolomics, transcriptomics

Abstrakt

Tomatenpflanzen (*Solanum lycopersicum*) haben, in dem sie ihre Abwehrmechanismen induzieren oder konstitutiv aktivieren, Strategien adaptiert, um eine ganze Reihe von Umweltfaktoren abzuschwächen und ein optimales Wachstum und eine optimale Entwicklung zu unterstützen. Tabakmottenschildläuse (*Bemisia tabaci*) führen zu schwerwiegenden wirtschaftlichen Folgen, da sie den Tomatenertrag beeinträchtigen, in dem sie sich von Blättern ernähren und Pflanzenviren übertragen. Die Züchtung durch Introgression von Chromosomenfragmenten aus resistenten Wildarten wie *Solanum habrochaites* bietet eine mögliche Lösung zur Entwicklung von Varietäten, die gegen Tabakmottenschildläuse resistent sind. In der vorliegenden Arbeit wurden die Auswirkungen einer Infektion mit *B. tabaci* auf zwei Populationen von *Solanum habrochaites sp. glabratum* x *S. lycopersicum* durch phänotypische und genotypische Beobachtung von Tomatenblättern, sowie biochemische Analysen unter Verwendung von ungezielter Metabolomik durch Flüssigkeit- und Gaschromatographie-Massenspektrometrie, untersucht. Diese Experimente führten zur Identifizierung von zwei QTL-Regionen in den Chromosomen 5 und 11, die positiv mit der Eiablage korrelieren und mit einem kürzlich entdeckten Kandidaten für einen Suszeptibilitätsfaktor kolokalisieren, welcher an der Polyamin- Biosynthese beteiligt ist. QTL11 beinhaltet eine Region mit dem Gen *Acylzucker Acyl Transferase 3 (ASAT3)*, welches mit einer erhöhten Mortalität für die Tabakmottenschildlaus, einer erhöhten Dichte von Trichomen vom Typ IV und einem höheren Vorkommen von Acylzuckern assoziiert wird. In *S habrochaites* sp. *glabratum* hingegen wurde eine Isoform von ASAT3 gefunden, für welche angenommen wird, dass sie regiospezifisch den Transfer von Acyl-Gruppen zu einer anderen Position vermittelt als die Enzyme, welche zuvor in *S. lycopersicum* charakterisiert wurden. In Hinblick auf die Acylzucker-Biosynthese wurden darüber hinaus zwei Gene und deren Produkte jeweils mittels Virus-induziertem Gen-Stummschaltung und Enzymtest charakterisiert, die für *Acyl-CoA-Synthetasen* kodieren, welche spezifisch für kurz- und mittelkettige Fettsäuren sind. Diese Studie zeigt, dass die Integration verschiedener Ansätze und Techniken einschließlich Metabolomik, Genomik und Transkriptomik erforderlich ist, um die Komplexität der Resistenz-/Anfälligkeitsmerkmale von Tomaten gegenüber Tabakmottenschildläuse zu entschlüsseln. Die hier beschriebenen Ergebnisse können Züchtern helfen, um Biomarker zur Feinabstimmung der metabolischen Zusammensetzung von Tomaten auszuwählen.

Schlüsselwörter: Tabakmottenschildlaus, Tomate, LC/GC-MS, Polyamine, Acylzucker, QTL-Analyse, Metabolomik, Transkriptomik

Abbreviations

AACS	Acyl-CoA synthetase
AAO	Abscisic aldehyde oxidase
ACD	Acyl-CoA dehydrogenase
ACOX	Acyl-CoA oxidase
AMD	Adenosyl methionine decarboxylase
AS	Acylsugar
ASAT	AS acyl transferase
ATP	Adenosine triphosphate
Bp	Base pair
BP	Biological process
CC	Cellular component
cDNA	Complimentary DNA
Chr.	Chromosome
CIM	Composite interval mapping
cM	Centi-Morgan
dcSAM	Decarboxylated s-adenosyl-methionine
ESI	Electrospray ionization
FAEE	Fatty acid ethyl esters
FD5	Phenylalanine-D5
Fig	Figure
FA	Formic acid adduct
FAs	Fatty acyls
Fmoc-Cl	Fluorenylmethoxycarbonyl-chloride
GBS	Genotyping by sequencing
GC-MS	Gas chromatography mass spectrometry
GLV	Green leaf volatiles
GT	Glandular trichomes
IBC	Inbred backcross
IQR	Interquartile range
JA	Jasmonic acid
LFC	Log fold change
MeSA	Methyl salicylate
Met	Methionine
MEP	Methylerythritol phosphate pathway
MEV	Mevalonic acid pathway

MF	Molecular function
MIM	Single trait multiple interval mapping
min	Minute(s)
mL	Mililiter
mQTL	metabolite quantitative trait loci
MRM	Multiple reaction monitoring
mRNA	messenger RNA
Neg	Negative mode
NIL	Near isogenic lines
NRG	Energy mode
QTL	Quantitative trait loci
LC-MS	Liquid chromatography mass spectrometry
PA	Polyamine
PCA	Principal component analysis
Pos	Positive mode
RF	Recombination frequency
Rf	Retention factors
RIN	RNA integrity number
RNAseq	RNA sequencing
RT	Retention time
Q-RT-PCR	Real time quantitative polymerase chain reaction
s	Second(s)
SA	Salicylic acid
SAH	S-Adenosyl-L-homocysteine
SAM	S-Adenosyl-methionine
SAMS	S-Adenosyl-methionine synthetase
SAMDC	S-adenosyl-methionine decarboxylase
SGA	Steroidal glycoalkaloids
SIM	Simple interval mapping
SNPs	Single nucleotide polymorphisms
SPE	Solid phase extraction
SWATH	Sequential window acquisition of all theoretical mass spectra
TPM	Transcript per million
TPS	Terpene synthase
TYLCV	Tomato yellow leaf curl virus
VIGS	Virus induced gene silencing
VOC	Volatile organic compounds

Table of tables

Table of figures

1. Introduction

1.1. *Bemisia tabaci* biology, damage, and control

Whiteflies are insects classified in the Aleyrodidae family and consists of more than 1,500 species (Martin *et al.,* 2000). Their developmental stages include egg, instar crawler, instar pupae and adult (Fig. 1). Female whiteflies originate from fertilized eggs whereas males originate from unfertilized eggs; typically, the sex ratio is 2:1, females to males (Tsai & Wang, 1996). Although their lifetime is limited to 24-32 days, it is sufficient to cause serious crop yield loss, *e.g.* damages by *Bemisia tabaci*, a species of whitefly that is one of the most invasive pests in today's agriculture (http://www.iucngisd.org/). They feed on the foliar phloem using their four interlocked stylets enclosing a food and a salivary canal allowing independent movements between plant mesophyll cells (Rosell *et al.*, 1995). During feeding, whiteflies secrete sheath saliva during the penetration of the stylets as lubricant and watery saliva containing enzymes and metabolites, thereby providing protection against plant's wound response. (Huang *et al.*, 2020). Whiteflies indirectly cause damages to plants by acting as a vector for viruses, including more than 100 plant viruses ranging from the *Begomovirus* genus like the famous tomato yellow leaf curl virus (TYLCV) to the *Crinivirus* genus like the tomato chlorosis virus (Jones, 2003). Moreover, *B. tabaci* has a wide host range such as tomato, cucumber, cotton, and sweet potato. The two most dominant biotypes, genetic groups Middle East-Minor Asia 1 (biotype B) and Mediterranean (biotype Q), have caused serious yield losses in more than 60 countries (Pan *et al.*, 2013) with annual losses of more than one billion dollars (Jiang et al., 2012).

The most common prevention strategy used by farmers against whiteflies is by agriculture cultivation method and application of pesticides. Chemical insecticides are a fast and effective means of killing whiteflies. However, the application is either governmentally regulated or under international conventions such as the Stockholm Convention on Persistent Organic Pollutants and the Rotterdam Convention on Prior Informed Consent, as they can cause health and environmental problems. For instance, in the European Union, only five neonicotinoids have been approved for the use in plant protection products (https://ec.europa.eu/food/pesticides). Other examples, glyphosate and chlorpyrifos/-methyl are active compounds that have been tightly regulated and authorized as herbicide and insecticide, respectively. Non-ecological application can result in increased resistance to pesticides; therefore, several types of preventive and curative biological control methods have been employed such as using natural predators and parasitoids (Gerling *et al.,* 2001). Integrated plant management is a prevention strategy that has become an alternative approach to control whiteflies by incorporating biological control, crop plant resistance, physical/mechanical control, and minimal chemical pesticide (Gilberson *et al.,* 2011). Additionally, using resistant plant varieties and high-quality seeds could contribute to cost-effective and environmental cost

pest management. To create resistant varieties by means of classical breeding, breeders face challenges to identify, characterize and introgress resistant alleles into a population. Modern tomato genetics have already used marker assisted selection (MAS) to determine regions and conducted functional analyses to identify genes underlying the selected traits (Foolad & Panthee, 2012). Moreover, the availability of genome sequences in combination with the development of other omics technologies will help accelerate the discovery of resistance factors that control whitefly resistance.
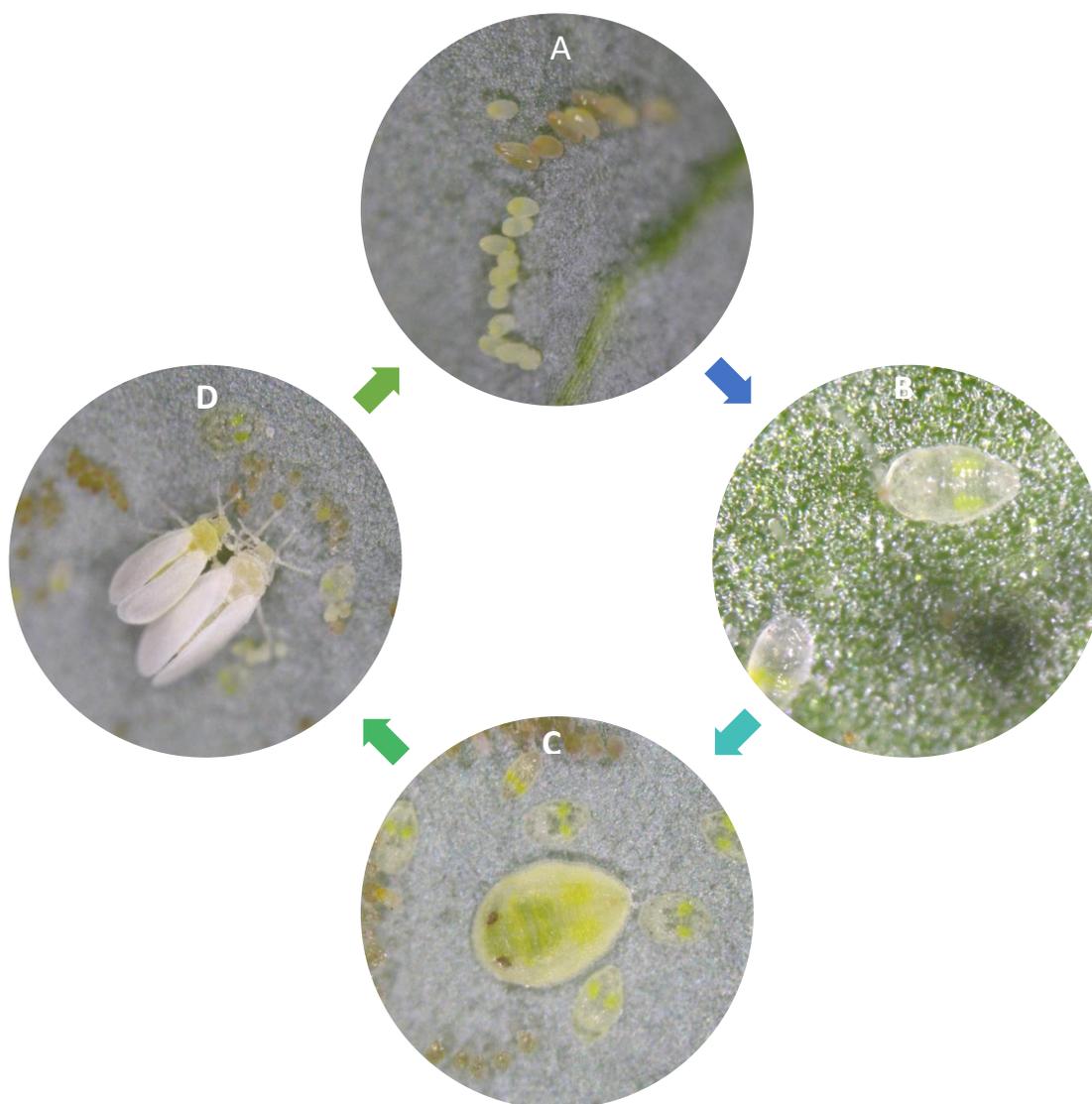
Figure 1. The lifecyle of *Bemisia tabaci.*
A) Eggs are oval-shaped, and they are attached vertically to the leaf by a small stalk that function as channel to take up fluid. B) Eggs hatched into nymphal instars, called crawlers, that are mobile and seeks suitable feeding site. C) The nymphal instars become red-eyed called the pupal phase. D) Adult whiteflies emerge from the red-eyed nymphs leaving their transparent shells (microscopy images provided by AVRDC).

1.2. Tomato systematics

Tomato, a member of the *Solanum* genus native to South America, is one of the most important vegetable crops in the world with an estimated production of over 16 million tonnes of fresh and processing tomatoes in 2019 just in the European Union alone (https://ec.europa.eu). Today, tomatoes are not only sold fresh as dessert snacks or condiments in dishes, but they also serve as model plants in research or are used in the production of therapeutic molecules which then are purified (Bergougnoux, 2014). However, most of the tomato industry still focuses on breeding to improve cultivated tomatoes (*Solanum lycopersicum*). Common cultivated tomatoes are products of domestication and continuous breeding effort. These, domesticated tomato varieties are the results of a co-evolution process not only driven by a mix of ecological and biological factors, but also human culture. Favourable traits have been naturally/synthetically selected to meet human preference, both quantitatively and qualitatively. Meanwhile, unfavourable traits such as bitterness in flavour or low sugar content which might be of importance to plant adaptation to (a)biotic stress became partially lost over the time.
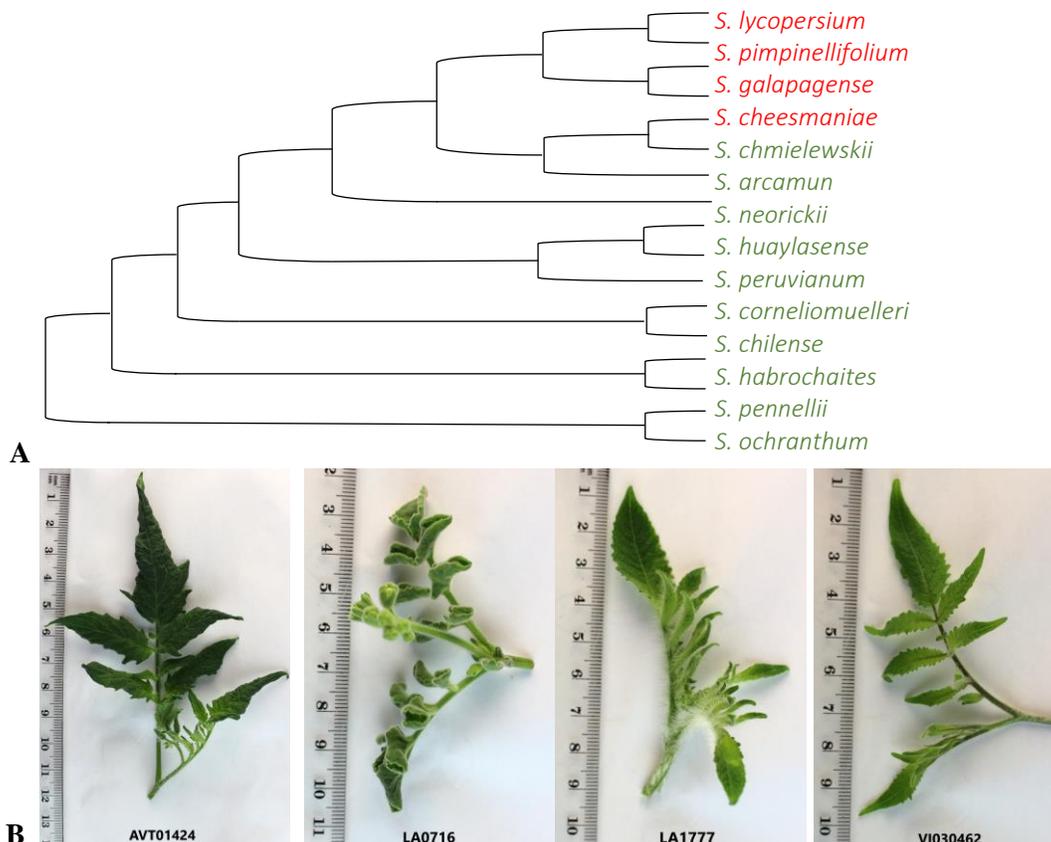
Figure 2. Phyolgeny and variation in Solanum species.
A) Phylogeny of Solanum species in the tomato clade based on conserved orthologous set nuclear loci (Rodriguez, et.al. 2009). Red: bear red/orange fruits. Green: bear green/purple fruits. B) Diversity of apical leaf (lef to right): *S. lycopersicum* (AVT01424), *S. pennellii* (LA0716), *S. habrochaites sp. typicum* (LA1777), and *S. habrochaites sp. glabratum* (VI030462).

To date, wild *Solanum* species, notably: *S. habrochaites*, *S. galapagense,* and *S. pennelli* have been shown to provide an abundant source of allelic variation valuable for breeding whitefly resistant tomatoes (Table 1). In my work, *S. habrochaites,* a species native to the highlands of Peru and southern parts of Ecuador, was selected as the male donor as it contains resistance factors which have not been fully explored such as genes regulating trichome density or biosynthesis of insecticidal compounds. As illustrated in Fig. 2A, the green-fruited *S. habrochaites* is distant to the red-fruited *S. lycopersicum*, and more like *S. pennellii* (Rodriguez *et al.,* 2009; Bolger *et al.,* 2014). *S. habrochaites* is divided into two subspecies. The *typicum* subspecies is characterized by densely haired fruits and a strongly exerted stigma which makes them difficult to self-pollinate. The *glabratum* subspecies shows less hairy leaves and stems and has smaller flowers compared to *S. habrochaites* sp. *typicum,* the outbreeder form (Fig. 2B). This subspecies is highly self and cross compatible with other cultivated tomatoes, which makes it more convenient to work with (Sifres *et al.,* 2011).

1.3. Recent advances in breeding and gene editing technologies

Genetic manipulation through mutations and plant transformation has been applied to induce genetic variation of genes/regions of interested traits. Induced mutagenesis by chemical and physical mutagens has become one of the most effective strategies for trait improvement (Chaudary *et al.,* 2019). Ethyl methane sulfonate is an example of a chemical mutagen that has been used to develop tomatoes resistant to whitefly (Gopalakrishnan & Selvanarayanan, 2009). Another chemical agent, colchicine, has been applied to eggplants (*S. melongena, S. integrifolium*) producing amphidiploid rootstocks that are more resistant to bacterial wilt disease (Ali *et al.,* 1992). Fast-neutron bombardment of seeds is an example of physical mutagenesis that induces deletions and reconstitutions of chromosomes. This method is relatively new in plant science and least explored due to high population and specialized equipment requirement (Menda *et al.,* 2004). However, the application of this tool led to the discovery of a deletion in the *Tangerine* locus that encodes a carotenoid isomerase essential to produce β-carotene that has antioxidant function, mediating plant-insect interaction (Isaacson *et al.,* 2002). Similarly, gamma ray irradiation has also been used to create populations which were used to identify brix-regulating genes in tomato (Matsukura *et al.,* 2007). Furthermore, insertional mutagenesis (e.g. T-DNA insertion mutagenesis) and targeted mutation (*e.g.* zinc finger nucleases, transcription activator-like effector nucleases, CRISPR), have provided opportunities to precisely mutate genes although their application is not easily implemented commercially due to restrictive genetically modified organism regulations in different countries. For instance, the worlds first gene edited tomato with increased gamma-aminobutyric acid called "The Sicilian Rouge High GABA" has been approved in Japan (https://sanatech-seed.com/). Suhag *et al.* (2020) have provided a summary of whitefly resistance factors which have been exploited using genetic engineering via nuclear or chloroplast transformation in transgenic plants as well as nontranstrangenic RNA interference approaches and double

stranded RNA delivery via fungal endosymbiont/entomopathogen. Alternatively, altering the genome of whitefly via Receptor-Mediated Ovary Transduction of Cargo CRISPR-Cas9-based targeting an ovary peptide ligand in *B. tabaci* is another strategy to control whitefly population (Heu *et al*., 2020).

1.4. Tomato introgression lines and exploring wild tomato genetics

An alternative effort to create variation, help broaden and enhance the basis of the gene pool is by restoring genetic diversity with exotic genetic resources in breeding programs. Challenges encountered when using wild accessions include interspecific crossability, limited data on species basic biology and characterization, unpredicted expression of allelic combination in different cultivated backgrounds, and bottlenecks in the pre-breeding process (Dempewolf *et al.,* 2017). Moreover, wild accessions have traits that are associated with poor agronomic performance (*e.g.*, low yield, smaller seed/fruits) that are difficult to break through repetitive backcrosses because they may tightly linked to a trait of interest, a phenomenon known as linkage-drag (Zamir, 2001). However, these difficulties can be overcome through research and development of analytical and molecular techniques for marker development based on the identification and characterization of genes/regions from exotic libraries. These exotic libraries consist of populations of introgression lines from a cross between wild accessions and cultivated tomatoes.

Regions of interest called quantitative trait loci (QTLs) are regions in the chromosome that correlate with variation of quantitative traits of a population. QTLs are identified by QTL analysis typically using molecular markers. This analysis is known as linkage analysis, which is a statistical test that compares the probability of obtaining the observed trait if two loci are linked. The availability of good quality phenotypic data and a sufficient number of informative markers are required to ensure the accuracy of the analysis. The results are measured by logarithm of odds (LOD) scores which are the log 10 value of the probability of an observed trait with a given linkage value over the probability with no linkage. The higher the LOD score, the higher likelihood of linkage occurring, and reciprocally. There are different algorithms that are commonly used and have been applied in this study to obtain LOD scores such as simple interval mapping (SIM), composite interval mapping (CIM) and single-trait multiple interval mapping (MIM), a technique called QTL mapping. SIM is based on simple regression of a predicted genotype at each tested QTL position on the map (Haley & Knot, 1992). CIM is a method that implements the least-squares method of Zeng (1994), like SIM but with cofactor selection. Cofactors or regressors are additional predicted markers that are selected to improve the prediction inside and outside the respective marker intervals (Broman, 1997). In MIM, cofactors are used to fit multiple putative QTL by using maximum likelihood for estimating genetic parameters (Kao & Zeng, 1999). Another method that is commonly used is multiple QTL mapping (MQM), which comprises of two step QTL mapping: 1) cofactor selection by multiple regression and backward elimination; 2) interval mapping by maximum likelihood (Arends *et al.,* 2010). These methods have been used in several QTL analyses in search for whitefly resistance (Table 1).

Table 1. Details of findings in relation to whitefly resistance in different tomato species.

Tomato accession	Findings	Characterization method	Phenotype	Reference
S. habrochaites x lycopersicum	QTL *Tv-1*, *Tv-2*, *TriIV-1* & *TriIV-2*	QTL analysis of intraspecific population	Affects oviposition and type IV trichome density	(Maliepaard *et al.,* 1995)
S. lycopersicum (Moneymaker)	*Mi-1.2* (Leucine Zipper Family)	*Agrobacterium tumefaciens*, pCGN1557, transformation	Possibly prevents whitefly stylet form reaching the phloem sieve elements.	(Milligan *et al.,* 1998); (Nombela *et al.,* 2003)
S. pennellii (LA0716)	QTL associated with AS accumulation	QTL analysis of intraspecific population	Acylsugar composition.	(Blauth *et al.,* 1998)
S. lycopersicum var. *cerasiforme*	Selection of resistant accessions	Whitefly assays	Variation in resistance due to trichome density	(Sánchez-Peña *et al.,* 2006)
S. habrochaites x lycopersicum	QTL TG313 on Chr.10, C2_At2g41680 on Chr.9, TG523/T0408 on Chr.11, & TG400/cLEG-37-G17 on Chr.11	QTL analysis of intraspecific population	Resistance to whitefly based on oviposition and type IV trichome.	(Momotaz *et al.,* 2010)
S. pimpinellifolium	Breeding line ABL 14-8	Electrical penetration graph studies & insect assay	Whitefly showed reduced ability to start probing/reach the phloem, resistance due to type IV trichomes.	(Rodríguez-López *et al.,* 2011; 2020)
S. galapagense, cheesmaniae, arcanum, glandulosum, habrochaites, neorickii, lycopersicum, pimpinellifolium	New resistance found in *S. galapagense* (PRI95004/PY-8027)	Free-choice assay, resistance screening	Resistance to whitefly associated with type IV trichome.	(Fidaus *et al.,* 2013)
S. habrochaites LYC4, *S. pennellii*, LA3791	Metabolite QTLs (mQTL)	Gas/liquid chromatography mass spectrometry (GC/LC-MS)	Resistance to whitefly.	(Van den Elsen, 2013)
S. galapagense	QTL *WF-1* Chr.2	QTL analysis of intraspecific population	Regulates the formation of type IV trichomes.	(Vosman *et al.,* 2019)

1.5. –Omics assisted tomato breeding

Next generation sequencing (NGS) has revolutionized biological science due to its ultra-high throughput, scalability, and speed to analyze whole genome and transcriptome. In the scope of tomato genomics, sequence alignment, and assembly utilized the first tomato genome (*S. lycopersicum* cv. Heinz 1706) (Bombarely *et al.*, 2011; Sato *et al.,* 2012). Since the first released version of the tomato genome, more projects were initiated not only to sequence the genomes of cultivated and wild tomato varieties but also to perform genotyping by sequencing (GBS) and generate transcriptomics data by RNA sequencing (RNAseq). GBS is a method to discover single nucleotide polymorphisms (SNP) useful for quantitative trait loci (QTL) analysis and genome-wide association studies (Poland & Rife, 2012). RNAseq is based on NGS tools and is utilized to quantify transcripts present in a biological sample at a given moment, it provides higher sensitivity and reproducibility compared to other transcriptomic methods (http://ted.bti.cornell.edu/). Analysis of differential co-expression complemented with metabolomics may assist in investigating key regulatory steps in metabolic pathways. Metabolomics studies provide biochemical profiles of tomato phenotypes by identification and quantification of small molecules that may associate with important nutritional and toxicological characteristic. Metabolomics experiments based on chromatographic separation techniques that relate to mass-spectrometry or nuclear magnetic resonance spectroscopy (NMR) are two metabolomics techniques that have been widely used to investigate metabolites beneficial for taste, fragrance, appearance, and resistance (Jorge *et al.*, 2016).

The association of QTL mapping, transcriptomics and metabolome profiling have contributed to finding tomato factors for whitefly resistance. Several volatile metabolite QTLs (mQTLs) that determine *B. tabaci* resistance in *S. pennellii* (LA791) and are related to the production of glandular trichomes (GT) were identified (van den Oever-van den Elsen *et al*., 2015). These mQTLs underlies the biosynthesis of sesquiterpenes: a) β-caryophyllene, α-humulene, and bicyclogermacrene, which co-localized with QTLs for whitefly susceptibility; and b) (E)-β-farnesene and guaia-6,9-diene, which co-localized with QTLs for resistance. In a more recent publication, Vosman *et al*. (2018) identified whitefly resistant QTL at the end of chromosome 2 that co-localizes with QTLs for type IV galndular trichomes (GT)as well as mQTLs for AS, methyl esters of flavonols myricetin, and quercetin from a recombinant inbred line population derived from a cultivated tomato and *S. galapagense*. The gene underlying this QTL was predicted to be a transcription factor regulating trichome density and is currently under investigation (personal discussion with Ben Vosman, 2019).

1.6. The role of tomato trichomes in insect resistance

Understanding plant-herbivore interaction is important to dissect the genetic basis of insect resistance, because there are antagonistic or synergistic effects when high quality fruit is bred. One can categorized the relationship between insects and plants based on the type of resistance: 1) non-preference or antixenosis is a mechanism employed by the plant to reduce colonization by insects; 2) antibiosis operates after insects have colonized and started utilizing the plant; lastly, 3) tolerance is the ability of the plant to withstand insect population. Due to their sessile nature, plants must adapt quickly for their survival and reproduction by engaging different strategies using their physical structures (*e.g.*, trichomes, wax/cuticles, thorns/spines) and by producing chemical compounds (secondary metabolites) for direct or indirect defense (*e.g.*, repelling pests or attracting their natural enemies) (Fig. 3), that can lead to induced resistance. These defense mechanisms can be present constitutively (*e.g.* phytoanticipins) or induced by stimuli (*e.g.* phytoalexins) from various elicitors, thereby altering gene expression and activating enzymes leading to the biosynthesis of volatile/non-volatile secondary metabolites.
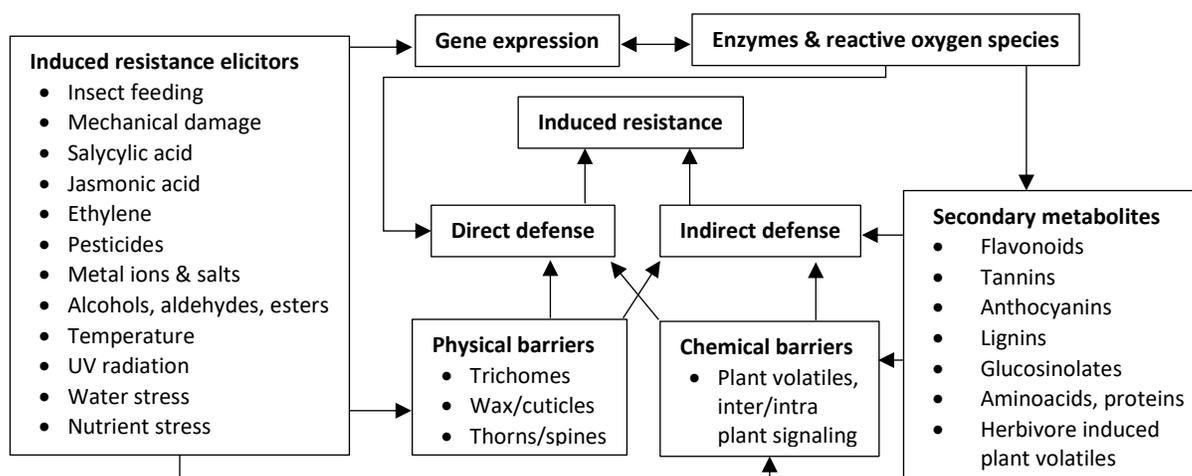
Figure 3. Mechanism of plant resistance adapted from War et al. (2012).

In this study, we focus on trichomes which are hair structures made of specialized epidermal cells protruding out of plant surfaces. They are morphologically and functionally diverse among plant species (Tissier, 2012). Non-GT are trichomes that do not possess secretory mechanism. These trichomes impede and prevent herbivores from moving, feeding and ovipositing (Xing *et al.,* 2017; Oriani, *et. al.,* 2010). Regarding abiotic stress defense, non-GT provide protection against drought and cold damage (Zhang *et al.,* 2020). They also accumulate large quantities of phenolic compounds providing protection as an optical filter against UV-B radiation (Karabourniotis *et al.,* 2019). GT, on the other hand, are trichomes with glands typically located at the peak of the trichome stalk and known to secrete blends of volatile and non-volatile compounds (Schuurink & Tissier, 2019). The bioactive compounds that are synthesized and stored in the GT have been shown to play a role in insect attraction/repulsion (Zhou *et al.,* 2017; Wang *et al.,* 2020),

toxicity/trapping/wounding (Kennedy, 2003), and inter/intra-specific interactions among plants and herbivores (Escobar-Bravo *et al.*, 2018). For example, GT have been widely studies in cannabis, where stalked GT accumulate higher levels of lipidic metabolites *i.e.* cannabinoids, compared to sessile and bulbous GT (Livingston *et al.*, 2020). Decarboxylation of cannabinoids produces cannabidiols, which have been reported to repel soilborne pathogens as well as increase the mortality rate of *Manduca sexta* larvae (Park *et al.,* 2019). Another example, the blend of volatiles (in this case terpenoids) secreted by the GT of rosemary reduces prevalence of major pest (*Frankliniella intonsa, Myzus persicae,* and *Bemisia tabaci*) populations of sweet pepper when intercropped (Li *et al.*, 2021). Overall, trichomes perform an important biological role not only as physical barrier but they also play a key role in plant adaptation to attract pollinators, protect against herbivores, and abiotic stresses by producing secondary metabolites (Bar & Shtein, 2019; Kariyat *et al.,* 2019).
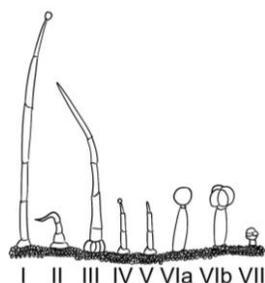
Type	Base	Stalk		Gland
I	Globular multicellular	6-10 cells	2-3 mm	Single, small, round
II	Globular multicellular	6-10 cells	0.2-1.0 mm	Non-glandular
III	Flat unicellular	4-8 cells	0.4-1.0 mm	Non-glandular
IV	Flat unicellular	Up to 3 cells	0.2-0.4 mm	Single, small, round
V	Flat unicellular	Up to 3 cells	0.2-0.4 mm	Non-glandular
VI	Flat unicellular	2 cells	0.1-0.2 mm	4 gland cells
VII	Flat unicellular	Unicellular	0.05 mm	4-8 gland cells

Figure 4. Tomato non-glandular (II, III) and glandular (I, IV, Via, VIb, VII) trichomes.
Type VIb GT are found in *S. lycopersicum* and consist of four secretory cells on one plane. Type VIa GT, found in S. habrochaites, also have 4 secretory cells but they are covered in a common envelope making it look like a single round ball. (Glas *et al.*, 2012)

Trichomes in the nightshade families come in various forms which can be categorized into seven distinct types all of which are either non-GT or GT (Fig. 4; McDowell *et al.*, 2011; Glas *et al.*, 2012). Of the GT, type VI trichomes are known to produce terpenes whereas type I and IV are known to produce AS (Blauth *et al.,* 1998, Fidaus *et al.,* 2013*;* Maliepaard *et al.,* 1995; Momotaz *et al.,* 2010). These trichomes have been shown to be involved in resistance against insects, including whitefly. Type VI trichomes in tomatoes can be categorized into round or clover shape depending on the species (Bergau *et al.,* 2015). They can be easily distinguished from other types of GT by their large gland size (approximately 50-60 µm). Although type I has a longer stalk compared to type IV, both have smaller glands and are more commonly found in *S. habrochaites* than in the cultivated species. Furthermore, type I constitutively secrete AS, which makes the leaf surface sticky, whereas type VI trichomes release volatile terpenes as signaling molecules as well as other defense compounds upon mechanical injury. Research in the scope of plant resistance has focused on increasing the density of GT such as increased type VI trichomes by the addition of methyl jasmonate (Escobar-Bravo *et al.,* 2017). Photosynthetically active radiation (Escobar-Bravo *et al.,* 2018) and mechanical damage also induce a higher density of leaf GT (Dalin *et al.*, 2008; Bloomer *et*

al., 2014). Furthermore, the search for resistance factors has led to the identification of QTLs and characterization of genes involved in the biosynthesis of volatile and non-volatile organic compounds.

1.7. Role of VOC produced in GT in plant-herbivore interaction

Volatile organic compounds (VOCs) are metabolites that plants release into the atmosphere not only to mediate intra- and inter- plant communication, but also as a response to insect infestation and function as signals for attraction of predators or parasitic insects as well as repulsion of herbivores (Li *et al.*, 2014). According to their biosynthetic origin and chemical structure, VOCs can be classified into fatty acid derivatives, tepenoids, benzenoids and phenolics (Fig. 5). These compounds can be induced, such as green leaf volatiles (GLV), or constitutively secreted such as volatiles derived from the isoprenoid/terpenoid pathway of plant flowers, leaves, and roots. In a study, priming tomato with GLV (z)-3-hexenol reduces TYLCV transmission by whiteflies by increasing flavonoid levels and inducing the transcript levels of jasmonic acid/salicylic acid biosynthetic genes (Su *et al.*, 2020). In leaves, many of the lipophilic VOCs are released through the epidermal tissues or trichomes where they are synthesized. Trichomes associated VOCs are induced by phytohormone jasmonic acid (JA) and have been reported to confer resistance against chewing-biting and cell content feeding insects (Walling, 2000; War *et al.*, 2012).
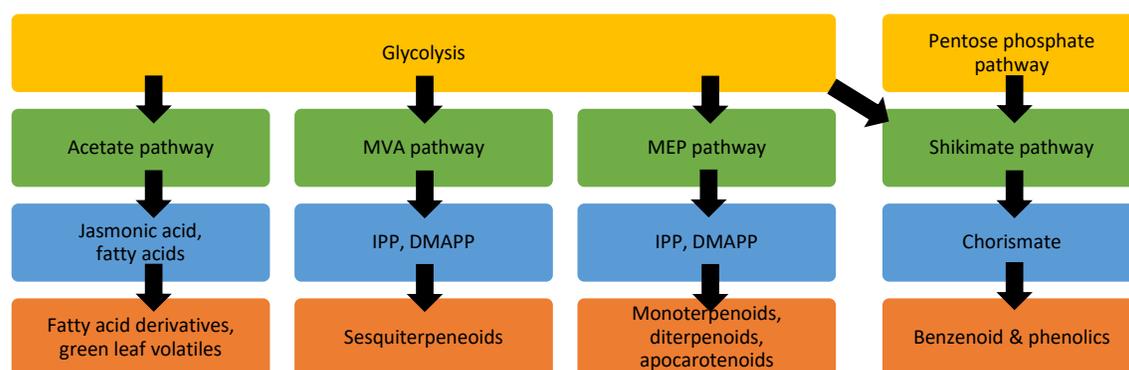
Figure 5. Biosynthetic pathways of major VOC classes in glandular trichome.
Fatty acid-derived VOCs are produced from glycolysis prior to the acetate pathway. Acetyl-coA from glycolysis is also the substrate for cytosolic mevalonic acid (MVA) pathway, while pyruvate from glycolysis which is for the plastidic methylerythritol phosphate (MEP) pathway producing different terpenoids. The benzenoids and phenylpropanoids are produced through the shikimate pathway that gets its substrates from glycolysis and pentose phosphate pathway (Bouwmeester *et al.*, 2019).

Type VI trichomes are known to produce terpenoids, specifically monoterpenes and sesquiterpenes. Terpenes are made of a five-carbon isoprene building unit and have evolved for specific ecological functions (Zhou & Pichersky, 2020). Researchers have studied how different terpenoid compositions influence whitefly choice behaviour as demonstrated by a whitefly free-choice assay on a collection of *S. pennellii* LA716 x *S. lycopersicum* Moneyberg introgression lines. The strongest repeling effect was observed from the sesquiterpenes zingiberene and curcumene, as well as the monoterpenes *p*-cymene, *α*-terpinene and *α*-phellandrene (Bleeker *et al.*, 2009). In addition to terpene composition, type VI morphology and density also effect whitefly preference. A histological study revealed that type VI trichome

of wild *S. habrochaites* with round shaped glands have big intercellular cavities, which can therefore accumulate higher concentration of secondary metabolites compared to *S. lycopersicum* with a clover shaped gland (Bergau *et al.*, 2015). In a different experiment, cultivated tomato carrying a recessive mutation called *odorless*-2 exhibited reduced GT density and impaired trichome development. As a result, there was a reduction in the quantity of secreted terpenoids as well as polyphenolic compounds. These polyphenolic compounds were detected via fluorescence microscopy in the intermediate cell between stalk and gland of type VI trichomes and have anti-oxidative activities, *e.g.* rutin, kaempferol/quercetin-glycosides, and 3-*O*-methylmyricetin (Kang *et al.*, 2010; Tohge *et al.*, 2017). Therefore, with the advancement of molecular biology and mass spectrometry techniques, breeding efforts to alter type VI trichome density/morphology and increase the emission of specific VOCs might decrease whitefly infestation.

1.8. Role of acylsugars in tomato resistance and recent findings on their biosynthesis

As previously mentioned, type I and type IV tomato trichomes are known to continuously secrete AS. AS have been studied not only in *Solanum species* but also in *Nicotiana benthamiana* (Chang *et al.*, 2020)*, Petunia axillaris* (Nadakuduti *et al.*, 2017) and *Salpiglossis sinuate* (Moghe *et al.*, 2017). AS consist of esters of sugar molecules, typically sucrose and glucose, with short-medium chain fatty acids (FAs). In the food and cosmetic industry, AS are used as emulsifier, fruit preservatives, and food additives as they are perceived as being more environmentally friendly than other synthetic surfactants (Neta *et al.*, 2015). In pharmaceutical research, AS are used as stabilizers or surfactant on vesicles for drug delivery systems (Szüts & Szabó-Révész, 2012). In plant adaptation to biotic stress, AS are secondary metabolites produced abundantly in GT. AS function as defense factors due to their stickiness/toxicity in trapping of, and the reduction of oviposition by, pests such as the sweet potato whiteflies *Bemisia tabaci* (Oriani & Vendramim, 2010), two-spotted spider mites *Tetranychus urticae* (Lucini *et al.*, 2015), as well as western flower thrips *Frankliniella occidentalis* (Ben-Mahmoud *et al.*, 2019). Weinhold & Baldwin (2011) have reported that neonate larvae of the specialist herbivore *Manduca sexta* were tagged with a distinctive odor, stemming from FAs that were hydrolyzed by hydrolases, after consuming AS, this allows predators to detect them. Free FAs have shown insect settling deterrent, antifeedant, nematicidal and toxic effects against *Bemisia tabaci* and *Myzus persicae* (Cruz-Estrada *et al.,* 2019). With regards to pathogen resistance, when the filamentous fungi *Fusarium brachygibbosum* was exposed to germination medium which contained AS, this reduced fungal spore germination, whereas removal of AS from leaf surfaces increased necrotic lesions caused by this fungus (Luu *et al.,* 2017*).*

In *Solanum* species, AS are sugar molecules commonly composed of a sucrose or glucose backbone that are accessorized with fatty acids of various chains lengths and numbers, typically C2-C12, at different positions (Leckie *et al.*, 2014; Ghosh *et al.,* 2014). The biosynthesis of AS involves a two-step process: 1)

production of a FA donor; and 2) esterification of the acyl groups to the sugar backbone (Fig. 6A). The synthesis of acyl groups in the form of acyl-Coenzyme A requires the activity of *Acyl-CoA Synthetases* (*AACS*). *AACS*, which requires Mg^{2+} as cofactor, catalyzes the ATP dependent conversion of free FAs into fatty acyl-CoA esters. Fan *et al.* (2020) have characterized a gene cluster located at chromosome 7 that was involved in the accumulation of medium chain AS in tomato trichomes. They found *Sl-AACS1* (Solyc07g043630), of cultivated M82 tomato plants to be one of the genes that specifically produced medium chain acyl-CoA (C6-C14) and confirmed this via virus induced gene silencing (VIGS) using tobacco rattle virus (TRV)-based silencing vectors. VIGS offers an attractive and quick alternative for knocking down expression of a gene without transforming the plant, silencing is induced when the viral vectors are degraded by post-transcriptional gene silencing triggered by double stranded RNA (Senthil-Kumar, 2001a; 2001b; 2014).
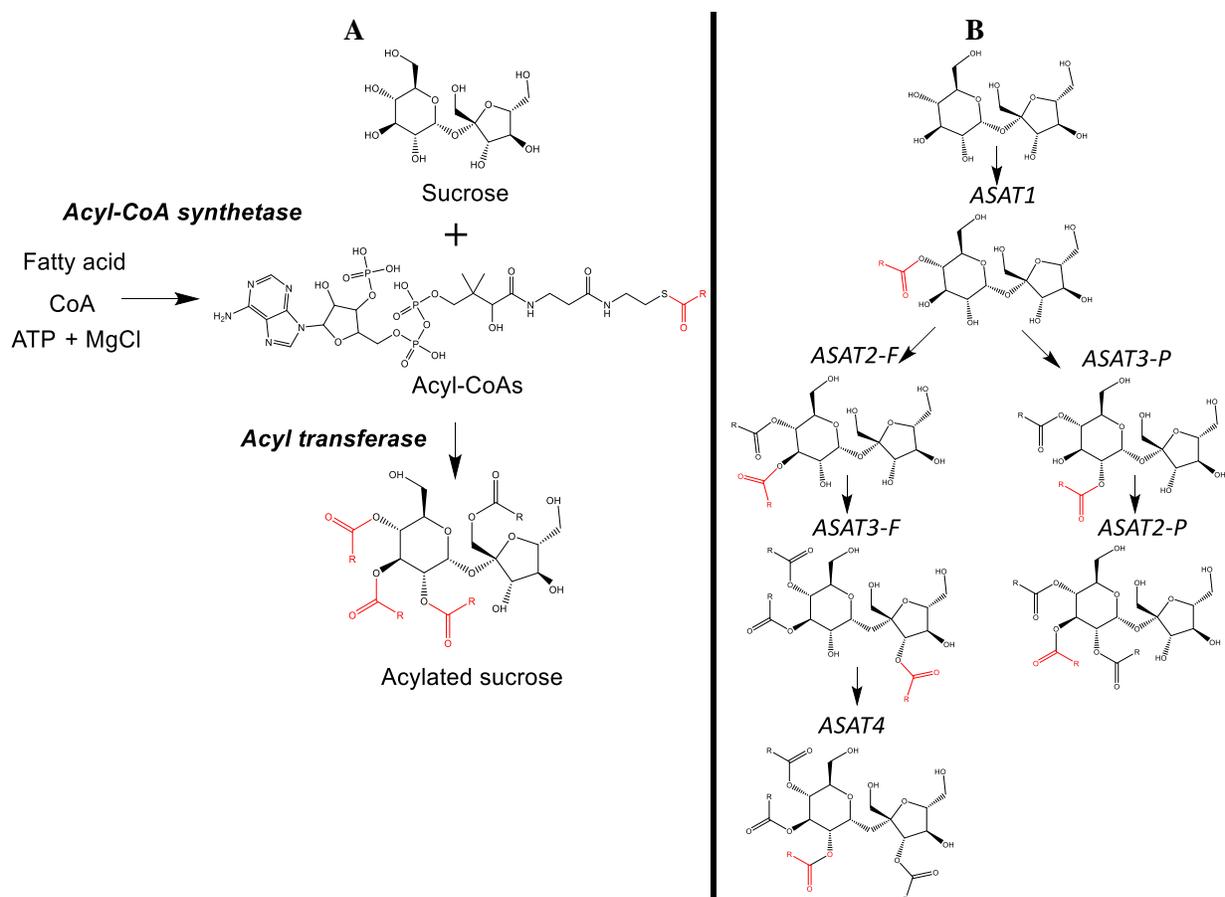
Figure 6. The biosynthesis of acylsugars.
A) The biosynthesis of acylsucrose requires acyl transferases to convert sucrose and acyl-CoAs into acylated sucrose. Acyl-CoA synthetase converts fatty acid and CoA into acyl-CoA in the presence of ATP and MgCl. R= various acyl groups. B) AS biosynthetic pathway involving different acyltransferases and their regiospecificity (Fan *et al,* 2012;2015).

The second step of the biosynthesis requires the BAHD family acyl transferases known as *Acylsugar Acyl Transferases* (*ASATs*) that are specifically expressed in type I/IV GT. BAHD (BEAT, AHCT, HCBT, and DAT) acyltransferases are Co-A dependent and they are known to be involved in the

production of secondary metabolites by transferring acyl groups supplied from the branched chain amino acid metabolism (Binder, 2010) to either O, N-, C-, or S- (St-Pierre & De Luca, 2006; D'Auria, 2000). Structural studies on the first BAHD enzyme to be crystalized, vinorine synthase, confirmed two characteristic signature motifs: 1) a HXXXDG domain located in the active center of the enzyme; and 2) a DFGWG motif located near the carboxyl terminus which fulfills a structural role (Ma *et al.*, 2005). In tomato, four BAHD acyltransferases involved in acylsucrose biosynthesis (*SlASAT1-4*) have been characterized. These four ASATs can accept fatty acids of C2 to C12 carbons and attach them at different positions of the sugar backbone creating a diverse AS composition between plant varieties and species (Fig. 6B). Thereby, the action of individual *ASAT*s follows a sequential order. *ASAT1* (Solyc12g006330) and *ASAT2-F/P* (Solyc04g012020) BAHD acyltransferases catalyze the first two steps of acylsucrose biosynthesis from sucrose at position 4 and 3, respectively (Fan *et al*., 2012; 2015). Next, *ASAT3* (Solyc11g067270) attaches acyl groups at position 3' (F type) and 2 (P type) followed by *ASAT4* (Solyc01g105580) that acylates in position 2 (Schilmiller *et al.*, 2012; 2015). Interestingly, another *ASAT* (Solyc09g014280) gene was also recently characterized by Balyan *et al.* (2020) that is a positive regulator of thermotolerance.

1.9. The role of polyamine in plant resistance

Polyamines (PAs) are aliphatic amines that are involved in various physiological processes and environmental stress responses (Roumani *et al.*, 2020). PA metabolism has been reported to be a trigger for disease tolerance as they accumulate during hypersentitive response as signalling molecules or conjugated with proteins or phenolic acids when infected by pathogens (Takahashi, 2016). Kaur *et al.* (2010) demonstrated that silencing *NaMYB8* led to the depletion of caffeoylputrescine and dicaffeoylspermidine, thereby increasing the performance of specialist and generalist caterpillars (*Manduca sexta* and *S. littoralis*). PA conjugates, hydroxycinnamic acid amides, synthesized by the formation of an amide linkage with cinnamic acids, mainly p-coumaric, ferulic, and caffeic acids have been reported to enhance plant resistance to pathogens by forming a phenolic barrier resistant to enzymatic hydrolysis (Zeiss *et al.*, 2020). Contrarily, Vilas *et al.* (2018) found that there was correlation between plant defense and *Pseudomonas syringae* colonization when external putrecine was added. As for insects, a study suggested enhanced susceptibility towards Hessian fly (*Mayetiola destructor*) in wheat (*Triticum aestivum*), this increased susceptibility was due to higher PA levels caused by low expression of an ornithine decarboxylase enzyme (Subramanyam *et al.*, 2015). In maize, induced accumulation of *p*-coumaroyltyramine in response to *S. littoralis* attack benefits the development of its larvae (Marti *et al.*, 2013). However, based on these fragmentary results, the roles of PAs in plant resistance and susceptibility have not been thoroughly elucidated to date.

Putrescine, spermidine, spermine and thermospermidine are the four most prominent PAs. Fig. 7 illustrates the biosynthesis of these four PAs. Carbon is removed by arginine decarboxylase to form

agmatine from arginine, then nitrogen is removed from agmatine to form N-carbamoyl putrescine, before being hydrolysed into putrescine. S-adenosyl-L-methionine (SAM) is a key molecule produced from L-methionine (Met) by SAM synthase that is involved in different pathways such as ethylene, and PA biosynthesis, transmethylation and transulfuration (Sauter *et al.*, 2013). Decarboxylated SAM (dcSAM), synthesized from SAM by SAM decarboxylase, donates an aminopropyl group to putrescine to synthesize spermidine, catalyzed by spermidine synthase. Another aminopropyl moiety from dcSAM reacts with spermidine to produce spermine or thermospermidine.
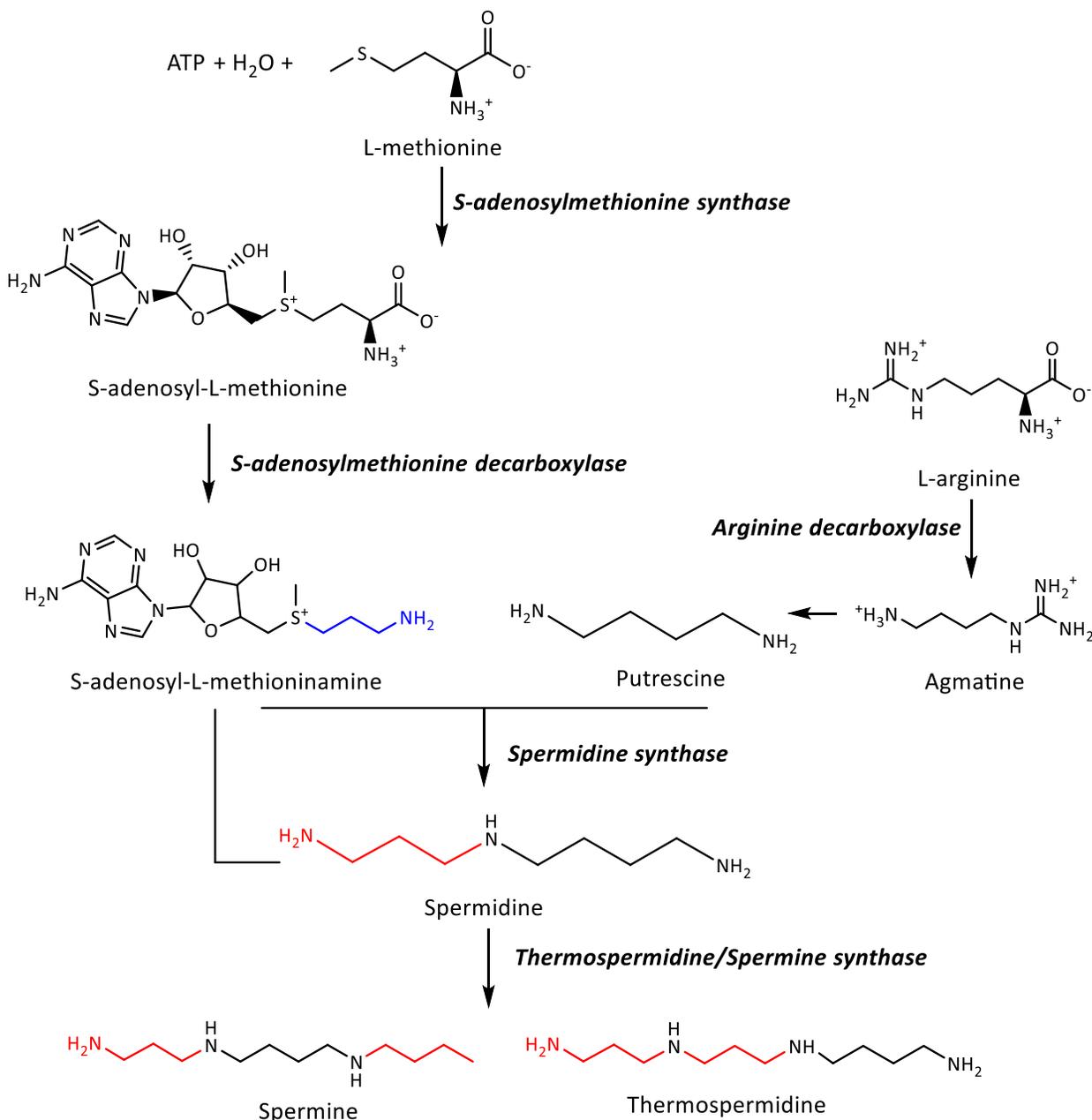
Figure 7. Biosynthetic pathway of the main polyamines.

1.10. Objectives

Most breeding efforts have focused on exploring new sources of resistance to whitefly. In this research, we do not only focus on resistance but also on susceptibility factors that can be used in marker assisted breeding in breeding programs. The main objective of this research is to characterize new sources of resistance and/or susceptibility to whitefly from crosses (F2 and BC1F2) created by AVRDC using wild *Solanum habrochaites sp. glabratum* (VI030462) and cultivated *S. lycopersicum* (AVT01424). The identification and characterization of these resistance and/or susceptibility factors involves genomics, metabolomics, and transcriptomics data analyses that will require to:

1. Develop and optimize a pipeline for untargeted metabolomics by gas/liquid chromatography mass spectrometry (sample extraction, analysis, data treatment).

2. Conduct QTL analyses of the population to find chromosomal regions that segregate with resistance and/or phenotypic traits stemming from biochemical analysis, trichome quantification, and insect assay data.

3. Perform RNAseq analyses on parental lines to help narrow down the number of candidate genes that play a role in the biosynthesis of secondary metabolites that correlate with insect resistance and/or attraction.

4. Characterize the most abundant AS and synthesize AS for *in-vitro* enzyme activity.

5. Characterize the function of candidate genes that play roles in AS (*AS acyl transferase* and *AS acyl-coA synthetase*) and polyamine (*S-adenosylmethionine decarboxylase* and *S-adenosylmethionine synthase*) biosynthesis via enzyme assays and virus-induced gene silencing.

Since this project is mainly funded by the Deutsche Gesellschaft für Internationale Zusammenarbet GmbH, the knowledge obtained, will be transferred to our project partners in developing countries.

2. Experimental procedures

2.1. Plant information and greenhouse settings

Seeds of wild tomato: (VI030462 - *S. habrochaites* sp. *glabratum)* and cultivated tomato (AVT01424 - *S. lycopersicum*) were obtained from the AVRDC World Vegetable Center (WorldVeg) and grown in the Leibniz Institute of Plant Biochemistry, Halle Saale (Germany). These plants were grown on soil in the greenhouse (65% humidity, 25°C) under long day condition (16 h light period, additional MT250DL metal halide lamps with an intensity of 165 μmol s^{-1} mm^2). Every week, plants were fertilized using (0.1% Kamasol Brilliant Blau, Compo Expert GmbH, Germany).

2.2. Whitefly feeding assay, trichome quantification and trait analysis

Whitefly feeding assay and trichome quantification were conducted by WorldVeg in Taiwan on F2 and BC1F2 population (Rakha *et al.*, 2017). The whitefly feeding assays were performed using a no-choice set up with clip-on cages. Plants were infected with five pairs of adult whiteflies per cage. Oviposition (number of eggs) and adult mortality (measured in percentage of dead whiteflies) were evaluated. Trichome quantification focuses on the density of Type IV and VI GT, where five sampling regions of approximately Ø1cm on the adaxial side of the leaf were selected randomly. Observations were made on four- and seven-weeks old plants on the F2, and only at four weeks on the BC1F2. Trait analysis hmisc, corrplot and ggpubr packages in R using was conducted to evaluate whether there are correlations between observed traits.

2.3. Genotyping, construction of genetic map and QTL analysis

Genotyping was conducted by WorldVeg in Taiwan. F2 and BC1F2 population were genotyped by sequencing. The data obtained were processed by WorldVeg and repeated in IPB with the following procedures. Non-segregating markers, markers with missing values, and markers showing skewed segregation ratio were removed leaving 1,019 markers for F2 and 2,146 markers for BC1F2 population using JoinMap 5.0 (Van Ooijen, 2006, 2009, 2011). Segregation ratio was visualized in R using the qtl package (Broman *et al.*, 2003). The construction of genetic map and Quantitative Trait Loci (QTL) analyses were performed in QGene 4.4 (Joehanes & Nelson, 2008). CIM and MIM were performed to detect regions in the chromosomes associated with variation in phenotypic trait in this study. Cofactors, functioning as genetic background control, were selected by forward cofactor selection to improve the estimation of QTLs. The significance of LOD scores obtained were tested by permutation with 1000 iterations taking α = 0.5, 0.1 and 0.05 (Churchill & Doerge, 1994).

2.4. Derivatization using fatty acid ethyl esters (FAEEs) and quantification by gas chromatography mass spectrometry

Derivatization method was adapted from (Ning *et al.*, 2016) with some modifications. Metabolites from wild and cultivated trichomes and trichome-less leaves (20mg, in triplicates) were extracted using 100% methanol and dried down under nitrogen. Extracts were then treated with 0.1 M sodium ethoxide (in ethanol) 500 µL for 20 minutes at room temperature. Following that, 0.25 mL KCl (0.9% w/v) was added, and pH was adjusted to 5 using HCl (5.25 µL, 37%). Product was partitioned with 400 µL hexane. The organic phase was inserted in glass vials and measured in GC-MS starting at 36°C for 4 minutes, ramping up to 150 °C with 10 °C/minutes, then 220 °C with 20 °C/minute and 300 °C for 2 minutes. One microliter of sample was injected to Trace GC Ultra gas chromatograph coupled to an ISQ mass spectrometer (Thermo Scientific), 30 m x 0.32 mm capillary with 0.25 µm film of ZB 5 ms (Phenomenex). Splitless mode was used for injection with the inlet temperature set to 250 °C. Helium was used as carrier gas at 1 mL/minute. Electron impact was recorded at 70 eV and MS data were collected from 50 to 450 m/z during temperature ramp. FAEE peaks were manually checked by Thermo Xcalibur Qual Browser and spectra similarity by NIST17 MS Search 2.3. Lastly, FAEE peaks were quantified using Thermo LCquant™ 2.8 and normalized with log2-fold change.

2.5. Identification of candidate genes via RNAseq

Trichomes and trichome-less leaves (by brushing method under liquid nitrogen) were collected from a pool of tomato plants (15 plants per variety) three times (with an interval of two weeks) starting with five weeks old plants. RNA was prepared from 50 mg of fine powdered samples using peqGOLD Total RNA Kit (buffer P) by VWR™, followed by DNAse treatment using DNA-free™ DNA Removal Kit by Invitrogen™. RNA quantity was quantified via NanoDrop, whereas quality was assessed using 2100 Bioanalyzer to ensure RIN value of above 8 for downstream application. Library preparation and sequencing of the samples were made by GATC Biotech AG using Illumina HiSeq 4000, 50bp strand-specific single-end reads. Read adapters were removed with cutadapt (version 1.33) and mapped by STAR (alignIntronMin 40–alignIntronMax 5000, version 2.5.2) against the reference genome assembly of tomato SL2.4 downloaded from Plant ENSEMBL database. Further data processing was performed similarly to Schubert *et al.* (2019). All RNAseq result is accessible through IPB intranet and visualized by Shiny app. Enrichment analysis was performed using gene set enrichment analysis which is a computational method that compares whether a cluster of genes are significantly different between two varying biological states (Subramanian *et al.,* 2005). Gene clusters were categorized into molecular function (MF), cellular component (CC), and biological process (BP). This analysis was followed by gene annotation by MapMan (https://mapman.gabipd.org/de). The results were then visualized in Cytoscape (Merico *et al.*, 2010). Nodes

were filtered using P-value cutoff of 0.05, Q-value of 0.25 and edge cutoff (similarity) of 0.5. Charts generated from the analysis was generated using normalized enrichment score columns and radial heat map.

2.6. Real time (RT) - qPCR

First-strand cDNA were synthesized using ProtoScript® First Strand cDNA Synthesis Kit by New England BioLabs Inc and was diluted ten times. Two qPCR primer pairs were designed to target intron spanning coding region of candidate genes using Primer3 Plus program and OligoCalc Northwestern (Appendix 1). One microliter of cDNA was mixed with 2 µL Eva Green® No Rox (Bio&Sell GmbH) and 2 pmol forward and reverse primers. Samples were processed using CFX Connect™ Real-Time PCR Detection System (Bio-Rad) under the following conditions: denaturation 95°C for 15 minutes, 45 cycles (primer annealing 95°C for 15 seconds, primer extension 58°C for 30 seconds), 95°C for 10 seconds, 65°C for 5 seconds, and 95°C for 50 seconds. Data were analyzed with CFX Manager Software (Bio-Rad) and relative gene expressions were calculated by comparative quantitation cycle method (Schmittgen *et al.,* 2008). EF1α was used as reference.

2.7. Polyamine measurement of cultivated and wild tomato leaves

Trichome and trichome-less leaves were collected from pools of 15 plants (starting from five weeks old plants, three times with an interval of two weeks) grown in IPB greenhouse. Samples (20 mg) were grinded (three 3 mm steel beads in 1.6 ml cryo-tubes for 2 x 30 s^{-1} x 30 seconds) and extracted with 100 µL of 20% MeOH, 200 mM NaCl, 10 mM KPO_4, pH 6 containing 200 pmol 1.7-diaminoheptane (DAH, internal standard). Samples were sonicated for 10 minutes and extracted for 20 minutes at 4°C before centrifuged for 5 minutes at 14,000 rpm. Supernatant was transferred to 1.5 mL tube and centrifuged.

Fluorenylmethoxycarbonyl-chloride (100 µL Fmoc-Cl 0.7 mg / 3 mL in acetone) and 0.5 M borate buffer (25 µL, pH 7.9) were added to the extracts (25 µL) followed by incubation for 10 minutes. The reaction was stopped using 20mM arginine (25 µL). Supernatant was transferred to HPLC vials after centrifugation and measured by Agilent 1260 Infinity II bioinert with quanternary pump, 1260 DAD (60 mm path length), 1260 FLD spectra using Nucleoshell RP18 2.7 µm 100/2, 250 mm (Macherey-Nagel) column at 30°C. Eluents C (water 0.2 % acetic acid) to D (acetonitrile 0.2 % acetic acid) were used starting from 40:60, then increased to 2:98 for 11 minutes before decreased to 40:60 for 22 minutes. Diode-Array Detection (DAD) was performed with 260 nm / 4 nm bandwidth, with 360 nm / 100 nm bandwidth as reference. Peaks were integrated with Agilent Open lab LC1200 software and calculation was done using the following formula:

$$\frac{\frac{Peak\ area\ Analyte}{Peak\ area\ Diaminoheptane} * amount\ (pmol)\ Diaminoheptane}{\frac{slope\ Analyte}{slope\ Diaminoheptane}} = \frac{pmol}{mg} fresh\ weight$$

2.8.   Protein purification

Candidate genes that were selected based on RNAseq data and confirmed via Q-RT-PCR were cloned in either pQE30 (Carbenicillin antibiotic selection) or pET28 (Kanamycin antibiotic selection) expression vector with different restriction sites (Appendix 2). The products were transformed into either M15Rep4 or Rosetta™-DE3 electro-competent *E. coli* strain. Colonies were checked via PCR and sequencing (Appendix 3). A single bacteria colony were grown overnight in 800 mL Luria Bertani media containing 100 µg/mL of antibiotics until 0.7 OD_{600} before subjecting to 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) and grown for 24 hours in 16°C with continuous shaking (400 rpm). Bacteria cells were harvested and washed with sterile ddH_2O by centrifugation at 10,000 g for 15 minutes at 4°C. Cells were resuspended in 20 mL buffer (TRIS 20 mM, 100mM NaCl, and 15% glycerine, adjusted to pH 7.3 with HCl). Lysozyme (10 mg/mL, Roche), DNAseI (10 mg/mL) and protease inhibitor cocktail tablet (cOmplete Tablets, Roche) was added to the culture before subjecting to cell lysis (High Pressure Cell Disruptors from Constant Systems). Lysed cells were centrifuged (10,000 rpm, 4°C, 30 minutes) and supernatant was subjected to purification via 1mL HisTrap™ HP, running 0.05 M Imidazol followed by 0.01-0.5 M Imidazol to elute the proteins. Two microliters from each fraction were resuspended in 40 µL of 2x Laemli sample buffer (Sigma) and incubated in 95°C for 5 minutes before running SDA PAGE.

A coomassie gel consists of samples and an Amersham ECL rainbow molecular weight marker (5 µL, GE Healthcare), whereas a western blot gel in addition has a 6 His western ladder (5µL, Qiagen) for a positive control (Appendix 3). Coomassie staining was performed using Roti®-Blue quick for 20 minutes and washed with water before visualized using Fusion FX6 (Vilber). Western blot was conducted using the Trans-Blot Turbo Transfer System (Bio-Rad) followed by overnight incubation in western blot buffer with the addition of anti-His6-peroxidase (Sigma). The membrane was washed 3 x 10 minutes with western blot buffer and PBS-1x, respectively. Amersham™ ECL™ prime western blotting detection reagent (GE Healthcare) was used to detect the targeted protein visualized in the Fusion. Confirmed fractions were desalted using PD-10 pre-packed columns (GE Healthcare) with equilibration buffer (Tris 20 mM, Glycerin 10 %, NaCL 20 mM, pH7.5). Desalted proteins were concentrated using Amicon® Ultracel 30K centrifugal filters and quantified using the NanoPhotometer® (extension coefficient: 1.050). The enzymes were stored at -80°C.

2.9.   Enzyme assay and kinetics of Sh-AACS

Enzyme activity experiments were performed in 50 mM Tris–HCl (pH 7.5), 100 µM CoA, 2.5 mM ATP, 2.5 mM $MgCl_2$ and 5 µg *Sh-ASCS* with a total volume of 30 µl. The reaction was started by the addition of fatty acid as substrate and incubated at 30°C with continuous shaking (300 rpm). Substrate specificity test was conducted with 100 µM of either C4, iC4, C5, iC5, aiC5, C6, C8, C10 or C12 fatty acids (FAs) for 30 minutes. Enzyme kinetics were performed on Sh-AACS2 with iC5-FA and aiC5-FA (10-100

µM) using iC5-CoA for standard curves. The reaction was stopped by the addition of 60 µL isopropanol: acetonitrile: formic acid (1: 1: 0.001) and centrifuged at 14,000 g for 15 minutes at 4°C. Unpurified (SN) pQE30 was used as control. Samples were separated with NRG mode and measured using targeted ESI-coupled to QTRAP (see 2.15, 2.17).

2.10. Enzyme assay ASAT3 and product characterization

The activity of ASAT3-LA1777, ASAT3-VI030462w, and ASAT3-VI030462c was checked by incubating each enzyme in a 30 µL reaction comprising 50 mM ammonium acetate (pH 6), and 100µM of CoA for 30 minutes at 30°C at 300 rpm (Leong *et al*., 2020). The reaction was stopped by the addition of 60 µL isopropanol: acetonitrile: formic acid (1: 1: 0.001) and centrifuged at 14,000 g for 15 minutes at 4°C. pET28 - SN was used as control. Next, a one pot enzyme reaction was conducted in a 30 µL reaction using 5 µg enzyme, 100 µM iC5-FA, 100µM MIC026-F1, 100 µM CoA, 2.5 mM ATP, 50 mM ammonium acetate (pH 6), 50 mM Tris–HCl (pH 7.5) and 2.5 mM $MgCl_2$. Since the AS product was minimal and due to limited synthesized ASubstrate, one pot enzyme reaction was not performed for product purification. Isovaleric acid-CoA (100 µM) was used as a substrate for ASAT3-VI030462w and ASAT3-VI030462c enzyme reaction. The reaction was performed, and the product was collected followed by 1: 5 (MeOH 50%) dilution. Samples were centrifuged at 14,000 g for 15 minutes at 4°C before injecting in the SPE cartridges followed by separation using Nucleoshell RP18 2.7 µm 100/2, 100 mm column (Macherey-Nagel) on SPARK coupled to QTOF. Fraction containing the AS product was collected and dried under rotary evaporator before being sent for NMR analysis.

2.11. Enzyme assay AMDs and PA measurement by HPLC

Enzyme activity experiment was performed in 30 µL reaction containing 200 mM sodium phosphate (pH 7.5), 10 mM $MgCl_2$, 200 µM thiamine pyrophosphate, and 10 µg protein in a total volume of 30 µl (modified from Holcomb & Shapiro, 1975). The reaction was started by the addition of substrate 200 µM SAM and incubated for 30 minutes at 37°C at 300 rpm. The second reaction was performed in 30 µL reaction containing 50 mM TRIS (HCl pH 7.5), 2.5 mM $MgCl_2$, 2.5 mM ATP, 200 µM L-methionine (Met) and 10 µg protein. Both reactions were stopped by the addition of 10 µL isopropanol: acetonitrile: formic acid (1: 1: 0.001) and centrifuged at 14,000 g for 15 minutes at 4°C. No enzyme and no substrate were used as negative control. Samples were derivatized using Fmoc-Cl and the solid-phase extracted as mentioned in Ziegler & Abel (2014). S-adenosyl-methionine (SAM), decarboxylated s-adenosyl-methionine (dcSAM) and S-adenosyl-L-homocysteine (SAH) were measured by Nucleoshell RP18 2.7 µm 100/2, 50 mm column (Macherey-Nagel) in HPLC with the flow rate of 500µL/min starting from 95% of solvent A (0.2% v/v acetic acid in water) for 5 minutes, isocratically increased to 50% solvent A for 1 minute, then elution with solvent B 95% (0.2% v/v acetic acid in acetonitrile and 2mL/L of glacial acetic

acid) for 1.5 min, lastly to 95% solvent A for 3 minutes. Data were extracted and T-test with α=5% was conducted to test the significance level between treatments.

2.12.   Virus-induced gene silencing (VIGS)

Tobacco rattle virus was (TRV) used for the VIGS assay (Fig. 8; Senthil-Kumar & Mysore, 2014). Fragments of approximately 500bp of candidate genes (Appendix 2) were cloned from tomato cDNA using primers with BsaI restriction sites and inserted into pTRV2-GG (ADDgene Plasmid #105349) using Golden Gate Cloning with BsaI (NEB) and T4 DNA ligase (Promega) as described in Gantner *et al.* (2018). pTRV1 and pTRV2-GOI plasmids were individually inserted into *Agrobacterium tumefaciens* GV3101 electro-competent strain. Partial fragment of P*hytoene Desaturase* (*PDS*) gene was used as positive control whereas empty vector pTRV2-GG was used as negative control. After growth in liquid culture overnight at 28°C, the cells were centrifuged and resuspended in infiltration medium (10 mM MES, 10 mM $MgCl_2$ and 0.2 mM acetosyringone, pH 5.8) with OD_{600} = 2.0. One TRV2 was mixed with one TRV1 and left at room temperature for 3 hours. The mix was then injected into two weeks old cotyledons using a syringe, five plants per treatment and the experiment was repeated in duplicate. All plant materials were grown in similar greenhouse setting environment as previously mentioned but with lower temperature at 20°C. Leaf discs located at the basal midrib of the second to fourth leaf sets were harvested after 56 days post infection (dpi) and stored at -80°C. Expression profile of individual plants was analyzed by Q-RT-PCR. AS levels were measured by LC-MS for the *Sh-AACS* and *ASAT3* silenced plants whereas, PAs were measured by direct injection and derivatization for AMD.
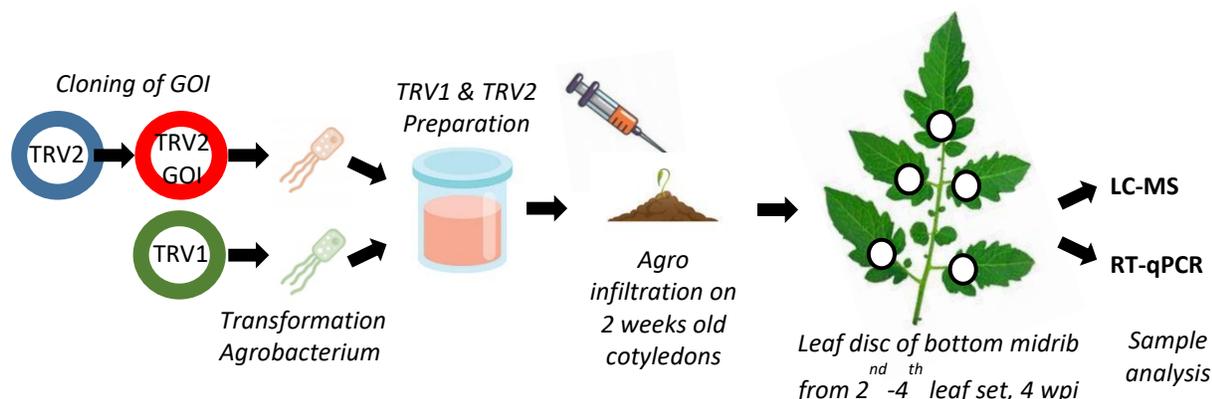
Figure 8. VIGS method from cloning to sample analysis.

2.13.   Two-phase metabolite extraction from leaf surface

Samples (70 mg) were inserted in 1.6 ml cryo-tubes (Precellys Steel Kit 2.8 mm, Peqlab Biotechnologie GmbH, Erlangen, Germany) filled with 200 mg glass beads (0.75-1.0 mm), one large steel bead (5 mm) and three small steel beads (3mm). Samples were ruptured in 900 µL dichloromethane: ethanol (2:1, -80°C) and 100µL hydrochloric acid (pH 1.5, 200mM) using FastPrep bead beating (3x20 s, speed 6.5

m/s, 1st round –80°C, 2^{nd,} and 3rd round room temperature; FastPrep24 instrument with cryo-adapter, MP Biomedicals LLC, Santa Ana, CA, USA). The aqueous phase (300 µL) was collected in a separate 2 mL tube after centrifugation (12,700 rpm, 4°C, 3 minutes). More hydrochloric acid (50 µL) was added, and 120 µL of the aqueous phase were discarded following the same procedure. Then, 600 µL of the organic phase was collected in a new Eppendorf tube (1.5 mL) and stored on ice. Tetrahydrofuran (500 µL) was added to the cryo-tubes and samples were disrupted and centrifuged. Supernatant (450 µL) was collected and combined with the first organic phase, then dried down in a stream of nitrogen gas (Turbovap LV, Biotage). After dried down, samples were resuspended in 80% methanol (180 µL) using a vortex and ultrasonic water bath for 5 minutes. Samples were then centrifuged for 5 minutes, and supernatant were inserted in glass measurement vials.

2.14. Separation of hydrophilic metabolites (NRG mode)

Separation of hydrophilic metabolites was performed on a Nucleoshell RP18 (2.1 x 150 mm, particle size 2.1 µm, Macherey & Nagel, GmbH, Düren, Germany) using a Waters ACQUITY UPLC System, equipped with an ACQUITY Binary Solvent Manager and ACQUITY Sample Manager (10 µL sample loop, partial loop injection mode, 5 µL injection volume, Waters GmbH Eschborn, Germany). Eluents A and B were aqueous 10 mmol/L tributyl amine (adjusted to pH 6.2 with glacial acetic acid) and acetonitrile, respectively. Elution was performed isocratically for 2 min at 2% eluent B, from 2-5 min at 7% eluent B, from 5-9 min at 15% eluent B, 9-15 min at 60% eluent B, 15-18 min at 80% eluent B, from 18-20 min to 95% B, and from 20-21 min at 2% B. The flow rate was set to 400 µL/min and the column temperature was maintained at 40 °C. Metabolites were detected by negative electrospray ionization and mass spectrometry.

2.15. Separation of medium polar metabolites (Pos and Neg mode)

Separation of medium polar metabolites was performed on a Nucleoshell RP18 (2.1 x 150 mm, particle size 2.1 µm, Macherey & Nagel, GmbH, Düren, Germany) using a Waters ACQUITY UPLC System, equipped with an ACQUITY Binary Solvent Manager and ACQUITY Sample Manager (20 µL sample loop, partial loop injection mode, 5 µL injection volume, Waters GmbH Eschborn, Germany). Eluents A and B were aqueous 0.3 mmol/L NH$_4$HCOO (adjusted to pH 3.5 with formic acid) and acetonitrile, respectively. Elution was performed isocratically for 2 min at 5% eluent B, from 2 to 19 min with linear gradient to 95% B, from 19-21 min isocratically at 95% B, and from 21.01 min to 24 min at 5% B. The flow rate was set to 400 µL/min and the column temperature was maintained at 40 °C. Metabolites were detected by negative and positive electrospray ionization and mass spectrometry. Extracted medium-polar metabolites were processed by reverse phase LC-ESI-MS negative and positive mode using Acquity UPLC (Waters) and TripleTOF 5600 mass spectrometer. Extracted volatiles from leaf surface were measured by injecting 1 µL of extract in a Trace GC Ultra gas chromatograph coupled to an ISQ mass

spectrometer (Thermo Scentific), 30 m x 0.32 mm capillary with 0.25 µm film of ZB 5 ms (Phenomenex). Quality of the data generated by LC-MS were checked using the Analyst 1.6 TF (Sciex, Toronto, Canada) software.

2.16.   Mass spectrometric analysis

Mass spectrometric analysis of small molecules was performed by two strategies: Targeted MS/MS via multiple reaction monitoring (QTRAP6500) and untargeted via MS-TOF-SWATH-MS/MS (TripleToF 5600, both AB Sciex GmbH, Darmstadt, Germany) operating in negative ion mode and controlled by Analyst 1.7.1 and 1.6 TF software (AB Sciex GmbH, Darmstadt, Germany). The source operation parameters were as the following: ion spray voltage, -4500 V / +5500 V; nebulizing gas, 60 psi; source temperature, 450 °C (QTRAP) 600 °C (600 TripleToF); drying gas, 70 psi; curtain gas, 35 psi. For APCI, a nebulizer current of 3 units was used. TripleToF instrument tuning and internal mass calibration were performed every 5 samples with the calibrant delivery system applying APCI negative tuning solution, respectively (AB Sciex GmbH, Darmstadt, Germany). MRM data acquisition parameters are listed in Appendix 4. TripleToF data acquisition was performed in MS^1-ToF mode and MS^2-SWATH mode. For MS^1 measurements, ToF masses were scanned between 65 and 1250 Dalton with an accumulation time of 50 ms and a collision energy of 10V (-10V). MS^2-SWATH-Experiments were divided into 26 Dalton segments of 20 ms accumulation time. Together the SWATH experiments covered the entire mass range from 65 to 1250 Dalton in 48 separate scan experiments, which allowed a cycle time of 1.1 s throughout all MS/MS scans a declustering potential of 35 (or -35 V) was applied. Collision energies for all SWATH-MS/MS were set to 35 V (-35) and a collision energy spread of ±25V, maximum sensitivity scanning, and elsewise default settings.

2.17.   MS non-targeted data processing

Tomato plants were categorized into resistant (R: more than 70% dead whiteflies), neutral (N: 40-70% dead whiteflies), and susceptible (S: less than 40% dead whiteflies), based on adult mortality. Moreover, wild parent (W), cultivated parent (C), F1 and BC1F1 were also included in the analysis. The pipeline for non-targeted metabolomics analysis involves several software (Fig.   9). Raw data generated from LC-MS and GC-MS were converted into abf format using Reifycs Abf converter. Converted data were then preprocessed (feature extraction, deconvolution of MS1 and MS/MS features, alignment, and precursor ion prediction) in MS-DIAL (Tsugawa *et al.,* 2015; https://youtu.be/SOJvPhCsubQ). Features with predicted identity were identified via a concatenated library which includes available open-source spectra (MassBank, GNPS, ReSpect), commercial (NIST17), in-house IPB library as well as WeizMass library for LC-MS and NIST17 for GC-MS. Alignment results which includes raw data matrix (height) and representative deconvoluted spectra (msp format) were imported to MetFamily (Treutler *et al.*, 2016) for principal component analysis (PCA; Pareto scaling), hierarchical component analysis (HCA; Jaccard

intensity-weighted distance function) and assignment of metabolite families (filtered using the default setting with min spectrum intensity of 1000, MS/MS peak proportion of 0.05, and neutral losses based on fragment versus precursor for LC-MS data and fragment versus fragment for GC-MS data). Data were also imported in MetaboAnalyst for additional data filtering to remove non-informative variables by interquartile range (IQR <25%), sample normalization by median, data transformation by generalized logarithm, and data scaling by mean centering. One-way ANOVA and Fisher's least significant difference tests (adjusted P-value FDR cutoff: 0.05) were performed to obtain significantly different features between lines that are resistant and susceptible. Significant features were revised in Thermo Xcalibur Quality Browser (GC-MS) or Peakview (LC-MS) and ontologies were re-annotated based on Classyfire output. Manual quantification as performed either in Thermo LCquant (GC-MS) or LC MultiQuant (LC-MS) software. Hypothetical structure elucidation of unknown features was performed in ChemDraw.
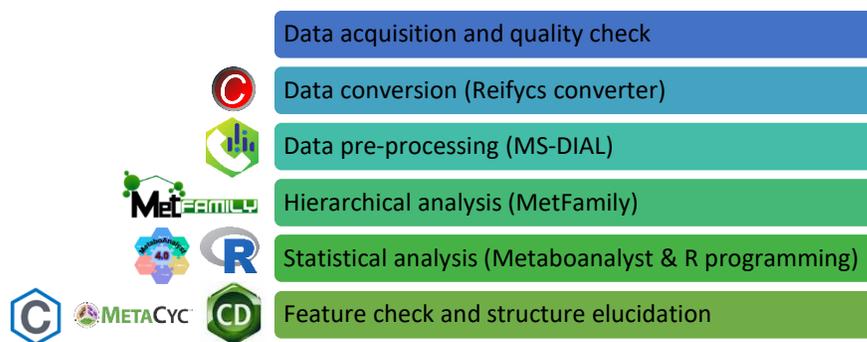
Figure 9. Data processing pipeline for non-targeted metabolomics using a combination of tools.

2.18.   Nuclear magnetic resonance (NMR)

Samples were measured via the Varian/Agilent VNMRS 600 NMR spectrometer operating at a proton NMR frequency of 599.829 MHz with a 5 mm inverse detection cryoprobe. Samples were dissolved in 0.75 mL CD_3OD (99.96%) with a final concentration of approximately 20 mmol/L. ^{1}H NMR spectra were recorded with a digital resolution of 0.37 Hz/point, pulse width of 2.2 µs (30°), relaxation delay of 0.27 s, acquisition time of 2.73 s, and number of transients of 40. Two-dimensional NMR (gHSQCAD and gHMBCAD) spectra were recorded using standard CHEMPACK 8.1 pulse sequences and parameter sets implemented in Varian VNMRJ 4.2A spectrometer software. The heteronuclear single quantum correlation (HSQC) experiments were optimized for [1]JCH = 146 Hz with distortion-less enhancement by polarization transfer-like editing and ^{13}C-decoupling during acquisition time. The heteronuclear multiple bond correlation (HMBC) experiment was optimized for a long-rang coupling of 8 Hz; a two-step [1]JCH filter was used (130-165 Hz). [1]H and ^{13}C chemical shifts were referenced to internal tetramehylsilane (0 ppm).

2.19.   Acylsugar purification, structure characterization and nomenclature

One kilogram of leaves harvested from fifteen wild tomato (VI030462) plants grown in the greenhouse were harvested. Dried methanol surface leaf extracts were collected and pre-purified via silica (SiOH) column. Hydrophobic compounds and chlorophylls were removed using 60:40 (dichloromethane [DCM]: ethyl acetate [EtoAc]), followed by 10:80:10 (DCM: EtoAc: methanol [MeOH]) to elute the AS (60mg). Three most abundant AS: AS2-S3:15(5,5,5), AS3-S3:20(5,5,5,5), and AS4-S3:22(5,5,12) was selected and purified by SPE cartridges on SPARK coupled to QTOF and PICO valve one fraction at a time in H_2O: MeOH, 1:1. The nomenclature system was based on Schillmiller *et al.*, (2012) for instance AS4, S represents sucrose, 4 indicates the total number of acyl chains, 22 is the sum of the number of carbon molecules in the acyl chains and (5,5,12) describes the length of each individual acyl chains. Uncollected AS from the SPARK and LC-MS wash were purified by two-phase extraction with DCM. The organic phase was dried down by rotatory evaporator and reinjected in the SPARK system. Purified extracts were quality checked via LC-MS. Structure of purified AS were confirmed via NMR analysis. The HSQC and HMBC experiments were conducted to characterize the AStructures.

2.20.   Synthesis of acylsucrose

Acylsucrose synthesis was performed by combining properly protected and activated glucose as glycosyl donor (MIC021) and properly protected fructose as acceptor (MIC010). MIC025-F1/F2 were used in this study. Detailed synthesis protocol is provided in Appendix 5.

3. Results

3.1. Characterization of *S. habrochaites sp. glabratum* (VI030462) and cultivated tomato (AVT01424) phenotype and breeding population

AVT01424 (*S. lycopersicum*) and VI030462 (*S. habrochaites sp. glabratum*) exhibit different phenotypes. Like *S. habrochaites*, the *glabratum* subspecies VI030462 are known to produce green fruited tomatoes approximately 1.5-2 cm in size with thick hard hairy skin on the outer layer of the fruit. This characteristic differs from cultivated tomatoes, which are red fruited and larger in size (Fig. 11 C-D). Fig. 11 A-B and C-F illustrates how both tomato species differ in trichome density and type. Type II and III non-GT were more abundant in the wild tomato especially around the midrib and edge of leaves compared to the cultivated tomatoes. Focusing on GT, type VI trichomes known to produce terpenes and flavonoids, dominated the surface of both species. Specifically, for *S. lycopersium*, the clover-shaped type VI trichomes were densely populated around the abaxial central side of the leaf surface and became sparse towards the edge of the leaf. On the other hand, fully round type VI trichomes of VI030462 were distributed homogenously across the leaf. Type I and IV GT, known to produce AS, were abundantly found in VI030462 but were difficult to detect in AVT01424.

Crosses were made by AVRDC WorldVeg in Taiwan using the two tomato varieties by pooling pollen from VI030462 and pollinating AVT01424 (Fig. 10). AVT01424 is a standard breeding line used for crossing by AVRDC whereas VI030462 has shown to harbor resistant factors to whitefly as well as spider mites in a no-choice assay (Rakha *et al.,* 2017). An F2 population was created from F1 sib-mate crossed plants that were selfed producing 172 plants. Pollen from the five most resistant F2 plants were collected and use to pollinate the cultivated parent parents, producing 180 BC1F2 plants. Both F2 and BC1F2 populations were phenotyped and metabolomics data was analyzed to correlate them to observed traits.
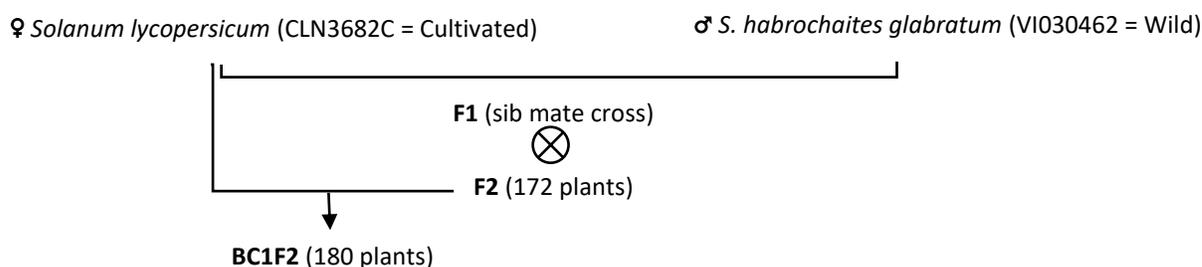
Figure 10. Breeding scheme showing population derived from VI030462 (male donor) and AVT01424 (female donor) cross produced by AVRDC in Taiwan.
F: population from a cross, BC: backcross population crossed to female cultivated parents, ⊗: selfing.
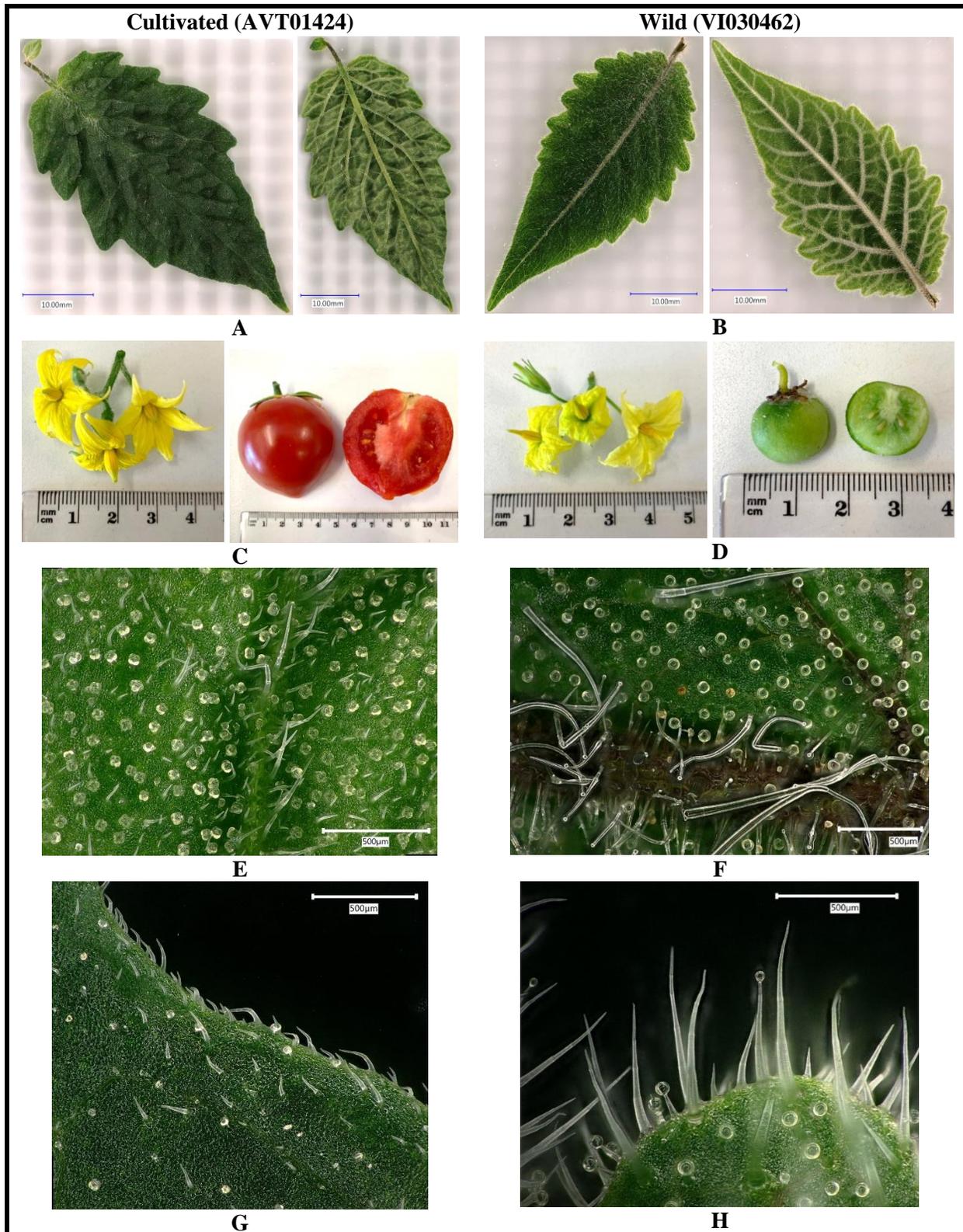
Figure 11. Morphological differences between *S. lycopersicum* - AVTO1424 and *S. habrochaites sp. glabratum* - VI030462.
A&B) Whole leaf images of adaxial and abaxial of leaf. C&D) Flower and ripen fruits. E&F) Adaxial surface of leaf bottom midrib.
G&H) Leaf edge. Captured using VHX-6000 Keyence digital microscope.

3.2. Brief overview of comparative transcriptomics clusters between cultivated and wild tomatoes

To provide an overview of differentially expressed genes between the cultivated and wild tomato species in this research, RNA from trichomes and trichome-less leaves of wild and cultivated tomatoes was sequenced. 35,216 genes passed through the initial filtering criteria for minimum expression levels and we ended up with 5,647 differentially expressed genes (false discovery rate-adjusted $P < 0.05$, $\log_2$ fold-change ±1). There were 456 genes that were upregulated, and 1,450 genes were downregulated in the wild tomato trichomes compared to cultivated trichomes by at least one log fold change (LFC). There were 1,355 genes upregulated, and 1,568 genes downregulated in the wild trichomes, compared to the wild leaves by at least one LFC. Whereas 1,027 genes were upregulated and 701 genes downregulated in the cultivated trichomes compared to the cultivated leaves by one LFC. Lastly, 1,155 genes were downregulated and 1,024 genes upregulated in the wild leaves compared to the cultivated leaves.

Table 2.Number of gene clusters compared between cultivated and wild, trichomes and leaves.

Data sets	Clusters
Cultivated leaves vs. cultivated trichomes	222
Cultivated leaves vs wild leaves	8
Cultivated trichomes vs. wild trichomes	8
Wild leaves vs. wild trichomes	245

Genes that were extracted were clustered and the expressions were compared between treatments via an enrichment analysis. When comparing gene clusters of leaves and trichomes between the two species respectively, there were no more than eight individual clusters each, displaying significantly different gene expression (Table 2). By contrast, there were more than 200 clusters that were significantly different when gene clusters between leaves and trichomes for each individual species were compared. All eight gene clusters that showed a higher significantly different expression in the cultivated leaves compared to the wild leaves include structural constituent of nuclear pores (molecular function - MF), cytoplasmic translation linked (biological process - BP) to cytosolic large ribosomal subunit (cellular component - CC), DNA replication initiation (BP) linked to DNA replication (BP), and cytosolic small ribosomal subunit (CC) linked to a gene cluster which was not named. Five of the eight gene clusters in trichomes were significantly upregulated in cultivated trichomes compared to the wild trichomes, these clusters relate to ammonia-lyase activity (MF), carboxylic metabolic process (BP), ferrous iron binding (MF), histone deacetylase activity (MF) and mitochondrial respiratory chain complex I (CC) (Appendix 6), conversely, gene clusters for phosphatidylcholine 1-acylhydrolase activity (MF), ionotropic glutamate receptor activity (MF) and ionotropic glutamate receptor signaling pathway (BP) were downregulated in the cultivated trichomes. Interestingly, one gene found within the ammonia-lyase activity cluster (MF), *Phenylalanine Ammonia-Lyase* (Solyc09g007910), was reported to catalyze the initial step in phenylpropanoid synthesis from L-phenylalanine into other derivatives which include salicylic acid as well as phenolics (Lin *et. al.,* 2019).

Salicylic acid plays an important role in plant development and is upregulated during infestation. Genes categorized under carboxylic metabolic processes (BP) that were identified include *Tryptophan Decarboxylase 2* (Solyc07g4280), which seems to play a vital role in plant flower and fruit development,. T*ryptophan Decarboxylase 5* (Solyc03g045020) was also detected via RNAseq but was not successfully characterized in their study (Pang *et al.*, 2019). The last two interesting genes in this cluster were *2-Isopropylmalate Synthase 1* & 3 (Solyc06g053400 & Solyc08g014230), which affect AS composition in cultivated and wild tomatoes (Ning *et al.*, 2015) (Smeda *et al.,* 2016).

3.3. Metabolite profiling of *S. habrochaites sp. glabratum* (VI030462) and cultivated tomato (AVT01424) F2 and BC1F2 population

3.3.1 Profiling volatile compounds of *S. habrochaites sp. glabratum* (VI030462) and cultivated tomato (AVT01424) F2 and BC1F2 population by GC-MS

Results from GC-MS measurements of leaf extracts enabled the isolation of a total of 2,207 and 1,428 unique masses (features) to retention times (RT) for the F2 and BC1F2 populations, respectively. Identifier masses (m/z) that were picked per feature were not based on the monoisotopic mass, but on the the most abundant ion found within each peak. Among the extracted features, 119 (F2) and 184 (BC1F2) features were annotated based on 75% similarity to the spectra library. Data imported in MetFamily resulted in 2,201 (F2) and 1,428 (BC1F2) features with quantitative information. Using MetFamily, those features were processed by principal component analysis (PCA) and Pareto scaling. PCA score describes the direction of the principal components in relation to the observation, whereas loading score describes the direction of the principal components. Typically, scores that are greater than 0.75 are considered as strong and lower than 0.50 are considered as weak. PCA scores and loadings for the F2 and BC1F2 showed that the model using only the first two principal components explained 20.2 %: 31.8 % and 15.4%: 30.9 % of the variance in the datasets with 37.1 %: 22.6 % and 25.9 %: 21.9 % (Q^2) as a measure of consistency between the original and the cross-validation predicted data (Fig. 12 & 13). Both PCAs of F2 and BC1F2 populations showed separation from the parental lines, however, this separation was not clear between the resistant and susceptible progenies. Metabolic families were manually annotated based on similarities of characteristics fragments to known metabolic families in MetFamily's PCA loadings plot, because heirarchical component analysis (HCA) failed to group features due to the high share of ubiquitous small fragments, which were produced extensively by fragmentation after hard electron impact ionization. Annotated features were dominated by alkane, fatty acyl, and terpene cluster.

ANOVA analyses were performed on the data output from MetFamily (quantitation ion of each feature) to compare features that were abundant but showed different intensities between the resistant (R) and susceptible (S) lines, revealing four and eleven features from the F2 and BC1F2 population respectively

(Table 3, Fig.   14). Most significant features were unknowns, and therefore re-annotated based on the highest percentage similarity in the Excalibur Quality Browser-NIST17. Annotated features were classified into alkane, carboxylic acid derivatives, fatty acyl, monoterpenes, sesquiterpenes, diterpenes, benzoic acid derivatives, phenol, sulfurous acid, or quinone/hydroquinone.
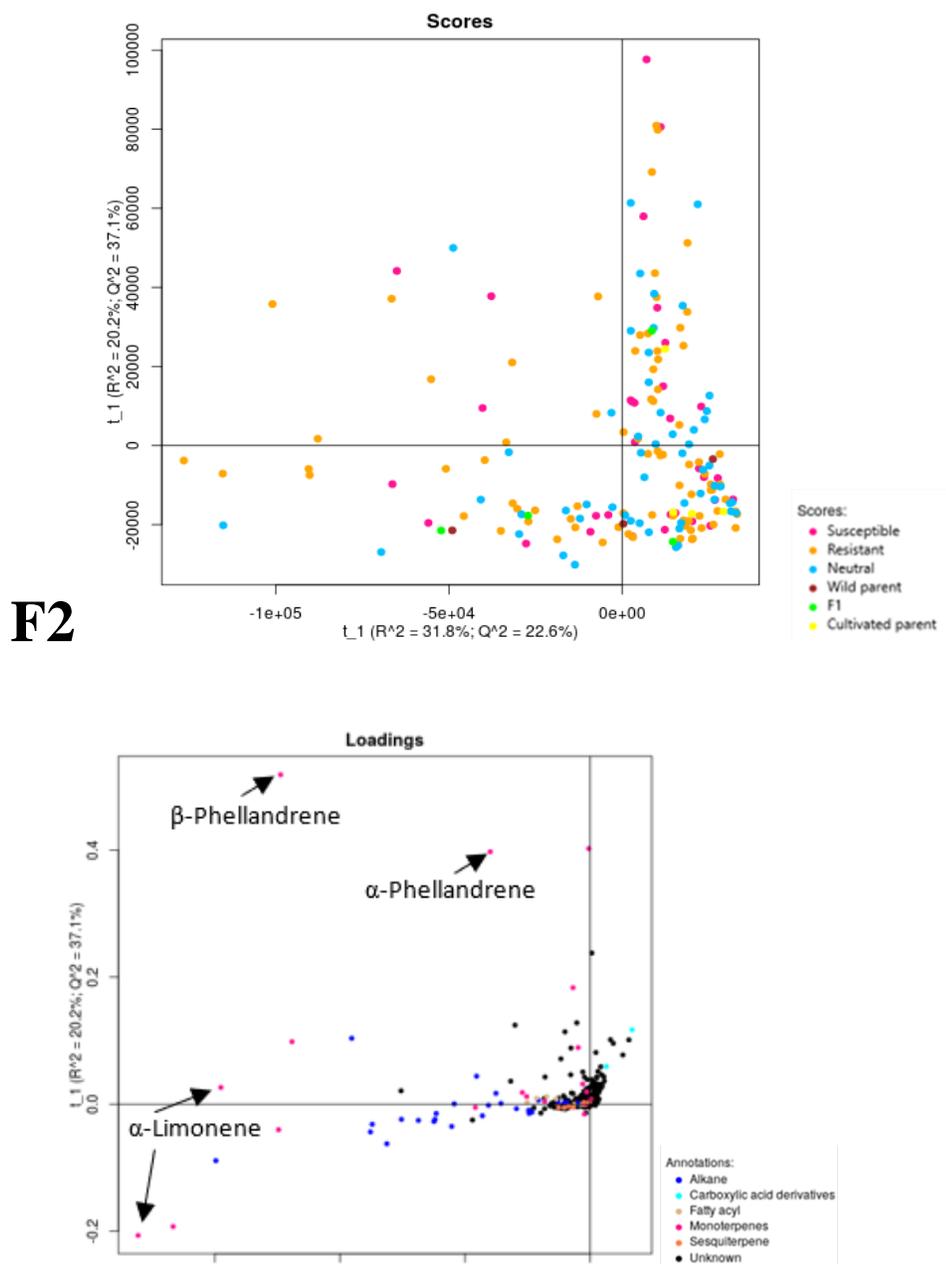
Figure 12. PCA of the F2 population obtained from GC-MS.
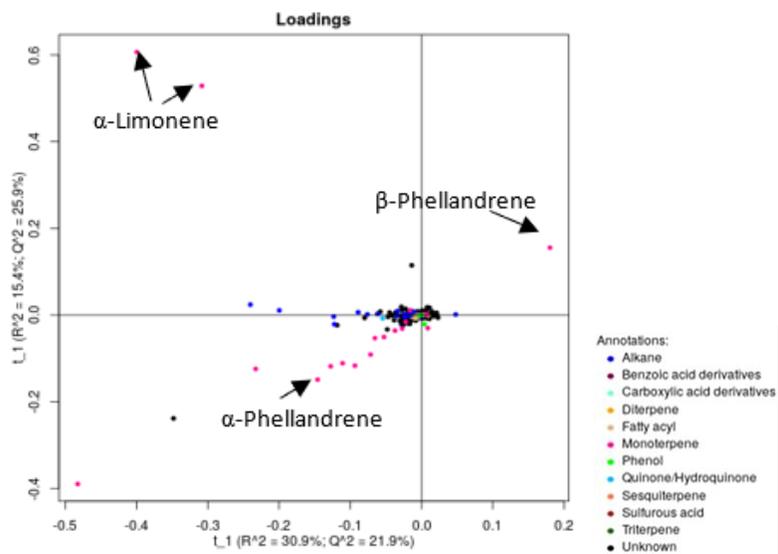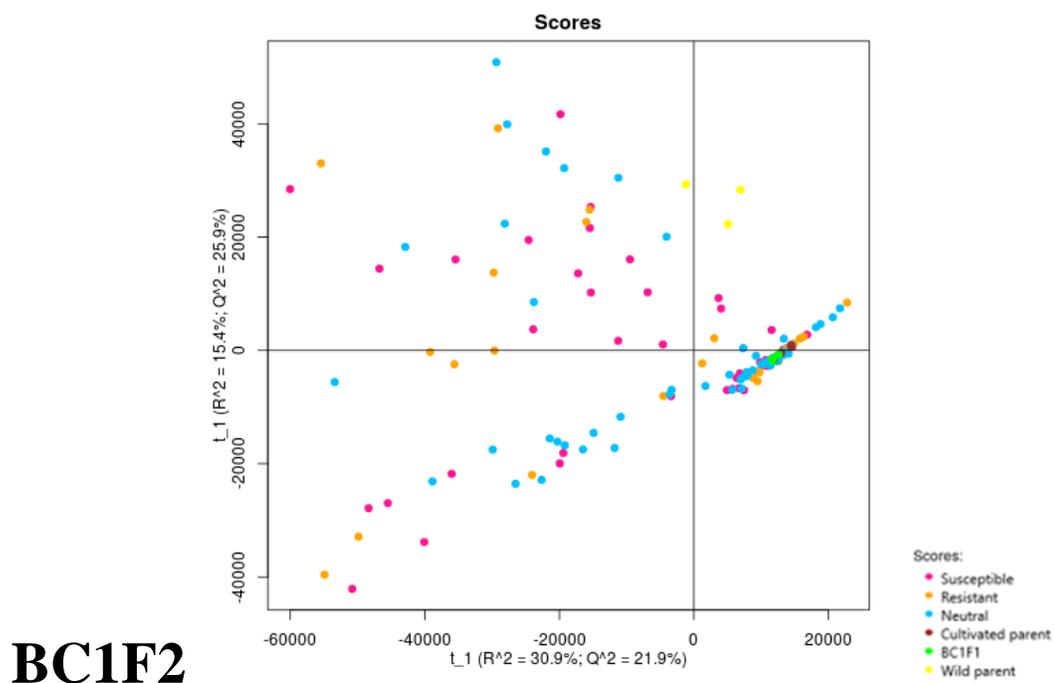
BC1F2

Figure 13. PCA of the BC1F2 population obtained from GCMS.

Table 3. GC-MS result of metabolic features that are significantly different between resistant and susceptible lines.
Features were manually checked and re-annotated based on highest percent similarity. Red: unlikely annotation or tomato metabolite. Identifier m/z was based on the most abundant ion.

RT	Identifier m/z	Putative metabolite name (% similarity to spectra database)	Ontology
F2			
14.162	60	N-decanoic acid (45.5%)	Fatty acid
16.75	51	Azulene (6.3%)	Sesquiterpene
17.971	85	Valeric anhydride (15.5%)	Fatty acid
18.573	85	Prop-2-yn-1-yl 2-methylbutanoate (23%)	Fatty acid
BC1F2			
5.062	93	3-Thujene (28.4%)	Monoterpene
5.208	136	α-Pinene (22.9%)	Monoterpene
7.042	119	P-Cymene (27.1%)	Monoterpene
7.586	71	Octane, 2,3,6,7-tetramethyl (14.1%)	Alkane
15.392	119	γ-Muurolene (6.97%)	Sesquiterpene
25.708	79	9-Octadecanoic acid (Z)-methyl ester (12.2%)	Fatty acid
31.176	57	N-Octacosane (13.4%)	Alkane
35.29	97	Uridine 5-monophosphate (16.3%)	Pyrimidin nucleotide
38.133	109	β-Amyrin (50.5%)	Triterpene
38.208	57	Dotriacontane-1-iodo (26.1%)	Alkane
38.309	57	Phenol-2,4-bis(1,1-dimethylethyl) phosphite-3:1 (27.2%)	Phosphite ester
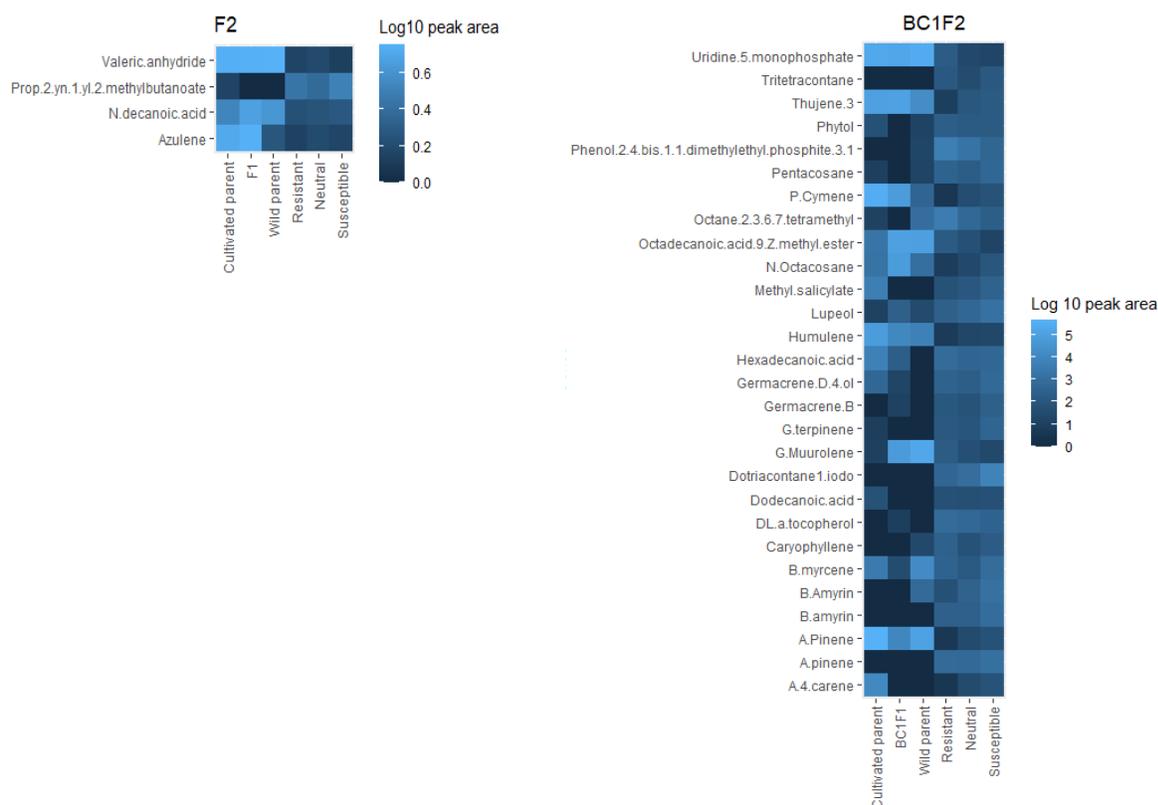
Figure 14. Heatmap of features from the F2 and BC1F2 population showing log10 relative abundance of peak area.

Alkanes and fatty acid derivatives

Alkanes are characteristic components of plant cuticle and epicuticle waxes and are amongst the most abundant in quantity and number of features found in surface extract of the leaves (Bauer et al., 2004). Some examples of alkanes are pentacosane (m/z 71; RT 26.08) and tetracotane (m/z 57; RT 33.27 min). Amongst the alkanes, two features that were lower in the resistant lines are n-octacosane and dotriacontane-1-iodo. Dotriacontane has been detected tomatoes, however when conjugated with iodine, it is an organic contaminant present on microplastics; from the sample tubes (Campanale *et al*., 2020). In pepper, it was shown that high levels of long chain fatty alkanes correlate with increased susceptibility to thrips *(Frankliniela occidetalis)*, a generalist with similar feeding strategy as whiteflies (Macel *et al.,* 2020).

FAs, especially the long and very-long chains, are also known to make up the cutin and the cuticular wax. MSDial identified fatty acyl features with C12 (dodecanoic; m/z 102; RT 18.39 min) and C16 (hexadecanoic acid; m/z 73; RT 23.39 min) that more abundant in cultivated parents compared to the wild parents. Putative N-decanoic acid (C10), valeric (C5) anhydride, and 9-octadecanoic acid (Z)-methyl ester (methyl oleate, C18), clustered under fatty acids and were more abundant in resistant lines. However, prop-2-yn-1-yl 2-methylbutanoate had lower relative abundance in the resistant compared to the susceptible lines. Short and medium chain fatty acids are also known to be the major compounds of AS secreted by type I and IV GT. Furthermore, AS are more abundant in wild tomatoes compared to cultivated tomatoes.

Terpenes

Terpenes found in the analysis were sub-categorized into monoterpenes, sesquiterpenes, ditepenes and triterpenes. An example of diterpene is phytol (m/z 71; RT 25.41 min), a precursor of tocopherol and chlorophyll (Valentin *et al*., 2006), which was shown to have an anti-herbivory affect to cabbage looper *(Trichoplusia ni)* larvae (Neupane & Norris, 1991). In data processing by MSDial, a synthetic tocopherol peak annotated as DL-α-tocopherol (m/z 165; RT 36.41 min) was identified and the family cluster was annotated as quinone/hydroquinone in MetFamily. Content of both compounds was on average slightly elevated in the resistant compared to the susceptible lines.

Four examples of sesquiterpenes are caryophyllene (m/z 93; RT 14.92 min), humulene (m/z 93; RT 15.54 min), germacrene D-4-ol (m/z 81.175; RT 18.40 min) and germacrene B (m/z 121; RT 17.37 min). Researchers have shown that *Sesquiterpene synthase 1 (SST1)* from *S. habrochaites* catalyzes the conversion of farnesyl diphosphate to germacrene B whereas *Sesquiterpene Synthase 2 (SST2)* to germacrene D (van der Hoeven *et al.¸*2000). *SST1* in *S. lycopersicum* produces β-caryophyllene and α-humulene whereas *SST2* is a non-functional gene. In this experiment, the abundance of both germacrenes and caryophyllene were low in the parental lines compared to the BC1F2. On the other hand, humulene levels were lower in the BC1F2 compared to the parental lines.

An example of tritepene is lupeol (m/z 95; RT 38.44 min). It was reported that lupeol was significantly higher in a cassava variety susceptible to whitefly infection (Perez-Fons *et al.,* 2019). A similar trend was observed in this research, in that susceptible lines had a slightly higher abundance compared to resistant lines. Amyrins are typically found in the intra-cuticular wax layer and can be distinguished into α-, β-, and γ-amyrins. The composition of the three amyrins differ between tomato varieties depending on which oxidosqualene cyclases are expressed (Wang *et al.*, 2011). A triterpene feature annotated as putative β-amyrin (109 m/z; 38.133 min) was identified and showed higher abundance in the susceptible than the resistant lines. α-amyrin (m/z 218; 38.15 min) was identified by MSDIAL (Fig. 14) and showed a similar trend even though the ANOVA result did not show a significant difference between resistant and susceptible lines. γ-amyrin was not detected, possibly due to low abundance.

Examples of monoterpenes found in the data set were Ɣ- terpinene (m/z 93; RT 7.77 min), β-myrcene (m/z 93; RT 6.33 min), α-4-carene (m/z 93; RT 8.31 min), and α-pinene (m/z 93; RT 5.25 min). A multi-choice assay was conducted involving cultivated tomato plants and eggplants (Darshanee *et al.*, 2017). Resistant tomato showed to have a low content of azulene and α-pinene but high amount of α-humulene and caryophyllene. In our study, we observed higher α-humulene and α-pinene in the susceptible lines, whereas higher caryophyllene was observed in the resistant lines. Features annotated as monoterpenes include putative azulene, 3-thujene, p-cymene, and α-pinene. All four monoterpenes were lower in resistant lines. For Azulene and α-pinene a similar result was described by Darshanee *et al.* (2017). 3-thujene is a monoterpene that was found to be present in the *S. pennellii*, LA0716 introgression lines but not to be detectable in cultivated tomatoes, M28 (Schilmiller *et al.*, 2010). In that research, monoterpene synthase genes that regulate the biosynthesis of 3-thujene were predicted to be on chromosomes 1 and 8. A putative γ-muurolene feature, also a monoterpene, was identified and showed higher abundance in the resistant compared to the susceptible lines. γ-muurolene was also identified in an experiment involving cultivated tomatoes infected with spider mite (*Tetranychus urticae*), however it did not show a significant difference (Kant *et al.*, 2004).

The two most prominent monoterpene peaks, with short retention time, were α-limonene (RT 7.15 min) and β-phellandrene (RT 7.20 min). α-limonene is specifically found in the wild tomatoes whereas, β-phellandrene is typically found in the cultivated tomato (Schillmiller *et al.*, 2009). It was evident from the PCA loading plot that both monoterpenes are distantly distributed (Fig. 12 & 13). Two limonene and phellandrene features were observed. The two limonene features located at the same retention time were picked by MSDial came from α–limonene and its degradation product having one hydrogen less. α- and β-phellandrene differ in retention times. β-phellandrene was more abundant compared to α-phellandrene and located nearly overlapping with α-limonene. Quantification of β-phellandrene and α-limonene peak areas were manually extracted from raw data and showed not much difference between the resistant and

susceptible lines of both populations, although there was slightly higher content in the resistant compared to the susceptible lines based on the median values for BC1F2 population (Fig 15). However, this result is due to the skewedness of the data set that is not normally distributed, even after logarithmic transformation. We also observe that there was a relatively wide distribution of the data points due to high variation in the population. There are several previously characterised *Terpene Synthases* (*TPS*): limonene is synthesized by *TPS 7* (Solyc01g105920) and β-phellandrene by *TPS4* (Solyc01g105880), and *TPS20* (Solyc08g005665) (Falara *et al.*, 2011; Zhou & Pichersky, 2020). Additionally, transcriptomics data captured the relative increase of *TPS4*, but not other *TPS* expression in wild compared to cultivated tomato trichomes and therefore we could provide further supporting evidence from the RNA expression level. From these results, although previous research has shown α-limonene and β-phellandrene to have a repellent activity to whiteflies (Du *et al.*, 2016; Bleeker *et al.,* 2009), these compounds do not seem to play a major role and there are therefore other determining factors for tomato resistance to whiteflies.
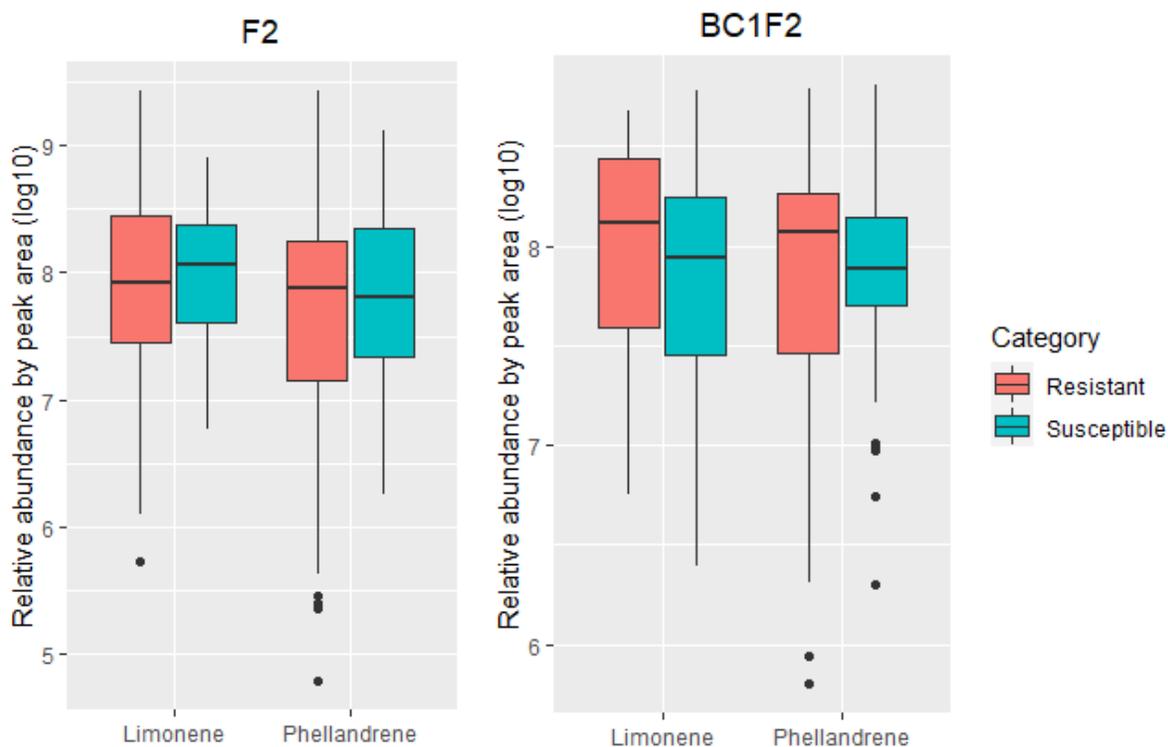
Figure 15. Boxplot of α-limonene and β-phellandrene in F2 and BC1F2 population.

MeSA

One feature that was annotated with at least 75% similarity to spectra library under benzoic acid derivative is methyl salicylate (m/z 120.095; RT 10.66 min). The quantity of methyl salicylate (MeSA) was high in the cultivated parents compared to the wild parents. The BC1F2 lines had levels that were in between the cultivated and wild parental lines. MeSA is a methyl ester of salicylic acid (SA) with s-adenosyl-l-methionine (SAM) as methyl donor, catalized by the enzyme Salicylate 1-O Methyltransferase (Tiemann *et al*., 2010). MeSA is an important volatile that contributes to taste and scent of many fruits and flowers. A choice assay involving greenhouse tomato plants sprayed with MeSA has shown to reduce greenhouse whitefly (*Trialeurodes vaporariorum*) population development and to increase tomato yield by immediately inducing plant defense (Conboy *et al*., 2020). Interestingly, it has been shown that whiteflies induce SA and reduce the expression of JA induced defenses as shown for *B. tabaci* (Zarate *et al.,* 2007; Zhang *et al.*, 2013). In earlier experiments using the BC5S2 population derived from *S. pimpinellifolium* acc. TO-937 x *S. lycopersicum* acc. Moneymaker, treating tomato plants with external methyl jasmonate resulted in a 60% increase in type IV trichome density and enhanced resistance to whiteflies (Escobar-Bravo *et al.*¸ 2016).

Other metabolites

Another feature that was higher in the resistant lines was a feature that was annotated as putative uridine-5-monophosphate (UMP). UMP is a pyrimidine nucleotide, a monomer in RNA that plays a central role in cellular regulation and metabolism. However, we doubt that this feature is UMP since it is prone to degradation and the recoveries of standard measurements via GC-MS are low (Koek *et al.*, 2006). Therefore, measurement of UMP by derivatization before GC-MS measurement or via LC-MS is preferable.

A putative phenol-2,4-bis(1,1-dimethylethyl)-phosphite-3:1, octane,2,3,6,7-tetramethyl feature, also known as tris(2,4-di-tert-butylphenyl) phosphite, was higher in the resistant compared to the susceptible lines. This compound was crystalized from *Vitex negundo*, a woody medicinal plant (Vinuchakkaravarthy *et al.*, 2010). Moreover, this organophosphorus compound function as antioxidants and is commonly used as stabilizer in polymers (Hermabessiere *et al.*, 2020).

3.3.2 Metabolite profiling of semi-hydrophobic metabolites of *S. habrochaites sp. glabratum* (VI030462) and cultivated tomato (AVT01424) F2 and BC1F2 population by LC-MS/MS

Semi-hydrophobic metabolites of the tomato populations were measured by negative and positive electrospray ionization in LC-MS. Results from the positive ionization mode non-targeted LC-MS analysis showed 16,887 and 12,506 features that were extracted from the F2 and BC1F2 population data set in MSDIAL, respectively. Among the F2 features, 629 were annotated based on 85% spectra similarity to our

spectra database, 6,610 were annotated without MS/MS spectra similarity, and 9,648 were unknown. As for BC1F2 population, 425 were annotated based on 85% spectra similarity to data base, 5,396 were annotated without MS/MS spectra similarity, and 6,685 were unknown. F2 and BC1F2 data that were exported to MetFamily upload resulted in 6,883 and 4,344 features. PCA scores and loadings plots for F2 and BC1F2 showed that the model accounts (PC1: PC2) for 11.9 %: 28.6 % and 17 %: 27 % (R^2) of the variance in the dataset (measure of model fit to the original data) with 30.2 %: 24.3% and 37.4 %: 22 % (Q^2) measure of consistency between the original and the cross-validation predicted data (Fig. 16 &17). PCA score plot of the F2 population showed clear separation between the wild and cultivated parental lines, however, not for BC1F2. Both datasets showed overlapping clustering for the susceptible and resistant lines. Therefore, we can conclude that the variation in secondary metabolite composition is not sufficient for the PCA analysis. AS, flavonoids, tomati(di)nes, amino acid, and quinic acid metabolic clusters were found back in the positive mode, showing a different fragmentation pattern to that of negative ionization. Moreover, additional metabolic families were annotated, which includes chalcones, phenylpropanoid, phosphocholine, nucleotide/nucleoside, and terpene glycoside.

Results from the negative ionization mode non-targeted LC-MS analysis showed 5,200 and 5,780 features that were extracted from the F2 and BC1F2 population data set in MSDIAL, respectively. Among the F2 features, 890 were annotated based on 85% spectra similarity to data base, 1,879 were annotated without sufficient MS/MS spectra similarity, and 2,432 were unknown. As for BC1F2 population, 1,698 were annotated based on 85% spectra similarity to data base, 2,455 were annotated without MS/MS spectra similarity, and 1,627 were unknown. F2 and BC1F2 data that were exported to MetFamily, which resulted in 4,405 and 3,409 features. PCA scores and loadings plots for F2 and BC1F2 showed that the model counts (PC1: PC2) for 23 %: 33.7% and 10.6 %: 32.8% (R^2), respectively of the variance in the dataset with 29.7 %: 28.9% and 37.5 %: 29% (Q^2) measure of consistency between the original and the cross-validation predicted data (Fig. 18 & 19). Both PCA of F2 and BC1F2 dataset showed clear separation between the wild and cultivated parental lines. However, high variation was found in the wild line. F2 lines that were categorized as susceptible cluster similarly to the F1, whereas those that were resistant distributed unevenly. The distribution of lines that were categorized as neutral were in between susceptible and resistant lines. The distribution of scores is reflected by the loadings of features. Features with similar MS/MS fragments cluster together in the HCA plot, and were manually annotated as a metabolite family. In the F2 dataset, we annotated seven metabolite classes (including its derivatives): AS, amino acid, fatty acid, flavonoid, oxylipins, quinic acid, and tomati(di)nes. However, in the BC1F2 dataset, with the same data analysis settings, we were able to annotate more metabolic families: AS, amino acid, fatty acids, flavonoid, phenols, quinic acids, sugar derivatives, sulfuric acid derivatives, and tomati(di)nes.

F2
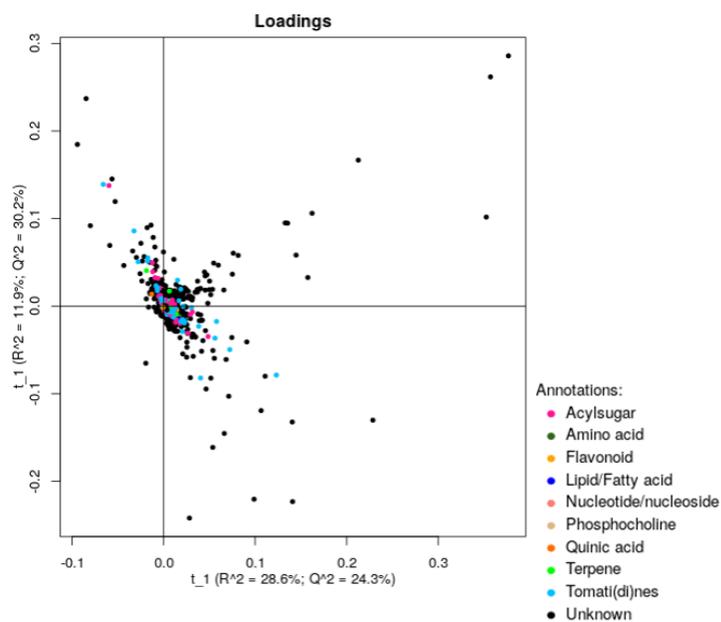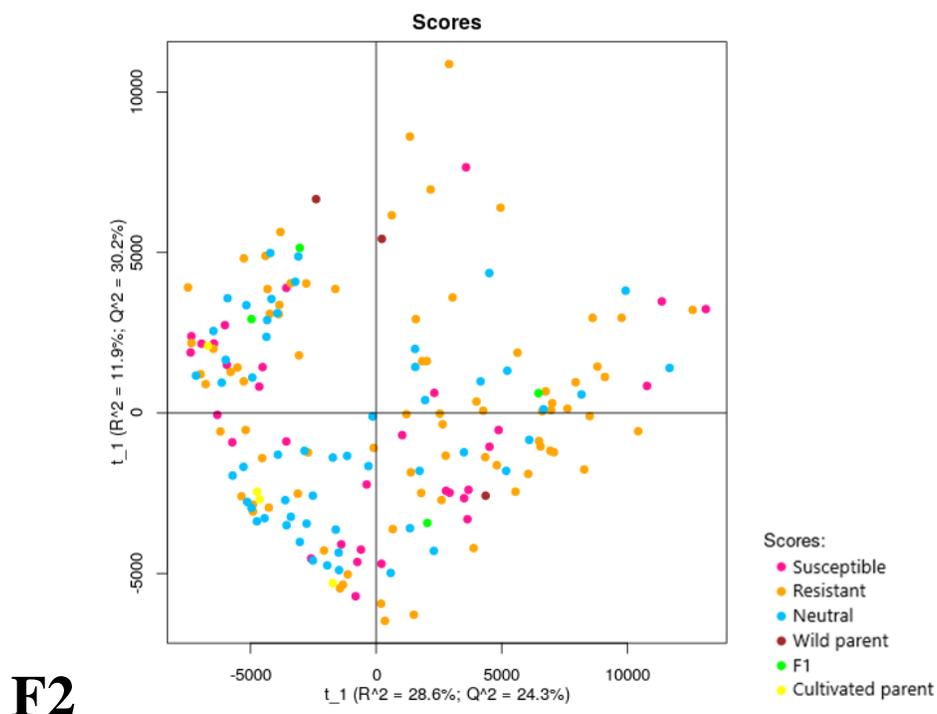
Figure 16. PCA of the F2 population obtained from LC-MS-pos.
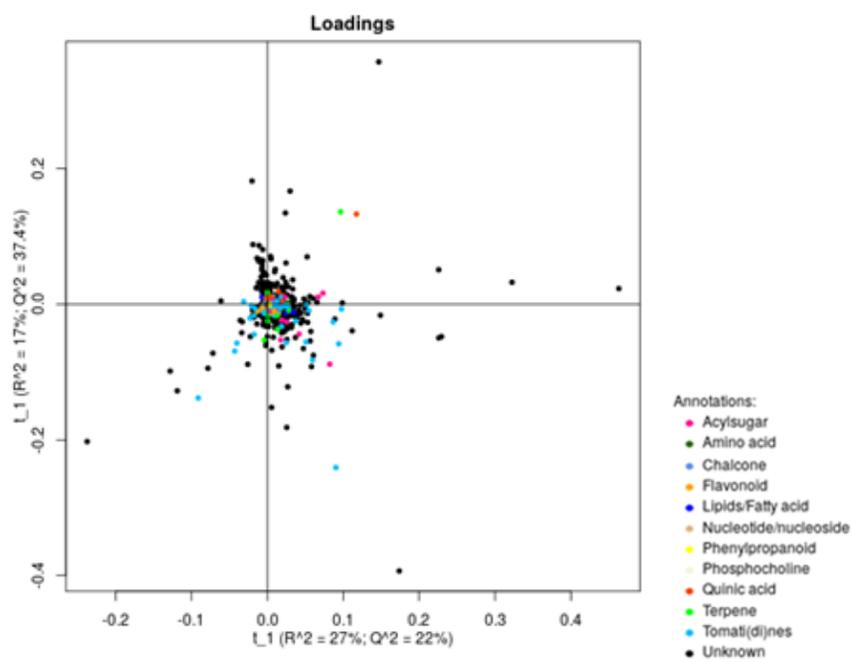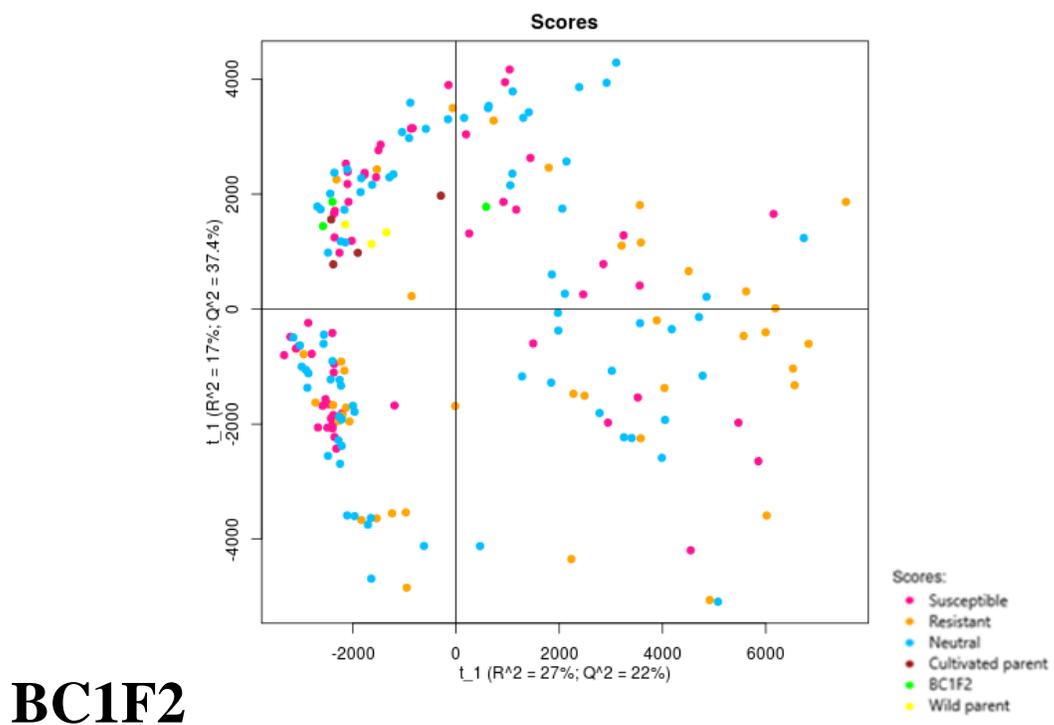
BC1F2

Figure 17. **PCA of the BC1F2 population obtained from LC-MS-pos.**
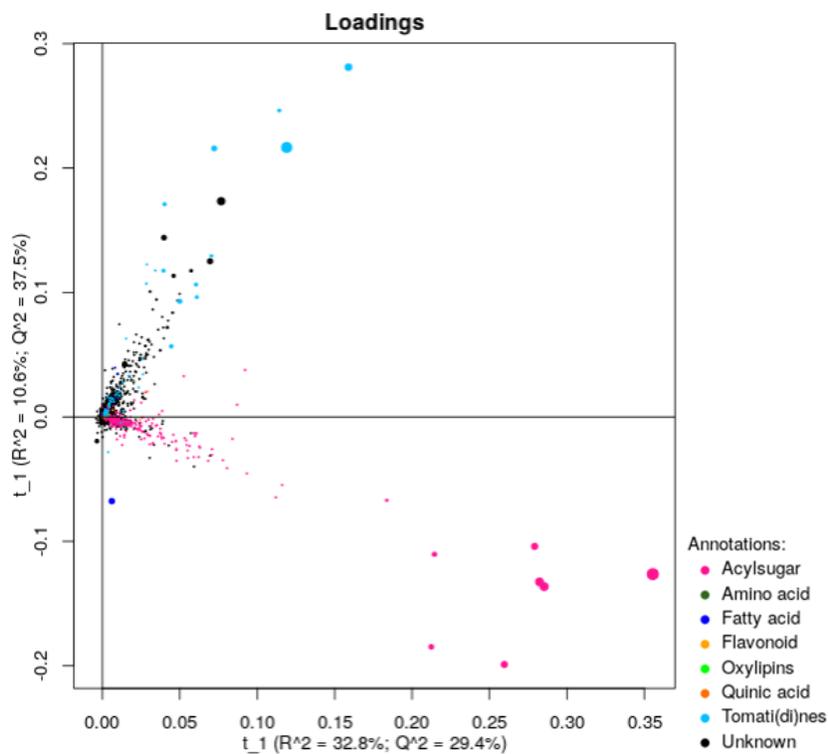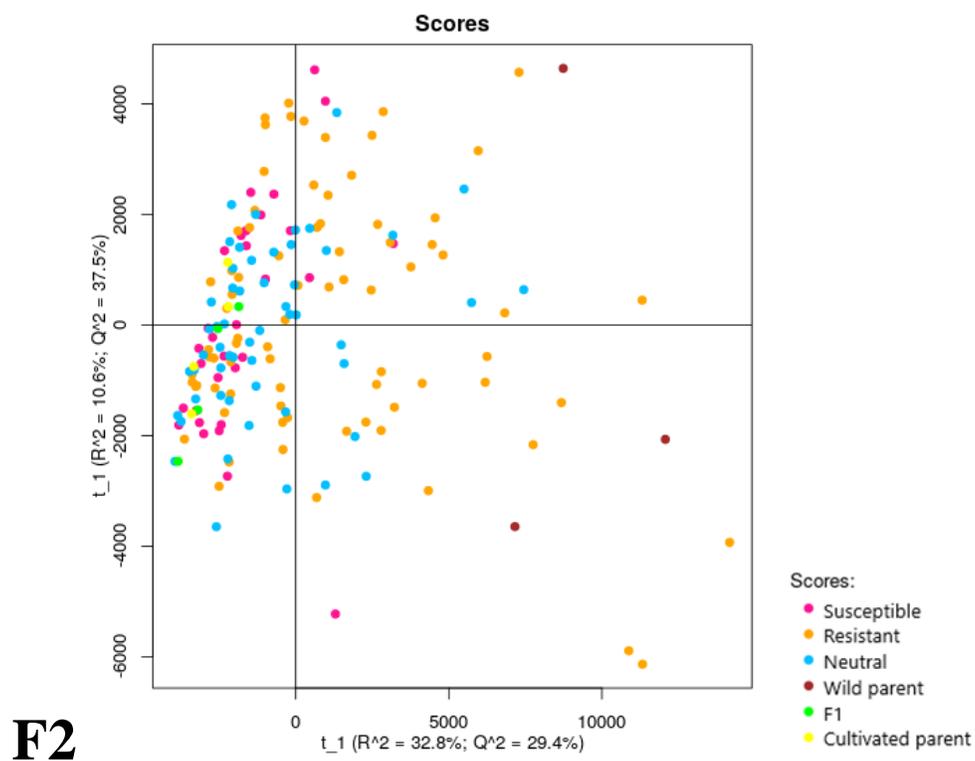
Figure 18. PCA of the F2 population obtained from LC-MS-neg.
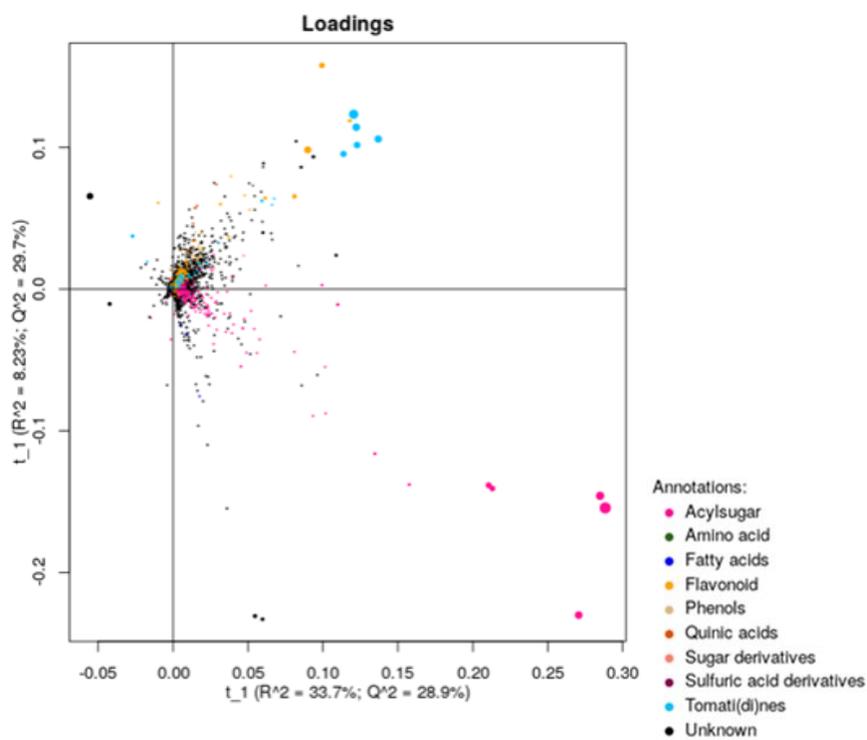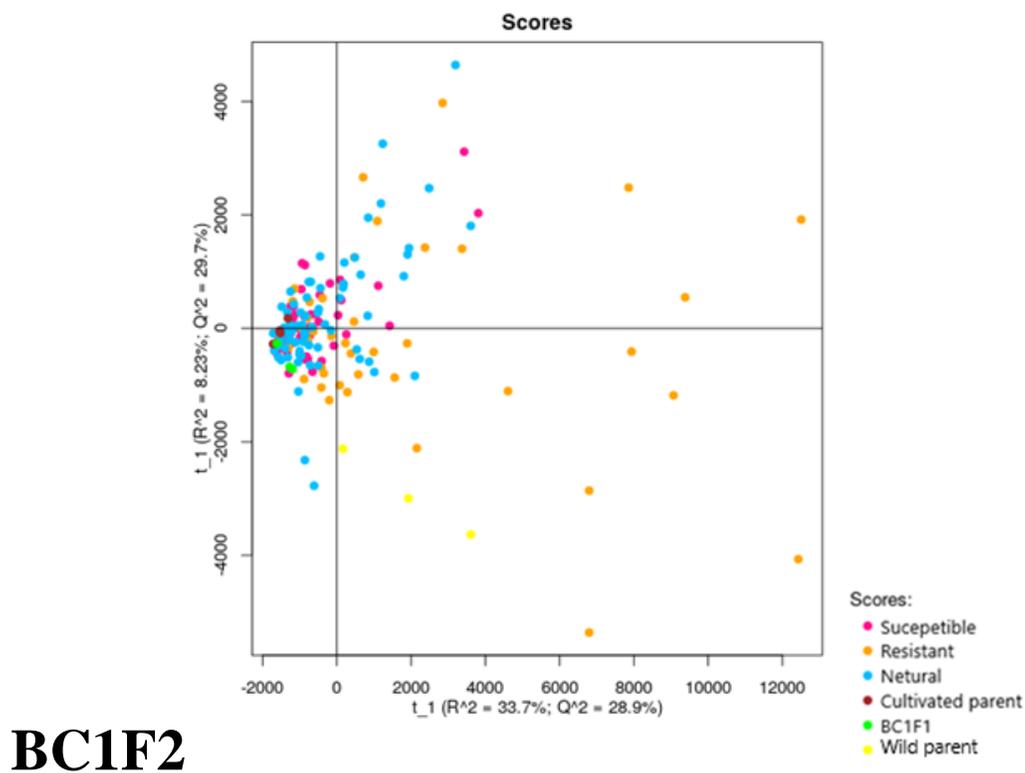
Figure 19. PCA of the BC1F2 population obtained from LC-MS-neg.
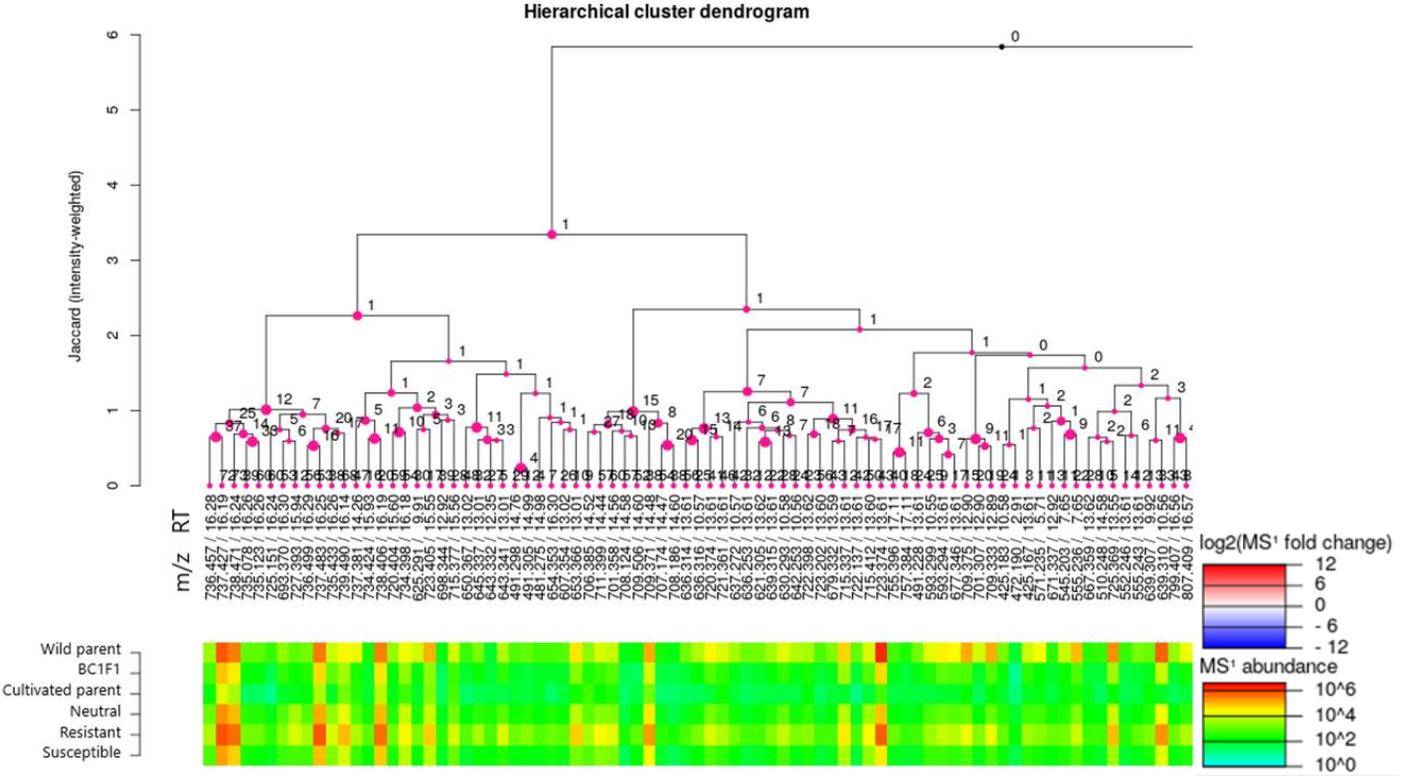
Figure 20. HCA of acylsugar cluster of BC1F2 MS^1 2 log$_2$-fold change. Masses represents [M - H$^-$ + FA]$^-$ / time in minutes / 100 mg of sample.

ANOVA analyses performed on the MetFamily output resulted in 137 (F2) and 464 (BC1F2) features that were significantly different between the resistant and susceptible lines for LC-MS-neg and 746 (F2) and 552 (BC1F2) features for LC-MS-pos (Table 4). We observed that not all significant features consistently appear in both populations, although all data processing procedures were conducted similarly. One reasons could be due to less genetic and phenotypic variation in the backcross lines compared to the F2 population, as some metabolic traits could be selected out. Furthermore, most of the annotated features that were significantly different between the resistant and susceptible lines were AS followed by flavonoids and tomati(di)nes.

Table 4. Number of significantly different features between resistant and susceptible lines.

Features	LC-MS-neg		LC-MS-pos	
	F2	BC1F2	F2	BC1F2
AS	215	81	20	20
Amino acid	-	2	-	-
Chalcone	-	-	-	3
Fatty acid/lipid	3	-	-	1
Flavonoid	3	16	2	1
Oxilipins	4	-	-	-
Phenolamide	-	1	-	-
Quinic acid	-	1	-	2
Terpene	-	-	1	10
Tomati(di)nes	5	3	1	4
Unknown	516	448	113	422
Total	746	552	137	463

Acylsugars

Among the annotated metabolic families, the AS cluster consisted of features that were abundantly found in the wild parent as well as resistant segregating F2 and BC1F2 lines (Fig. 20). All AS that showed $\log_2$-fold change above 2 and were more abundant in the resistant compared to the susceptible lines, were AS with a sucrose backbone and with at least 2 x C5 acyl groups attached (C5 FA - 101.0617 m/z) at unknown position. Different AS composition resulted in variation in hydrophobicity ranging from the most hydrophilic (S1:5(5) – 426.17 m/z) at 2.91 min to the most hydrophobic (S3:22(5,5,12) – 692.40 m/z) at 16.25 min. Identifier peaks (m/z) typically found in negative mode for AS includes: 323.098 and 425.166 for the sucrose backbone, whereas 101.061 and 199.170 for the C5 and C12 fatty acid. In positive mode, the fragmentation pattern is relatively more difficult to elucidate because the ionization causes the compound to split differently between the fructose and glucose ring, causing a more complex pattern.

Extracts of trichomes and trichome-less leaves were used to evaluate the different FA composition between the two tomato species. Twenty-one FAs were detected and identified from the fatty acid ethyl ester (FAEE) derivatization experiment (Table. 5). FAs with carbon chains less than six acyl carbon chain were categorized as short chain FAs, FAs with six to twelve acyl carbon chains were categorized as medium chain FAs, and FAs with more than twelve carbon chains were categorized as long chain FAs. Data showed

that short-medium FAs were higher in trichomes than leaves. This result confirms that tomato AS produced by the trichomes consist of short-medium FAs, typically C2-C12 FAs (Ghosh *et al.,* 2014). When we compared the two lines, *S. habrochaites* sp. *glabratum* - VI030462 trichomes contain more short-medium FAs than *S. lycopersicum* - AVT01424 trichomes. This can be explained by higher AS content in the wild tomato (Fig. 20). Particulary for C5 FA, we observed that short-chain FA iso-C5 (iC5) was more abundant than its isoform ante-iso-C5 (aiC5) (Fig. 21). Long and very-long chained FAs are not only essential components that make up triacylglycerol, waxes, phospholipids, sphingolipids, and plant cell membranes, they are also signaling molecules in plant defense (Lim *et al.,* 2017). The outcome of this experiment also showed some differences in the amount of long FAs such as C19-FA at 22.46 min and C19 at 22.52 min, which are higher in the trichomes, but most FAs are in a similar range.

Table 5. FAEE heat map result of *S. lycopersicum - AVT01424* and *S. habrochaites sp. glabratum -* VI030462 trichomes and trichome-less leaves based on average peak area.

Fatty acid	RT min	m/z	AVT01424 Leaf	AVT01424 Trichome	VI030462 Leaf	VI030462 Trichome
C2	5.59	102.068	22910	314034	79593	1270521
C4	6.7	116.084	28848	25423	22248	38577
aiC5	7.87	130.099	17753	85741	697172	10766702
iC5	7.99	130.099	18765	193053	2823553	43651615
C6	10.35	144.115	7674	57130	387347	5300547
C8	14.66	172.146	443170	566305	326427	1351275
C10	17.32	200.178	634556	1583978	2618098	39632862
C12	19.14	228.209	333524	1004130	36734462	122422355
C22	19.66	368.365	262404	1488803	127650	1801361
C22	20.06	368.365	5515	25440	12822	35367
C15	20.88	270.256	324660	1319980	281409	403405
C16	21.31	284.272	308777918	337177388	206604266	313221921
C16	21.32	284.272	290934227	337177388	206604266	323617484
C20	21.64	340.334	42689220	34025956	51151219	9296969
C17	21.71	298.287	15393557	55185348	16755277	48910192
C20	21.95	340.334	33301470	15383271	40956034	34273463
C20	22.01	340.334	20917547	65115779	22801603	30615652
C18	22.09	312.303	152808342	86082290	133736873	116825099
C19	22.46	326.318	1755880	18008679	1189077	8898827
C20	22.52	340.334	1636334	4947408	1711601	10306927
C20	22.85	340.334	1609926	1639047	1764939	2029966
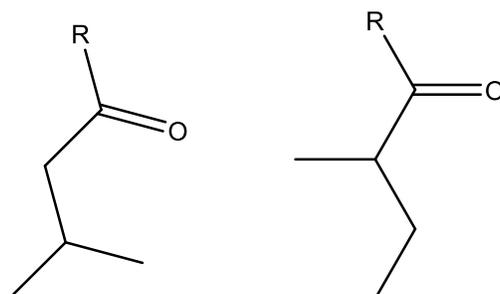
Figure 21. iC5 (left) vs aiC5 (right) fatty acid.

Oxylipins

Oxylipins are derived from mono- and poly-unsaturated FAs which include the phytohormone JA and related jasmonate metabolites that are not only involved in plant growth and development but also regulate trichome development and protection against (a)biotic stresses (Deboever *et al.*, 2019). Treating tomato plants with methyl jasmonate resulted in a 60% increase in type IV trichome density (Escobar-Bravo *et al*., 2017). High oxylipins content has already been reported in GT of *S. habrochaites* (Balcke *et al.*, 2017). Annotated oxylipin features contain a 295.231 m/z [M - H]⁻ fragment located around 13.5 min that was at least 1 $\log_2$-fold change higher in the resistant compared to the susceptible plants (Fig. 22). A list of known (*tomlox B-D*) and putative lipoxygenases, key enzymes in the biosynthesis of oxylipins, was compiled from transcriptomics data. Moreover, T*omloxC*, a lipoxygenase responsible for the synthesis of C5 volatiles independent of hydroperoxide lyase, showed high expression in wild tomato trichomes based on transcriptomics data (Shen *et al*., 2014). Notably, these C5 are not of the same origin as C5 acids that makes up AS, those are derived from the branched-chain amino acid metabolism (Walters & Steffens, 1990; Kroumova *et al*., 1994).
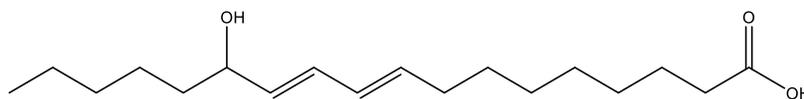
Figure 22. Structure of 13S-HODE-(d4) (m/z 296.24) found in tomato.

Flavonoids

Flavonols, flavones, anthocyanins, catechins, chalcones, flavanones are among the flavonoids in the classs of phenolic compounds that have important functions in plant defense against herbivores as antioxidants (Tohge *et al*., 2017). A cluster of flavonoids was observed. Two main flavonols in tomatoes, quercetin 3-rutinoside (rutin) and kaempferol-3,7-di-O-glucoside, have the same monoisotopic mass but differ in a single fragment in negative mode 284.032 m/z and 300.035 m/z (Fig. 23). It was reported that elevation of rutin and quercetin trisaccharide levels deter whitefly probing and salivation (O'Neill *et al.,* 1990; Kang *et al.*, 2014; Tohge *et al*., 2015; Ballester *et al.,* 2016; Su *et al.*, 2020). Hanson *et al.*¸ (2014) observed, that in an introgression line of *S. habrochaites* and *S. lycopersicum*, a QTL for high levels of rutin was found in chromosome 5. Several genes that are directly involved in flavonoid biosynthesis have already been characterized on chromosome 5, such as *Chalcone Isomerases* (*SlCHI1*-Solyc05g10320 or *SlCHIL*-Solyc05g052240) and *Chalcone Synthases* (Solyc05g053170 or *SlCHS1*-Solyc05g053550) (Su *et al.*, 2020)*.*
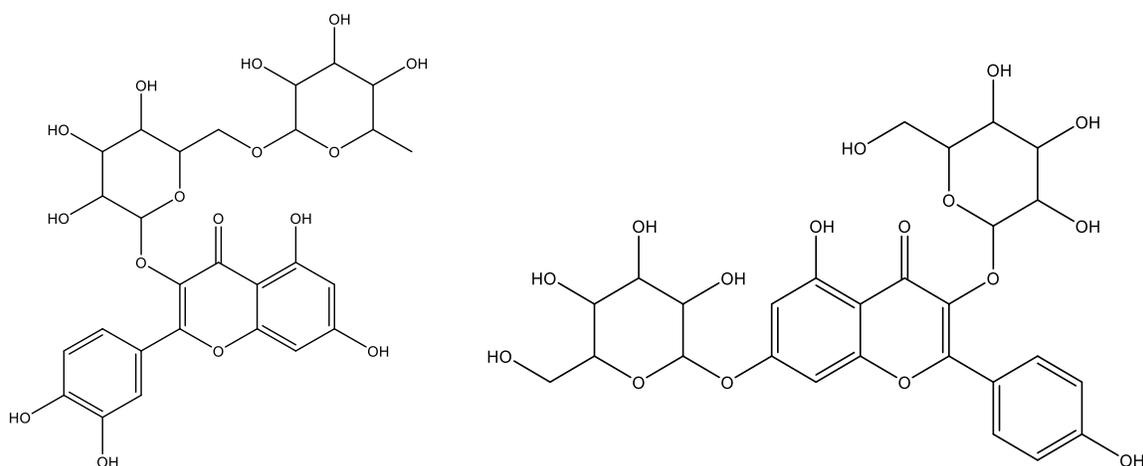
Figure 23. Quercetin 3-rutinoside (m/z 610.153 - left) and kaempferol-3,7-di-O-glucoside (m/z 610.153 - right) structures.

Steroidal gylcoalkaloids

Steroidal glycoalkaloids (SGAs) are present as toxic compounds against biotic threats. They comprise of steroidal backbone from cholesterol and are regulated by glycoalkaloid metabolism genes (Itkin *et al.*, 2013) as well as JA– ethylene responsive transcription factor (Thagun *et al.*, 2016). In tomatoes, the most common glycoalkaloids found derive from the group of tomatines, tomatidines, solasodine, lycoperoside, esculeoside and their derivatives. Tomato SGAs fragment well in both positive and negative ionization. SGA cluster annotated as tomati(di)nes based on some identifier peaks such as 253.193 m/z and 1034.53 m/z $[M+H]^+$ located around 6-7 min. Due to the diversity of chemical structures, feature identification poses a challenge and it would require additional work to decipher individual structures (Schwahn *et al.*, 2014). For example, Iijima *et al*. conducted a study in 2013 that showed *S. habrochaites* LA1777 contained high levels of SGAs, typically α-tomatine and habrochaitoside A (both having 1032.537 94 m/z $[M-H]^-$) a compound specific for *S. habrochaites* (Fig. 24). In this study, we found that some SGA features detected in both ionization methods were more abundant in the resistant compared to the susceptible lines for example solasodine (414.334 m/z $[M+H]^+$; 7.15 min) and a putative tomatine (1032.557 m/z $[M+H]^+$; 7.242 min) in the F2 population. However, there are also features that were more abundant in the susceptible parent such as features at 7.17 min with 1032.604 m/z and at 7.199 min with 1078.483 m/z $[M+H]^+$.
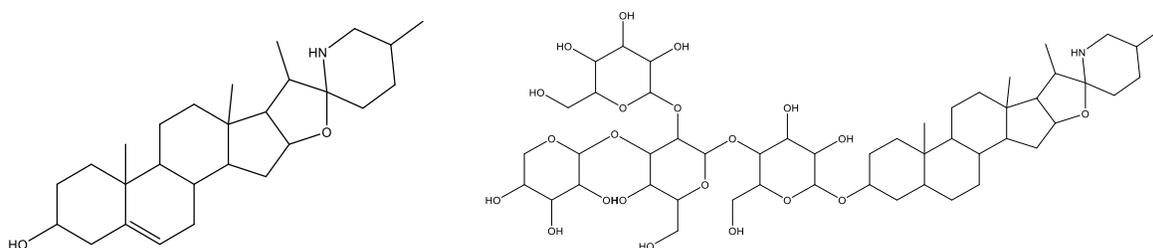
Figure 24. Structures of solasodine (m/z 413.329 - left) and α-tomatine (m/z 1033.546 - right).

Quinic acid and its derivatives

Quinic acid and its derivatives are compounds containing a quinic acid moiety (or derivative) as well as a carboxylic acid at position 1. They are derived from the phenylpropanoid pathway. Quinic acid and its derivatives have shown negative effects to not only whiteflies but also thrips by acting as feeding deterrents, growth inhibitors or toxins (Zhang *et al.*, 2017). A quinic acid cluster present in the BC1F2 population is comprised of features with common fragments of either 134.034m/z [ferulic acid–H–CO_2–CH_3]$^-$, 135.046 m/z [caffeic acid-H-CO_2]$^-$, 173.050 m/z [M-H]-, 191.057 m/z [*p*-coumaric acid–H–CO_2]$^-$, 193.055 m/z [ferulic acid–H]$^-$, and 353.085 m/z [M-H]- (Masike *et al.*, 2017). This cluster, that consisted of putative chlorogenic acid, 4,5-dicaffeoylquinic acid, cynarin and feruroyl quinic acid showed a higher relative abundance in resistant compared to susceptible lines by at least 1 $\log_2$ fold change (Fig. 25). Chlorogenic acid is one of the most abundant beneficial polyphenols and has been shown to play a significant role in plant-herbivore interaction (Kundu & Vadassery, 2018). During herbivory, *Hydroxycinnamoyl-CoA Quinate Hydroxycinnamoyl Transferase* is induced to transfer the quinic group to caffoyl-CoA forming chlorogenic acid. It is then oxidized to chloroquinone by the plant's peroxidase enzyme that binds amino acids and reduces amino acid bioavailabiliy, which inhibits insect growth (Kundu *et al.*, 2019). Therefore, targeting this transferase as a biomarker might be beneficial to produce insect resistant tomatoes.
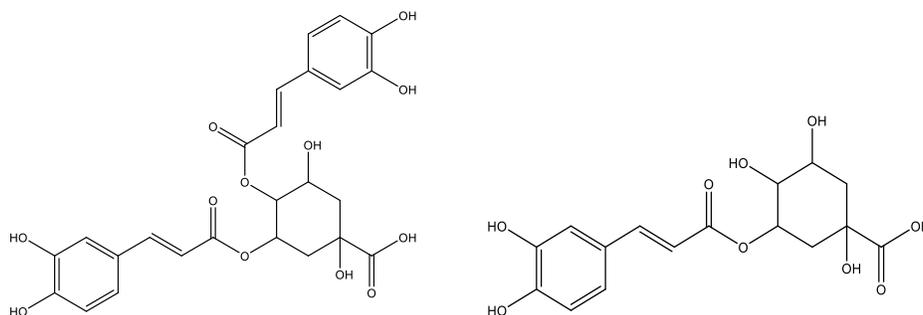
Figure 25. Structures of 4,5-dicaffeoylquinic acid (m/z 516.127 - left) and chlorogenic acid (m/z 354.095 - right).

Phenolamides are referred to as hydroxycinnamic acid amides or phenylamides, which consists of a phenolic moiety conjugated to polyamines (PAs) or deaminated with aromatic aminoacids (Roumani *et al.*, 2020). For plant defense against pathogens, phenolamides have often been described as bioactive compounds with radical scavenging activities and their accumulation is related to hypersensitive response and wounding (Bassard *et al.* 2010). In response to insects, phenolamides such as p-coumaroyl putrescine act as an ovipositioning-decreasing factor towards *Iriomyza trifolii* leaf miner in sweet pepper (Tebayashi *et al.*, 2007). In this study, a phenolamide feature annotated as n-caffeoyl-o-methyltyramine (312.133 m/z [M-H]-, 5.409 min) from the BC1F2 population was significantly higher in resistant compared to the susceptible lines (Fig. 26). N-caffeoyl-o-methyltyramine was first isolated from *Cuscuta reflexa* and

showed an inhibitory activity against α-glucosidase, a carbohydrate hydrolase (Anis *et al.*, 2002). An *in-vitro* enzyme experiment conducted by Sun *et al.* (2019) showed that n-caffeoyl-o-methyltyramine is the favored substrate by two glycosyltransferases from *Nicotiana benthamiana* for the formation of glucosides which are involved in plant resistance. Additionally, some phenolamides were reported to be susceptibility factors. Marti *et al.* (2013) showed that the accumulation of E-p-coumaroyltyramine in insect diet is beneficial for larval growth (*Spodoptera littoralis*) in infested maize leaves. Pearce *et al.* (1998) reported that the accumulation of E-p-coumaroyltyramine is affected by wounding in cultivated tomatoes. In this study, the E-p-coumaroyltyramine (284.118 m/z [M+H], 6.861 min) feature was detected and showed a higher abundance in the cultivated parent compared to the wild parent upon whitefly infection, by at least 1 $\log_2$-fold change. From these results, different combinations of phenolamides are recommended to enhance tomato defense against whitefly and other herbivores (Ongkokesung, *et. al*, 2012).
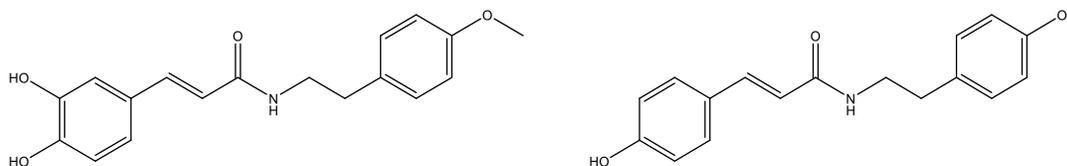
Figure 26. Structure of n-caffeoyl-o-methyltyramine (m/z 313.131- left) and E-p-coumaroyltyramine (m/z 283.121 - right).

3.4. Trait analysis of F2 and BC1F2 population

Trait analysis conducted for F2 and BC1F2 population included phenotypic data collected by AVRDC and quantification of four acylsugars (AS 1-4) by LC-MS-neg. Correlation analysis showed a positive correlation of the AS to adult mortality and type IV trichome density for both populations (Fig. 27). Since type IV trichomes produce AS that are toxic to whiteflies, increased trichome density leads to higher adult mortality. Furthermore, the number of deposited eggs correlates negatively with adult mortality. Negative or inverse correlation describes a relationship between two variables in which an increase in one variable causes the other to decreases. In this case, high densities of type IV trichomes producing AS leads to an increase in adult mortality and consequentically also fewer eggs laid by the whiteflies. With respect to type VI trichome density, no significant correlation was observed to other traits except for a weak correlation (<-0.2) between type VI trichome densities on seven-week-old plants to mortality in the F2 population. This means that the variation within and between traits tested is not sufficient to make any correlation and that we cannot model a linear relationship between different variables. In these two populations, the low variation of the type VI trichome density in addition to the small number of samples was insufficient to allow for correlation analysis. Therefore, we suggest that correlation should be tested for individual metabolites produced by type VI trichome.
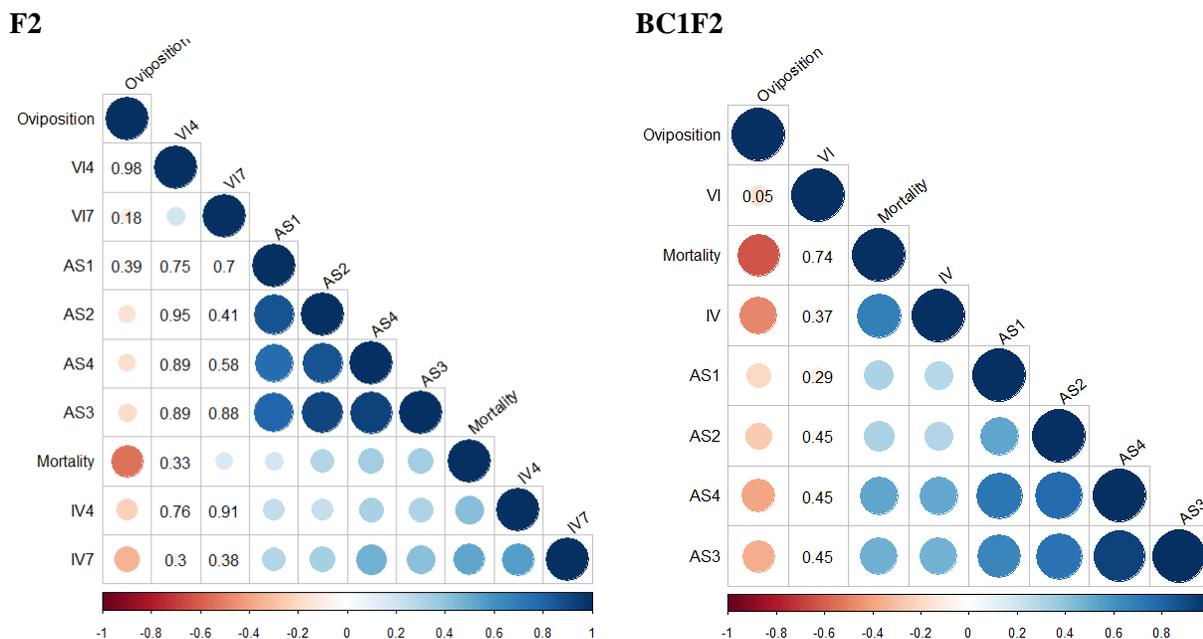
Figure 27. Correlation analysis of traits and AS of F2 and BC1F2 population.
Colored bar represents correlation value. Numbers in circles show p-value < 0.05. AS1- S2:10 (5,5): 493.233 m/z, 7.83 min AS2- S2:15 (5,5,5): 594.289 m/z, 10.56 min AS3- S4:25 (5,5,5,5): 678.346 m/z, 13.78 min AS4- S3:22 (5,5,12): 776.456 m/z, 16.52 min

3.5.   Genetic analysis of F2 and BC1F2 population

All the samples in both populations were genotyped by sequencing. Single nucleotide polymorphisms (SNPs) were identified and used as markers. Markers that segregate and with a skewed segregation ratio of less than 5% of the whole population, based on chi-square test, were retained. Following that, individual plants in each population were tested, to test if they showed genotype frequencies based on the expected segregation pattern. A mendelian segregation ratio of 1:2:1 was expected for the F2 population, however we observed a slightly lower distribution of the cultivated compared to the wild allele frequency. This was likely due to low number of plants used in the experiment (Fig. 28). Furthermore, pollen used to make the F1 that started both populations came from a pool of the wild accession (5 plants) that was not homogeneous, thereby increasing the variation. An inbred backcross population was created by crossing the most resistant tomatoes (5 plants) from the F2 population that were selected by the breeder to the recurrent cultivated parent creating the BC1F2 population. For this population, higher genotype frequency was observed for homozygous cultivated allele compared to the homozygous wild and heterozygous alleles (Fig. 28). However, there are several individuals that showed no variability (~0 % AA/AB or ~100 % BB genotypes). This mean that the plants are likely to be self-fertilized cultivated plants, so they were removed from the analysis.

Prior to performing QTL analysis, genetic maps, also known as linkage maps, were constructed from F2 and BC1F2 population based on ITAG 2.8 tomato genome. Genetic maps analyse genetic marker frequency relative to each other, which is a measure of recombination frequency (RF), that is, the frequency of recombination between markers during crossover of homologous chromosomes. Higher RF means larger physical distance, and *vice versa*. This information was used to assemble linkage groups, which are sets of genes that are linked and in their maximum group size cover entire chromosomes of the tomato genome. Two maps consisting of 12 linkage groups/chromosomes each of 7,566 cM (centi-Morgan) for the F2 and 8,055 cM for the BC1F2 were created.

F2
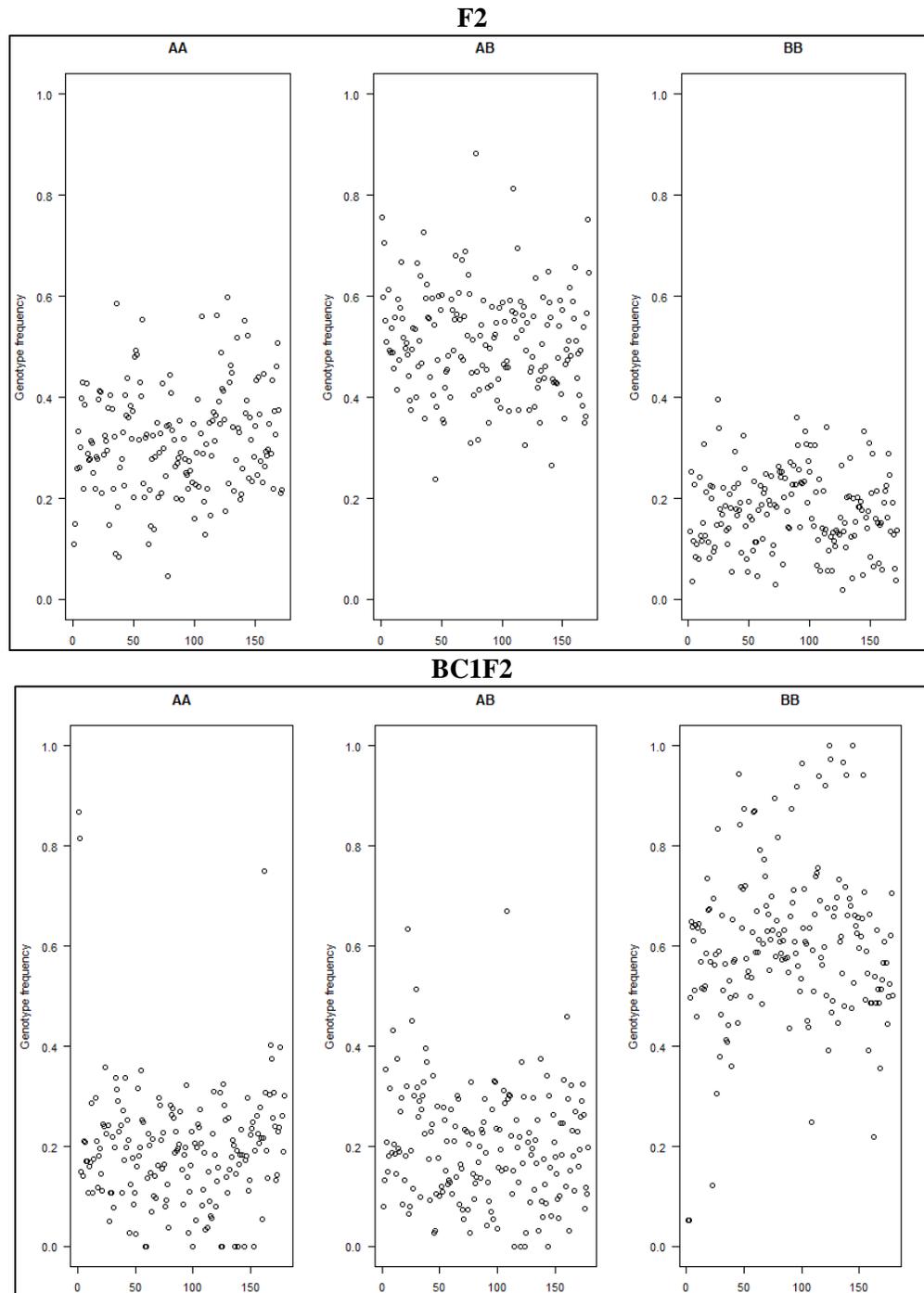
BC1F2

Figure 28. Genotypic frequency per individual of F2 and BC1F2 population.
A: wild-*S. habrochaites sp. glabratum* allele, B: cultivated-*S. lycopersicum* allele.

Table 6. Selected QTL regions of F2 and BC1F2 population obtained from insect assay and trichome quantification.

Mortality, TMortality (log 10 mortality), egg (number of whitefly eggs), egg_sqrt (square root of number of whitefly eggs), IV (density of type IV trichomes), 7VI (density of type VI trichomes observed 7 weeks after whitefly infection), 4IV (density of trichome type IV observed 4 weeks after treatment), 7IV (density of trichome type IV observed 7 weeks after treatment). CIM: composite interval mapping, MIM: single-trait multiple interval mapping. ITAG 2.8 tomato genome was used as reference.

Institute	Pop.	Chr.	Map (cM)	Trait	Start	End	Method LOD	P-value			LOD	Additive effect	Phenotypic variance explained (Generalized R^2)	Characterized genes
								α=0.10	α=0.05	α=0.01				
IPB	F2	1	38-48	Egg	2208979	77005468	MIM	3.053			4.11-7.63	-(0.67-0.72)	10-19%	ASAT4 Solyc01g105580
WorldVeg	F2	3	584.7	4IV	57878344	59478643	CIM				4.16	-2.67	11%	
IPB	F2	3	585.8	4IV	1781383	63389292	CIM 4.84, α=0.05	3.77	4.32		4.84	3.08	12%	
IPB	F2	3	585.8	4IV	1781383	63389292	MIM	5.47	6.49		3.8	2.44	10%	
WorldVeg	F2	3	588.7	7IV	57878344	59478643	CIM				6.69	-5.76	16%	
IPB	F2	3	585.8	7IV	1781383	63389292	CIM	3.88			4.48	3.92	11%	
IPB	F2	3	129.8-381.8	Egg	1781383	63389292	MIM	3.053			3.32-7.57	-(0.66-0.73)	8-18%	
IPB	F2	3	585.8	TMortality	1781383	63389292	CIM	4.01	4.831		4.84	3.08	12%	
WorldVeg	F2	5	580.4	4IV	7896412	60376099	CIM				3.87	-2.42	10%	
IPB	F2	5	432.6-436.6	7IV	41681163	45415075	CIM	3.88	4.57		7.5-8.5	5.18-5.53	18-20%	
IPB	F2	5	633.3	7IV	63257963	63703094	MIM	5.96			6.36	4.58	16%	ASH1 Solyc05g051660, ASH2 Solyc05g051670
WorldVeg	BC1F2	5	56-60	Egg	5551533	6154477	MIM	4.558	5.708		8.62-8.99	-(14.69-15.62)	20-21%	
WorldVeg	BC1F2	5	56-60	Egg_sqrt	5551533	6154477	MIM	5.013	6.02		9.83-9.98	-(1.38-1.44)	22-23%	
WorldVeg	F2	7	573.1-581.1	Tegg	7896412	60376099	CIM				3.71-4.55	-(0.28-0.46)	10-12%	
WorldVeg	F2	6	434.5	4IV	41343427	43604648	CIM				7.65	-3.35	19%	
IPB	F2	6	432.6-436.6	4IV	43191908	43631157	CIM	3.77	4.32		4.3-4.6	2.6	11-12%	
IPB	F2	6	432.6-436.6	7IV	43191908	43631157	CIM	3.88	4.57		4.93-5.07	3.66-3.68	12-13%	
WorldVeg	F2	6	450.5	7IV	41343427	45415939	CIM				4.46	-3.45	11%	
IPB	BC1F2	6	448.5	Mortality	44254826	49452509	CIM	4.37			4.43	8.64	11%	
IPB	F2	6	432.6-436.6	TMortality	43191908	43631157	CIM	4.01			4.29-4.61	2.6	11-12%	
IPB	F2	7	89.1-159.1	Egg	7714789	49316436	MIM	3.053			4.39-9.81	-(0.66-0.73)	11-23%	
IPB	F2	7	587.1-423.1	Egg	54118145	60717576	CIM	4.01			4.32-4.34	3.85-3.89	11%	
WorldVeg	F2	7	391.1-405.1	TMortality	4940528	54998827	CIM				4.68-6.15	0.85-2.69	12-15%	
IPB	F2	8	109.9-353.9	Egg	397946	60992215	MIM	3.053			4.37-8.14	0.67-0.71	11-20%	
WorldVeg	F2	9	58.1	7VI	2483212	4707473	CIM 7.2, α=0.05				7.18	8.41	18%	

Pop	Gen	Chr	Position	Trait	Start	End	Method				Range	Additive	%	Gene
IPB	F2	9	102.4-214.4	Egg	8041945	35645651	MIM	3.053			3.89-7.05	0.67-0.72	10-17%	
IPB	F2	9	26.4-30.4	Mortality	2441686	5844347	MIM	3.912			4.03-4.15	11.81-12.32	10-11%	
IPB	F2	9	24.4-30.4	TMortality	2441686	5844347	MIM	3.254	4.338		4.21-4.73	0.91-0.95	11-12%	
WorldVeg	F2	10	22.4	7VI	1955773	2922592	CIM 4.3, α=0.05				4.34	4.22	11%	
WorldVeg	BC1F2	10	270	IV	22679499	58655306	CIM	5.42	6.61		6.67	3.51	17%	
IPB	BC1F2	10	271.3	IV	22679499	58655306	CIM	4.612	5.116		5.55	3.042	14%	
IPB	F2	11	460.4-484.4	4IV	45846280	52443490	MIM	5.47	6.49		3.92-4.43	2.44	10-11%	
WorldVeg	F2	11	44.7	7IV	3731263	4522770	CIM				5.79	-3.39	14%	
WorldVeg	BC1F2	11	33-35	Egg_sqrt	3327957	3737658	CIM 6.7-7.1, α=0.01	4.99	6.38		6.43-7.13	0.08	15-17%	
IPB	F2	11	170.4-178.4	Egg	1449775	45846280	MIM	3.053			5.14-7.84	-0.72	13-19%	
IPB	BC1F2	11	33.7-35.7	Egg	175076	6171810	CIM	4.36	5.197		4.85-5.36	-(11.36-12.31)	12-13%	
IPB	BC1F2	11	185.7	IV	14039987	18715388	CIM	4.61	5.12		4.67-5.58	3.42-4.07	13%	
IPB	BC1F2	11	509.7	IV	6171810	56179997	MIM	7.19			7.54	0.26	18%	ASAT3
WorldVeg	BC1F2	11	545-551	Mortality	54369244	55295152	MIM	4.96	6.08		8.81-10.45	16.33-17.97	20-24%	Solyc11g067270
IPB	BC1F2	11	553.7	Mortality	6171810	56179997	MIM	4.43	4.98		5.88	12.08	14%	
IPB	BC1F2	11	553.7	Mortality	6171810	56179997	CIM	4.37	4.86	6.18	6.21	13.95	15%	
IPB	F2	11	498.4	Mortality	45846280	52443490	CIM	3.791			4.356	9.158	11%	
WorldVeg	BC1F2	11	545.7-551.7	Mortality	54369244	55295152	CIM 8.8-10.0, α=0.01	4.87	6.27		8.85-10.03	9.21-19.61	20-23%	
WorldVeg	F2	11	478.7	TMortality	47516991	49221325	CIM				3.25	-0.69	8%	

QTL analysis on F2 and BC1F2 population was conducted independently by us and colleagues from AVRDC WorldVeg using the same genotypic and phenotype information, which includes: trichome type IV and VI density, number of whitefly eggs (oviposition) and adult whitefly mortality (Table 6). Data was compared and only QTLs with LOD scores above the threshold based on a permutation test were said to be significantly different. Some of these QTLs coincide with regions that contain previously characterized genes related to AS biosynthesis such as *AS Acyl Transferase 4* (Solyc01g105580) on chromosome 1 as well as *AS Acyl Hydrolase (ASH) 1 and 2* (Solyc05g051660 and Solyc05g051670) on chromosome 5. ASH are enyzmes that remove acyl chains from specific positions of certain types of acylsugars (Schilmiller *et al.*, 2016).

Among the QTLs found, this study focused on QTL5 and QTL11 (Fig. 29). QTL5 was selected because a marker co-segregated with egg number at the top of chromosome 5 (56-60 cM region) with LOD > 5.5. The second QTL was located at the bottom end of chromosome 11 (498-551 cM region) and co-segregated with increased number of adult whitefly mortality LOD > 6.2) and density of Type IV trichome (LOD > 7.5).

In search of candidate genes, an RNAseq experiment was conducted to evaluate the differences between the wild (VI030462, *S. habrochaites sp. glabratum*) and cultivated (AVT01424-CLN3682C, *S. lycopersicum*) transcriptome. Two candidate genes, one in each QTL, were found to be significantly differently expressed: 1) *S-Adenosylmethionine Decarboxylase* (Solyc05g010420) on chromosome 5, and 2) *Acyl Transferase* (Solyc11g067270) on chromosome 11. Further experiments were conducted, results of which will be presented at chapter 3.4 and 3.5.
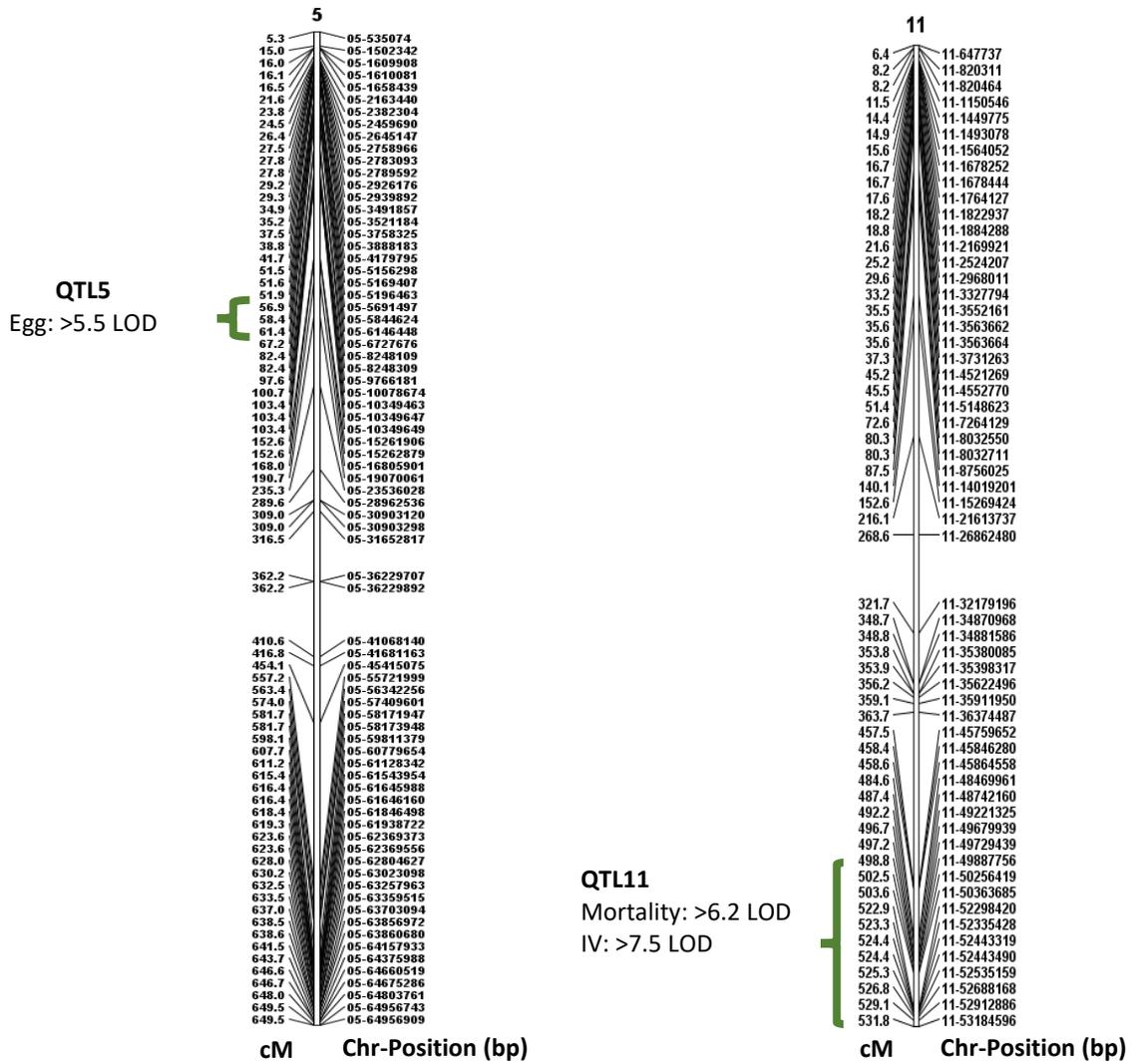
Figure 29. Location of QTL5 and QTL11 in the tomato genome.

3.6. QTL5, which segregates with increased oviposition, led to the identification of a gene involved in the biosynthesis of polyamines

3.6.1 Identification of *AMD and SAMS* candidate genes

QTL mapping results provided a hint that candidate genes that are in QTL5 should be genes that play a role in ovipositionings; in other words, genes that render plants more susceptible to *B. tabacci*. Within QTL5, transcriptomics results did not show many differentially expressed genes (Appendix 6). There was only one gene, Solyc05g010420 (*AMD5*) that codes for *S-Adenosylmethionine Decarboxylase Proenzyme* (*AMD*), with higher expression in the cultivated line. AMD proteins converts S-adenosylmethionine (SAM) to decarboxylated S-adenosylmethionine (dcSAM) which is one of the substrates for polyamine biosynthesis. This was confirmed by q-RT-PCR (Fig. 30 A). AMD5 has 70% identity to S-adenosylmethionine decarboxylases: At-SAMDC1 (AT3G02470), At-SAMDC2 (AT5G15950), and At-SAMDC3 (AT3G25570). The expression of *At-SAMDC1* was particularly high in the siliques, *At-SAMDC2* was more strongly expressed in the roots, leaves and flowers, whereas *At-SAMDC3* was weakly expressed in all organs of Arabidopsis (Ge *et al.,* 2006). The *At-SAMDC*1 promoter exhibited high activity, whereas *At-SAMDC*2 & 3 were moderate to low in seedlings (Majumdar *et al.*, 2017). Overexpression of *At-SAMDC1* increased expression of defense-related/jasmonic acid metabolism genes and resistance to *Pseudomonas syringae* and *Hyaloperonospora arabidopsidis* (Marco *et al.,* 2014).

Six other genes from the transcriptomics data were available with the same annotation (Appendix 6). A phylogenic tree was constructed to represent the evolutionary relationship of the tomato AMD genes to other species on the protein level (Fig. 30 B). Surprisingly, three out of seven candidate genes that were annotated as *SAM Decarboxylase* were more like *SAM Synthetase*. SAM synthetase synthesize SAM from methionine. Among the candidate genes, only Solyc12g099000 (*SAMS12*) showed similar pattern to *AMD5*, although the overall expression was lower in the transcriptomics data. SAMS12 has 80% identity to S-adenosylmethionine synthetases: At-MAT1 (AT1G02500) and At-MAT2 (AT4G01850), At-MAT3 (AT2G36880) and At-MAT4 (AT3G17390) in *Arabidopsis thaliana* (Peleman *et al.,* 1989; Goto *et al.,* 2002; Sekula *et al.*, 2020). In addition, AMD2, located on chromosome 2 and with similarity to At-SAMDC1 (Majumdar *et al.*, 2017), displayed higher expression in cultivated than in wild plants. The expression profile from qRT-PCR was not like the transcriptome data however, showing wild leaves to have higher expression than cultivated leaves. In the end, we did not go forward with this gene and selected *AMD5* and *SAMS12* to be characterized.

A

B
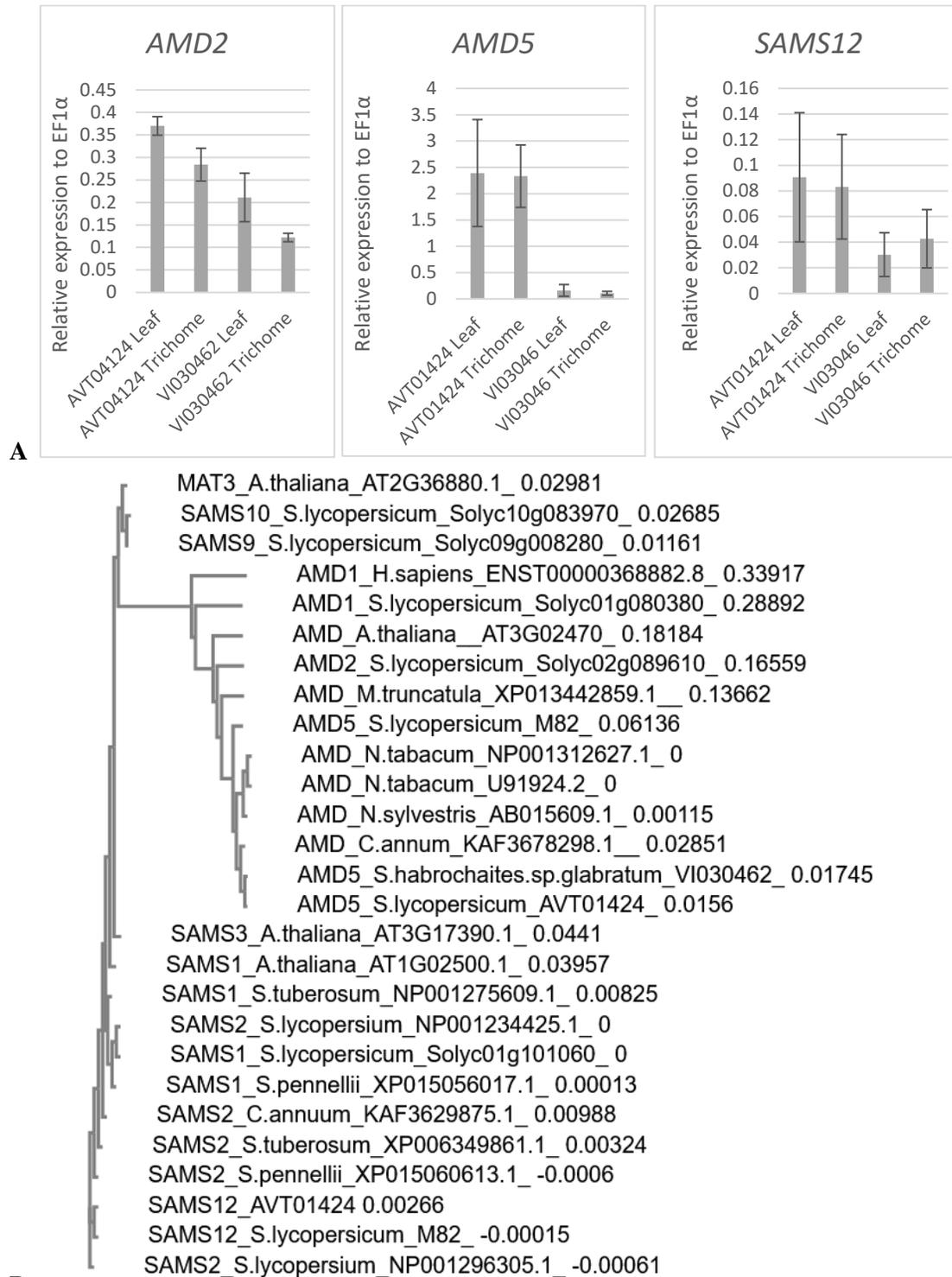
Figure 30. Q-RT-PCR result of candidate genes related to polyamine biosynthesis.

A) *AMD2* (Solyc02g089610), *AMD5* (Solyc05g010420) & *SAMS12* (Solyc12g099000). TPM: Transcript per million. Error bars respresent standard error.

B) Phylogenic tree of AMD and SAMS proteins. Protein name_species_id_values from neighbour-joining tree without distance corrections.

3.6.2 Polyamine measurement via FMOC confirmed high conversion of putrescine into spermidine in cultivated tomatoes

Before characterizing the function of the AMD candidate genes, the levels of polyamines (PAs) were measured for cultivated and wild trichome and trichome-less leaves to see whether there are significant differences between the samples and whether we can explain that different expression levels in different samples affect the production of PAs (Fig. 31). This experiment was conducted in collaboration with Jörg Ziegler based on the protocol mentioned in Ziegler & Abel (2014). PAs were extracted from powdered trichome and trichome-less leaf samples of cultivated – AT01424 and wild – VI030462 tomato in three biological triplicates and were targetedly measured by HPLC. Standards were used to quantify the absolute quantity of individual PAs. Putrescine was observed with the same level in both wild and cultivated leaves. However, the level was significantly lower in the wild trichome compared to cultivated trichomes. As for the product of decarboxylated SAM (dcSAM) and putrescine, spermidine was highly abundant in leaves and trichomes of cultivated compared to the wild tomatoes. Based on these observations, we suspected that lower amount of spermidine are due to a smaller amount of dcSAM being supplied as substrate. However, the method for measuring dcSAM was not optimized and these measurements could not be carried out. The peaks detected for thermospermidine and spermine were very low, therefore it was difficult to draw any conclusions about small metabolic differences.
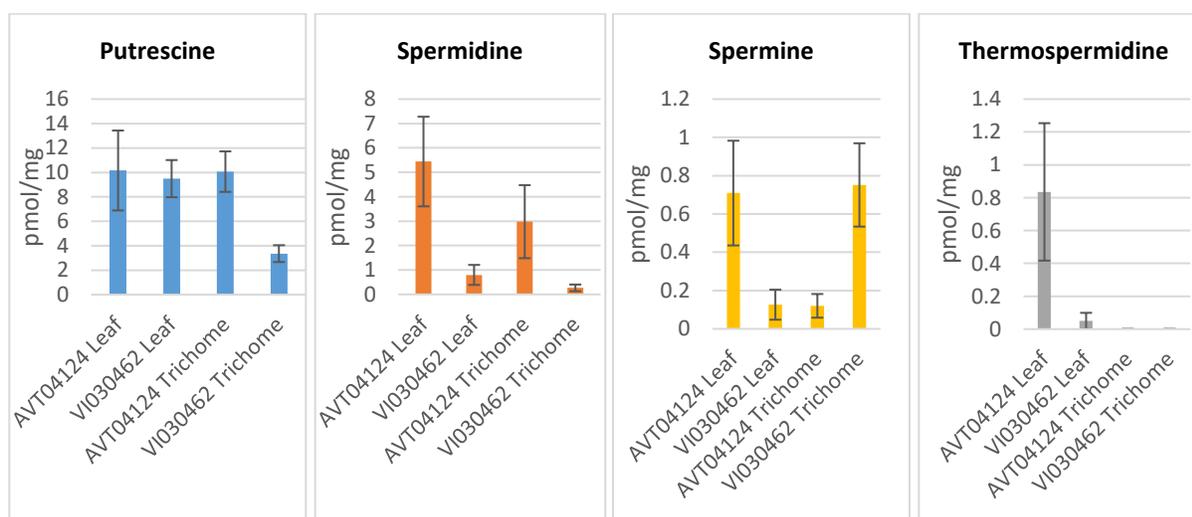
Figure 31. Putrescine, spermidine, thermospermidine and spermine measurement result in wild and cultivated leaf and trichomes.
Error bars represent standard error.

3.6.3 AMD5 and SAMS12 enzyme activity assay

Enzyme assays were conducted with His-tag purified AMD5 and SAMS12 cloned from cultivated tomatoes (Appendix 3). In the experiment empty vector pET28 SN (supernatant: non-purified) and no enzyme were used as control. We expect that functional AMD protein would decarboxylate SAM, whereas SAMS would synthesize SAM from L-Methionine (Met). When SAM was added as substrate, no dcSAM was formed in all treatments. No products were observed in the pET28 SN control with endogeneous AMD proteins from the bacteria. This may mean that our experimental setup is not suitable for the enzyme. Furthermore, we suspect that AMD5 could be inactive because the amount of external cofactor (thiamine pyrophosphate) added in the reaction was not sufficient to activate the enzyme. When Met was added as substrate in another set of reaction, as expected, high amount of SAM was observed for SAMS12 but not for AMD5 (Fig. 32). SAM was also observed for pET28 SN which confirm that there are endogenous SAM proteins from the *E. coli,* even though the activity is lower compared to SAMS12. This result showed that SAMS12 was miss-annotated in the transcriptomic analysis and needs to be corrected as a SAM synthase which are ATP dependent synthetases that synthesizes Met into SAM.
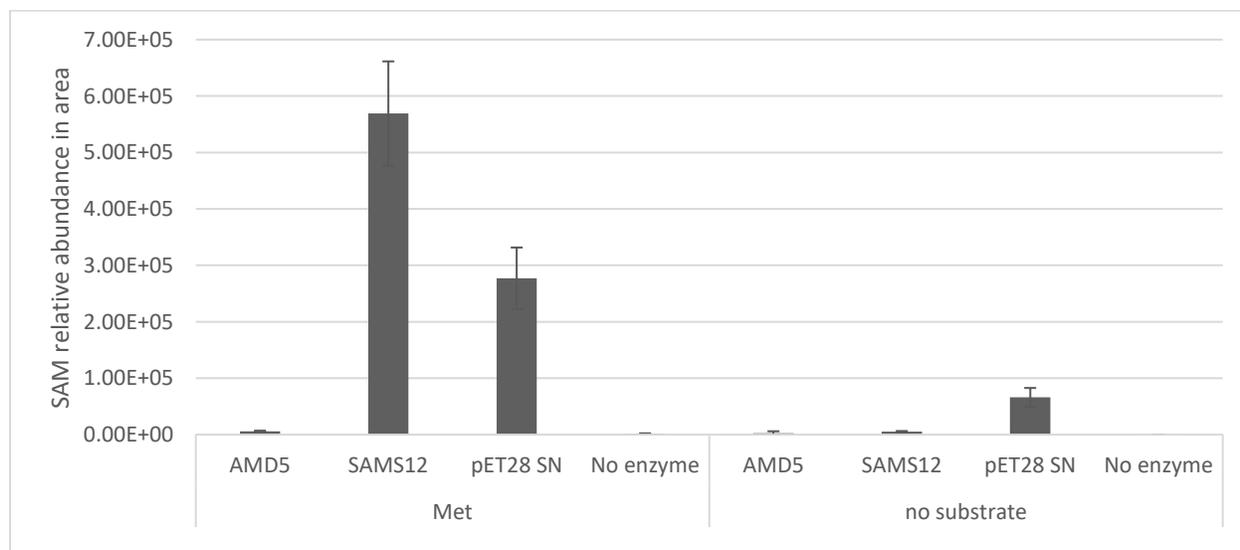
Figure 32. Quantification of SAM produced via *in-vitro* enzyme assay.
Error bars represent standard error.

3.6.4 VIGS silencing of AMD5 and SAMS12 results in lower *in planta* levels of Pas in cultivated tomato plants

VIGS was conducted to characterize the function of both AMD5 and SAMS12 in cultivated tomatoes. Cotyledons of two week old plants were infiltrated and samples were taken from the 2-4th leaves, five weeks post-infection. We expected to see lower level of PAs when measured by HPLC as well as lower expression in silenced plants via Q-RT-PCR. TRV2 was used as empty vector control and should not show any differences in PA levels and gene expression. Lower expression was observed in *AMD5* and *SAMS12* silenced plants compared to empty TRV2 (Fig. 33 B). Furthermore, PA measurement showed decreasing trend of putrescine, spermidine, spermine and thermospermidine for the *AMD5* silenced plants compared to TRV2 (Fig. 33 C). However, the values were not significantly different based on T-test (α=0.05). This could be due to other *AMD* and *SAMS* homologous genes that are redundantly active. Furthermore, non-homogenous chlorotic symptoms on leaves were observed in plants that were silenced with AMD5 (Fig. 33 A).
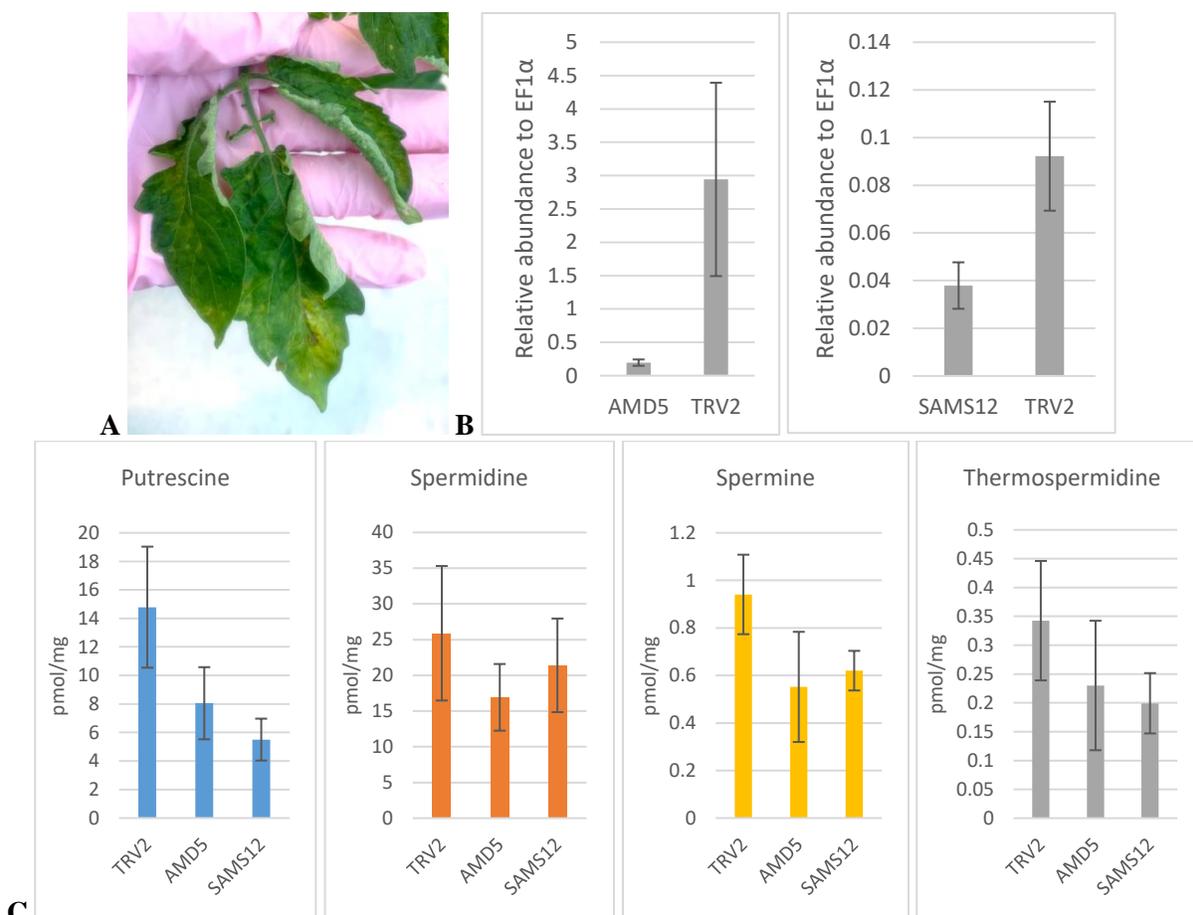
Figure 33. VIGS experiment result.
A) Chlorotic symptoms on leaves. B) Q-RT-PCR VIGS result. C) PA measurement of VIGS treated cultivated tomato plants. Error bars represent standard error.

3.7. QTL11

3.7.1 Identification of candidate genes located on QTL11

QTL11 showed correlation with whitefly mortality and type IV trichome density, similarly had us search for candidate genes whose expression is higher in the *S. habrochaites sp. glabratum* - VI030462 trichome and that are related to a production of insecticidal compounds. Based on this criterion, the list of candidate genes was narrowed down to seven (Appendix 6).

Two predicted esterase (annotated as ovarian cancer-associated gene 2), Solyc11g010420 and Solyc11g010430, are 70% identical to to *Serine Hydrolase 1* (*FSH1/UP7;* AT5G65400) which is a *serine alpha/beta-hydrolases* superfamily protein. *UP7* localized at peroxisomes and was suggested to have potential roles in β-oxidation (Cassin-Ross & Hu, 2014). Two predicted *Xanthine Dehydrogenase/Oxidase*, Solyc11g071610/Solyc11g071620, have 67% similar identity to abscisic aldehyde oxidases: *AAO1* (AT5G20960), *AAO2* (AT3G43600), *AAO3* (AT2G27150), and *AAO4* (AT1G04580). Aldehyde oxidases are involved in the oxidation of aromatic and aliphatic aldehydes such as abscisic aldehyde into abscisic acid. *AAO3* expression increase during infection of pectinolytic enterobacterium, *Dickeya dadantii* (van Gijsegem *et al.,* 2017). AAO4 is involved in the accumulation of benzoic acid in Arabidopsis siliques, which helps to protect from, and delay senescence by catalyzing aldehyde detoxification (Srivastava *et al.*, 2017). Solyc11g071800 is annotated as strictosidine synthase which catalyzes the stereospecific condensation of tryptamine with secologanin to form strictosidine. Strictosidine is a key intermediate of indole alkaloid biosynthesis. However, no strictosidine or related alkaloids have been reported yet in tomato, indicating the function of this gene is unlikely to be in strictosidine biosynthesis. BLAST result in the TAIR database similarity led to a predicted *Strictosidine Synthetase-Like 1 & 3* (AT1G08470 & AT2G41300) which lacks the conserved catalytic glumate residue found in active enzymes (Sohani *et al.*, 2008). The most strongly expressed gene in the wild trichome was Solyc11g067270, an acyltransferase that was identified and previously characterized as *ASAT3* involved in AS biosynthesis. Another gene Solyc11g067290 that showed low expression was annotated as *ASAT3-like* due to its high sequence similarity. Based on these annotations, *ASAT3* was considered to be the most likely candidate and was chose for further characterization.

3.7.2 Characterization of AS from VI030462

A method to obtain purified individual AS was established in this study. Fresh leaves (978 g) of VI030462 wild tomato were washed with 100% MeOH, filtered with filter paper and dried down via rotary evaporator leaving 4.60 g of dried leaf extracts. The extract was fractionation on silica column as mentioned in the method section resulting in 60 mg of pre-purified AS fraction. This fraction was then re-fractionated via solid phase extraction cartridges in an analytical UHPLC coupled to LC-MS system through a 10 cm RP18-column. Four AS were selected based on abundance (Fig.   34), hydrophobicity and acyl-group attachments, resulted in: 1) acylsucrose (AS) 1 with 2 x C5 with m/z 555.229 $[M - H^- + FA]^-$ eluting at 7.83 minutes; 2) AS2 with m/z 639.287 $[M - H^- + FA]^-$ eluting at 10.56 minutes composed of 3 x C5 acyl chains; 3) AS3 with m/z 723.344 $[M - H^- + FA]^-$ eluting at 13.78 minutes composed of 4 x C5 acyl chains; and 4) AS4 with m/z 737.397 $[M - H^- + FA]^-$ eluting at 16.52 minutes composed of 2 x C5 as well as 1 x C12 acyl chain. Sufficient AS2 (5 mg), AS3 (6 mg), and AS4 (26 mg) were obtained for NMR but not for AS1 (<1 mg).

NMR results showed that all acylsugars (AS2, AS3 and AS4) were a combination of at least two similar AS that differ in position and type of C5 acid (Fig.   34). In congruence with the derivatization result, there was only C12, iC5 and aiC5 present in all analyzed acylsucroses. AS2 consists of three main compounds: 1) iC5 at position 2,3,1'; 2) iC5 at position 3,4 and aiC5 at position 2; and 3) iC5 at position 3,1' and aiC5 at position 2; with approximate molar ratio of 1.00:0.57:0.49. The structure of AS3 was relatively easier to characterize because there are only two different acylsucroses that have either iC5 or aiC5 at position 2 of the glucose ring. One common acylation position that was observed in both acylsucroses was the acylation at the 1' position of the fructose ring. This analysis showed that the enzyme acylating on the fructose ring is different from the previously characterized ASAT3 protein, that acylates at the 3' position (Fan *et al*., 2015; Schilmiller *et al*., 2012; Schilmiller *et al*., 2015). However, this was not the case for AS4, that contained four main compounds with variation aiC5 and iC5 at position 2 as well as nC12 at position 3 with a molar ratio of 1.00:0.96:0.50:0.36. Due to strong overlapping signals, assignment of various acyl groups to individual compounds was not possible using COSY, HSQC and HMBC spectra information.
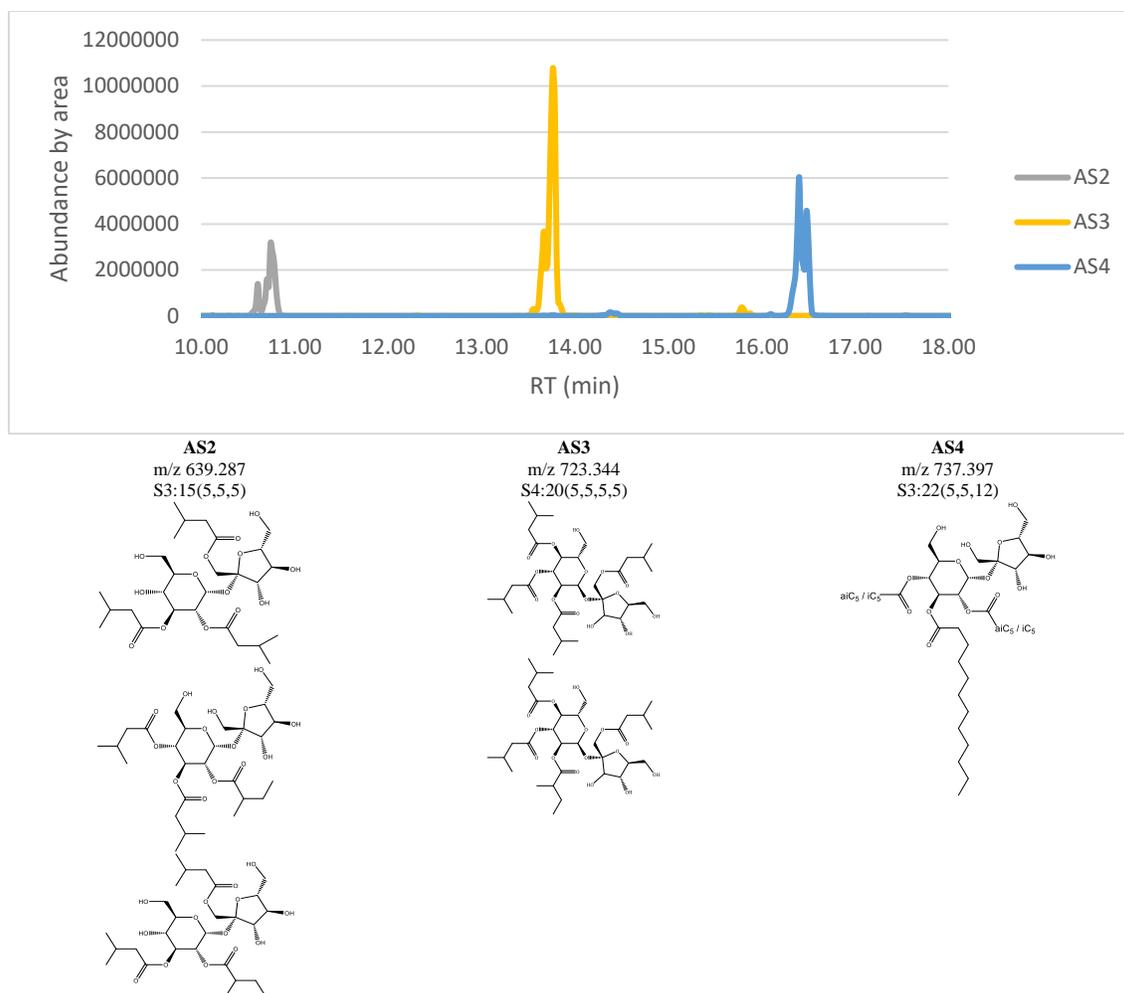
Figure 34. Purified AS LC-MS-neg chromatogram and characterized structure by NMR (Appendix 7).
Masses represents [M - H⁻ + FA]⁻.

3.7.3   Functional characterization of different ASAT3 isoforms via *in-vitro* enzyme assay

ASAT3, as characterized by Schilmiller *et al*. (2015), encodes an acyl-CoA-dependent acyltransferase that catalyzes the transfer of short chain branched acyl chains to the 3' position of the furanose ring. ASAT enzymes can also perform the reverse reaction and remove acyl chains rather than catalyzing the transfer of acyl chains (Fan *et al.*, 2016). We purified two different ASAT3 isoforms that were detected from *S. habrochaites sp. glabratum* - VI030462: 1) wild type (VI030462-W), and 2) cultivated type (VI030462-C). In addition, we have also isolated ASAT3 from *S. habrochaites* – LA1777, which has been well characterized by Schilmiller *et al*. (2015), as control. We performed a reverse reaction for 30 minutes on these three enyzmes, pET28 SN (without purification through HIS-tag column) as empty vector control, and no enzyme by supplying CoA to purified AS2 and AS3 to produce FA-CoA and AS with reduced acyl groups. Since AS is easily quantifed because of its stability, we expected to see AS with at least one C5 acyl group less starting from the 1' position via LC-MS. Both LC-MS spectra measured in positive and negative mode showed product formations with one C5 acyl group less, but it was difficult to determine whether the loss is located at position 1' or 3' based on fragmentation pattern. Moreover, we observed multiple peaks located at different retention times with one C5 acyl group removed (Fig.   35). Different products were formed due to the composition of the substrate that consisted of more than one isoform. A relatively high background peak product was visible for pET28 SN and the no enzyme control, when given AS2 as substrate possibly due to the impurity of the substrate. A small background peak was also observed when AS3 was given as a substrate. No additional significant peaks were observed when ASAT3 LA1777 was added to the reaction with AS3. It could be that the enzyme had lower activity than the ASAT3s from VI030462 or that we failed to produce active enzymes. However, in the case of ASAT3 VI030462-W and -C, five peaks were observed when given AS3 as substrate (Fig.   36). This means that the enzyme might not be specific towards certain C5 position. In the next example, a single product peak at 632.34 m/z, 13.89 min was found when ASAT3 VI030462-W was treated with AS4 (Fig.   37). This peak was composed of a sucrose backbone, with one C5 and one C12 acid attached. Since the NMR result showed that there was no acylation of the fructose ring of AS4, ASAT3 VI030462-W removed a C5 acyl group from the glucose ring. This result shows that the enzyme had higher affinity towards the type of acyl as to the regio-specificity of the acyl group in reverse reactions.
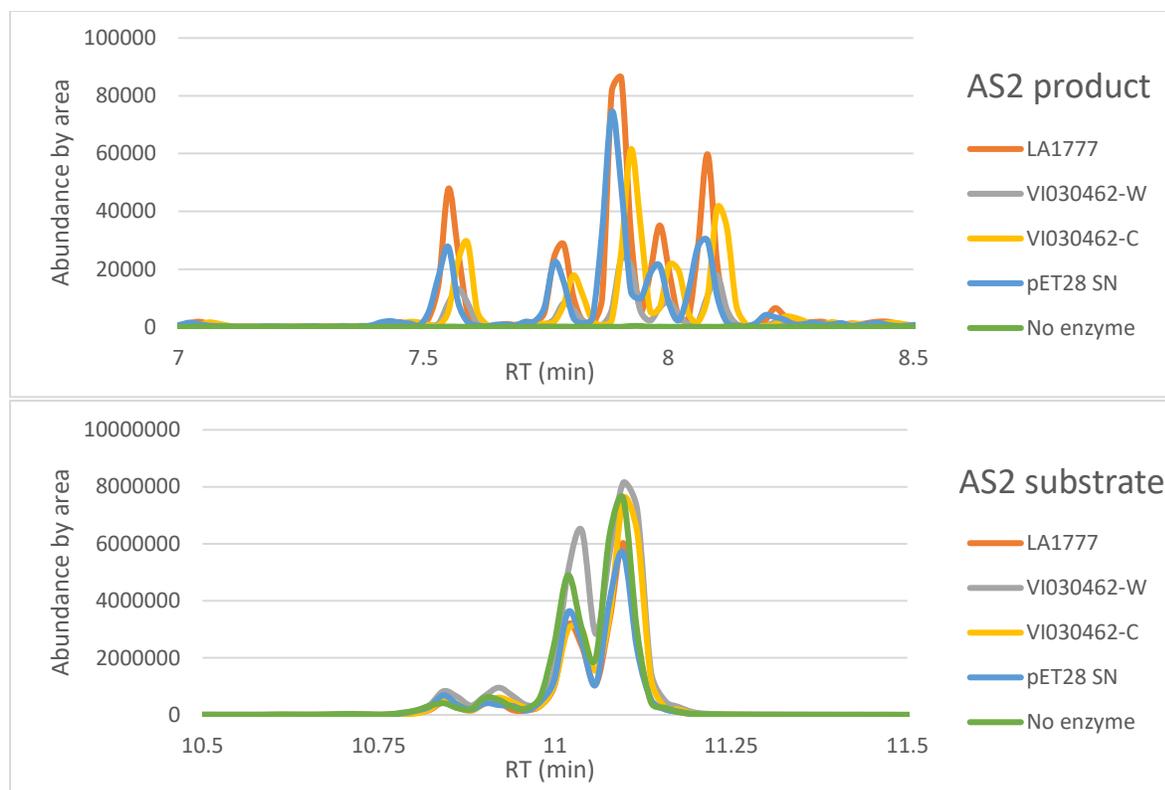
Figure 35. Spectra information and LC-MS result of ASAT3 invitro enyme assay of AS2 product (top) and substrate (bottom).
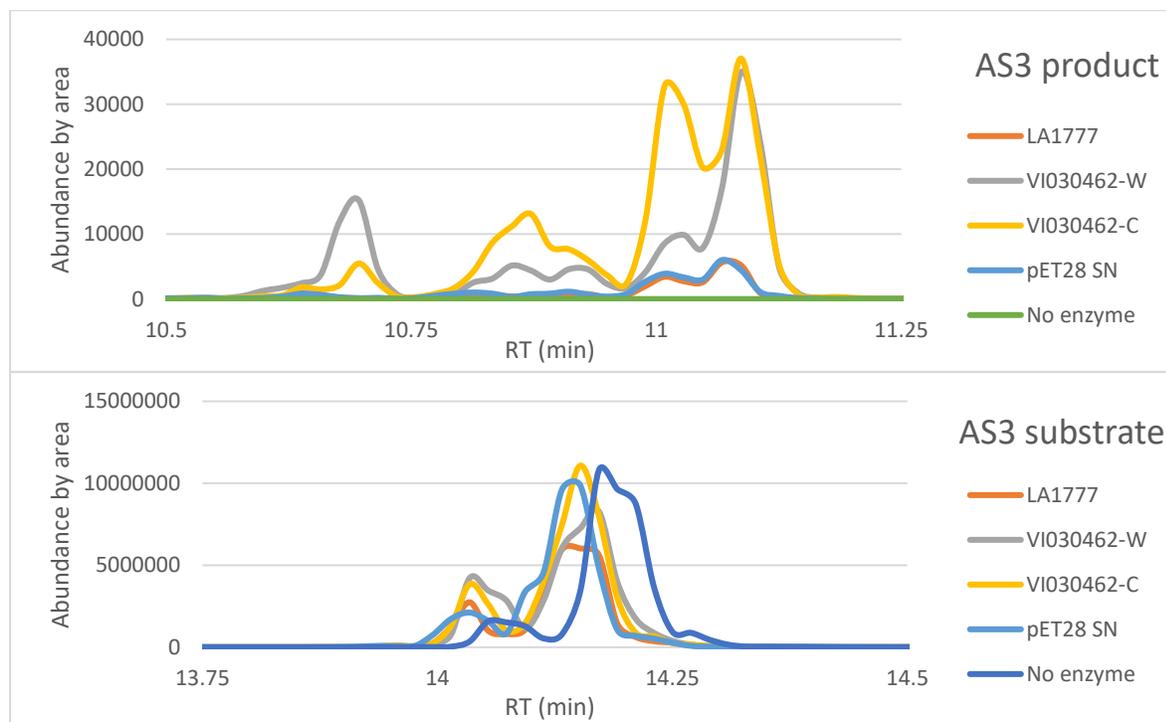
Figure 36. Spectra information and LC-MS result of ASAT3 invitro enyme assay of AS3 product (top) and substrate (bottom).
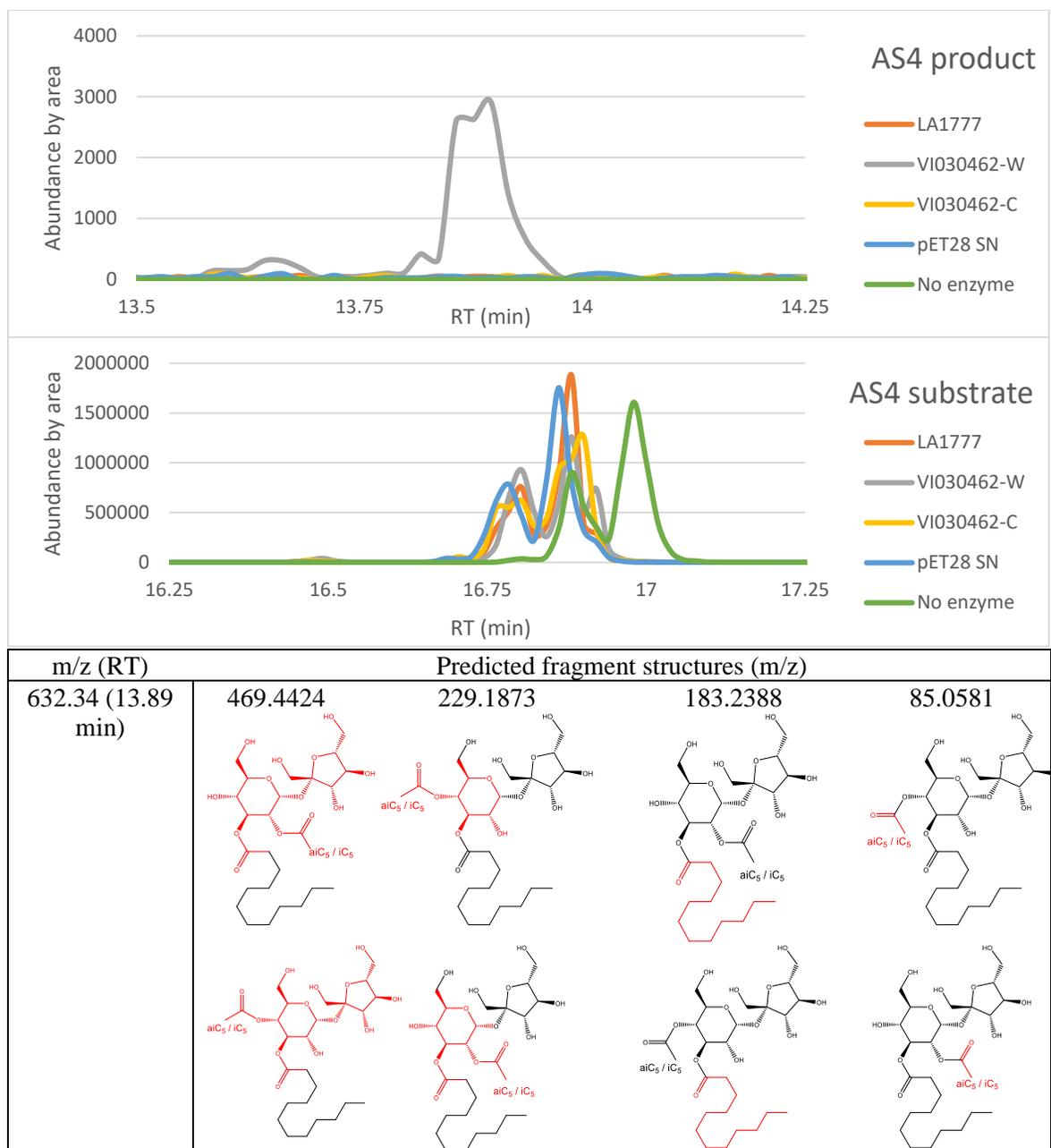
Figure 37. LC-MS result of ASAT3 invitro enyme assay of AS4 product (top) and substrate (center) as well predicted structures of AS fragment product (bottom) via SMpos.

To test whether ASAT3 -VI030462 attaches an acyl group at position 1', we conducted a one pot reaction which consisted of two different reactions that run consecutively. The first reaction is the biosynthesis of iC5-Coa from iC5 free fatty acid and Coa by the enzyme Sh-AACS2. Sh-AACS2, obtained from *S. habrochaites sp. glabratum* – VI030462, was characterized in this study and found to specifically biosynthesize short-chain FA-CoA (see 3.6). The second reaction is the transfer of the FA-CoA, which was produced from the first reaction to the AS substrate. For this reaction, we initially wanted to use synthetic AS with three iC5 acyls located at position 2, 3, 4 and no acylation at the fructose ring. However, we were only able to produce AS with acylations at position 2 and 6 of the glucose ring: 1) MIC026 - F1, and 2) - F2 (Appendix 5), produced with the help of synthetic chemists the the IPB, Yanira Mendez Gomez and Aldrin Vasco Vidal. Both MIC026 -F1 and -F2 are products of the same synthesis reactions, but differ in orientation of the sucrose backbone. Furthermore, they show a different LC-MS fragmentation pattern and elution time (Fig. 38). MIC026-F1 had the most similar stereochemistry to natural AS with an α, α - configuration. This means that the oxygen on the anomeric carbon of the cyclic sugar is on the opposite site of the ring relative to the substituent on the other carbon flanking the ring oxygen. Meanwhile, MIC026 - F2 has an α, β –configuration, where the oxygen on the anomeric carbon is on the same face of the ring as the substituent on the other carbon flanking the ring oxygen. Furthermore, MIC026 - F2 contains higher impurity even after additional purification through a normal phase column as depicted by the yellow/brown color. With this information, we suggest to use MIC026 - F1 as substrate to produce acylsugars for characterizing ASAT3 isoforms.
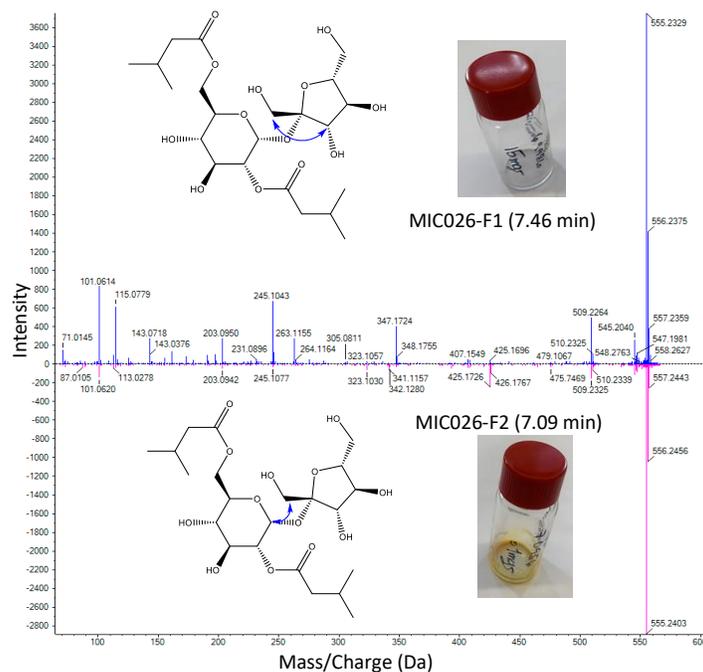
Figure 38. LC-MS spectra of synthesized acylsugar MIC026 – F1 and MIC026 – F2.

Fig. 39 A shows S3:15(5,5,5) formation at 10.70 min, the abundance area of each treatment was extracted and Fig. 39 B shows highest MIC026-F1 with an additional iC5 formation by *ASAT3* VI030462-W, compared to other isoforms and a control. An additional iC5 was also attached when MIC026-F2 was reacted with *ASAT3* VI030462-W. There was less product formed for MIC0026-F2 compared to MIC026-F1. The fragment masses of the MS2 peaks did not show specific differences when both products were overlaid (Fig. 39 C). Therefore, it is still difficult to determine the difference between both products (617.271 m/z) based on fragment similarity (Fig. 40). We propose to conduct MS3 for more in depth analysis. Moreover, the peaks at 10.70 min from *ASAT3* VI030462-W and C using MIC026-F1 as substrate were collected and measured with NMR. However, the quantity (< 100 μg) was not sufficient for characterization.
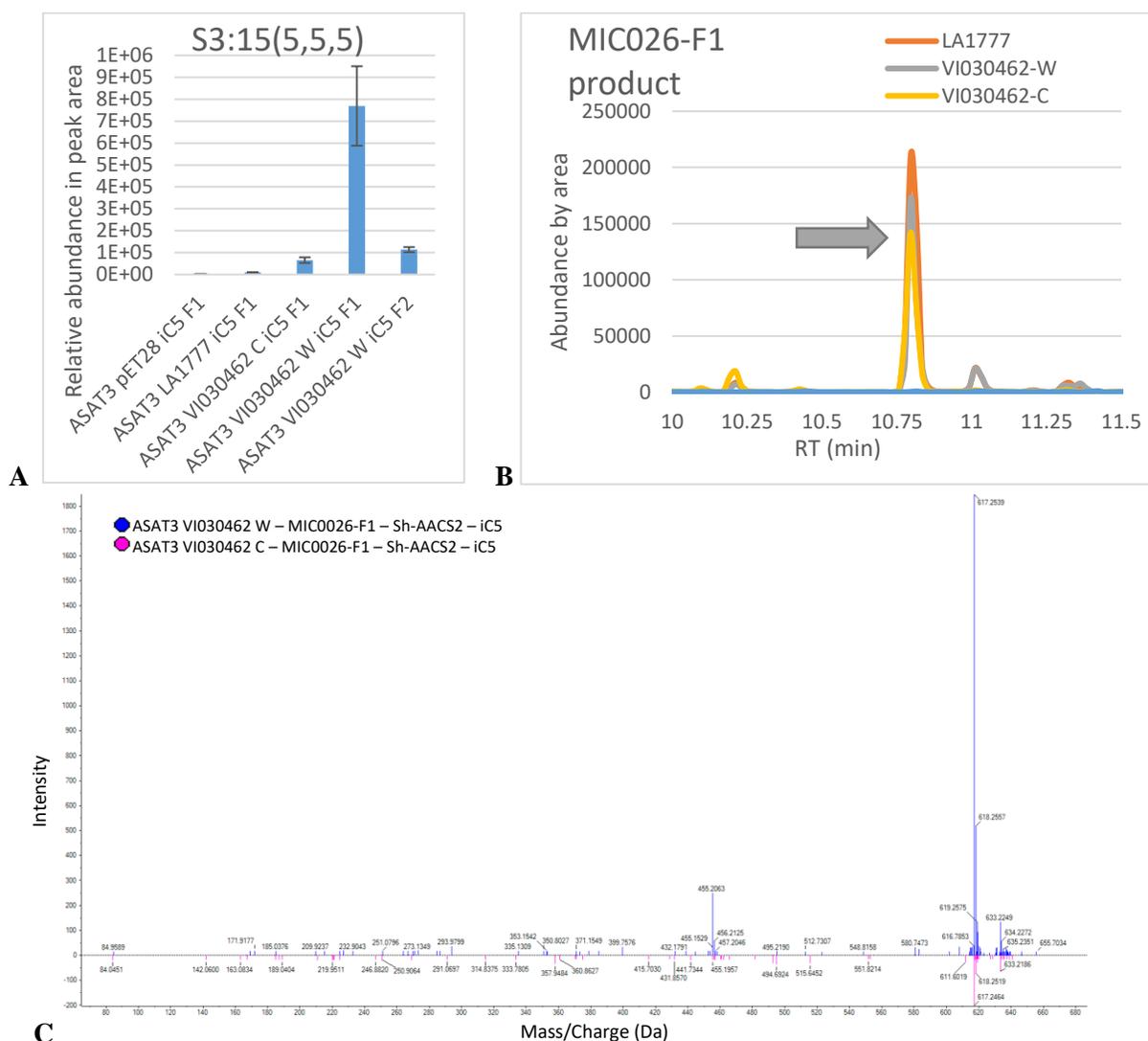
Figure 39. One pot enzyme result of different ASAT3 isoforms.

A) LC-MS measurement of enzyme reaction product, AS with three iC5 acylations. B) Location of collected product peak. C) Comparison of ASAT3 VI030462-W and -C product spectra.

Figure 40. Predicted structure of synthetic AS fragment product via SMpos.

m/z (RT)	Predicted fragment structures (m/z)
617.271 (10.69 min)	
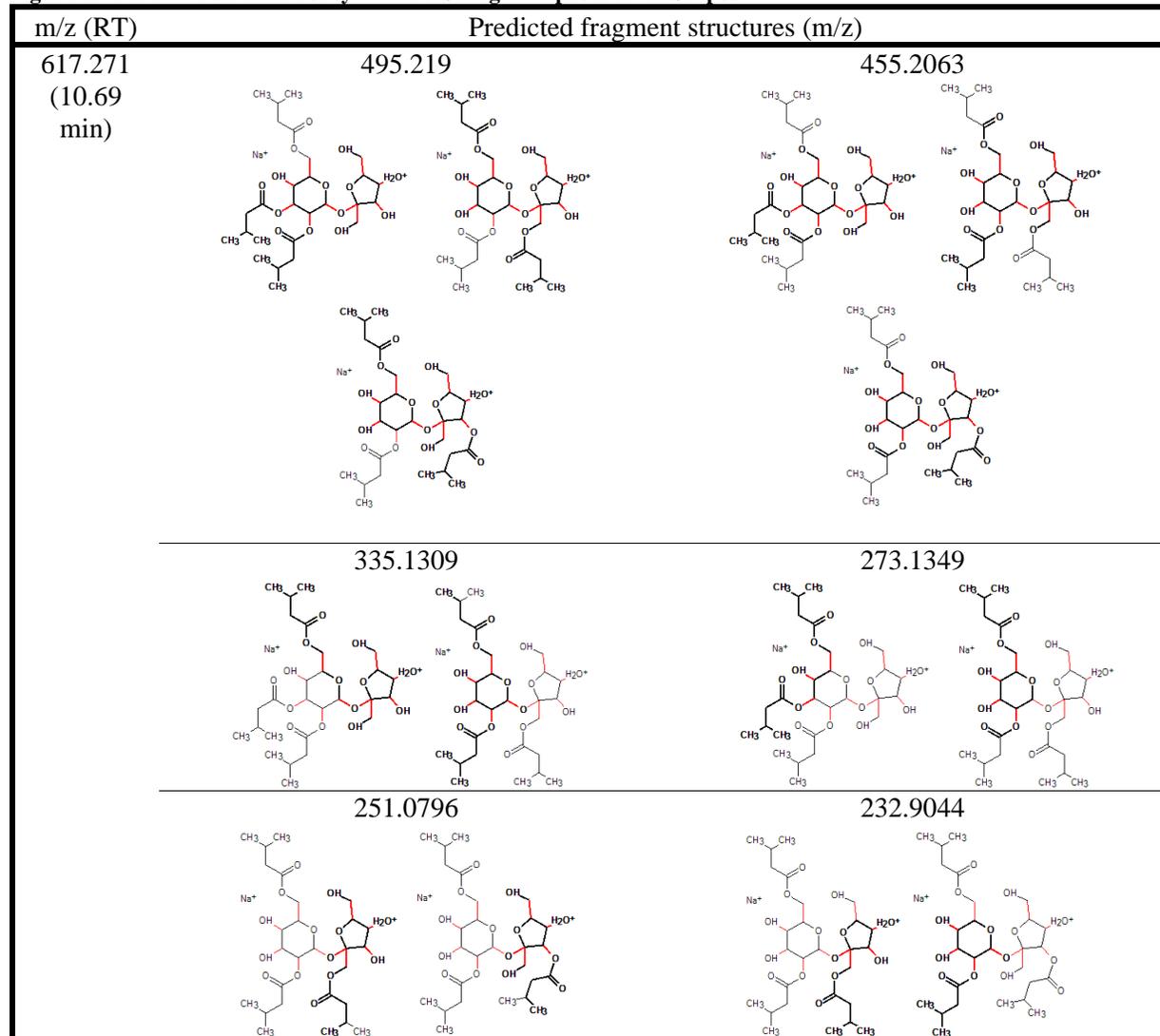
3.7.4. VIGS assay of different isoforms of ASAT3

To prove the function of ASAT3 *in-vivo*, we performed VIGS on *S. habrochaites sp. glabratum –* *VI030462-W, - C* and *S. habrochaites – LA1777*. To confirm that targeted genes are silenced, the relative expression of different ASAT3 isoforms were quantified in the plants by Q-RT-PCR. Q-RT-PCR result of VIGS treated plants showed no significant different between plants where *ASAT3* was silenced and the empty TRV2 control (Fig. 41 A). However, there was a slight trend of decreased expression in the *ASAT3* *VI030462 -C* and *-W* silenced plants compared to TRV2. AS quantification by LC-MS-neg also showed no significant differences between treatments via T-test, but again did show a slightly decreasing trend of *ASAT3* expression in *VI030462 -C* and *-W* compared to empty TRV2 (Fig. 41 B). Especially for AS1, the levels are already exceptionally low therefore small differences are not easily measured. When we compared the ratio between AS2, which has one C5 group less than AS3, we observed slightly higher ratio in *ASAT3* silenced plants compared to the empty TRV2 control (Fig. 41 C). This could mean that there was less acylation of AS2, producing less AS3 in the silenced plants, compared to the empty TRV2 control.
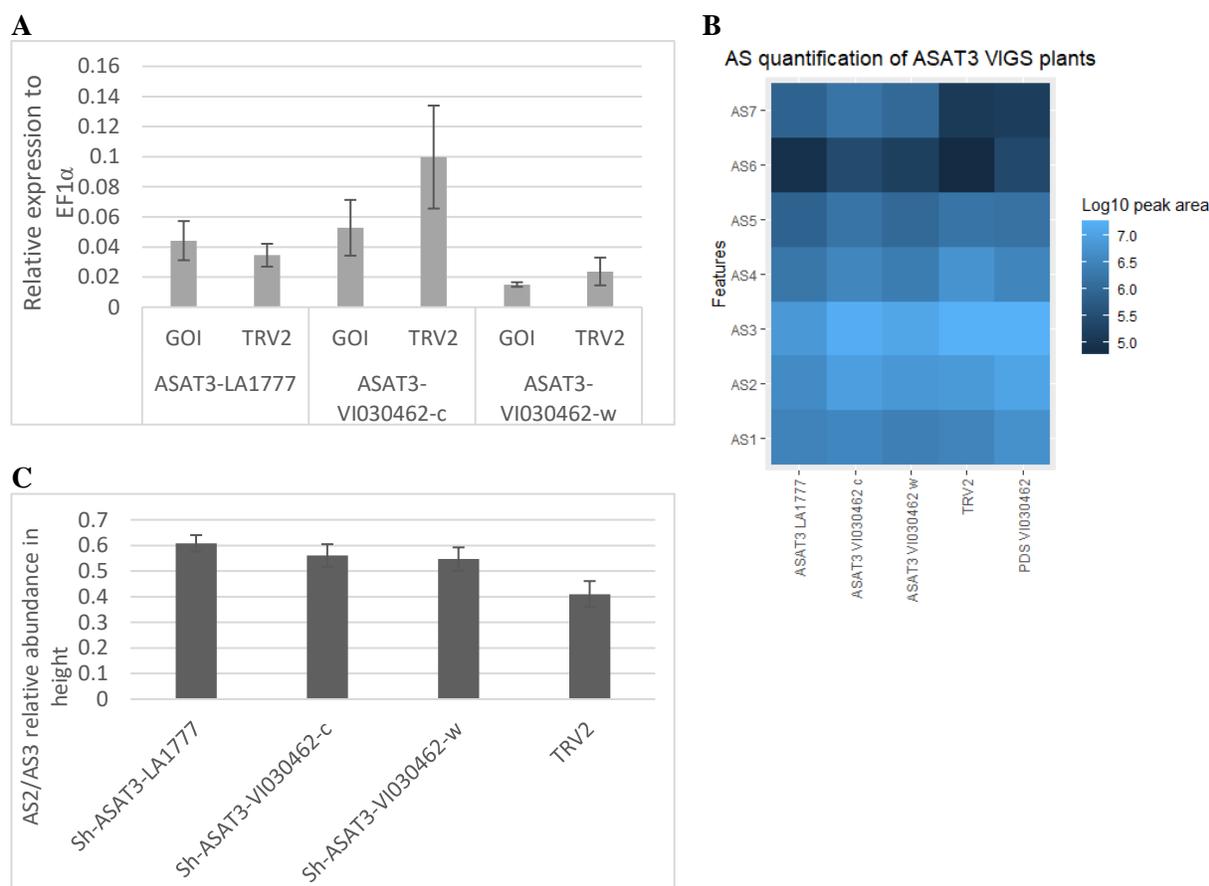
Figure 41. VIGS results of different ASAT3 isoforms.
A) Relative expression of different ASAT3 isoforms in VIGS treated plants obtained from Q-RT-PCR. B) AS quantification by LC-MS-neg of VIGS treated plants. C) Ratio of AS2/AS3 relative area abundance. GOI: gene of interest (silencing target *ASAT3*). TRV2: empty vector. AS1- S2:10 (5,5): 493.233 m/z, 7.83 min AS2- S2:15 (5,5,5): 594.289 m/z, 10.56 min AS3- S4:25 (5,5,5,5): 678.346 m/z, 13.78 min AS4- S3:22 (5,5,12): 776.456 m/z, 16.52 min. Error bar represent standard error.

3.8. Characterization of *Acyl-CoA Synthetases* (*AACS*)

3.8.1 Identification of *Acyl-CoA Synthetases* (*Sh-AACS*) from *S. habrochaites* sp. *glabratum* (VI030462)

Acyl-CoA is a group of coenzymes that metabolized fatty acids, which is required for the biosynthesis of many secondary metabolites including acylsugars. In our previous experiment, to characterize the function of different ASAT3 isoforms, we have conducted a one pot enzyme reaction which includes a reaction that produces fatty acid (FA)-coa that will be the substrate for synthesizing different acylsugars. In order to produce FA-coa, acyl-CoA synthetase (AACS) is required. Fourteen candidate acyl-CoA biosynthesis related genes were identified from the transcriptomics data (Appendix 6). These genes were annotated as either acyl-CoA synthetase, dehydrogenase or oxidase. The majority of genes were expressed higher in the trichomes compared to leaves. Although we will discuss each of these genes in brief, since they might be useful for other research, in this study we focused on genes annotated as synthetases.
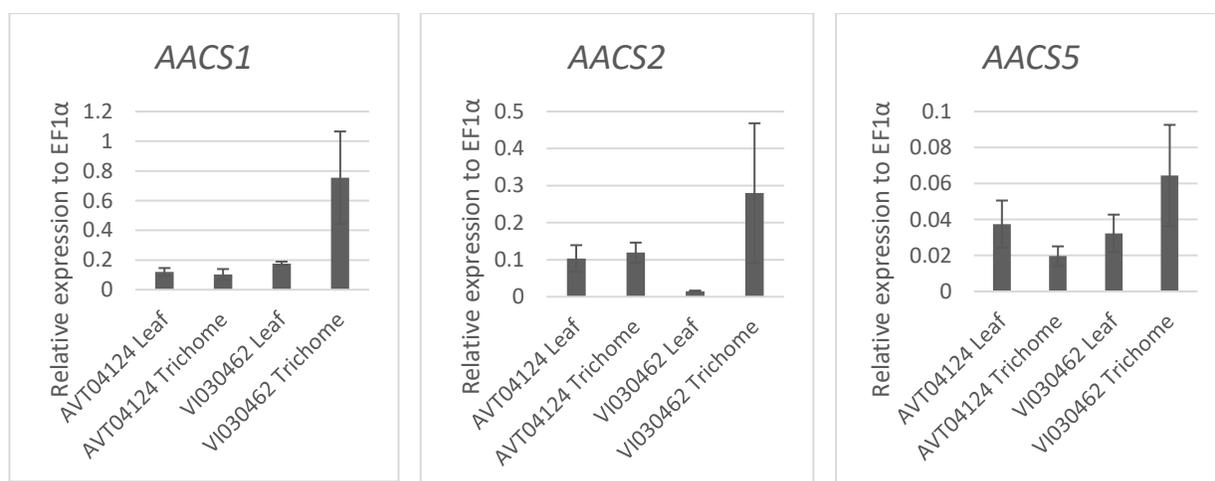
Figure 42. Q-RT-PCR result of acyl-coA biosynthetic related genes of wild – VI030462 and cultivated – AVT04124 tomatoes, that showed similar expression pattern as the RNAseq result.

AACS1 (Solyc07g043630), AACS2 (Solyc02g082880) and AACS3 (Solyc02g082870) are homologs of an ATP-dependent synthetase and ligase family protein (AT2G17650) from Arabidopsis (Kliebenstein et al., 2007). AT2G17650 is in clade VI of this protein family which is plant specific and showed short-chain or medium-chain acyl-CoA ligase activity (Schokey, *et al.*, 2003). AACS1 (Solyc07g043630) from *S. lycopersicum, S. pennellii*, and *S. quitoense* have recently been characterized by Fan *et al.* (2020) and shown to play a role in the production of medium-chain FA-CoAs. *AACS1* has two homologs based on our transcriptomics dataset, Solyc07g043640 and Solyc07g043660. Solyc07g043640 has very low expression in all tissue compared to Solyc07g043660 and *AACS1*. Solyc07g043660 is derived from recent duplication, however, its deletion did not have effect on trichome AS quantity (Fan *et al.,* 2020).

The expression of *AACS1* and *AACS2* was measured by Q-RT-PCR and showed a higher expression in the wild compared to cultivated plants (Fig. 42). Moreover, both genes were expressed higher in trichomes compared to leaves. The transcript of *AACS3* was not detected by Q-RT-PCR.

Two other candidate genes annotated as *Acyl-CoA Synthetases* are: *AACS4* (Solyc08g075810) and *AACS5* (Solyc12g044300). *AACS4* is homologous to *Acyl Activating Enzyme 1 (AAE1,* AT1G20560.1) of Arabidopsis. It was shown that *AAE1* is involved in the peroxisomal activation of long-chain fatty acids before they enter the β-oxidation cycle (Cassin-Ross & Hu, 2014). *AACS5* was predicted to show medium-chain FA-CoA ligase and 4-coumarate-CoA ligase activity based on the UniProt Knowledgebase. However, Al-Abdallat *et al.* (2014) reported that when they overexpressed the *Wax Inducer1/SHINE1* (*Sl-SHN1/WIN1*) gene in the Moneymaker cultivar background, higher cuticular wax deposition was observed under drought stress. *AACS5* was one of the cutin synthesis-related genes that was significantly upregulated in that study. Based on these reports, we speculate that these two genes are more specific to long-chain FAs. Q-RT-PCR results were only obtained for *AACS5* and the result showed similar pattern to *AACS1* and *AACS2*. However, we decided not to go forth with this gene in this study as we were interested in short-medium chain FAs.

Sh-AACS1 and *Sh-AACS2* genes were isolated and cloned from *S. habrochaites sp. glabratum -* VI030462 cDNA. Sequencing (Fig. 43) showed that the amino acid sequence of Sh-AACS1 was like AACS1 found in *S. lycopersicum*, *S. pennellii* and *S. quitoense*, which were characterized by Fan *et al.* (2020). Therefore, the sequence of *Sh-AACS1* is relatively conserved between different tomato species. Sh-AACS2 on the other hand showed a rather different amino acid sequence not only based on the fifty amino acids at the beginning of the protein. This provided a hint that differences in sequence may leads to different protein function and affinity towards different substrates.

```
Sh-AACS2   1  MNPFFNISRFNGLLHALNRVRVHPILSQRSRYLSQIIDK-NVETHPWESMEGLMRCSANY
Sq-AACS1   1  MNKFFQRSTMALR--F-NRSVQLTAHAQRVRKMCQHAGGIEPMDESQKLLEGLVTSPTNY
Sl-AACS1   1  MNKFFKTSNIALR--FFNGSVQLPAPTHRVRQLCQLAGSIESTDESRKLLEGVVTSPANY
Sh-AACS1   1  MNKFFKTSNIAVR--FFNGSVQLAAPTHRVRQLCQLAGSIESTDESQKLLEGVVTSPANY
Sp-AACS1   1  MNKFFKTSNIALR--FFNGSVQLPAPTHRVRQLCQLAGSIESTDESRKLLEGVVTSAANY

Sh-AACS2  60  FPLTPISFLDRAAKVFRDRTSVVYGSSVKFTWEETHNRCLKMASALSQLGISRGDVVATL
Sq-AACS1  58  VPLTPLSFLERAAKVFHDRTSVVFGSSVKYTWEETHSRCLKLASALVHLGISRGDVVATL
Sl-AACS1  59  VPLTPISYLERAADVFGDRTSVVFGSSVKYTWEETHSRCLKLASALIQLGISRGDVVATL
Sh-AACS1  59  VPLTPISFLERAADVFGDRTSVVFGSSVKYTWEETHSRCLKLASALIQLGISRGDVVVTL
Sp-AACS1  59  VPLTPISFLERAADVFGDRTSVVFGSSVKYTWEVSHSRCLKLASALIQLGISRGDVVATL

Sh-AACS2 120
APNVPAVQELHFAVPMAGAVLCTLNTRHDSAMVSVLLRHSEAKIIFVDQQLFDVAQGALD
Sq-AACS1 118  APNVPAMQELHFAVPMAGALLCTLNTRLDSSMVAQLLKHSETKIVFVDQQLLQIAQGALN
Sl-AACS1 119  APNVPAMQELHFAVPMAGAVLCTLNTRLDSSMVAYLLKHSETKMIFVDQQFLQIAQQALS
Sh-AACS1 119
APNVPAMQELHFAVPMAGAVLCTLNTRLDSSMVADLLKHSETKMIFVDQQFLQIAQQALS
Sp-AACS1 119
APNVPAMQELHFAVPMAGAVLCTLNTRLDSSMVADLLKHSETKMIFVDQQFLQIAQQALS
```

Sh-AACS2 180 LLADAKT--RPLLILIPDPENLPPPVAAPNVHEYETLLASGRDDFAIKWPLTEFDPISVN
Sq-AACS1 178 LLSKDKIIKPPILVLITESENFAP-----NVHEYENLLTSGSSNFTIRWPKTEMDPISIN
Sl-AACS1 179 LLSKDKTIKPPILILIPESNDSSPPV--SNIHEYENLLSSGSSNFTIRWPKSEFDPISIN
Sh-AACS1 179 LLSKDKTIKPPILILIPKSNNSSPPA--SNIHEYENLLSSGSSNFTIRWPKSEFDPISIN
Sp-AACS1 179 LLSKDKTIKPPILILIPKSNNSSPPA--SNIHEYENLLSSGSSNFTIRWPKSEFDPISIN

Sh-AACS2 238
YTSGTTSHPKGVVYNHRGAYLNAIATPYTHEMGSMPTYLWTVPMFHCNGWSLTWGVAALG
Sq-AACS1 233
YTSGTTSSPKGVVYSHRGAYLNTIASFWSQGMGTMPTYLWTLPMFHCNGWCMIWGLAAIG
Sl-AACS1 237
YTSGTTSSPKGVVYNHRGAYLNSISAFLCHGMALMPTYLWTLPMFHCNGWCMNWGVAALG
Sh-AACS1 237
YTSGTTSSPKGVVYNHRGAYLNSISAFLCHGMGPMPTYLWTLPMFHCNGWCMNWGMAAIG
Sp-AACS1 237
YTSGTTSSPKGVVYNHRGAYLNSISAFLCHGMAPMPTYLWTLPMFHCNGWCMNWGMAAIG

Sh-AACS2 298 GTNVCLRRVSPKDIFENISLHKVTHMSAAPTVMNMIVNSPKSDRKPLPHKVEITTGGSPP
Sq-AACS1 293 GTSICLRHVTAKHIFESISLYQVTHMGAVPTVLSMIANCPPNDRKPLPHKVQIVTGGSAP
Sl-AACS1 297 GTNVCLRHVSAKDIFESISVNKVTHMSAAPIVLSMMANASPNDRKPLLHKVEIMTGGSPP
Sh-AACS1 297 GTNVCLRHVSAKDIFESISVNKVTHMSAAPIVLSMMANASPNDRKPLPHKVEIMTGGSPP
Sp-AACS1 297 GTNVCLRHVSAKDIFESISINKVTHMSAAPIVLSMMANASPNDRKPLPHKVEIMTGGSPP

Sh-AACS2 358 PPHIISKMEELGFSVSHIYGLTEIHGPCMSCLHQPEWESLPPDERFALKARQGVEHYFTQ
Sq-AACS1 353 PPQILSKMEELGFGVIHGYGLTETYSAATSCVWKPEWDSLPLEERAVIKSRQGVQHLCIE
Sl-AACS1 357 PPQILSKMEQLGFGVSHGYGLTETYSGATTCLWKPEWDSLPLEERAALKSRQGVKVLCIE
Sh-AACS1 357 PPQILSKMEQLGFGVSHGYGLTETYSGATTCLWKPEWDSLPLEERAALKSRQGVQVLCIE
Sp-AACS1 357 PPQILSKMEQLGFGVSHGYGLTETYSGATTCLWKPEWDSLPLEERAVLKSRQGVQVLCIE

Sh-AACS2 418 GIDIRDPDTMERVPDDGKTLGEIMIKGNTVMSGYLKNIKATEEVFRGGWFHTGDLAVRHP
Sq-AACS1 413
EVDVRDPETMEKVPADGKAIGEIVCRGNTVMNGYLKDVEATKEAFKGGWFHTGDLAVKHP
Sl-AACS1 417
RVDVRDPETMENVPADGKSIGEIVCRGNTVMSGYLKDVKSTEEAFKGGWFHTGDVAVKHP
Sh-AACS1 417
KVDVRDPETMENVPADGKSIGEIVCRGNTVMSGYLKDVKATEEAFKGGWFHTGDVAVKHP
Sp-AACS1 417
KVDVRDPETMENVPADGKSIGEIVCRGNTVMSGYLKDVKATEEAFKGGWFHTGDVAVKHP

Sh-AACS2 478 DGYIEVKDRMKDIIISGGENICSVEVERVLVSHPAVLEAAVVARPDDHWGQTPCAFVKLK
Sq-AACS1 473 DGYIEIKDRLKDIIISGGENISTIEVERVLYSHPAVVHAAVVARLDDHWGQVPCAFVQRK
Sl-AACS1 477 DGYIEIKDRLKDIIISGGENISTLEVEGVLHSHPAVVEAAVVARPDDHWGQTPCAFVKLK
Sh-AACS1 477 DGYIEIKDRLKDIIISGGENISTLEVEGVLHSHPAVVEAAVVARPDDHWGQTPCAFVKLK
Sp-AACS1 477 DGYIEIKDRLKDIIISGGENISTLEVEGVLHSHPAVVEAAVVARPDDHWGQTPCAFVKLK

Sh-AACS2 538 EGFS-LGSEDIINYCRDHLPHYMAPQTVIFEDLPTTSTGKIQKFVLREKAKALGSVCEIK
Sq-AACS1 533 EGFEEITSDEIIKFCRDHLPHYMAPRAVLFEDLPMTSTGKVQKFILREKAKALPNLFNNE
Sl-AACS1 537 EGSEEITSDEIIKYCRDHLPHYMVPRAVVFQDLPTTSTGKVQKFILREKAKALASLFNTD
Sh-AACS1 537 EGSEEITSDEIIKYCRDHLPHYMVPRAVVFQDLPRTSTGKVQKFILREKAKALASLFNTD
Sp-AACS1 537 EGYEEITSDEIIKYCRDHLPHYMVPRAVVFQDLPRTSTGKVQKFILREKAKALASLFNTD

Sh-AACS2 597 REIAV
Sq-AACS1 593 KQV--
Sl-AACS1 597 RKV--

Sh-AACS1 597 `RKV`--
Sp-AACS1 597 `RKV`—

Figure 43. Multiple alignment of amino acid sequences of AACS isoforms from different species.
Sh-AACS1, Sh-AACS2, Sl-AACS1: MT078737.1 (*S. lycopersicum*), *Sp-AACS1*: MT078735.1 (*S. pennellii*), and *Sq-AACS1*: MT078732.1 (*S. quitoense*).

3.8.2 Functional expression, detection of activity and substrate compatibility of *Sh-AACS*

Based on expression and sequence results, the cDNA of *Sh-AACS1* and *Sh-AACS2* was inserted into an expression vector containing an N-terminal His-tag. The active enzymes are monomers with a calculated mass of 66.89 kDa and 65.75 kDA for *Sh-AACS1* and *Sh-AACS2,* respectively. Although the enzymes were soluble and were purified by affinity chromatography, we could not obtain very pure enzymes (Appendix 3). This can mean that the current vector is not suitable enough for purification and optimization of the protocol could still be conducted to increase production. However, the amount of purified enzyme was sufficient to perform activity assays.

Abundance by area
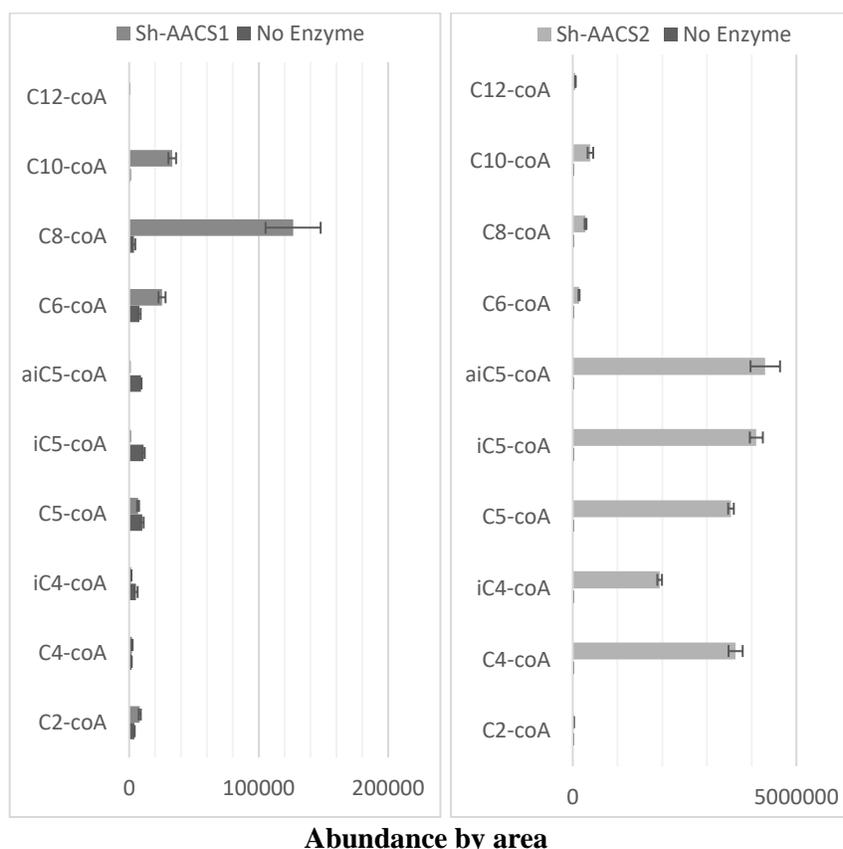
Figure 44. Sh-AACS1and Sh-AACS2 substrate specificity assay using different short-medium chain fatty acids.

I*n-vitro* enzyme assays were carried out using the previously isolated proteins. The aim was to determine whether the enzymes are active and whether the difference in sequence might explain protein substrate specificity for different FAs (Fig.   44). Sh-AACS2 is an active protein and results showed that there was relatively high product formation when C4, iC4, C5, iC5, and aiC5 FAs were added as substrates. Low product formation was observed for C6, C8, and C10. This showed that this enzyme has high affinity towards short chain FAs. Sh-AACS1 is also an active protein and the results showed that it was highly active when using C6, C8, and C10 FAs as substrate. There was a background level of product formation when using C2, C4, iC4, C5, iC5, aiC5 FAs, andC12 as substrate. This was not the case for Sl-AACS1 (Fan

et al., 2020), an isoform of Sh-AACS2, which was able to also form C12-CoA. These results showed that a small difference in protein sequence was sufficient to affect protein specificity.

An enzyme kinetic experiment was conducted with Sh-AACS2 for iC5 and aiC5 FAs as these two short chain FAs were the most abundant short-chain FAs found in VI030462 (Fig. 45 A). The Michaelis-Menten constant (K_m) was calculated using 10-100 µM FAs as substrate over thirty minutes. Observations showed that Sh-AACS2 had higher affinity towards iC5 (K_m 159.020 µM) than for aiC5 FA (K_m 283.000 µM) (Fig. 45 B). k_{cat} values were calculated to describe the turnover rate of enzyme substrate complex to product, note however that these values were not based on pure enzymes. A high catalytic efficiency (k_{cat}/K_m) was also observed nonetheless, a measure for the proportion of substrate converted into product.

A

B	Substrate	K_m (µM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
	iC5	159	0.596	3747
	aiC5	283	0.407	1437
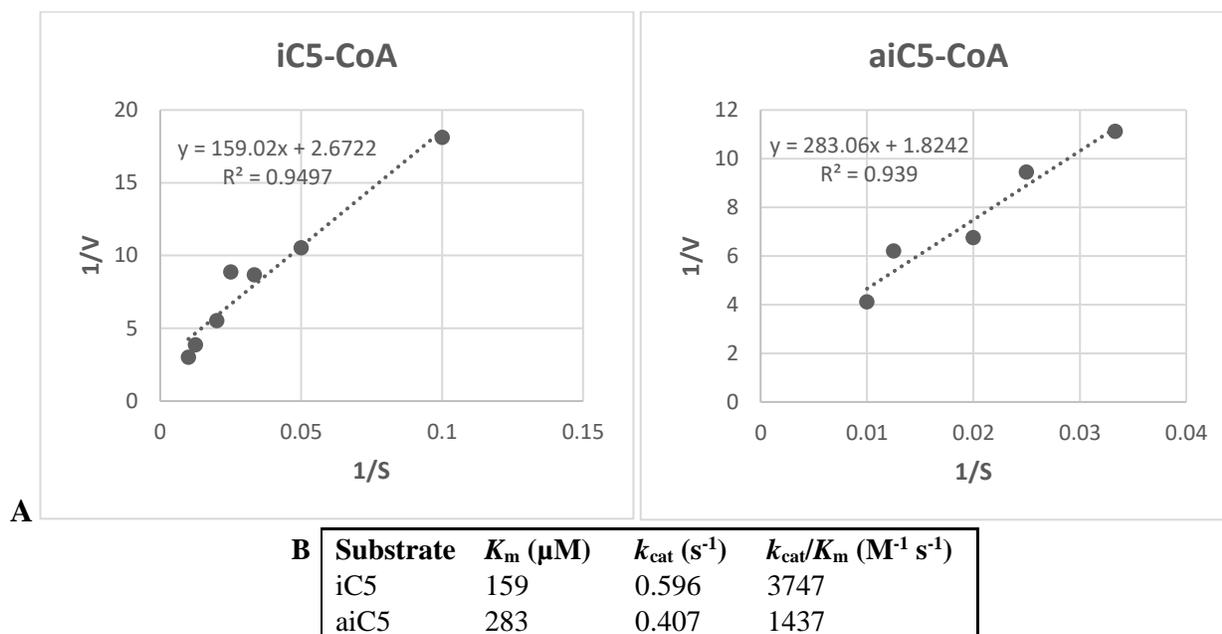
Figure 45. AACS2 enzyme kinetic results.
A) Enzyme activity of Sh-AACS2 using iC5 and aiC5 FAs as substrate. B) Enzyme kinetic results indicating K_m, k_{cat}, and k_{cat}/K_m. Note that these values are not based on very pure proteins. V: µM/min. S: µM.

3.8.3 Sh-AACS VIGS experiment

VIGS was conducted to characterize the function of both AACS1 and AACS2 in *S. habrochaites sp. glabratum* – VI030462 tomato plants. Infiltration was conducted on the cotyledons two weeks after germination and samples were taken from the 2-4th leaves, five weeks post-infection. We expected to see lower level of acylsugars (AS) when measured by LC-MS as well as lower expression of *AACS1* and *AACS2* via Q-RT-PCR in silenced plants. Empty vector TRV2 was used as negative control and should show no differences in AS levels and related gene expression. The expression of both genes was relatively low even though they were not significantly different to the control when assessed via T-test (Fig. 46). However, there is a trend that showed lower expression in VIGS construct targeting the gene of interest (GOI) and the not in the TRV2 control. AS were expected to be reduced in AACS silenced plants. Results showed that both *Sh-AACS1* and *Sh-AACS2* silenced plants showed a trend of reduced AS content compared to the empty vector control (Fig. 47). In addition a positive control, PDS silenced plants, also showed a reduction in AS. PDS converts phytoene to colored ξ-carotene in a two-step desaturation reaction that is important in carotenoid metabolism and photosynthesis. Photosynthesis does not only take place in leaves but also in the secretory cells of glandular trichomes that are involved in the secretion of specialized metabolite (Laterre *et al.*, 2017). Therefore when photosynthesis is affected, this also affects the production of AS produced by the type IV trichomes.
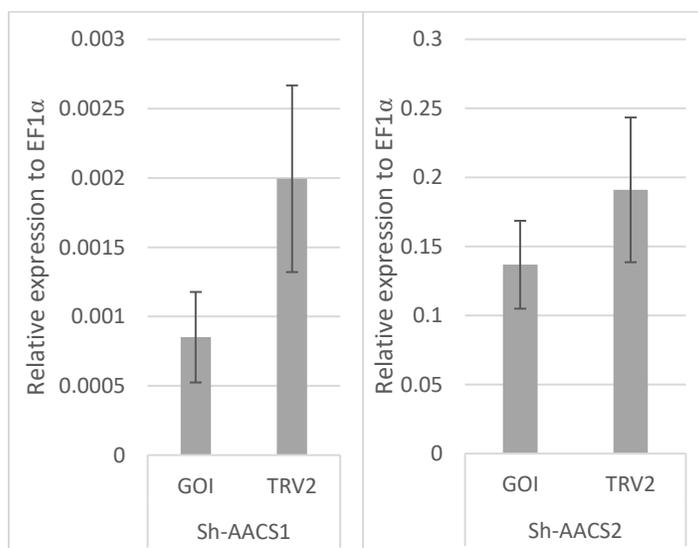
Figure 46. Q-RT-PCR result of VIGS silenced *Sh-AACS1* and *Sh-AACS2* genes in *S. habrochaites sp. glabratum* – VI030462 tomato plants.
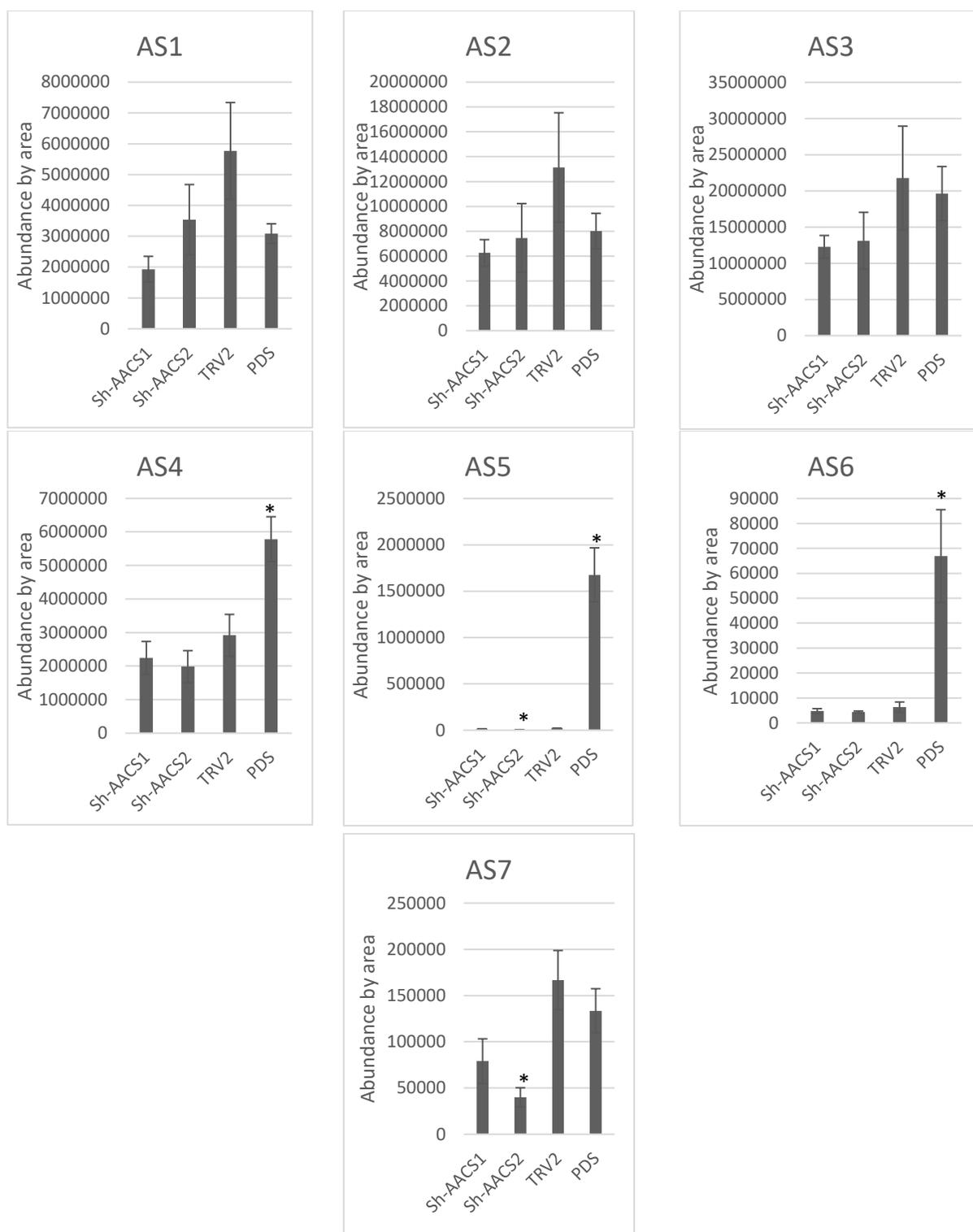
Figure 47. Acylsugar quantification of VIGS silenced *Sh-AACS1* and *Sh-AACS2* genes in *S. habrochaites sp. glabratum* – VI030462 tomato plants.
AS1- S2:10 (5,5): m/z 493.233, 7.83 min; AS2- S2:15 (5,5,5): m/z 594.289, 10.56 min; AS3- S4:25 (5,5,5,5): m/z 678.346, 13.78 min; AS4- S3:22 (5,5,12): m/z 776.456, 16.52 min; AS5- S3:20 (5,5,10): m/z 664.367, 15.22 min; AS6- S3:19 (4,5,10): m/z 650.351, 14.51 min; AS7- S3:16 (5,5,6): m/z 608.304, 11.8 min.

4. Discussion

Natural selection has resulted in the evolution of morphological, behavioral, and biochemical diversity among plants and herbivores. Plants have intricate and dynamic defense systems in response to different insects and other pests. This ongoing co-evolution resulted in development of morphological and biochemical defensive traits to dodge each other's strategies (Fig. 48). Some defensive traits that evolved include the formation of a physical barrier like trichomes, thick waxy cuticles and secondary metabolites such as AS, flavonoids, and VOC. These traits can be constitutively present (*e.g.* phytoanticipins) or induced through different stresses (*e.g.* phytoalexins), thereby affecting the plant's phenotype morphologically and biochemically. These responses affect whitefly probing, as well as post-penetration, pre-phloem resistance mechanism and phloem-located factors (McDaniel *et al.*, 2016). Due to the complexity of plant-herbivore interaction, this study could only capture a small part of the tomato defense against *B. tabaci* and explore potential resistance. The focus here lays on acylsugars and susceptibility factors such as polyamines from *S. habrochaites* sp. *glabratum* population x *S. lycopersicum* introgression lines. Acylsugars are often targeted for insect resistance breeding since they play a role as phytoanticipins that are secreted by tomato trichomes prior to insect attack (Priani & Vendramim, 2010; Lucini *et al.*, 2015; Ben-Mahmoud *et al*., 2019). Polyamines on the other hand have not been well explored by breeders, but several reports have shown their role as susceptibility factors (Subramanyam *et al.*, 2015; Marti *et al.*, 2013).
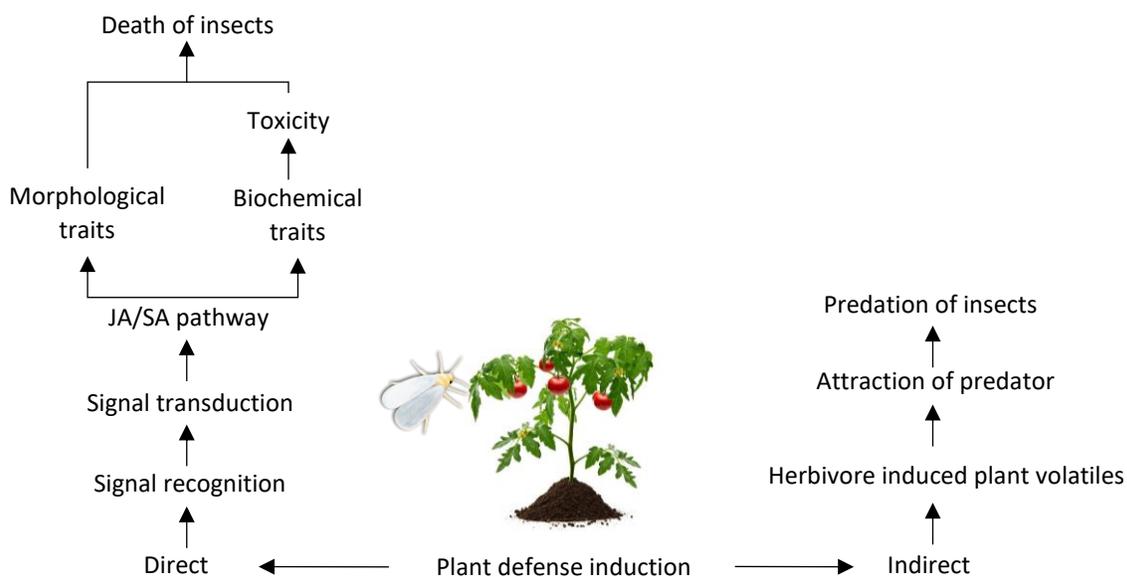
Figure 48. An overview of plant-herbivore interactions.
Plant defense may directly affect insect growth and development through morphological traits or toxic secondary metabolites as well as by luring natural enemies through herbivore induced plant volatiles.

4.1. Phenotyping and genetics of whitefly resistance from *S. habrochaites sp. glabratum* x *S. lycopersicum* introgression lines

Crop domestication can be described as an ongoing process of selection that has resulted in the crops that we feed on today. However, successive selection for favorable traits has led to the reduction of genetic variation compared to their wild progenitors. Over the years, breeders have favored high crop productivity and uniformity. Because of recurrent selection, potentially valuable genetic variation that is associated with (a)biotic stress responses have been filtered out of the gene pool. One of its implication is the loss of insect resistance related genes. Moreover, a lot of crop producers rely on the application of chemical pesticides which is quick, easy, and inexpensive solution for controlling insect pests, but pollute and contaminate the environment. Alternatively, producing resistant varieties through breeding has been conducted to restore genetic diversity of crops by reintroducing genetic variation from wild accessions/species which had been selected out. These efforts have not only been made in Solanum species against whitefly - *Bemisia tabaci* (Vosman *et al.,* 2019) but also in other crops such as in Capsicum species (Firdaus *et al*., 2011) and in Brassica species against cabbage whitefly - *Aleyrodes proletella* (Pelgrom *et al*., 2014).

Herbivores such as whiteflies are sap-sucking insect that are also vector for a lot of plant viruses transmitted through the insects from plant to plant. Therefore, exploring the diversity in leaf surface morphology and the chemical composition of leaf/trichome secondary metabolites of wild tomato species offers a feasible approach to increase tomato resistance against herbivores (Lucatti *et al.,* 2013; Bar & Shtein, 2019; Mandal *et al.,* 2020). The wild species *S. habrochaites* sp. *glabratum* (VI030462) that was chosen for this study differs from the green fruited tomato *S. habrochaites* (LA1777), which was widely utilized for trichome research (Momotaz *et al.,* 2010; Kim *et al.,* 2012; Bennewitz *et al.*, 2018). One main difference is that this *S. habrochaites glabratum* sub-species is partially self-compatible in contrast to LA1777, allowing self-pollination and thereby facilitating breeding programs (Bedinger *et al.*, 2010; Covey *et al.,* 2010). Moreover, differences can also be measured based on leaf and trichome specific metabolite , composition, such as acylsugar (AS) (Gosh *et al.,* 2014; Schilmiller *et al.,* 2015). The acylsugars of the *glabratum* species are mainly acylated with C5 and C12 chains. This reduces the complexity when characterizing substrate specificity of enzymes involved in acylsugar biosynthesis. Another difference is that, unlike LA1777, this wild species does not contain a large amount of sesquiterpene carboxylic acids produced in glandular trichomes that enhance tomato host plant resistance towards fruitworm-*Helicoverpa zea* and beet armyworm- *Spodoptera exigua* (Frelichowski & Juvik, 2001). This means that the *glabratum* species have other resistance factors besides sesquiterpene carboxylic acids to be resistant towards whiteflies.

When we compared between the cultivated *S. lycopersicum* (AVT04124) and wild *S. habrochaites.* sp. *glabratum* (VI030462) tomato accessions, the differences could be seen microscopically from the density of type I and IV glandular trichomes, as well as in the metabolite level secreted on the leaf surface. Furthermore, these phenotypic differences were shown to correlate with insect mortality and oviposition. Higher number of insect mortality and lower oviposition were observed when there were higher type IV trichome density and more secreted AS were present, similarly to results from an association study conducted by Andrade *et al.* (2017) on a different wild accession, *S. galapagense* (LA1401). In this study, we also confirmed that type VI trichomes, that are known to produce terpenes, showed no correlation to either insect mortality or oviposition as was also found in the choice and no-choice assay of 22 different wild tomato accessions (Rakha *et al.*, 2017). Based on these findings, we could conclude that the specific blend of secondary metabolites, especially terpenoids, produced by the type VI trichomes of VI030462 does not play a significant role in whitefly attraction or repellent for this specific genetic background. This could be due to the high density of type IV trichomes, which might masks the potential effect of type VI trichomes. Therefore, we decided to look at individual metabolic features extracted from the GC-MS analysis. The two most abundant monoterpenes with known insect repelling action, limonene and phellandrene (Falara *et al.*, 2011; Zhou & Pichersky, 2020), did not not show significant difference between susceptible and resistant lines in the segregating population. Furthermore, we have identified several putative features as mono- and sesquiterpenes that were significantly different using ANOVA between resistant and susceptible lines of the F2 and BC1F2 population (Table 3). Unfortunately, we did not manage to characterize these features to see whether they have repellent action, like zingiberene and other sesquiterpene derivatives have against whiteflies (Bleeker *et al*., 2011; Zabel *et al.,* 2021).

In addition to genetic and metabolic differences, VI030462 was shown to be resistant not only to whiteflies (*Bemisia tabaci*) but also to spider mites (*Tuta absoluta*) in a no-choice insect assay (Rakha *et al.,* 2017); making this accession a good parental line for an inbred cross (IBC). The advantage of IBC is that unlinked donor fragments are separated by segregation and the linked donor fragments are minimized due to the recombination with the recurrent parent. To reduce the number and size of donor fragments, backcrossing is repeated, generating advanced backcross lines. IBC populations is also called an immortal population because it is inbred and can be propagated by self-pollination if they are self-compatible. This allows replication of the evaluation in multiple environments, hereby increasing the precision of trait measurements and single factor analysis. However, there are disadvantages of IBC populations. They work poorly for quantitative traits and for epistatic interaction studies - the interaction of multiple and unlinked genes from the donor line. Furthermore, it is more difficult to select for recessive traits because there is a lower frequency for homozygous recessive alleles. Lastly, many backcrosses are required to produce a new

commercial variety because of the difficulty in removing unwanted genes of the donor from the elite cultivar due to linkage drag.

In this study, we have integrated several omics approaches such as genomics, metabolomics, and transcriptomics to investigate the metabolic networks aiming involved in plant-insect interactions. These approaches facilitated the discovery of QTLs and locating of genes regulating the biosynthesis of metabolites correlating with the traits of interest. However, we faced several challenges when working with this population. VI030462 is a heterogeneous wild tomato accession, which results in a few drawbacks which includes high variability, false positives/negatives due to limited sample size, and the requirement of more accurate phenotyping method. Sib-mate crossing was used to generate the F2 population and preselection of the most resistant plants was made to generate BC1F2 population. This intentionally adds bias, also called preselection bias, thereby resulting in confounding associations due to the population structure and linkage disequilibrium.

To overcome these bottlenecks, tested lines were subjected to genotyping by sequencing for the identification of genetic variation in the form of single nucleotide (SNPs) polymorphisms and rapidly genotype samples. This SNPs are used as markers to track the QTL regions that cosegregate with the trait of interest. Generally, QTL regions that segregates for a phenotype producing a LOD score greater than 3 which indicates a 1000 to 1 odds that observed linkage observed did not occur by random chance. However, we have adapted permutation / resampling to avoid asymptotic approximations by replicating and reshuffling the original trait data, while leaving the marker unchanged (Churchill & Doerge, 1994). This approach accounts for missing marker data, actual marker densities and non-random segregation of marker alleles, thereby establishing higher statistically significant QTLs. Using this approach, we have selected significant markers and focused on two main QTLs on chromosome 5 and 11, which is discussed in the preceeding sub-chapters.

In conclusion, we have collected comprehensive information on resistance to whitefly, consisting of biochemical, and genetic analysis as well as transcriptomics of VI030462. This information will be useful for future research. Furthermore, the population can be used to investigate resistance to other herbivores or pathogen.

4.2. Comparative untargeted metabolomics and transcriptomics reveals metabolic families and features associated with whitefly resistance

To identify which metabolic features and genes have a role in plant insect interaction during whitefly infection, we used a multi-omics approach: transcriptomics and metabolomics. Genome-wide gene expression analysis by RNAseq has become a standard technique in molecular biology not only to understand genes that are affected in tomatoes due to whitefly infection (Li *et al.,* 2021; Chen *et al.,* 2016; Ding *et al.*, 2019), but also in uncovering tomato genes that are hijacked by whiteflies to counteract the plant defence systems (Xia *et al*., 2021). RNAseq has also been emplyed for the identification of genes involved in the biosynthesis of methylketones of *S. habrochaites* sp. *Glabratum*, which were shown to have insecticidal activity (Ben-Israel *et al*., 2009; Yu *et al*., 2010). In our study, RNAseq was conducted to select for candidate genes involved in whitefly resistance from trichomes and trichomeless leaves of the wild *S. habrochaites* sp. *glabratum-* VI030462 and cultivated *S. lycopersicum-* AVT01424. There were several challenges encountered during sample preparation. Cultivated tomatoes do not produce a lot of glandular trichomes (GT), therefore samples were collected using the brush method in liquid nitrogen from 15 five-weeks old greenhouse grown tomato plants every two weeks during the summer period and then pooled. Individual pooled samples were stored at -80°C. Next, we faced difficulties in obtaining enough RNA above RIN value 8, especially from the wild tomatoes. High amount of acylsugars caused the extraction column of standard RNA extraction kits that were available in the lab to clog. Therefore, we suggest using the eqGOLD RNA extraction kit with buffer P by VWR™ as an alternative.

In the last few decades, a significant amount of data has been generated via metabolomics technologies, allowing researchers to understand the metabolism of many biological systems. To successfully translate the data generated from a metabolomics experiment, one must start with a robust experimental design followed by data acquisition, mining, interpretation, and validation of candidate features. Several challenges faced in the development of metabolomics include 1) the absence of a universal method to measure the complete metabolome; 2) difficulties to obtain accurate quantification due to impurities, systematic matrix effects, and reproducibility; and 3) availability of validated spectra information for metabolite identification and pathway mapping. One primary bottleneck is the amount of time and manual effort required for analyzing the data, especially in selecting and prioritizing several thousand of detected features to identify and characterize, as well as determining the biological importance of the hundreds of statistically significant metabolites. Furthermore, categorizing and determining the function of the identified features with respect to biological importance and chemical characteristics requires knowledge of biology, analytical chemistry, statistics, and bioinformatics. Therefore, to handle these issues, we have established a pipeline for non-targeted analysis of LC/GC-MS data using a

combination of commercial and opensource software packages that do not require extensive knowledge of coding expertise as they are mostly built on graphical user interface platforms.

Table 7. Summary of MS pipeline analysis.
85% spectra similarity to LC-MS library, and 75% spectra similarity to GC-MS library. 85% similarity was used for LC-MS data and 75% similarity was used for GC-MS data.

Population	Output	LC-MS-neg	LC-MS-pos	GC-MS
F2	MSDial	5,200	16,887	2,207
	-% similarity	890	629	119
	-MS1 only	1,879	6,610	-
	MetFamily	4,405	6,883	2,201
	Metaboanalyst	818	515	22
	Resistant-Susceptible	746	137	4
BC1F2	MSDial	5,780	12,506	2,201
	-% similarity	890	425	184
	-MS1 only	1,698	5,396	-
	MetFamily	3,409	4,344	1,428
	Metaboanalyst	1,070	1,054	54
	Resistant-Susceptible	552	464	11

In this study, the pipeline that we have developed for untargeted metabolomics was applied to narrow down the number of candidate features that have a potential role in plant-insect interaction. Over the period of this study, we communicated closely with Hiroshi Tsugawa to optimize the MSDial software. This software does not only automatically select features and aligns them between samples but can also predict and annotate features based on spectra similarity to MS-spectra that we have collected from commercial and non-commercial libraries (in .msp format). Extracted height and .msp files were imported to MetFamily to perform principal component analysis (PCA) and hierarchichal component analysis (HCA). PCA reduces the dimensionality in complex polydimensional data by binning individual m/z;RT features to a new space of lower dimensionality. Individual m/z;RT features (loadings) contribute to the separation of groups (scores), which is observed as characteristic pattern and trend of m/z;RT features. HCA was applied to find similarity between fragment mass spectra in the data set and cluster features based on fragment similarities. HCA was applied to LC-MS, but not to GC-MS data because there we would require an additional filtering script to remove small fragments prior to the analysis. The exported files from MetFamily were imported to Metaboanalyst to conduct ANOVA, followed by T-test for resistant and susceptible lines. Table 9 shows that we were able to reduce the number of selected features using parameter settings that were adjusted to the type and condition of the datasets, removing low abundance peaks, artifacts, and non-significantly different features between treatments. Even though most of these features were annotated as unknowns, we were able to estimate predicted structures or metabolic families via literature search for known metabolites that have a role on insect resistance as well as by correlating them with gene expressions data from the transcriptomic experiment. For example, we observed higher quantity of flavonoids (rutin and kaempferol-3,7-di-O-glucoside) in resistant lines compared to susceptible lines of

the F2 population (Fig. 49- left). Supporting that, the relative expression of flavonoid biosynthesis related genes in tomato trichomes compared to leaves were higher, which means that the biosynthesis of flavonoids was more active in trichomes than in the leaves (Fig. 49- right).
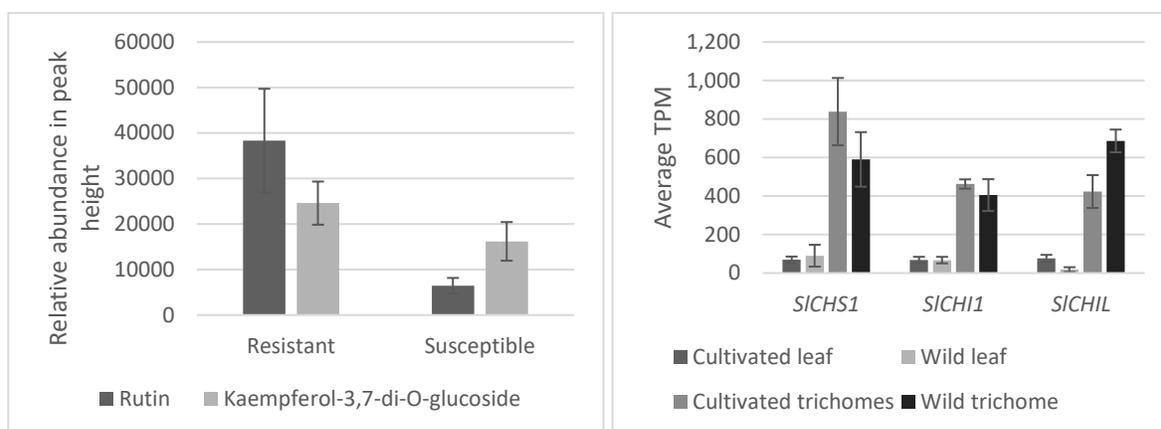
Figure 49. Relative abundance of rutin and kaempferol-3,7-di-O-glucoside (left), and expression of genes related to flavonoid biosynthesis (right) in the F2 population.
Resistant and susceptible lines were obtained from the F2 population of cultivated x wild. Cultivated: *S. lycopersicum* – AVT01424. Wild: *S. habrochaites sp. glabratum* – VI030462. TPM: Transcript per million. *SlCHS1*: Solyc05g053550. *SlCHI1*: Solyc05g010320. *SlCHIL*: Solyc05g052240. Resistant: F2 with mortality >70%. Cultivated: BC1FF22 with mortality <40%.

Resistant and susceptible lines responded differently to whitefly, as seen from the number of significantly different secondary metabolic features, including: terpenes, flavonoids, steroidal glycoalkaloids, acylsugars, oxylipins, phenolamides, alkanes, and fatty acid derivatives and lastly quinic acid and its derivatives. These metabolic features are affected by the salicylic acid (SA)/jasmonic acid (JA) ratio in the plant. In the cultivated tomato, we found that either the SA biosynthetic pathway was more active or that enzymes from whitefly saliva might have contributed to the suppression of the JA pathway (Zarate *et al.*, 2007). This was supported by data from the transcriptomics analysis that showed a cluster containing increased p*henylalanine ammonia-lyase* (Solyc09g007910) in the cultivated trichomes. This enzyme was reported to catalyze the initial step in phenylpropanoid synthesis from L-phenylalanine into other derivatives which includes SA (Lin *et. al.,* 2019). In our study, we observed higher methyl salycilate (MeSA) content in cultivated compared to wild tomato (Fig. 50). MeSA is the biosynthetic product of SA and S-adenosylmethionine (SAM), whereas SAM is required for the biosynthesis of polyamines (PAs), ethylene, transmethylation, transsulfuration, and lignin/alkanes biosynthesis. Interestingly, this was consistent with our observation that higher PAs and alkanes were present in the cultivated tomato. However, insect experiments have shown that triggering the JA pathway provides a better strategy against *Bemisia tabaci* (Escobar-Bravo *et al.,* 2016). Activation of the JA defense pathway is known to induce the production of anti-oxidative enzymes, protease inhibitors, volatile organic compounds (VOCs), alkaloid production, secretion of extra floral nectar, and GT formation (War *et al.*, 2012). This was seen by us from

higher density of GT and content of metabolic features such as acylsugars (AS) and terpenes, which are higher in wild- *S. habrochaites* sp. *glabratum* tomatoes (Fig. 11).
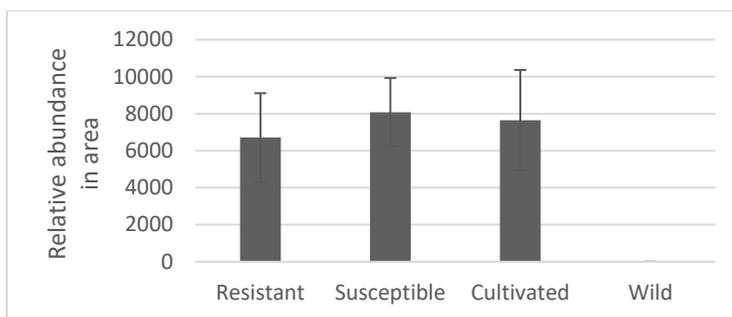
Figure 50. Relative abundance of methyl salicylate (m/z 120; RT 10.665) measured via GC-MS of the BC1F2 population. Wild: *S. habrochaites sp. glabratum*- VI030462. Cultivated: *S.lycopersicum* – AVT01424. Resistant: BC1F2 with mortality >70%. Cultivated: BC1F2 with mortality <40.

Unfortunately, our experimental setup could only compare the relative abundance of metabolic features between resistant and susceptible tomato lines that were infected with whiteflies, but could not answer which features are induced or reduced upon infection and to what degree it is toxic or beneficial to the whiteflies. Su *et al.*, (2017) compared results of *S. lycopersicum* that was infected or non-infected with whiteflies and reported that the infection did not affect the levels of quinic acid and its derivatives, steroidal glycoalkaloids and AS. However, a different response was observed by Rossouw *et al.* (2018), who reported that quinic acid and derivative levels such as chlorogenic acid and caffeoyl quinic acid in *S. habrochaites* introgressions lines infected with whiteflies were decreased significantly compared to the untreated plants. This strategy is used by the resistant plants to mount stronger defense response.

AS are known to be constitutively produced and highly abundant in *S. habrochaites*. However, they can be induced in *S. lycopersicum* by the addition of exogenous JA (Escobar-Bravo *et al.,* 2016). Although much research has focused on exploring genes regulating AS biosynthesis such as *ASATs*, *ASHs*, and *2-isopropylmalate synthases,* recently, Santegoets *et al.* (2021) identified a QTL region on chromosome 12 from *S. galapagense* that dominantly affects whitefly mortality but not oviposition rate. Interestingly, this QTL does not correlate with high AS, which means there are other factors affecting resistance, such as antioxidant levels. Antioxidants are compounds preventing damage of reactive oxygen species by scavenging them and primarily consist of phenolics, such as flavonoids. Su *et al.*, (2017) found that flavonoids (anthocyanin, rutin and other flavonol derivatives) were lower in whitefly-infected compared to non-infected leaves. Yao *et al.* (2019) showed that flavonoids deter oviposition by lengthening probing duration and phloem ingestion.

Lastly, Su *et al.*, (2017) also reported that the terpenes α-phellandrene and α-terpinene, known to be synthesized in type VI GT, affected whitefly preference, but not oviposition when added exogenously. In a subsequent study, Su *et al.* (2018) showed that essential amino acid concentrations and sugar levels

were not influenced by whitefly infection, but terpenoids emission was affected. Terpenoids, which includes α-pinene, β-myrcene, α-phellandrene, 2-carene, limonene, ρ-cymene, and α-terpinene were affected, which was supported by lower expression levels of several *TPS* genes. In our transcriptomics studies, 31 different *TPS* genes were captured and shown to be expressed differenitially between cultivated and wild trichomes. These genes are responsible for synthesizing the backbone of monoterpenes, sesquiterpenes, and diterpenes (Falara *et al.*, 2011). The *S. habrochaites* wild species has been used to functionally characterize sesquiterpene synthases and several main sesquiterpenes (7-epi-zingiberene, R-curcumene, R-α-santalene, R-endo-β-bergamotene, and R-endo-α-bergamotene) responsible for resistance against whitefly were detected (Sallaud *et al.*, 2009; Bleeker *et al.*, 2011; Li *et al.,* 2021). In a recent study, Zabel *et al.* (2021) characterized a cytochrome P450 oxygenase that is highly expressed in trichomes and co-segregates with the presence of 7-epi-zingiberene, which exhibits toxicity against *B. tabaci*. Although these reports have shown how terpenoids affects whitefly preference, we did not see signficiant correlation between insect mortality and the density of type VI trichomes in our population. Since each chemotype may have different terpenes and thus different impact on different organisms, it is challenging to determine which combination are most effective to reduce whitefly infection.

These findings provide valuable insights into the underlying biochemical mechanisms associated with whitefly resistance, while providing characterised parental materials for future breeding programmes directed towards conferring whitefly resistance on cultivated tomatoes. The challenge is to select for potential biomarkers conferring whitefly resistance and to create a blend that functions as repellent against herbivores.

4.3. Characterization of *AS acyl-coA synthetases* specific for short and medium chained fatty acids

Focusing on genes regulating acyl-CoA biosynthesis, we have confirmed the role of *Sh-AACS1* and *Sh-AACS2* in short and medium chain acylsugar (AS) biosynthesis. The characterization of these two genes was based on the transcriptomics data where the expression is higher in *S. habrochaites* sp. *glabratum-*VI030462 compared to *S. lycopersicum* – AVT01424 tomato trichomes. S*h-AACS2* is located on chromosome 2 which is not in the vicinity of known AS biosynthesis characterized genes. Compared to Sh-AACS1 and other AACS characterized by Fan *et al.* (2020), Sh-AACS2 has strong divergence in approximately fifty N-terminal amino acid, in addition to smaller amino acids changes further downstream in the protein. An *In-vitro* enzyme assay revealed that these differences are important for substrate specificity, since AACS2 showed only short-chain FA-coA products. Furthermore, kinetic experiments showed that Sh-AACS2 has higher affinity towards iC5 than aiC5. This result could be explained by the fact that iC5 was more abundant than aiC5 in the FAEE experiment. We observed only minor amino acid differences in the chromosome 7 localized Sh-AACS1 compared to Sl-AACS1, of which it is an isofrom (Fan *et al.,* 2020). These small variations in amino acid sequence led to differences in substrate specificity limiting Sh-AACS1 to C6, C8, and C10 FAs. Based on this finding, we hypothesize that there could be novel C12 FA specific *Sh-AACS* listed among the candidate genes.

Stable transformation of wild *S. habrochaites* has been challenging and difficult due to low transformation efficiency and poor regeneration of mutants. VIGS is a rapid tool that has been proven to be a robust method for silencing genes transiently in wild tomato plants (Leong *et al.,* 2020). There are limitations to this technique such as non-homogeneous silencing effect in plant tissues and possible phenotypic effects not solely due to the silencing of the targeted genes, but due to plant immune responses to the viral RNA (Senthil-Kumar & Mysore, 2011a&b). Silencing genes expressed specifically in trichomes is also more challenging, since the viral RNA needs to be transported up to the trichomes. Moreover, the effect and efficiency of VIGS can also be influenced by the age and tissue of plants, as well as different environmental condition (Broderick & Jones, 2014). In this study, VIGS experiments were conducted under a lower temperature (20°C) on two week old cotyledons for better infection efficiency. Leaf material was harvested (five weeks post infection) based on the severity of infection of the PDS positive control treated plants. Silencing *Sh-AACS2* showed decrease in AS, predominantly consisted of iC5 and aiC5 FAs. A similar observation was also made for *Sh-AACS1,* which is specific for medium chain fatty acid. Even though there was a significant difference in AS5 and AS7, which contain C10 FAs, there is a possibility that the selected 500 bp silenced region were not precise enough to target *Sh-AACS1,* specific for C6, C8, and C10. This suggests another limitation of the VIGS system is off-target silencing.

The genes that were characterized in this study are just two of many *Acyl-CoA Synthetases* in wild tomatoes, which may be of interest for evolutionary studies. Future research could focus deeper on how plants evolved and how selective pressure effects the genetic and metabolic composition of different solanum species. With respect to adaptation, we could ask the question how plants have evolved to produce a certain blend of AS required for resistance against biotic stresses. Conclusively, introducing these genes in cultivated varieties could potentially help breeders to produce resistant tomatoes.

4.4. The regiospecificity of different ASAT3 isoforms

The QTL for insect mortality and trichome density at chromosome 11 contains several genes whose expression level is higher in wild tomatoes compared to cultivated tomatoes. One well characterized gene, *AS acyltransferase* (*ASAT3*) was strikingly expressed at higher levels in wild tomato trichomes compared to other samples (Schilmiller *et al.*, 2012; 2015). *ASAT3* can be distinguished into F or P type based on a amino acid difference, both catalyzing the acylation of (acyl)-sugars, typically forming di-acylsucroses or tri-acylsucroses on position 2 or 3', respectively. In the *S. habrochaites* sp. *glabratum* – VI0301462 wild tomato, we have identified two ASAT3-F isoforms, one highly similar to the cultivated *S. lycopersicum* – AVT01424 (*ASAT3-VI030462-c*) and another more distant that we qualify as wild allele (*ASAT3-VI030462-w*). In our study, we only found AS that were acylated not on the 3' but in the 1' of the furanose ring based on the characterized AS via NMR that we purified from the *S. habrochaites* sp. *glabratum* surface extracts (AS2 &AS3). We expected that both isoforms might have similar regiospecificity. AS with acylation on the 1' position have also been reported by King *et al.* (1990) who elucidated major complexes of acylsucroses and found that acylation occured at position 2,3,4,1' in *S. habrochaites* – LA1353, previously called *Lycoperson hirsutum*, via reverse phase HPLC and NMR. LA1353 originated from Contumaza, Cajamarca, Peru and is located distantly to LA1777 that was collected from from Rio Casma in the Ancash province, a region closer to the coast of Peru. Since the exact identity of VI030462 is unknown, due to miss-identification/mixed up of seed batches in the AVRDC gene bank, we hypothesized that this variety may come from a region that is more on the central-northern part of Peru. This was inferred from the association studies between geographical distribution and phylogenetic relationships of *ASAT3* loci by Shillmiller *et al.* (2015) who showed distribution of ASAT3-F isoform that produces detectable acylsucrose with short chain chain ($\leq$ C5) fatty acid on the furanose ring predominantly in this region. However, the comparison of the sequences of different ASAT3 isoforms did not allow for the identification of amino acids responsible for the specificity of the enyzmes. Targeted mutagenesis based on 3D structure would heave helped, but 3D structure of these proteins are not available.

In this study, we have isolated both *ASAT3-VI030462* isoforms and conducted a VIGS experiment using contructs that contain fragments (~500bp) specific to each isoform. We could only provide information that showed a trend of reduced AS content compared to the control and reduced expression of the targetted genes using specific primers via qRT-PCR (Fig. 41). We also observed that there was a higher proportion of AS2 that has one fewer C5 group than AS3 on *ASAT3* silenced plants compared to the control. This could mean that there is lower activity of enzymes that transfer acyl groups to the sugar backbone.

We also conducted *in-vitro* enzyme assays using reverse enzyme reaction on purified AS as well as non-reverse enzyme reaction on synthetic AS. Synthetic AS were synthesized with the help of Yanira Mendez Gomez and Aldrin Vasco Vidal from the chemistry department of IPB. However, we were

unsuccessful in providing enough product from the synthetic AS lacking acylgroups in the furanose ring for the NMR measurement. Alternatively, we could try to measure further fragments (MS^3) and compare the fragmentation pattern of the fragments between samples and see whether we could detect differences. Even though we were unable to determine the structure, we still managed to synthesize AS with specific acyl groups attached to defined positions, which requires a lot of steps to deprotect and protect the acyl groups – as to sequencially build the molecule. Puterka *et al.* (2003) used synthetic AS with a range of sugar backbones and one acyl group attachments on a selection of arthropod pests. They found that a synthetic acylsucrose with a C8 acyl group had the highest insecticidal and miticidal activity compared to other combinations using either xylitol or sorbitol as the sugar backbone and C6, C10, or C12 as the acyl groups. However, they did not test C5 acid on a sucrose backbone.

To conclude this part of the study, we were unsuccessful in providing substantial evidence that different *ASAT3* isoforms are regiospecific in transferring acylgroups to synthesize acylsugars. However, purified AS (AS1-3) obtained from VI030462 surface extracts are currently being tested via whitefly feeding assays by AVRDC. It would be interesting to examine whether individual AS have different efficacy against whitefly infestation. Unfortunately, we did not supply any purified acyl-glucoses for this experiment since it was not detected in *S. habrochaites*, but only in *S. pennellii*. Sasaki *et al.* (2014) reported that acylglucoses have other roles in anthocyanin modifications. In the LC-MS studies, we have observed high quantities of flavonoids in resistant plants, however we did not further investigate . Additionally, scientists still have not fully uncovered the mechanism of how AS not only act as a trapping mechanism, but also entoxicate whiteflies affecting insect mortality and oviposition rate. We hypothesize that during infestation, there are enzymes that cleave fatty acids from the sugar backbone, thereby releasing sugars and fatty acids. There have been experimental reports of insecticidal activity of some reduced sugars against whiteflies (Hu *et al.*, 2010) as well as repellent action of a medium-chain free fatty acids towards houseflies *Musca domestica* (Ralston & Barrett, 1941). Thus, for future studies, we suggest not only to look at the diversity of *ASAT3* among different tomato species, but also the diversity and activity of invertases (Leong *et al.*, 2019) as well as *ASH* genes that cleave off acylgroups during infection.

4.5. Adenosyl methionine decarboxylase & s-adenosyl methionine synthetase candidate susceptibility factors to whiteflies

Susceptibility and resistance are opposite sides of the same coin. Most breeding programs have focused on resistance as breeding targets, which is commonly dominantly inherited. On the contrary, recessive resistance traits have long been studied and can be broad and durable as exemplified by the powdery *Mildew Resistance Locus O* (*Mlo*) allele (Bai *et al.*, 2008). As of the time of writing, most susceptibility factors have been investigated to confer susceptibility towards microbes/pathogens and only few are explored for herbivores (Engelhardt *et al.*, 2019; van Schie *et al.*, 2014). These genes can act as defence suppressors or genes regulating metabolites that are attractant or favourable to herbivores. Practically speaking, it is still challenging to breed for susceptibility traits, because they are often quantitative traits and whether loss of a susceptibility factor leads to full or partial recessive resistance. Furthermore, when susceptibility genes are silenced, there could be pleiotropic effects or negative feedback on the production of important metabolites.

Polyamine (PA) biosynthesis has been thoroughly investigated in relation to physiological processes. In this study, we have reported that and provided preliminary evidence for candidate genes related to PA biosynthesis acting as susceptibility factors toward whiteflies. Among the selected candidate genes, *S-Adenosylmethioinine Decarboxylase (AMD) 5* is positioned at chromosome 5 and catalyzes the conversion of S-adenosylmethionine (SAM) to decarboxylated SAM (dcSAM). *AMD5* was identified from the QTL analysis as correlating with increased ovipositioning. The second candidate gene was *S-adenosylmethionine synthetase* (*SAMS) 12* that is positioned on chromosome 12. *SAMS12* was obtained from the transcriptomics data and is known to convert L-methioinine (Met) into SAM. Both SAMS and AMD proteins are localized in the cytosol of plant cells and are required for the synthesis of PAs (spermidine, spermine and thermospermidine) from putrescine (Bale & Ealick *et al.,* 2010; Bennett *et al.,* 2002).

We have confirmed the function of SAMS12 via an *in-vitro* enzyme assay in *E. coli.* SAM12 was able to produce SAM when given Met as substrate. However, enzyme kinetic experiments still needed to be conducted for this enzyme. In the case for AMD5, we were not able to successfully provide evidence via an *in-vitro* enzyme assay. It could be that the experimental setup for *AMD5* is not adequate due to the lack of cofactors even though we have obtained relatively pure and high amounts of protein as shown in the western blot and coomassie result (Appendix 3). Therefore, we suggest complementing a yeast *AMD* (*SPE2*) knockout strain (Balasundaram *et al.*, 1991) with the cultivated tomato *AMD5* gene. Knocking out *SPE2* in yeast is lethal, therefore complementing with functional *AMD5* may restore the yeast its capability to synthesize spermidine and spermine which are essential for the aerobic growth of yeast (*S. cerevisiase*).

VIGS experiment was performed to transiently silence *AMD5* and *AMD12* genes in cultivated tomatoes. Trend of reduced PAs was observed in cultivated tomatoes for both genes. Meanwhile, *in-planta* experiment via VIGS assay showed reproducible chlorotic symptoms on *AMD5* but not on *SAMS12* silenced plants which suggests the importance of this gene in plant development. However due to the limitation of VIGS experiment and facility at the IPB, we were not able to conduct insect assays and evaluate plant resistance using whiteflies.

To answer whether PAs play a role in whitefly attraction, we propose to conduct whitefly choice feeding assays on resistant wild tomatoes in addition to an artificial diet treated with external PAs. Polyamines are a source of nitrogen, which is an important nutritient. Oota *et al*. (2020), showed that root-knot nematodes are attracted to specific polyamines which posess three to five methylene groups between two terminal amino groups. These polyamines, including cadaverine, putrescine and 1,3-diaminopropare, were detected in soybean root cortex cells. In lima bean (*Phaseolus lunatus*), the addition of external PAs, specifically spermine, increases the production of plant jasmonic acid and VOCs, thereby attracting a natural predator (*Phytoseiulus persimilis*) of mites (*Tetranychus urticae*) (Ozawa *et al.*, 2009). Moreover, induction of calcium influx and ROS production was observed. When PA are catabolized, it produces reactive oxygen species such as H_2O_2, which not only orchestrates programmed cell death and enhances resistant mechanisms, but also functions as signaling molecule that activates other signal molecules, for instance salycilic acid, jasmonic acid, ethylene and abscisic acid (Wang *et al.*, 2019). However, we hypothesize that *B. tabacci* might be able to circumvent and manipulate the tomato defence system through effector proteins secreted from their saliva. Whitefly evolved salivary effectors such as ferritin that suppresses herbivore associated molecular pattern–mediate plant defences (H_2O_2-generated oxidative signals) in tomatoes (Su *et al.*, 2019). A recent review reported that nearly half of the omics studies on aphids, a phloem-sucking herbivore like whiteflies, showed they may modify their host to their advantage (Åhman *et al.*, 2019). They suggested several pathways such as the ethylene, SA and JA pathway to be targeted, additionally components of watery saliva such as proteases, peroxidases, polyphenoloxidases and exodoreductases may facilitate host use. Similarly, whiteflies can suppress JA-mediated plant defences by inducing the SA signalling pathway and inhibit the production of JA-regulated defensive compounds (Zarate *et al.*, 2007). Therefore, we recommend constructing knockdown mutants of *AMD5* and *SAMS12* in cultivated tomato and observing phenotypic changes regarding plant development and insect response.

In the long run, researchers can explore different tomato *AMD* and *SAMS* homologs and provide marker information for breeders to fine tune PA levels in commercial tomato varieties, that could be not only resistant to whiteflies but also contain favorable quality traits for consumption by targeting specific isoforms. Thus, the use of susceptibility factors in plant breeding presents a promising alternative due to its durable and broad-spectrum characteristics.

5. References

Åhman, I. *et al.* 2019. Plant genes benefitting aphids – potential for exploitation in resistance breeding. Fron. Plant Sci. 10:1452. doi: 10.3389/fpls.2019.01452.

Al-Abdallat, A. *et al.* 2014. Overexpression of SlSHN1 gene improves drought tolerance by increasing cuticular wax accummulation in tomato. Int. J. Mol. Sci. 15(11):19499-19515. doi: 10.3390/ijms151119499.

Ali, M. *et al.* 1992. Production and characterization of Solanum amphidiploids and their resistance to bacterial wilt. Scientia Hortic. 49:181-196.

Andrade, M.C *et al*. 2017. Inheritance of type IV GT density and its association with whitefly resistance from *S. galapagense* accession LA1401. Euphytica. 213:52. doi: 10.1007/s10681-016-1792-1.

Anis, E. *et al.* 2002. α-Glucosidase inhibitory constituents from *Cuscuta reflexa*. Chem. Pharm. Bull. 50(1):112-114. doi. 10.1248/cpb.50.112.

Arends, D. *et al.* 2010. R/qtl: high-throughput multiple QTL mapping. Bioinformatics. 26(23): 2990-2992. doi: 10.1093/bioinformatics/btq565.

Bale, S. & Ealick, S. 2010. Structural biology of S-adenosyylmethionine decarboxylase. Amino acids. 38(2):451-460. doi: 10.1007/s00726-009-0404-y.

Bai, Y. *et al.* 2008. Naturally occurring broad-spectrum powdery mildew resistance in a Central American tomato accession is caused by loss of Mlo function. Mol. Plant Microbe Interact. 21(1):30-39. doi: 10.1094/MPMI-21-1-0030.

Balamsundaram, D. *et al.* 1991. Spermidine or spermine is essential for the aerobic growth of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. 88:5872-5876. doi: 10.1073/pnas.88.13.5872.

Ballester, A.R. *et al.* 2016. Identification of loci affecting accumulation of secondary metabolites in tomato fruit of a *Solanum lycopersicum* x *Solanum chmielewskii* introgression line population. Front. Plant Sci. 7:1428. doi:10.33.89/fpls.2016.01428

Balyan, S. *et al.* 2020. Characterization of novel regulators for heat stress tolerance in tomato from Indian sub-continent. Plant Biotechnol. J. 18(10):2118-2132. doi: 10.1111/pbi.13371

Bar, M. & Shtein, I. 2019. Plant trichomes and the biomechanics of defense in various stems, with Solanaceae as a model. Botany. 97:651-660. doi:10.1139/cjb-2019-0144.

Bauer, S. *et al.* 2004. Composition of the surface wax from tomatoes. Eur. Food Res. Technol. 219:223-228. doi: 10.1007/s00217-004-0943-0.

Bedinger, P.A. *et al.* 2010. Interpecific reproductive barriers in the tomato clade: opportunities to decipher mechanisms of reproductive isolation. Sex Plant. Reprod. 24(3):171-187. doi: 10.1007/s00497-010-0155-7.

Bennet, E.M. *et al.* 2002. Monomeric S-adenosylmethionine decarboxylase from plants provides an alternative to putrescine stimulation. Biochem. 41:14509-14517. doi: 10.1021/bi026710u

Bennewitz, S. *et al.* 2018. QTL mappin of the shape of type VI GT in tomato. Front Plant Sci. 9(1421). doi: 10.3389/fpls.2018.01421.

Ben-Israel, I. *et al.* 2009. Multiple biochemical and morphological factors underlie the production of methylketones in tomato trichomes. Plant Physiol. 151(4):1952-1964. doi: 10.1104/pp.109.146415

Ben-Mahmoud, S. *et al.* 2019. A thrips vector of tomato spotted wilt virus responds to tomato AS chemical diversity with reduced oviposition and virus inoculation. Sci. Reports 9(17157). doi: 10.1038/s41598-019-53473-y.

Bergau, N. *et al.* 2015. The development of type VI GT in the cultivated tomato *Solanum lycopersicum* and a related wild species *S. habrochaites*. BMC Plant Biol. 15(289). doi:10.1186/s12870-015-0678-z.

Bergougnoux, V. 2014. The history of tomato: From domestication to biopharming. Biotech. Adv. 32:170-189. doi: 10.1016/j.biotechadv.2013.11.003.

Binder, S. 2010. Branched-chain amino acid metabolism in *Arabidopsis thaliana*. Arabidopsis Book. doi: 10.1199/tab.0137.

Blauth, S.L. *et al.* 1998. Identification of quantitative trait loci associated with AS accumulation using intraspecific populations of the wild tomato *Lycopersicon pennelii*. Theor. Appl. Genet. 96:458-467. doi: 10.1007/s001220050762.

Bleeker, P.B. *et al.* 2009. The role of specific tomato volatiles in tomato-whitefly interaction. Plant Physiol. 151(2):925-935. doi: 10.1104/pp.109.142661.

Bleeker, P.M. *et al.* 2011. RNA-seq discovery, functional characterization, and comparison of sesquiterpene synthases from *Solanum lycopresicum* and *Solanum habrochaites* trichomes. Plant Mol. Biol. 77(4-5):323-36. doi: 10.1007/s11103-011-9813-x

Bloomer, R.H. *et al.* 2014. The genetic architecture of constitutive and induced trichome density in two new recombinant inbred line populations of Arabidopsis thaliana: phenotypic plasticity, epistasis, and bidirectional leaf damage response. BMC Plant Biol. 14(119). doi: 10.1186/1471-2229-14-119.

Bolger, A. *et al.* 2014. The genome of the stress-tolerant wild tomato species *Solanum pennellii*. Nature Genetics. 46: 1034–1038. doi:10.1038/ng.3046.

Bouwmeester, H. *et al.* 2019. The role of volatiles in plant communication. Plant J. 100:892-907. doi: 10.1111/tpj.14496.

Broderick, S.R. & Jones, M.L. 2014. An optimized protocol to increase Virus-Induced Gene Silencing efficiency and minimize viral symptoms in Petunia. Plant Mol. Biol. Report. 32(1):219-233. doi: 10.1007/s11105-013-0647-3.

Broman, K.W. 1997. Identifying quantitative trait loci in experimental crosses. PhD Dissertation. University of California. Berkeley.

Broman, K.W. *et al.* 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889-890. doi: 10.1093/bioinformatics/btg112

Bombarely, A. *et al.* The Sol Genomics Network (solgenomics.net): growing tomatoes using Perl. Nucleic Acids Res. 39: D1149–D1155. doi: 10.1093/nar/gkq86B6.

Campanale *et al.* 2020. A relevant screening of organic contaminants present on freshwater and pre-production microplastics. Toxics. 8(4):100. doi: 10.3390/toxics8040100.

Cassin-Ross, G. & Hu, J. 2014. Systematic phenotypic screen of arabidopsis peroxisomal mutants identifies proteins involved in the β-oxidation. Plant Physiol. 166:1546-1559. doi: 10.1104/pp.114.250183.

Chang, A.X. *et al.*, 2020. The trichome-specific acetoalctate synthatse NtALS1 gene, is involved in acylsugar biosynthesis in tobacco (*Nicotiana tabacum L.*) Planta. 252:13. doi: 10.1007/s00425-020-03418-x.

Churchill, G.A. & Doerge, R.W. 1994. Empirical threshold values for quantitative trait mapping. Genetics. 138(3):963-971.

Chaudhary, J. *et al.* 2019. Mutation breeding in tomato: Advances, applicability, and challenges. Plants. 8(128). doi:10.3390/plants8050128.

Chen, W. *et al.* 2016. The draft genome of whitefly *Bemisia tabaci* MEAM1, a global crop pest, provides novel insights into virus transmission, host adaptation, and insecticide resistance. BMC Biol. 14:110. doi: 10.1186/s12915-016-0321-y.

Conboy, N.J.A. *et al.* 2020. Volatile organic compounds as insect repellents and plant elicitors: an integrated pest management (IPM) strategy for glasshouse whitefly (*Trialeurodes vaporariorum*). J. Chem. Ecol. 46:1090-1104. doi:10.1007/s10886-020-01229-8.

Covey, P.A. *et al.* 2010. Multiple features that distinguish unilateral incongruity and self-incompatibility in the tomato clade. The Plant J. 64(3). doi: 10.1111/j.1365-313X.2010.04340.x.

Cruz-Estrada, A. *et al.* 2017. Medium-chained fatty acids from *Euginia winzelingii* leaves causing insect settling deterrent, nematicidal and phytotoxic effects. Molecules. 24(9):1724. doi: 10.3390/molecules24091724.

Dalin, P., Ågren, J., Björkman, C., Huttunen, P., Kärkkäinen, K. 2008. Leaf trichome formation and plant resistance to herbivory. Ed. Schaller, A. Induced plant resistance to herbivory. Pg. 89-105.

D'auria, J.C. 2006. Acyltransferases in plants: a good time to be BAHD. Curr. Op. Plant. Biol. 9(3):331-340. doi: 10.1016/j.pbi.2006.03.016.

Darshanee, H.L.C. *et al.* 2017. Volatile-mediated attraction of greenhouse whitefly *Trialeurodes vaporarium* to tomato and eggplant. Front. Plant Sci. 8:1285. doi: 10.3389/fpls.2017.01285.

Deboever, E. *et al.* 2019. Plant-pathogen interactions: underestimated roles of phyto-oxylipins. Trend Plant Sci. 25(1):22-34. doi: 10.1016/j.tplants.2019.09.009.

Dempewolf, H. *et al.* 2017. Past and future use of wild relatives in crop breeding. Crop Scie. 57(3). doi: 10.2135/cropsci2016.10.0885.

Ding, T.B. *et al.* 2019. Transcriptome profiling of the whitefly *Bemisia tabaci* MED in response to single infection of *Tomato yellow leaf curl virus, Tomato chlorosis virus*, and their co-infetion. 2019. Front. Physiol. doi: 10.3389/fphys.2019.00302.

Du, w. *et al.* 2016. A primary screening and applying of plant volatiles as repellents to control whitefly *Bemisia tabaci* (Gennadius) on tomato. S. Rep. 6:22140. doi: 10.1038/srep22140.

Engelhardt, S. *et al.* 2019. Good riddance? Breaking disease susceptibility in the era of new breeding technologies. Agronomy. 8:114. doi: 10.3390/agronomy8070114

Escobar-Bravo, R. *et al.* 2016. A jasmonate-inducible defense trait transferred from wild into cultivated tomato establishes increased whitefly resistance and reduced viral disease incidence. Front. Plant Sci. 7:1732. doi:10.3389/fpls.2016.0173

Escobar-Bravo, R. *et al.* 2017. Induction of jasmonic acid-associated defenses by thrips alters host suitability for conspecifics and correlates with increased trichome densities in tomato. Plant Cell Physiol. 58(3):622-634. doi: 10.1093/pcp/pcx014.

Escobar-Bravo, R. *et al.* 2018. Light intensity-mediated induction of trichome-associated allelochemicals increases resistance against thrips in tomato. Plant Cell Phsiol. 59(12): 2462–247. doi:10.1093/pcp/pcy166.

Falara, V. *et al.* 2011. The tomato terpene synthase gene family. Plant Physiol. 157:770-789. doi:10.1104/pp.111.179648.

Fan, P. *et al.* 2015. In vitro reconstruction and analysis of evolutionary variation of the tomato acylsucrose metabolic network. PNAS. doi: 10.1073/pnas.1517930113.

Fan, P. *et al.* 2019. Tip of the trichome: evolution of acylsugar metabolic diversity in Solanaceae. Curr. Opin. Plant. Biol. 49:8-16. doi: 10.1016/j.pbi.2019.03.005

Fan, P. *et al.* 2020. Evolution of a plant gene cluster in Solanaceae and emergence of metabolic diversity. eLife 9. doi: 10.7554/eLife.56717.

Firdaus, S. *et al.* 2011. Identification of silverleaf whitefly resistance in pepper. Plant Breeding. 130(6):708-714. doi:10.1111/j.1439-0523.2011.01894.x

Firdaus, S. *et al.* 2012. Resistance to *Bemicia tabaci* in tomato wild relatives. Euphytica. 187:31-45. doi: 10.1007/s00122-013-2067-z.

Foolad, M.R. & Panthee, D.R. 2012. Marker-assisted selection in tomato breeding. Crit. Rev. Plant Sci. 31:93-123. doi: 10.1080/07352689.2011.616057.

Frelichowski, J.E. & Juvik, J.J.A. 2001. Sesquiterpene carboxylic acids from a wild tomato species affect larval feeding behavior and survival of *Helicoverpa zea* and *Spodoptera exigua* (Lepidoptera: Noctuidae). J. Econom. Entomol. 94(5):1249-1259. doi: 10.1603/0022-0493-94.5.1249.

Gantner, J. *et al.* 2018. Peripheral infrastructure vectors and an extended set of plant parts for the modular cloning system. PLoS ONE. doi: 13(5): e0197185.

Ge, C. *et al.* 2006. BUD2, encoding an S-adenosylmethionine decarboxylase, is required for Arabidopsis growth and development. Cell Res. 16(5):446-456. doi: 10.1038/sj.cr.7310056.

Gerling, D. *et al.* 2001. Biological control of *Bemisia tabaci*. 20 (9):779-799. doi: 10.1016/S0261-2194(01)00111-9.

Ghosh, B. *et al.* 2014. Comparative structural profiling of trichome specialized metabolites in tomato (Solanumlycopersicum) and *S. habrochaites*: AS profiles revealed by UHPLC/MS and NMR. Metab. 10(3):496-507. doi: 10.1007/s11306-013-0585-y.

Glas, J.J. *et al.* 2012. Plant GT as targets for breeding or engineering of resistance to herbivore. Int. J. Mol. Sci. 13:17077-17103. doi: 10.3390/ijms131217077.

Gopalakrishnan, R. &Selvanarayanan, R. 2009. Preliminary evaluation of mutant tomato accessions for resistance against whitefly *Bemisia tabaci* Gennadius (Aleyrodidae: Hemiptera). Advan. Biol. Res. 3(5-6):159-161.

Goto, D.B. *et al.* 2002. A single-nucleotide mutation in a gene encoding S-adenosylmethionine synthetase is associated with methionine over-accumulation phenotype in *Arabidopsis thaliana*. Genes. Genet. Syst. 77:89-95. doi: 10.1266/ggs.77.89

Haley, C.S. & Knott, S.A. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity. 69:315-324. doi: 10.1038/hdy.1992.131.

Hanson, P. *et al.* 2014. Characterization and mapping of a QTL derived from Solanum habrochaites associated with elevated rutin content (quercetin-3-rutinoside) in tomato. Euphytica. 200:441-454. doi: 10.1007/s10681-014-1180-7.

Heu, C.C. *et al.* 2020. CRISPR-Cas9-based genome editing in the silverleaf whitefly (*Bemisia tabaci*). The CRISPR J. 3(2):89-96. doi: 10.1089/crispr.2019.0067.

Holcomb, E.R. & Shapiro, S.K. 1975. Assay and regulation of S-adenosylmethionine synthetase in *Saccharomyces cerevisaiae* and *Candida utilis*. J. Bacteriol. 121(1):267-271. doi: 10.1128/JB.121.1.267-271.1975.

Hu, J.S. *et al.* 2010. Insecticidal activity of some reducing sugars against the sweet potato whitefly, *Bemisia tabaci*, biotype B. J. Insect Sci. 10:203. doi: 10.1673/031.010.20301.

Huang, H.J. *et al.* 2020. Identification of salivary proteins in the whitefly *Bemisia tabaci* by transcriptomic and LC–MS/MS analyses. Insect Sci. doi: 10.1111/1744-7917.12856

Iijima, Y. *et al.* 2013. Steroidal glycoalkaloid profiling and structures of glycoalkaloids in wild tomato fruit. Phytochem. 95:145-157. doi: 10.1016/j.phytochem.2013.07.016.

Isaacson, T. *et al.* 2002. Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of β-carotene and xanthophylls in Plants. Plant Cell. 14:133-342. doi: 10.1105/tpc.010303.

Itkin, M. *et al.* 2013. Biosynthesis of antinutritional alkaloids in Solanaceous crops is mediated by clustered genes. Science. 341(6142):175-179. doi: 10.1126/science.1240230

Jiang, Z.F. *et al.* 2012. Genome sequences of the primary endosymbiont "*Candatus portiere aleyrodidarum*" in the whitefly *Bemisia tabaci* B and Q biotypes. J. Bacteriol. 194: 6678-6679. doi: 10.1128/JB.01841-12.

Joehanes R. & Nelson J.C. 2008. QGene 4.0, an extensible Java QTL-analysis platform. Bioinformatics 24: 2788-2789.

Jones, D.R. 2003. Plant viruses transmitted by whiteflies. Europ. J. Plant. Pathol. 109:195-219.

Jorge, T.F. *et al.* 2016. Mass spectrometry as a quantitative tool in plant metabolomics. Phil. Trans. R. Soc. A. 374:20153070. doi: 10.1098/rsta.2015.0370.

Kang, J.H. *et al.* 2010. The tomato *odorless-2* mutant is defective in trichome-based production of diverse specialized metabolites and broad-spectrum resistance to insect herbivores. Plant Physiol. 154:262-272. doi: 10.1104/pp.110.160192.

Kang, J.H. *et al.* 2014. The flavonoid biosynthetic enzyme chalcone isomerase modulates terpenoid production in glandular trichome of tomato. Plant Physiol. 164:1161–1174. doi: 10.1104/pp.113.233395.

Kant, M.R. *et al.* 2001. Differential timing of spider mite-induced direct and indirect defenses in tomato plants. Plant Physiol. 135:483-495. doi: 10.1104/pp.103.038315.

Kao, C.H. & Zeng, Z.B. 1999. Multiple interval mapping for quantitative trait loci. Genetics 152:1203-1216.

Karabourniotis, G. *et al.* 2020. Protective and defensive roles of non-glandular trichome against multiple stresses: structure-function coordination. J. For. Res. 31:1-12. doi: 10.1007s11676-019-01034-4.

Kariyat, R.R. *et al.* 2019. Feeding on glandular and non-glandular leaf trichomes negatively affect growth and development in tobacco hornworm (*Manduca sexta*) caterpillars. Athropod-Plant Interaction. 12:321-333. doi: 10.1007/s11829-019-09678-z.

Kennedy, G.G. 2003. Tomato, pests, parasitorids, and predators, tritrophic interactions involving the genus *Lycopersicon*. Annu. Rev. Entomol. 48:51-72. doi: 10.1146/annurev.ento.48.091801.112733.

Kim, J. *et al.* 2012. Striking natural diversity in glandular trichome AS composition is shaped by variation at the acyltransferase2 locus in the wild tomato, *S. habrochaites*. 2012. Plant Physiol. 160:1854-1870. doi: 10.1104/pp.112.204735.

King, R.R. *et al.* 1990. Sucrose esters associated with the glandular trichome of wild *Lycopersicon* species. Phytochem. 29(7):2115-2118. doi: 10.1016/0031-9422(90)83017-U.

Kliebenstein, D.J. *et al.* 2007. Characterization of seed-specific benzoyloxyglucosinolate mutations in *Arabidopsis thaliana*. The Plant J. 51:1062-1076. doi: 10.1111/j.1365-313X.2007.03205.x.

Koek, M.M. *et al.* 2006. Microbial metabolomics with gas chromatography/mass spectrometry. Anal. Chem. 78:1272-1281. doi: 10.1021/ac051683

Kroumova, A.B. *et al.* 1994. A pathway for the biosynthesis of straight and branched, odd- and even-length, medium-chain fatty acids in plants. Proc. Natl. Acad. Sci. USA. 91:11437–11441.

Kundu, A. & Vadassery, J. 2018. Chlorogenic acid-mediated chemical defence of plants against insect herbivores. Plant Biol. 21:185-189. doi: 10.1111/plb.12947.

Laterre, R. *et al*. 2017. Photsynthetic trichomes contain a specific rubisco with a modified pH-dependent activity. Plant Physiol. 173:2110-2120. doi: 10.1104/pp.17.00062.

Leong, L. *et al.* 2019. Evolution of metabolic novelty: A trichome-expressed invertase creates specialized metabolic diversity in wild tomato. Science Advances. 5(4). doi: 10.1126/sciadv.aaw3754.

Leong, B.J. *et al.* 2020. Specialized metabolism in a nonmodel nightshade: trichome acylinositol biosynthesis. Plant Physiol. 183:915-924. doi: doi.org/10.1104/pp.20.00276.

Li, Y. *et al.* 2014. Identification of plant chemicals attracting and repelling whiteflies. Arthropod Plant Interact. 8:183-190. doi: 10.1007/s11829-014-9302-7

Li, F. *et al.* 2021. Response of whitefly to the wild tomato *Solanum habrochaites*. BioRxiv. doi: 10.1101/2021.02.26.432993

Li, X.W. 2021. Intercropping rosemary (*Rosmarinus officinalis*) with sweet pepper (*Capsicum annuum*) reduces major pest population densities without impacting natural enemy populations. Insects. 12:74. doi:10.3390/insects12010074.

Lim, G.H. *et al.* 2017. Fatty acid- and lipid-mediated signaling in plant defense. Annu. Rev. Phyto. 55:505-536. doi: 10.1146/annurev-phyto-080516-035406.

Livingston, S.J. 2020. Cannabis glandular trichomes alter morphology and metabolite content during flower maturation. Plant J. 101:37-56. doi: 0.1111/tpj.14516.

Lucatti, A.F. *et al.* 2013. Diffferences in insect resistance between tomato species endemic to the Galapagos Islands. BMC Evol. Biol. 13(175). doi: 10.1186/1471-2148-13-175.

Lucini, T. *et al.* 2015. AS and the role of trichomes in tomato genotypes resistance to *Tetranychus urticae*. Arthrop-Plant Interact. 9:45-53. doi: 10.1007/s11829-014-9347-7.

Luu, V.T. *et al.* 2017. O-AS protect a wild tobacco from both native fungal pathogens and a specialist herbivore. Plant Physiol. doi:10.1104/pp.16.01904.

Ma, X. *et al.* 2005. Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. J. Biol. Chem. 280(14):13576-13583. doi: 10.1074/jbc.M414508200.

Macel, M. *et al.* 2020. High concentrations of very long chain leaf wax, alkanes of thrips susceptible pepper accessions (*Capsicum spp*). J. Chem. Ecol. 46:1082-1089. doi: 10.1007/s10886-020-01226-x.

Malierpaard, C. *et al.* 1995. Mapping of QTLs for glandular trichome densities and *Trialeurodes vaporum* (greenhouse whitefly resistance in an F2 from *Lycopersicon esculentum* x *Lycopersicon hirsutum f. glabratum*. Heredity, 75:425-433.

Majumdar, R. *et al.* 2017. Polyamines in the life of Arabidopsis: profiling the expression of S-adenosylmethionine decarboxylase (*SAMDC*) gene family during its life cycle. BMC Plant Biol. 17(264). doi: 10.1186/s12870-017-1208-y.

Mandal, S. *et al.* 2020. Candidate gene networks for acylsugar metabolism and plant defense in wild tomato *Solanum pennellii*. The Plant Cell. 32:81-99. doi: 10.1105/tpc.19.00552.

Marco, F. *et al.* 2014. Overexpression of *SAMDC1* gene in Arabidopsis thaliana increases expression of defense-related genes as well as resistance to *Pseudomonas syringae* and *Hyaloperonospora arabidopsidis*. Plant Sci.

Marti , G. *et al.* 2013. Metabolomics reveals herbivore-induced metabolites of resistance and susceptibility in maize leaves and roots. Plant Cell Environ. 36(3):621–639. doi: doi.org/10.1111/pce.12002.

Martin, J.H. *et al.* 2000. The whiteflies (Hemiptera: Aleyrodidae) of Europe and the Mediterranean Basin. Bul. Entomol. Res. 90:407-448. doi: 10.1017/S0007485300000547.

Masike, K. *et al.* 2017. Highlighting mass spectrometric fragmentation differences and similarities between hydroxycinnamoyl-quinic acids and hydroxycinnamoyl-isocitric acids. Chem. Cent. J. 11:29. doi: 10.1186/s13065-017-0262-8.

Matsukura, C. *et al.* 2007. Generation of gamma irradiation-induced mutant lines of the miniature tomato (Solanum lycopersicum L.) cultivar 'Micro-Tom'. Plant Biotech. 24:39-44. doi: 10.5511/plantbiotechnology.24.39.

McDaniel, T. *et al.* 2016. Novel resistance mechanisms of a wild tomato against the glasshouse whitefly. Agron. Sustain. Dev. 36:14. doi:10.1007/s13593-016-0351-4.

McDowell, E.T. *et al.* 2011. Comparative functional genomic analysis of *Solanum* glandular trichome types. Plant Physiol. 155:524-539. doi: 10.1104/pp.110.167114.

Mehta, A.P. *et al.* 2015. Radical SAM enzymes in cofactor biosynthesis: A treasure trove of complex organic radical rearrangement reactions. J. Biol. Chem. 290(7):3980-3986. doi: 10.1074/jbc.R114.623793.

Menda, N. *et al.* 2004. *In silico* screening of saturated mutation library of tomato. The Plant J. 38:861-672. doi: 10.1111/j.1365-313x.2004.02088.x.

Merico, D. *et al.* 2011. Visualizing gene-set enrichment results using the Cytoscape plug-in enrichment map. Methods Mol Biol. 781:257-77. doi: 10.1007/978-1-61779-276-2_12.

Milligan, S.B. *et al.* 1998. The root knot nematode resistance gene *Mi* from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. The Plant Cell. 10:1307-1319. doi: 10.1105/tpc.10.8.1307.

Moghe, G.D. *et al.*2017. Evolutionary routes to biochemical innovation revealed by integrative analysis of plant-defense related specialized metabolic pathway. eLife. 6:e28468. doi: 10.7554/eLife.28468.

Momotaz, A. *et al.* 2010. Identification of quantitative trait loci conferring resistance to *Bemisia tabaci* in an F2 Population of *Solanum lycopersicum x Solanum habrochaites* accession LA1777. J. Amer. Soc. Sci. 135(2):134-142. doi: 10.21273/JASHS.135.2.134.

Nadakuduti, S.S., Uebler, J.B., Liu, X., Jones, A.D., Barry, C.S. 2017. Characterization of trichome-expressed BAHD acyltransferases in *Petunia axillaris* reveals distinct acylsugar assembly mechanisms within the Solanaceae. Plant Physiol. 175(1):36-50. doi: 10.1104/pp.17.00538.

Neta, N.S. *et al.* 2015. Sugar ester surfactants: Enzymatic synthesis and applications in food industry. Crit. Rev. Food. Sci. Nutri. 55:595-610. doi:10.1080/10408398.2012.667461.

Neupane, F.P. & Norris, D.M. 1991. Alpha-tocopherol alteration of soybean antiherbivory to Trichoplusia ni larvae. J. Chem. Ecol. 17:1941-1951.

Ning, J. *et al.* 2015. A feedback-insensitive isopropylmalate synthase affects acylsugar composition in cultivated and wild tomato. Plant Physiol. 169:1821-1835. doi: 10.1104/pp.15.00474.

Nombela, G. *et al.* 2003. The root-knot nematode resistance gene Mi1.2 of tomato is responsible for resistance against the whitefly *Bemisia tabaci*. MPMI. 16(7):645-649. doi: 10.1094/MPMI.2003.16.7.645.

O'Neill, S.D. *et al.* 1990. Molecular genetic analysis of chalcone synthase in *Lycopersicon esculentum* and an anthocyanin-deficient mutant. Molecular & General Genetics. 224: 279–288. doi: 10.1007/BF00271562.

Onkokesung, N. *et al.* 2012. MYB8 controls inducible phenolamide levels by activating three novel hydroxycinnamoyl-coenzyme A: polyamine transferases in *Nicotiana attenuata*. Plant Physiol. 158: 389–407. doi: 10.1104/pp.111.187229.

Oota, M. *et al*. 2021. Identification of naturally occurring polyamines as root-knot nematode attractants. Mol. Plant. 13(4):658-665. doi: 10.1016/j.molp.2019.12.010.

Oriani, M.A.G. & Vendramim, J.D. 2010. Influence of trichomes on attractiveness and ovipositional preference of *Bemisia tabaci* (Genn.) B biotype (Hemiptera: Aleyrodidae) on tomato genotypes. Neotropi. Entomol. 39(6):1002-1007. doi: s1519-566x2010000600024.

Ozawa, R. *et al.* 2009. Exogenous polyamines elicit herbivore-induced volatiles in lima bean leaves: involvement of calcium, H_2O_2 and jasmonic acid. 50(12):2183-2199. doi: 10.1093/pcp/pcp153.

Pan, H.P. *et al.* 2013. Relative amount of symbionts in insect host changes with host-plant adaptation and insect hosts changes with host-plant adaptation and insecticide resistance. Environ. Entomol. 42:74-78. doi: 10.1603/EN12114.

Pang, X. *et al.* 2018. The tryptophan decarboxylase in *Solanum lycopersicum*. Molecules. 23(5):22-25. doi: 10.3390/molecules23050998.

Park, S.H., et al. 2019. Contrasting roles of cannabidiol as an insecticide and rescuing agent for ethanol-induced death in the tobacco hornworm *Manduca sexta*. Sci. Rep. 9:10481. doi:10.1038/ s41598-019-47017-7.

Pearce, G. *et al.* 1998. Phytochem. 47(4):659-664. doi: 10.1016/S0031-9422(97)00620-1.

Peleman, J. *et al.* 1989. Structure and expression analyses of the S-adenosylmethionine synthetase gene family in *Arabidopsis thaliana*. Gene. 84:359-369. doi: 10.1016/0378-1119(89)90510-6.

Pelgrom, K.T.B. *et al.* 2014. Host plant resistance towards the cabbage whitefly in *Brassica oleracea* and its wild relatives. Euphytica. 202:297-306. doi:10.1007/s10681-014-1306-y.

Perez-Fons, L. *et al.* 2019. A metabolomics characterization of natural variation in the resistance of cassava to whitefly. BMC Plant Biol. 19:518. doi: 10.1186/s12870-019-2107-1.

Poland, J.A. & Rife, T.W. 2012. Genotyping-by-sequencing for plant breeding and genetics. Plant Genome. 5:92-102. doi: 10.3835/plantgenome2012.05.0005.

Rakha, M. *et al.* 2017. Screening recently identified whitefly/spider mite-resistant wild tomato accessions for resistance to *Tuta absoluta*. Plant Breeding. 136(4). doi: 10.1111/pbr.12503.

Ralston, A.W. & Barrett, J.P. 1941. Insect repellent activity of fatty acid derivatives. J. Americ. Oil Chem. Society. 18(4):89-91. doi: 10.1007/BF02543621.

Rodríguez-López, M.J. *et al.* 2011. Whitefly resistance traits derived from the wild tomato *Solanum pimpinellifolium* affect the preference and feeding behavior of *Bemisia tabaci* and reduce the spread of tomato yellow leaf curl virus. Americ. Phtopathol. Soc. 101(10):1191-1201. doi: 10.1094/ PHYTO-01-11-0028

Rodríguez-López, M.J. *et al.* 2020. An acylsucrose-producing tomato line derived from the wild species *Solanum pimpinellifolium* decreases fitness of the whitefly *Trialeurodes vaporariorum*. Insects. 11:616. doi: 10.3390/insects11090616

Rosell, R.C. *et al.* 1995. Ultrastructure of the mouthparts of adult sweetpotato whitefly, *Bemisia tabaci* Gennadius (Homoptera: Aleyrodidae). Int. J. Insect Morphol. & Embryol. 24(3):297-306. doi: 10.1016/0020-7322(94)00026-M

Rossouw, L.T. *et al.* 2019. Deciphering the resistance mechanism of tomato plants against whitefly-mediated tomato curly stunt virus infection through ultra-high-performance liquid chromatography coupled to mass spectrometry (UHPLC-MS)-based metabolomics approaches. Metabo. 19:9. doi:10.3390/metabo9040060.

Roumani, M. *et al.* 2020. Phenolamides in plants: an update on their function, regulation, and origin of their biosynthetic enzymes. J. Exp. Bot. eraa582. doi:10.1093/jxb/eraa582.

Salinas, M. *et al.* 2013. Genetic mapping of two QTL from the wild tomato *Solanum pimpinellifolium* L. controlling resistance against two-spotted spider mite (*Tetranychus urticae* Koch). Theor. Appl. Genet. 126(1):83-92. doi: 10.1007/s00122-012-1961-0.

Sallaud, C. *et al.* 2009. A novel pathway for sesquiterpene biosynthesis from Z,Z-farnesyl pyrophosphate in the wild tomato *Solanum habrochaites*. Plant Cell. 21(1):301-317. doi: 10.1105/tpc.107.057885.

Sánchez-Peña, P. *et al.* 2006. Sources of resistance to whitefly (*Bemisia spp.*) in wild populations of *Solanum lycopersicum* var. *cerasiforme* (Dunal) spooner G.J. Anderson et R.K. Jansen in Northwestern Mexico. Gen. Resource Crop Evol. 53:711-719. doi: 10.1007/s10722-004-3943-9.

Santegoets, J. *et al.* 2021. A novel non-trichome based whitefly resistance QTL in *Solanum galapagense*. Euphytica. 217(43). doi: 10.1007/s10681-021-02770-7.

Sasaki, N. *et al.* 2014. The role of acyl-glucose in anthocyanin modifications. Molecules. 19(11):18747-18766. doi: 10.3390/molecules191118747.

Sato, S., *et al.* 2012. The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485: 635–641. doi:10.1038/nature11119

Sauter, M. *et al.* 2013. Methionine salvage and S-adenosylmethionine: essential links between sulfur, ethylene and polyamine biosynthesis. Biochem. J. 451:145-154. doi: 10.1042/BJ20121744.

Sekula, B. *et al.* 2020. S-adenosylmethionine synthases in plants: Structural characterization of type I and II isoenzymes from Arabidopsis thaliana and Medicago truncatula. Int. J. Biol. Macromol. 151:554-565. doi: 10.1016/j.ijbiomac.2020.02.100.

Senthil-Kumar, M. & Mysore, K.S. 2011a. Virus-induced gene silencing can persist for more than 2 years and also be transmitted to progeny seedlings in *Nicotiana benthamiana* and tomato. Plant Biotech J. 9(7). doi: 10.1111/j.1467-7652.2011.00589.x.

Senthil-Kumar, M. & Mysore, K.S. 2011b. New dimensions fo VIGS in plant functional genomics. Trends in Plant Sci. 16(12). doi: 10.1016/j.tplants.2011.08.006

Senthil-Kumar, M. & Mysore, K.S. 2014. Tobacco rattle virus-based virus-induced gene silencing in *Nicotiana benthamiana*. Nat. Protocol. doi: 10.1038/nprot.2014.092.

Schillmiller, A.L. *et al.* 2009. Monoterpenes in the glandular trichome of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. PNAS. 106(6): 10865-10870. doi:10.1073/pnas.0904113106.

Schilmiller, A. *et al.* 2010. Mass spectrometry screening reveals widespread diversity in trichome specialized metabolites of tomato chromosomal substitution lines. Plant J. 62:391-403. doi: 10.1111/j.1365-313X.2010.04154.x.

Schilmiller, A.L. *et al.* 2012. Indentification of a BHD acetyltransferase that produces protective acylsugar in tomato trichomes. PNAS. 109(40):16377-16382. doi: 10.1073/pnas.1207906109.

Schilmiller, A.L. *et al.* 2015. Functionally divergent alleles and duplicated loci encoding an acyltransferase contribute to acylsugar metabolite diversity in *Solanum* trichomes. The Plant Cell. 27:1002-1017. doi: 10.1105/tpc.15.00087.

Schillmiller, A.L. *et al.* 2016. Acylsugar acylhydrolases: carboxylesterase-catalyzed hydrolysis of acylsugar in tomato trichomes. Plant Physiol. 170(3):1331-1344. doi: 10.1104/pp.15.01348.

Schmittgen, T.D. & Livak, K.J. 2008. Analyzing real-time PCR data by the comparative C_T method. Nat. Protoc. 3: 1101–1108.

Schubert, R. *et al.* 2019. Tomato MYB21 acts in ovules to mediate jasmonate-regulated fertility. Plant Cell. doi: 10.1105/tpc.18.00978.

Schuurink, R. & Tissier, A. 2019. GT: micro-organs with model status? New Phyto. 225(6). doi: 10.1111/nph.16283.

Schwahn, K. *et al.* 2014. Metabolomics-assisted refinement of the pathways of steroidal glycoalkaloid biosynthesis in the tomato clade. J. Integ. Plant Biol. 56(9):864-875. doi: 10.1111/jipb.12274.

Sifres, A. *et al.* 2011. Pattern of genetic variability of Solanum habrochaites in its natural area of distribution. Genet. Resour. Crop. Evol. 58:347-360. doi: 10.1007/s10722-010-9578-0.

Silvia, K.F.A.S. *et al.* 2014. Resistance to *Bemisia tabaci* biotype B of *Solanum pimpinellifolium* is associated with higher densities of type IV glandular trichomes and acylsugar accumulation. Entom. Exp. Et. Aplicata. 151(3):218-230. doi: 10.1111/eea.12189.

Smeda, J.R. *et al.* 2016. Introgression of acylsuar chemistry QTl modifies the composition and structure of acylsugar produced by high-accumulating tomato lines. Mol. Breeding. 36:160. Doi: 10.1007/s11032-016-0584-6.

Sohani, M.M. *et al.* 2008. Phylogenetic and transcriptional analysis of a strictosidine synthase-like gene family in *Arabidopsis thaliana* reveals involvement in plant defence responses. Plant Biol. 11:105-117. doi:10.1111/j.1438-8677.2008.00139.x.

Srivastava, S. *et al.* 2017. Aldehyde oxidase 4 plays a critical role in delaying silique senescence by catalyzing aldehyde detoxification. Plant Physiol. 173(4):1977-1997. doi: 10.1104/pp.16.01939.

St-Pierre, B. & De Luca, V. 2000. Evolution of acyltransferase genes: Origin and diversification of the BAHD superfamily of acyltransferases involved in secondary metabolism. Rec. Adv. Phytochem. 34:285-315. doi: 10.1016/S0079-9920(00)80010-6.

Su, Q. *et al.* 2017. Whitefly aggregation on tomato is mediated by feeding-induced changes in plant metabolites that influence the behaviour and performance of conspecifics. Funct. Ecol. 32:1180-1193. doi: 10.1111/1365-2435.1305

Su, Q. *et al.* 2019. A salivary ferritin in the whitefly suppresses plant defenses and facilitates host exploitation. J. Exp. Bot. 70(12):3343-3355. doi: 10.1093/jxb/erz152.

Su, Q. *et al.* 2020. A non-vector herbivore indirectly increases the transmission of a vector-borne virus by reducing plant chemical defense. Funct. Ecol. 34(5):1091-1101. doi: 10.1111/1365-2435.13535.

Su *et al.* 2020. Defence priming in tomato by the green leaf volatile (Z)-3-hexenol reduces whitefly transmission of a plant virus. Plant Cell Environ. 43(11):2797-2811. doi: 10.1111/pce.13885

Subramanyan, S. *et al.* 2015. Hessian fly larval feeding triggers enhanced polyamine levels in susceptible but not resistant wheat. BMC Plant Biol. 15:3. doi: 10.1186/s12870-014-0396-y.

Subramanian, A. *et al.* 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA. 102: 15545–15550. doi: 10.1073/pnas.0506580102

Suhag, A. *et al.* 2020. Biotechnological interventions for the sustainable management of a global pest, whitefly (*Bemisia tabaci*). Insect Sci. 00:1-22. doi:10.1111/1744-7917.12853.

Sun, G. *et al.* 2019. Glucosylation of the phytoalexin N-feruloyl tyramine modulates the levels of pathogen-responsive metabolites in *Nicotiana benthamiana*. doi: 10.1111/tpj.14420.

Szüts, A. & Szabó-Révész, P. 2012. Sucrose esters as natural surfactants in drug delivery systems—A mini-review. Int. J. Pharmaceut. 43:1-9. doi: 10.1016/j.ijpharm.2012.04.076.

Takahashi, Y. 2016. The role of polyamines in plant disease resistance. Environ. Control. Biol. 54(1):17-´21. doi: https://doi.org/10.2525/ecb.54.17.

Tebayashi , S. *et al.* 2007. Induction of resistance against the leafminer, *Liriomyza trifolii*, by jasmonic acid in sweet pepper. Biosci. Biotech. Biochem. 71:1521–1526. doi:10.1271/bbb.70033.

Thagun, C. *et al.* 2016. Jasmonate-responsive ERF transcription factor regulate steroidal glycoalkaloid biosynthesis in tomato. Plant Cell Physiol. 57(5):961-975. doi: 10.1093/pcp/pcw067.

Tieman, D. *et al.* 2010. Functional analysis of a tomato salicylic acid methyl transferase and its role in synthesis of the flavor volatile methyl salicylate. Plant J. 62:113-123. doi:1 0.1111/j.1365-313X.2010.04128.x

Tissier, A. 2012. Trichome specific expression: promoters and their applications. Transgenic plants – advances and limitations. IntechOpen. doi:10.5772/32101.

Tohge, T. *et al.* 2015. Ectopic expression of snapdragon transcription factors facilitates the identification of genes encoding enzymes of anthocyanin decoration in tomato. Plant J. 83:686-704. doi: 10.1111/tpj.12920.

Toghe, T. *et al.* 2017. Current understanding of the pathways of flavonoid biosynthesis in model and crop plants. J. Exp. Bot. 68(15):413-4028. doi: 10.1093/jxb/erx177.

Treutler, H. *et al.*, 2017. Discovering regulated metabolite families in untargeted metabolomics studies. Anal. Chem. 88(16): 8082-8090. doi: 10.1021/acs.analchem.6b01569.

Tsai, J.H. & Wang, K. 1996. Development and reproduction of *Bemisia argentifolii* (Homoptera : Aleyrodidae) on five host plants. Pop. Ecol.

Tsugawa, H. *et al.* 2015. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nature Methods. 12:523-526. doi: 10.1038/nmeth.3393.

Valentin, H.E. *et al.* 2006. The Arabidopsis vitamin E pathway gene5-1 mutant reveals a critical role for phytol kinase in seed tocopherol biosynthesis. Plant Cell. 18(1):212-224. doi: 10.1105/tpc.105.037077.

Van den Elsen, F. 2013. Resistance menchanisms against *Bemisia tabaci* in wild relatives of tomato. PhD Dissertation. Wageningen University. Wageningen.

van den Oever-van den Elsen, F. *et al.* 2015. Quantitative resistance against *Bemisia tabaci* in *Solanum pennellii*: Genetics and metabolomics. J. Integ. Plant. Biol. 58(4):397-412. doi: 10.1111/jipb.12449.

Van der Hoeven, R.S. *et al.* 2000. Genetic Control and Evolution of Sesquiterpene Biosynthesis in *Lycopersicon esculentum* and *L. hirsutum*. Plant Cell. 12(11):2283-2294. doi: 10.1105/tpc.12.11.2283.

Van Gijsegem, F. *et al.* 2017. Manipulation of ABA content in Arabidopsis thaliana modifies sensitivity and oxidative stress response to *Dickeya dadantii* and influences peroxidase activity. Front. Plant Sci. 8(456). doi: 10.3389/fpls.2017.00456.

Van Ooijen, J. 2006. JoinMap 4, software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V, Wageningen, Netherlands.

Van Ooijen, J. 2009. MapQTL 6, Software for the mapping of quantitative trait loci in experimental populations of diploid species. Kyazma BV; Wageningen, Netherlands.

Van Ooijen, J.W. 2011. Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. Genetics Research. 93(5):343-349. doi: 10.1017/S0016672311000279.

Van Schie, C.C.N. & Takken, F.L.W. 2014. Susceptibility genes 101: How to be a good host. Annu. Rev. Phytopathol. 52:551-581. doi: 10.1146/annurev-phyto-102313-045854.

Vinuchakkaravarthy, T. *et al.* 2010. Tris(2,4-di-tert-butylphenyl) phosphate. Acta Cryst. 66(9):2207-2208. doi: 10.1107/S1600536810029673.

Vosman, B. *et al.* 2019. QTL mapping of insect resistance components of *Solanum galapagense*. Theo. Appl. Gen. 132:531-541. doi: 10.1007/s00122-018-3239-7.

Walling, L.L. 2000. The myriad plant responses to herbivores. J. Plant Growth Regul. 19:195-216. doi: 10.1007/s003440000026.

Walters, D.S. & Steffens, J.C. 1990. Branched chain amino acid metabolism in the biosynthesis of Lycopersicon pennellii glucose esters. Plant Physiol. 93:1544–1551.

Wang, Z. *et al.* 2011. Two oxidosqualene cyclases responsible for biosynthesis of tomato fruit cuticular triterpenoids. Plant Phsysiol. 155:540-552. doi: 10.1104/pp.110.162883.

Wang, W. *et al.* 2019. Polyamine catabolism in plants: A universal process with diverse function. Front. Plant Sci. 10:561. doi: 10.3389/fpls.2019.00561

Wang, F. *et al.* 2020. Glandular trichome-derived sesquiterpenes of wild tomato accessions (*Solanum habrochaites*) affect aphid performance and feeding behavior. Phytochem. 180(112532). doi: 10.1016/j.phytochem.2020.112532.

War, A.R. *et al.* 2012. Mechanisms of plant defense against insect herbivores. Plant Signaling & Behavior. 7(10):1306-1320. doi: 10.4161/psb.21663.

Weinhold, A. & Baldwin, I.T. 2011. Trichome-derived O-acylsugar are a first meal for caterpillars that tags them for predation. PNAS. 108(19):7855-7859. doi:10.1073/pnas.1101306108.

Xia, J. *et al.* 2021. Whitefly hijacks a plant detoxification gene that neutralizes plant toxins. Cell. 184(7):1693-1705. doi: 10.1016/j.cell.2021.02.014

Xing, Z. *et al.* 2017. Efficiency of trichome-based plant defense in *Phaseolus vulgaris* depends on insect behavior, plant ontogeny and structure. Front. Plant Sci. 8. doi: 0.3389/fpls.2017.0200.

Yu, G. *et al.* 2010. Enzymatic functions of wild tomato methylketone synthases 1 and 2. Plant Physiol. 154:67-77. doi:10.1104/pp.11 0.157073.

Zabel, S. *et al.* 2021. A single cytochrome P450 oxidase from *Solanum habrochaites* sequencially oxidizes 7-*epi*–zingiberene to derivatives toxic to whiteflies and various microorgansms. Plant J. 105(5):1309-1325. doi:10.1111/tpj.15113.

Zamir, D. 2001. Improving plant breeding with exotic genetic libraries. Nat. Rev. Gen. 2:983-989.

Zarate, S.I. *et al.* 2007. Silverleaf whitefly induces salicylic acid defenses and suppresses effectual jasmonic acid defenses. Plant Physiol. 143:866-875. doi: 10.1104/pp.106.090035

Zeng, Z.B. 1994. Precision mapping of quantitative trait loci. Genetics 136:1457-1468.

Zhang, P.J. *et al.* 2013. Feeding by whiteflies suppresses downstream jasmonic acid signaling by eliciting salicylic acid signaling. J. Chem. Ecol. 39:612-619. doi: 10.1007/s10886-013-0283-2.

Zhang, X. *et al.* 2017. Phenolic compounds induced by *Bemisia tabaci* and *Trialeurodes vaporariorum* in *Nicotiana tabacum* L. and their relationship with the salicylic acid signaling pathway.

Zhang, Y. *et al.* 2020. The roles of different types of trichomes in tomato resistance to cold, drought, whiteflies, and *Botrytis*. Agronomy. 10(3):411. doi: 10.3390/agronomy10030411.

Zhou, W. *et al.* 2017. Tissue-specific emission of (*E*)-α-bergamotene helps resolve the dilemma when pollinators are also herbivores. Curr. Biol. Rep. 27(9):1336-1341. doi: 10.1016/j.cub.2017.03.017.

Zhou, F. & Pichersky, E. 2020. The complete functional characterization of the terpene synthase family in tomato. New Phytol. 226:1341-1360 . doi: 10.1111/nph.16431.

Ziegler, J. & Abel, S. 2014. Analysis of amino acids by HPLC/electrospray negative ion tandem mass spectrometry using 9-fluorenylmethoxycarbonyl chloride (Fmoc-Cl) derivatization. Amino acids. 46:2799-2808. doi: 10.1007/s00726-014-1837-5.

Zou, F. & Pichersky, E. 2020. The complete functional characterization of the terpene synthase family in tomato. New Phytol. 226(5):1341-1360. doi: 10.1111/nph.16431.

6. Appendix

Appendix 1. Primers

Topic	Primer name	Solyc-ID	Sequence
Cloning	pQE-Forward	-	GTATCACGAGGCCCTTTCGTCT
Cloning	pQE-Reverse	-	CATTACTGGATCTATCAACAGGAG
Cloning	Sh-AACS1-Forward	Solyc07g043630	CATGAGCTCATGAACAAATTCTTCAAAACCTCTAATATTGC
Cloning	Sh-AACS1-Reverse	Solyc07g043630	GTCGTCGACGTTACACTTTTCTGTCAGTGTTGAAAAGAC
Cloning	Sh-AACS2-Forward	Solyc02g082880	CATGAGCTCATGAATCCTTTTTTCAATATTTCAAGATTCAATG
Cloning	Sh-AACS2-Reverse	Solyc02g082880	GTCGTCGACGTCAGACTGCAATCTCTCTTTTGATTTC
Cloning	ASAT3-Forward	Solyc11g067270	CATGCTAGCATGGCATCATCAACAATTATATCTAGAAAAATG
Cloning	ASAT3-Reverse	Solyc11g067270	GTCCTCGAGTCCTTATTTGGTTGATTCAACAACTGGAGAAG
Cloning	AMD5-Forward	Solyc05g010420	CATGCTAGCATGGAAATGGACTTGCCAGTTTCTG
Cloning	AMD5-Reverse	Solyc05g010420	GTCCTCGAGGCTACTCCTTTCCTTCTTTCTCTTCCTC
Cloning	SAMS12-Forward	Solyc12g099000	CATGCTAGCATGGAGACTTTCTTATTCACATCTGAATC
Cloning	SAMS12-Reverse	Solyc12g099000	GTCCTCGAGGTTAAGCTTCAGGCTTGTCCCACTTG
Cloning	AMD5-5'UTR-Forward	Solyc05g010420	CCTAGTCATAGCCACATACCTACCG
Cloning	AMD5-5'UTR-Reverse	Solyc05g010420	CTGTGCCTTTGTGAGAGATCGAAG
qPCR	EF1α-Forward	Solyc06g005060	CAAATGATCTGCTGCTGTAACAAGATGG
qPCR	EF1α-Reverse	Solyc06g005060	GTCAGGGGTTGTAACCAACCTTCTTGAGG
qPCR	Sh-AACS1-Forward	Solyc07g043630	GTGATGTTGTTGCAACGCTGGC
qPCR	Sh-AACS1-Reverse	Solyc07g043630	CTTCTGAATGCCTCAGGAGAACTG
qPCR	Sh-AACS2-Forward	Solyc02g082880	GGTTGTAACTCTGGCTCCGAATGTAC
qPCR	Sh-AACS2-Reverse	Solyc02g082880	AAATATCATCTTGGTTTCAGAATGTTTGAGC
qPCR	ASAT3-LA1777-Forward	Solyc11g067270	TGAACCTGTAAGAGTAACCCTAGCAAC
qPCR	ASAT3-LA1777-Reverse	Solyc11g067270	GATTCAACAACTGGAGAAGCAAACTCC
qPCR	ASAT3-VI030462w-Forward	Solyc11g067270	AACCTGTAAGAGTAACCCTAGCAACTTC
qPCR	ASAT3-VI030462w-Reverse	Solyc11g067270	GATTCAACAACTGGAGAAGCAAACTCC
qPCR	ASAT3-VI030462c-Forward	Solyc11g067270	CCTGTAAGAGTAACCCTAGCAACACATC
qPCR	ASAT3-VI030462c-Reverse	Solyc11g067270	GATTCAACAACTGGAGAAGCAAACTCC
qPCR	AMD5-Forward	Solyc05g010420	GGTCTGTTTGCTGATCCTAATG
qPCR	AMD5-Reverse	Solyc05g010420	GAATAAACGAAGAGGCTCGAC
qPCR	SAMS12-Forward	Solyc12g099000	ACCATTTTCCACCTCAACCC
qPCR	SAMS12-Reverse	Solyc12g099000	ACCTTGGTAGGATCCTTCCC
VIGS	Sh-AACS1-Forward (MIC239)	Solyc07g043630	GGTCTCATATGAATTTCGCGGGGAGATGTGGTTG
VIGS	Sh-AACS1-Reverse (MIC240)	Solyc07g043630	GGTCTCACACCCAGCCATTCCCCAATTCATGCACC
VIGS	Sh-AACS2-Forward (MIC235)	Solyc02g082880	GGTCTCATATGTGGTGATGTTGTTGCAACGCTGG
VIGS	Sh-AACS2-Reverse (MIC236)	Solyc02g082880	GGTCTCACACCCGCCCCAAGTAAGGCTCCATCC
VIGS	ASAT3-LA1777-Forward	Solyc11g067270	GGTCTCATATGGTACCCTACCCACAACGCATCCATC
VIGS	ASAT3-LA1777-Reverse	Solyc11g067270	GGTCTCACACCGATGTGTTGCTAGGGTTACTCTTACAGG
VIGS	ASAT3-VI030462w-Forward	Solyc11g067270	GGTCTCATATGCTACCCACAACGCATCCATCAAAATG
VIGS	ASAT3-VI030462w-Reverse	Solyc11g067270	GGTCTCACACCTGAAGTTGCTAGGGTTACTCTTACAGG
VIGS	ASAT3-VI030462c-Forward	Solyc11g067270	GGTCTCATATGCTACCTGCAACGCGTCCATCAC
VIGS	ASAT3-VI030462c-Reverse	Solyc11g067270	GGTCTCACACCGATGTGTTGCTAGGGTTACTCTTACAGG
qPCR	AACS3	Solyc02g082870	TCAGAAGCGTCGCCATTATATGTG
qPCR	AACS3	Solyc02g082870	GGAGAGAGCAGCAAATGCATTTAG
qPCR	AACS4-Forward	Solyc08g075810	ATGGAGGGCACAGTGAAATGCTC
qPCR	AACS4-Reverse	Solyc08g075810	GCAAGTTGGATACACCTTTGACGAG
qPCR	AACS5-Forward	Solyc12g044300	CGATGCGAGTAGCAATGACCAAAC
qPCR	AACS5-Reverse	Solyc12g044300	TACCCCTCTTCGTCAACAGTAACAG
qPCR	ACOX1-Forward	Solyc08g078400	AATGCTGATGAGGGTTGCACAAG
qPCR	ACOX1-Reverse	Solyc08g078400	AGCCCGAGACAAAGCACCAG
qPCR	ACD1-Forward	Solyc06g073560	TGGAAGACCAATCGGCGAATTTC
qPCR	ACD1-Reverse	Solyc06g073560	TGCCAGTAGTATAGTCCCAGAACAATC
qPCR	ACOX4-Forward	Solyc10g076600	AATGACTCCCGGTCAGGCTAGC
qPCR	ACOX4-Reverse	Solyc10g076600	ATGGGCTCCAAGTCACAGAATGC
qPCR	ACOX4-tv2-Forward	Solyc10g085200	GGCGGAATTTGACACTGTAGCTTG
qPCR	ACOX4-tv2-Reverse	Solyc10g085200	ACCAATATGTCGGCAAAAGTGCTG
qPCR	AMD2	Solyc02g089610	GCTGTTGAAGGCTCCGCAATC
qPCR	AMD2	Solyc02g089610	CACGGAGAACTCTGCAGGTCC
qPCR	ACOX4-Reverse	Solyc10g076600	ATGGGCTCCAAGTCACAGAATGC

Appendix 2. Sequences

>VI030462_*Sh-AACS1*_cDNA_ Solyc02g082880
ATGAACAAATTCTTCAAAACCTCTAATATTGCAGTAAGATTTTTTAATGGATCAGTCCAGTTAGCCGCACCTACTCATAGGGTCCG
GCAACTGTGCCAACTTGCTGGGAGTATTGAATCAACCGATGAGTCCCAAAAATTATTGGAGGGTGTTGTTACTAGTCCAGCAAAT
TATGTTCCATTGACACCTATAAGTTTTTTGGAGAGAGCAGCAGATGTTTTTGGCGATAGAACTTCCGTTGTGTTTGGGTCTAGTGTG
AAGTATACTTGGGAAGAGACACACTCTAGGTGTCTAAAACTTGCCTCTGCTTTAATTCAGCTTGGAATTTCGCGGGGAGATGTGGT
TGTAACTCTGGCTCCGAATGTACCAGCAATGCAGGAATTGCATTTTGCAGTACCAATGGCTGGAGCTGTTCTATGTACGTTAAATA
CTCGTCTTGATTCATCAATGGTGGCTGATTTGCTCAAACATTCTGAAACCAAGATGATATTTGTTGATCAACAATTCCTCCAAATTG
CTCAACAAGCACTTTCTCTTCTTTCAAAAGACAAAACAATTAAACCACCAATTCTTATATTAATCCCCAAATCTAACAATTCATCT
CCTCCTGCATCTAACATTCACGAATACGAAAATCTTTTGTCAAGTGGGAGTAGCAATTTTACAATAAGATGGCCAAAAAGTGAAT
TTGATCCGATCAGTATCAATTATACTTCCGGTACAACGTCCTCGCCCAAAGGGGTGGTGTACAATCACAGAGGCGCATATCTCAAT
TCAATTTCTGCTTTCCTCTGTCATGGCATGGGTCCGATGCCAACGTACCTTTGGACACTTCCAATGTTTCACTGCAATGGATGGTGC
ATGAATTGGGGAATGGCTGCAATTGGTGGCACAAATGTTTGCCTTAGACATGTCTCAGCGAAAGATATATTTGAGAGTATTTCTGT
CAACAAGGTCACACATATGAGTGCAGCACCAATTGTCTTGAGTATGATGGCAAATGCTTCACCAAATGACAGGAAGCCACTTCCC
CATAAGGTCGAAATAATGACAGGCGGATCACCACCGCCTCCACAAATTCTTTCCAAAATGGAGCAACTTGGATTTGGAGTATCCC
ATGGATATGGACTAACAGAAACTTATAGTGGTGCTACGACTTGCTTGTGGAAGCCTGAGTGGGATTCTTTGCCCCTGGAGGAACG
AGCTGCGCTTAAATCAAGACAAGGGGTACAAGTTCTTTGTATAGAAAAAGTTGACGTTAGAGACCCGGAGACCATGGAAAATGTT
CCAGCTGATGGAAAGAGCATTGGTGAGATTGTGTGCAGAGGAAATACTGTGATGAGTGGATATTTGAAAGATGTTAAAGCAACTG
AAGAAGCTTTTAAAGGCGGATGGTTTCATACTGGTGATGTTGCAGTGAAACATCCAGATGGATATATAGAAATTAAGGATCGGTT
GAAAGATATTATAATATCTGGAGGTGAAAACATAAGCACACTCGAAGTGGAAGGAGTATTACATAGTCATCCTGCAGTTGTTGAG
GCAGCAGTAGTCGCACGACCAGATGATCATTGGGGACAAACACCTTGTGCATTTGTGAAGCTGAAAGAAGGATCTGAAGAAATA
ACTTCAGATGAAATAATCAAATATTGTAGGGATCATTTGCCACATTACATGGTGCCTCGAGCAGTCGTTTTTCAAGATTTACCAAG
AACTTCAACTGGCAAGGTACAAAAAATTCATCTTGAGAGAGAAAGCAAAAGCTTTGGCCAGTCTTTTCAACACTGACAGAAAGTG
TAA

>VI030462_*Sh-AACS2*_cDNA_ Solyc07g043630
ATGAATCCTTTTTTTCAATATTTCAAGATTCAATGGGTTGTTACATGCTCTAAACAGAGTTCGAGTTCATCCCATTTTGAGTCAGAGG
TCTCGTTATTTGTCCCAAATTATTGATAAGAATGTTGAAACTCACCCATGGGAATCTATGGAGGGACTAATGAGGTGTTCAGCTAA
TTATTTCCCTTTAACACCCATTAGTTTCTTGGACAGAGCTGCTAAGGTTTTTAGAGACAGGACTTCTGTTGTGTATGGTTCTTCTGT
TAAATTCACTTGGGAAGAGACACATAATAGGTGTCTAAAGATGGCTTCTGCTCTGTCTCAGTTGGGTATCTCTCGTGGTGATGTTG
TTGCAACGCTGGCCCCTAATGTACCTGCAGTGCAAGAGTTGCATTTTGCGGTGCCAATGGCAGGAGCAGTGCTTTGTACGTTGAAT
ACACGTCATGATTCAGCTATGGTATCAGTTCTCCTGAGGCATTCAGAAGCTAAGATCATTTTCGTTGACCAGCAGTTGTTTGATGT
TGCTCAAGGAGCACTGGATCTTCTTGCTGATGCTAAAACAAGACCTCTTCTAATATTAATCCCTGATCCTGAAAATCTACCGCCTC
CTGTTGCTGCTCCCAACGTTCATGAATATGAAACTCTTCTGGCAAGTGGGCGCGATGATTTTGCTATAAAGTGGCCGTTAACTGAA
TTTGACCCTATCAGTGTCAACTATACTTCTGGGACAACGTCACATCCTAAAGGAGTCGTTTACAATCATAGAGGTGCATATCTCAA
TGCTATTGCAACTCCCTATACTCATGAGATGGGCTCTATGCCTACTTATCTTTGGACTGTTCCAATGTTTCACTGTAATGGATGGAG
CCTTACTTGGGGCGTGGCTGCACTTGGTGGCACGAATGTATGCCTAAGACGTGTCTCTCCTAAGGACATTTTTGAGAATATATCCC
TCCACAAGGTCACACATATGAGTGCTGCACCAACGGTCATGAATATGATTGTAAATTCACCAAAAAGTGACCGCAAACCACTTCC
TCACAAGGTTGAAATAACGACAGGTGGTTCACCACCGCCTCCTCATATCATTTCCAAGATGGAGGAGTTAGGCTTTTCGGTATCTC
ATATATATGGTCTCACAGAGATTCATGGTCCATGTATGTCTTGTCTTCATCAGCCAGAGTGGGAATCATTACCTCCTGATGAACGA
TTTGCTCTGAAAGCAAGACAAGGAGTGGAGCACTATTTTACGCAGGGAATTGACATAAGAGATCCTGATACAATGGAGAGGGTTC
CGGATGATGGAAAGACCCTTGGTGAAATTATGATTAAAGGGAACACTGTGATGAGTGGATACTTGAAAAATATTAAAGCAACAG
AAGAAGTGTTCAGAGGTGGATGGTTTCATACTGGCGATCTTGCTGTGAGACATCCAGATGGCTACATAGAAGTTAAGGACCGGAT
GAAGGACATTATAATCTCTGGAGGTGAAAATATTTGCTCGGTTGAGGTGGAAAGAGTTTTGGTCAGTCATCCTGCAGTCCTTGAA
GCAGCTGTAGTAGCAAGACCAGATGATCACTGGGGGCAGACACCTTGTGCTTTCGTGAAGTTGAAGGAGGGATTTAGTCTTGGAT
CTGAAGATATAATCAACTATTGTCGGGATCATTTGCCTCATTATATGGCTCCCCAGACAGTCATATTTGAAGATCTACCAACAACT
TCTACGGGCAAGATACAAAAGTTTGTCCTAAGGGAAAAAGCAAAAGCCTTAGGCAGTGTTTGTGAAATCAAAAGAGAGATTGCA
GTCTGA

>LA1777_*ASAT3*_cDNA_ Solyc11g067270
ATGGCATCATCAACAATTATATCTAGAAAAATGATTAAACTTTTATCCCCAACTCCTCCTTCACTTAGATGTCACAAACTCTCTTTT
ATGGATCACATAAATCTTCCTCTACATTCTCCATGTGCCTTCTTCTACCCTAAAATACCTCAAAATTATAGTAACAAAATATCACA
AATACTTGAAAATTCCCTTTCAAAAATATTATCCTTTTATTATCCCTTAGCGGGAAAAATCAATAATAATTACACCTATGTCGATTG
TAATGACACAGGTGCTGAGTATTTAAACGTCCGTATCAATTGTCCAATGTCTCAAATTATCAACAACCCTTATAATGATGCTGTGG
GTGTAGTCTTCCCACAAGATTTGCCTTGGAGTAGTAGCTTGAATCGAAGTCCACTAGTGGTTCAATTAAGTCATTTTGATTGTGGT
GGAATAGCAGTCAGCGCATGTACATCACATACAATTTTTGATGGATATTGTGTCTCTAAATTTATAAATGATTGGGCGTCTACAGC
TCGAAACATGGATTTCAAACCATCTCCTCAGTTCAGTGCATCTACTTTCTTCCCTTTACCGTCTGAAACTAATTTGAGTAGTACCCT
ACCCACAACGCATCCATCAAAACGTCATGTCTCAAGAATGTACAATTTCTCATCCTCGAATTTGACAAGACTCAAGGATATCGTAA
CAAAAGAATCACATGTAAAGAATCCAACTCGCATTGAAGTTGCCTCAGCACTTGTTCATAAATGTGGGGTGGCTATGTCAATGGA
GAAATCAGGCATATTCAAACCAACTCTAATGAGCCATGCTATGAATTTACGCCCACCAATTCCACTAAACACAATGGGAAATGCA
ACATGTATCATTCTCACAACAACAATGACAGAAGATGAGGTAAAACTTCCAAACTTTGTTGCTAAACTACAGAAGGATAAACAAC
AACTTCGAGACAAGTTGAAGGATATGAAAAAAGATATGATGCCCTTGTATACACTTGAACTAGCTAAAAACGCGATGAACATAAT
AGAGAAGGATACACATGATGTTTATCTTTGCTCAGGCATGACCAATACTGGATTACATAAGATCGATTTCGGATGGGGTGAACCT
GTAAGAGTAACCCTAGCAACACATCCAAATAAGAACAACTTCATTTTCATGGATGAACAAAGTGGAGATGGGCTAAATGTACTTA
TCACTTTAACAAAAGATGATATGCTGAAGTTTCAGAGCAACAAGGAGCTTCTAGAGTTTGCTTCTCCAGTTGTTGAATCAACCAAA
TAAGCTT

>VI030462_*ASAT3w*_cDNA_ Solyc11g067270
ATGGCATCATCAACAATTATATCTAGAAAAATGATTAAACTTTTATCCCCAACTCCTCCTTCACTTAGATGTCACAAACTCTCTTTT
ATGGATCACATAAATTTTCATCTGCATTCTCCATATGCCTTCTTCTACCCTAAAATACCTCAAAATTATAGTAGCAAAATATCACA
AATACTTGAAAATTCCCTTTCAAAAGTATTATCCTTTTATTATCCCTTAGCGGGAAAAATCAATAATAATTACACCTATGTCGATTG
TAATGATACAGGTGTTGAGTATTTAAACGTCCGTATCGATTGTCCAATGTCTCAAATTATCAACGACCCTTATAATGATGCTGTGG
GTGTAGTTTTCCCACAAGATTTGCCTTGGAGTACTAGCTTGAATCGAAGTCCACTAGTGGTTCAATTAAGTCATTTTGATTGTGGTG
GAATAGCAGTCAGCGCATGTACATCACATACAATTTTTGATGGACATTCTCTCTCTAAATTCATAAATGATTGGGCGTCTACAGCT
CGAAATATGGATTTCAAACCATCTCCTCAGTTTAATGCATCTACTTTCTTCCCTTTACCGTCTGAAACTAATTTGAATAGTACCCTA
CCCACAACGCATCCATCAAAATGTCATGTCTCAAGAATGTACAATTTCTCATCCTCGAATTTGACAAGACTCAAGGATATCGTAAC
AAAAGAATCGCATGTAAAGAATCCAACTCGCATTGAAGTTGCCTCAGCACTTGTTCATAAATGTGGGGTGACTATGTCAATGGAG
AAATCAGGCATGTTCAAACCAACTCTAATGATCCATGCTATGAATTTACGCCCACCAATTCCACTAAACACAATGGGAAATGCAG
TATGTATCATTCTCACAACAACAGTGACAGAAGATGAGGTAAAACTTCCACACTTTGTTGCTAAACTACAGAAGGATAAACAACA
ACTTCGTGACAAGTTGAAGGATATGAAAAAAGATATGATGCCCTTATATACACTTGAACTAGGTAAAAACGCGATGAACATAATA
GAGAAGGATACACATGATGTTTATGTTTGCTCAGGCATGACCAATACTGGATTACATAAGATCGATTTCGGATGGGGTGAACCTG
TAAGAGTAACCCTAGCAACTTCAAATAAGAACAACTTCATTTTCATGGATGAACAAAGTGGAGATGGGCTAAATGTACTTATCAC
TTTAACAAAAGATGATATGCTGAAGTTTCAGAGCAACAAGGAGCTTCTAGAGTTTGCTTCTCCAGTTGTTGAATCAACCAAATAA

>VI030462_*ASAT3c*_cDNA_ Solyc11g067270
ATGGCATCATCAACAATTATATCTAGAAAAATGATTAAACTTTTATCCCCAACTCCATCTTCACTTAGATGTCACAAACTCTCTTTC
ATGGATCACATAAATTTCCCTCTACATTCTCCATATGCCTTCTTCTACCCTAAAATACCTCAAAATTATAGTAACAAAATATCACA
AGTACTTGAAAATTCCCTTTCAAAAGTATTGTCCTTTTATTATCCCTTAGCTGGAAAAATCAATAATAATTATACCTACGTCGATTG
TAACGACACAGGTGCTGAGTATTTAAACGTCCGTATAGATTGTCCAATGTCTCAAATTCTCAACCACCCTTATAACGATGTTGTGG
ATGTAGTTTTCCCACAAGATTTGCCTTGGAGTAGTAGTAGTAGCTTGACTCGAAGTCCACTAGTGGTTCAATTAAGCCATTTTGATTGT
GGTGGAGTTGCAGTGAGTGCATGTACATCACATACAATTTTTGATGGATATTGTCTCTCTAAATTCATAAACGATTGGGCGTCTAC
AGCTCGAAACATGGAGTTCAAACCATCTCCTCAGTTCAATGCATCAACTTTCTTCCCTTTACCGTCTGAAACTAATTTGAGTAGCA
CCCTACCTGCAACGCGTCCATCACAACGTCATGTCTCAAGAATGTATAACTTCTCATCCTCGAATTTGACAAGACTCAAGGACATC
GTAACAAAAGAATCACATGTGAAGAATCCAACTCGCGTTGAAGTTGCCTCAGCACTTGTTCATAAATGTGGGGTGACTATGTCAA
TGGAGAGTTCAGGCATGTTCAAACCAACTCTGATGAGCCATGCTATGAATTTACGCCCACCAATTCCACTGAACACAATGGGAAA
TGCAACATGTATCATTCTCACAACATCAATGACAGAAGATGAGGTAAAACTTCCAAACTTTGTTGCTAAACTACAGAAGGATAAA
CAACAACTTCGAGACAAGTTGAAGGATATGAAAGAAGATAGGATGCCCTTATATACACTTGAACTAGGTAAAAACGCGATGAAC
ATAATAGAGAAGGATACACATGATGTTTATCTTTGCTCAGGCATGACCAATACTGGATTACATAAGATCGATTTCGGATGGGGTG
AACCTGTAAGAGTAACCCTAGCAACACATCCAAATAAGAACAACTTCATTTTCATGGATGAACAAAGTGGAGATGGGCTAAATGT
ACTTATCACTTTAACAAAAGATGATATGCTCAAGTTTCAGAGCAACAAGGAGCTTCTGGAGTTTGCTTCTCCAGTTGTTGAATCAA
CCAAATAA

>AVT01424_*AMD5*_cDNA_Solyc05g010420
ATGGCTAGCATGGAAATGGACTTGCCAGTTTCTGCCATTGGTTTTGAAGGTTTCGAAAAGAGGCTCGAAATTTCTTTCGTCGAGCC
TGGTCTGTTTGCTGATCCTAATGGAAAAGGACTTCGATCTCTCACAAAGGCACAGTTGGATGAAATTCTCGGACCTGCTGAGTGCA
CCATTGTTGATAACCTGTCAAATGACTATGTTGATTCCTATGTGCTGTCCGAGTCGAGCCTCTTCGTTTATTCTTACAAGATAATCA
TCAAAACATGTGGTACCACAAAGCTGCTTCTTGCAATTCCGCCCATTCTGAGGTTGGCTGAGACCTTGTCTCTCAAAGTACAAGAC
GTGAGGTATACCCGTGGGAGCTTCATTTTCCCTGGTGCTCAATCGTTTCCTCACCGCCACTTTTCTGAAGAAGTTGCTGTCCTCGAT
GGATATTTTGGAAAGCTCGCTGCCGGTAGCAAGGCTGTGATTATGGGAAATCCCGACAAAACACAGAAATGGCATGTCTACTCTG
CCTCAGCTGGGACTGTTCAGTGTAATGACCCTGTTTACACTCTTGAGATGTGTATGACTGGTTTGAACAGGGAGAAGGCATCTGTC
TTCTACAAAACTGAAGAAAGTTCGGCTGCTCACATGACTGTTAGATCTGGCATCAGGAAGATCCTCCCCAAGTCTGAGATATGTG
ATTTTGAGTTTGAACCCTGTGGTTATTCTATGAATTCTATTGAAGGAGCTGCTGTTTCAACCATTCACATTACCCCGAGGACGGCT
TTAGCTATGCCAGCTTTGAATCTGTTGGATATGATCCTAAAACCAATGAGTTGGGTCCCCTGGTTGAGAGGGTGCTTGCATGTTTT
GAGCCAGCTGAGTTCTCTATTGCTCTGCATGCTGATGTTGCTACCAAGTTACTGGAGCATGTTTGCTCTGTTGATGTTAAGGGCTAC
TCTCTTGCTGAGTGGAGTCCAGAAGAGTTTGGCAAAGGCGGTTCCATTGTCTACCAGAAGTTCACTAGAACTCCTTACTGTGAATC
TCCCAAGTCCGTTCTGAAGGGCTGCTGGAAGGAGGAAGAGAAAGAAGGAAAGGAGTAGCCTCGAGCAC

>M82_*AMD5*_cDNA_Solyc05g010420
ATGGCTAGCATGGAAATGGACTTGCCAGTTTCTGCCATTGGTTTTGAAGGTTTCGAAGAGAGGCTCGAAATTTCTTTCGTCGAGCC
TGGTCTGTTTGCTGATCCTAATGGAAAAGGACTTCGATCTCTCACAAAGGCACAGTTGGATGAAATTCTCGGACCTGCTGAGTGCA
CCATTGTTGATAACCTGTCAAATGACTATGTTGATTCCTATGTGCTGTCCGAGTCGAGCCTCTTCGTTTATTCTTACAAGATAATCA
TCAAAACATGTGGTACCACAAAGCTGCTTCTTGCAATTCCGCCCATTCTGAGGTTGGCTGAGACCTTGTCTCTCAAAGTACAAGAC
GTGAGGTATACCCGTGGGAGCTTCATTTTCCCTGGTGCTCAATCGTTTCCTCACCGCCACTTTTCTGAAGAAGTTGCTGTCCTCGAT
GGATATTTTGGAAAGCTTGCTGCCGGTAGCAAGGCTGTGATTATGGGAAATCCCGACAAAACACAGAAATGGCATGTCTACTCTG
CCTCAGCTGGGACTGTTCAGTGTAATGACCCTGTTTACACTCTTGAGATGTGTATGACTGGTTTGAACAGGGAGAAGGCATCTGTC
TTCTACAAAACTGAAGAAAGTTCGGCTGCTCACATGACTGTTAGATCTGGCATCAGGAAGATCCTCCCCAAGTCTGAGATATGTG
ATTTTGAGTTTGAACCCTGTGGTTATTCTATGAATTCTATTGAAGGAGCTGCTGTTTCAACCATTCACATTACCCCGGAGGACGGCT
TTAGCTATGCCAGCTGTGAATCTGTTGGATATGATCCTAAAACCAATGAGTTGGGTCCCCTGCTTGAGAGGGTGCTTGCATGTTTT
GAGCCAGCTGAGTTCTCTATTGCTCTGCATGCTGATGTTGCTACCAACTTACTGGAGCATGTTCGCTCTGTTGATGTTAAGGGCTAC
TCTCTTGCTGAGTGGAGTCCAGAAGAGTTTGGCAAAGGCGGCTCCATTGTCTACCAGAAGTCCACTAGAATCTCCTTACTGTGAAT
CTGCCCAAGTCCGTTCTGAAGGGCTGCTGCAAGGAGGAAGAGCTAGACCGAAAGGAGTAGCCTCGAGCACC

>VI030462_*AMD5*_cDNA_Solyc05g010420
ATCATGGTGGCATTGCCATTTTCTGCCATTGGTTTTGAAGGTTTTGAAAAGAGGCTCGAAATTTCTTTCGTCGAGCCTGGTCTGTTT
GCTGATCCTAATGGAAAAGGACTTCGATCTCTCACAAAGGCACAGTTGGATGAAATTCTCGGACCTGCTGAGTGCACCATTGTTG
ATAACCTGTCAAATGACTATGTTGATTCCTATGTGCTGTCCGAGTCGAGCCTCTTCGTTTATTCTTACAAGATAATCATCAAAACAT
GTGGCACCACAAAGCTGCTTCTTGCAATTCCGCCCATTCTGAGGTTGGCTGAGACCTTGTCTCTCAAAGTACAAGACGTGAGGTAT
ACCCGTGGGAGCTTCATTTTCCCTGGTGCTCAATCGTTTCCTCACCGCCACTTTTCTGAAGAAGTTGCTGTCCTCGATGGATATTTT
GGAAAGCTTGCTGCCGGTAGCAAGGCTGTAATTATGGGAAATCCCGACAAAACACAGAAATGGCATGTTTACTCTGCCTCAGCTG
GGACTGTTCAGTGTAATGACCCTGTTTACACTCTTGAGATGTGTATGACTGGTTTGGACAGGGAGAAGGCATCTGTCTTCTACAAA
ACTGAAGAAAGTTCGGCTGCTCACATGACTGTTAGATCTGGCATCAGGAAGATCCTCCCCAAGTCTGAGATATGTGATTTTGAGTT
TGAACCCTGTGGTTATTCTATGAATTCTATTGAAGGAGCTGCTGTTTCAACCATTCACATTACCCCGGAGGACGGCTTTAGCTATG
CCAGCTTTGAATCTGTTGGATATGATCCTAAAACCACTGAGTTGGGTCCCCTGGTTGAGAGGGTGCTTGCATGTTTTGAGCCAGCT
GAGTTCTCTATTGCTCTGCATGCTGATGTTGCTACCAAGTTACTGGATCGTGTTTGCTCTGTTGATGTTAAGGGCTACTCTCTTGCT
GAGTGGAGTCCAGAAGAGTTTGGCAAAGGCGGTTCCATTGTCTACCAGAAGTTCACTAGAACTCCTTACTGTGAATCTCCCAAGT
CCGTTCTGAAGGGCTGCTGGAAGGAGGAAGAGAAAGAAGGAAAGGAGTAGTGTTGTCTTGAGGGTCGTATTGTTGTTTTTATTTC
AGTGTCTGTTGTTTGCTCAGAATAATGGACTTATACGTCCAAAACTTGTGTCTGTTGGGATTTGCAACGTCTGTGTGCAATTTCTCA
ACTAGTCTTGCCTTTTGGTGTTCCACCAGAAGCCTTTTATGTGTCTGCACTTTGAATTGTGTCATGTTGTTGGTTTCTGTTCCTGTTG
CGTCTATTTATAAAAGTGTGTTTGTTTTGTGTAAAAAAAAAC

>AVT01424_*SAMS12*_cDNA_ Solyc12g099000
ATGGCTAGCATGGAGACTTTCTTATTCACATCTGAATCAGTCAACGAGGGACATCCCGACAAGCTCTGTGACCAGGTCTCTGATGC
AGTGCTCGATGCCTGTCTAGCTCAGGACCCTGAAAGCAAAGTTGCATGTGAGACTTGTACCAAGACCAACTTAGTTATGGTCTTTG
GAGAGATCACCACCAAGGCCAATATTGATTATGAGAAGATTGTACGTGACACTTGCCGGGAAATTGGATTTGTGTCCCCTGATGT
TGGTTTGGATGCTGACAATTGCAGAGTCCTTGTGAACATTGAGCAGCAGAGCCCTGATATTGCTCAAGGTGTTCATGGTCATTTGA
CTAAGCGACCTGAGGAGATTGGTGCTGGTGACCAGGGTCACATGTTTGGCTATGCCACTGACGAGACACCGGAGTTGATGCCCCT
TAGCCATGTTCTTGCTACCAAACTTGGAGCTCGCCTTACTGAGGTCCGCAAGAATGGAACTTGCTCTTGGCTTAGACCTGATGGTA
AAACACAAGTGACTGTTGAGTATCACAATGACAATGGTGCTATGGTTCCTCTACGTGTTCACACTGTTTTAATCTCCACTCAGCAT
GATGAGACTGTTACCAATGATGAAATTGCTCGTGATCTCAAAGAGCATGTCATCAAGCCTGTCATCCCCGAGAAGTATCTTGATG
AGAACACCATTTTCCACCTCAACCCTTCAGGCCGTTTTGTCATTGGTGGACCTCACGGTGATGCTGGTCTCACTGGCCGTAAGATC
ATCATTGACACTTACGGAGGTTGGGGTGCTCACGGAGGTGGTGCTTTCTCTGGGAAGGATCCTACCAAGGTCGACAGAAGTGGAG
CCTATATCGTGAGGCAAGCAGCTAAGAGCATTGTTGCTAACGGTCTTGCCAGAAGGTGCATTGTTCAGGTTTCATATGCCATCGGT
GTGCCTGAGCCATTGTCCGTCTTTGTGGACACTTATGGAACTGGGAAGATCCCTGACAAGGAGATTCTCAACATTGTGAAGGAGA
ACTTTGATTTCAGGCCCCGGAATGATTTCCATTAACCTCGATCTACTGAGGGGTGGCAATGGTCGATTCTTGAAGACTGCTGCCTAT
GGGCATTTTGGTAGAGATGATCCTGACTTCACATGGGAAGTAGTTAAGCCTCTCAAGTGGGACAAGCGGGAAGCTTAACCTCGAG
CAC

>M82_*SAMS12*_cDNA_ Solyc12g099000
ATGGCTAGCATGGAGACTTTCTTATTCACATCTGAATCAGTCAACGAGGGACATCCCGACAAGCTCTGTGACCAGGTCTCTGATGC
AGTGCTCGATGCCTGTCTAGCTCAGGACCCTGAAAGCAAAGTTGCATGTGAGACTTGTACCAAGACCAACTTAGTTATGGTCTTTG
GAGAGATCACCACCAAGGCCAATATTGATTATGAGAAGATTGTACGTGACACTTGCCGGGAAATTGGATTTGTGTCCCCTGATGT
TGGTTTGGATGCTGACAATTGCAGAGTCCTTGTGAACATTGAGCAGCAGAGCCCTGATATTGCTCAAGGTGTTCATGGTCATTTGA
CTAAGCGACCTGAGGAGATTGGTGCTGGTGACCAGGGTCACATGTTTGGCTATGCCACTGACGAGACACCGGAGTTGATGCCCCT
TAGCCATGTTCTTGCTACCAAACTTGGAGCTCGCCTTACTGAGGTCCGCAAGAATGGAACTTGCTCTTGGCTTAGACCTGATGGTA
AAACACAAGTGACTGTTGAGTATCACAATGACAATGGTGCTATGGTTCCTCTACGTGTTCACACTGTTTTAATCTCCACTCAGCAT
GATGAGACTGTTACCAATGATGAAATTGCTCGTGATCTCAAAGAGCATGTCATCAAGCCTGTCATCCCCGAGAAGTATCTTGATG
AGAACACCATTTTCCACCTCAACCCTTCAGGCCGTTTTGTCATTGGTGGACCTCACGGTGATGCTGGTCTCACTGGCCGTAAGATC
ATCATTGACACTTACGGAGGTTGGGGTGCTCACGGAGGTGGTGCTTTCTCTGGGAAGGATCCTACCAAGGTCGACAGAAGTGGAG
CCTATATCGTGAGGCAAGCAGCTAAGAGCATTGTTGCTAACGGTCTTGCCAGAAGGTGCATTGTTCAGGTTTCATATGCCATCGGT
GTGCCTGAGCCATTGTCCGTCTTTGTGGACACTTATGGAACTGGGAAGATCCCTGACAAGGAGATTCTCAACATTGTGAAGGAGA
ACTTTGATTTCAGGCCCGGAATGATTTCCATTAACCTCGATCTACTGAGGGGTGGCAATGGTCGATTCTTGAAGACTGCTGCCTAT
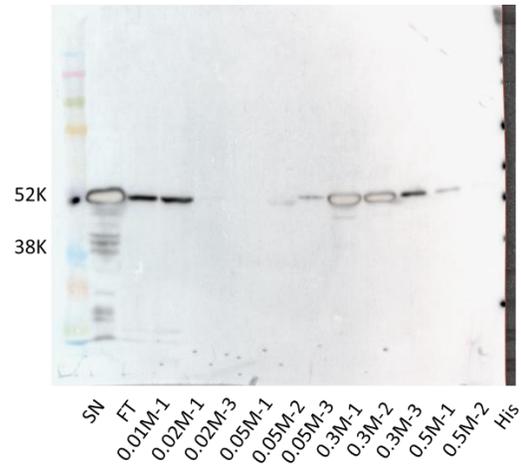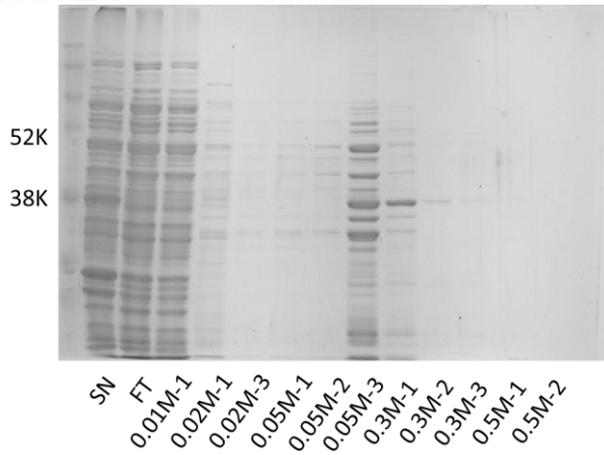GGGCATTTTGGTAGAGATGATCCTGACTTCACATGGGAAGTAGTTAAGCCTCTCAAGTGGGACAAGCCTGAAGCTTAACCTCGAG
CAC

Appendix 3. Western blot and coomassie gel

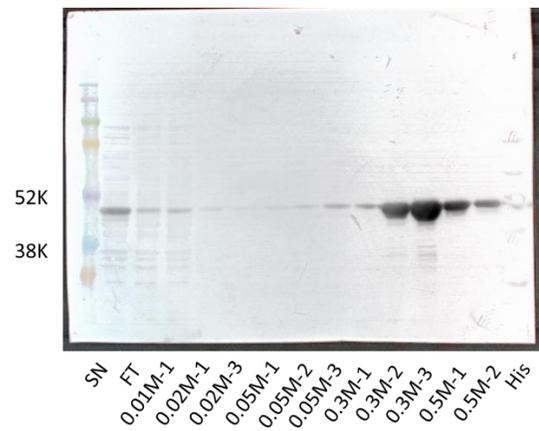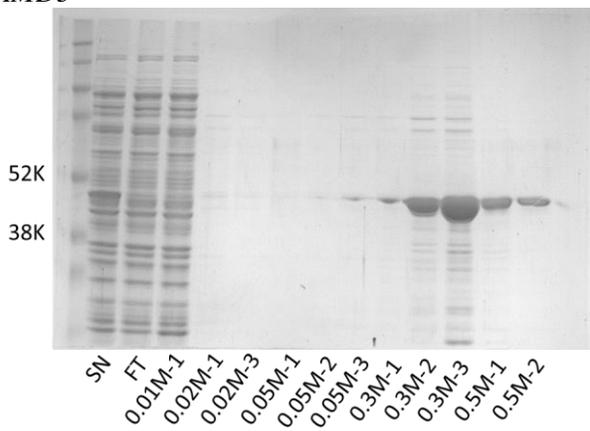| Coomassie	Western Blot

Sh-AACS1

Sh-AACS2
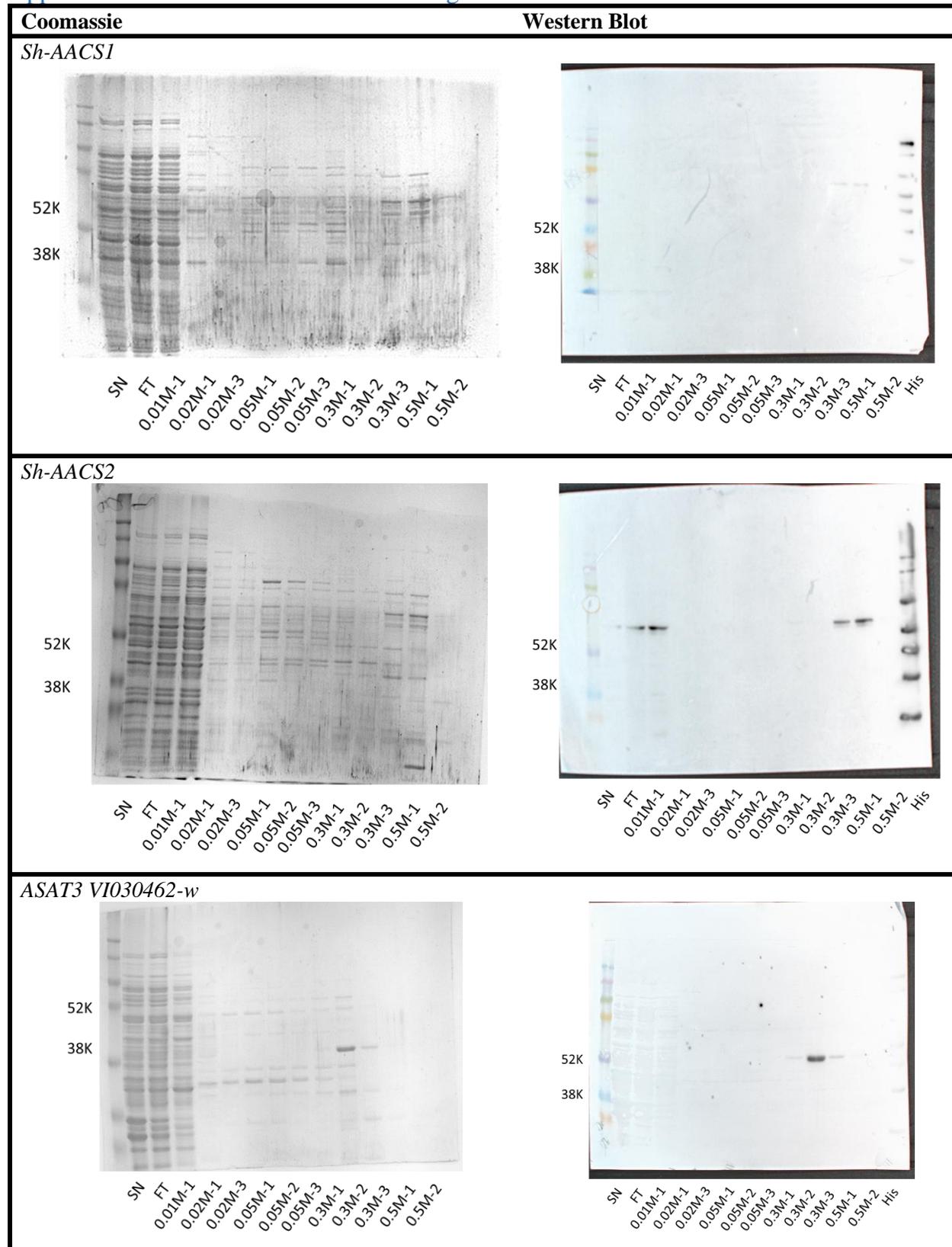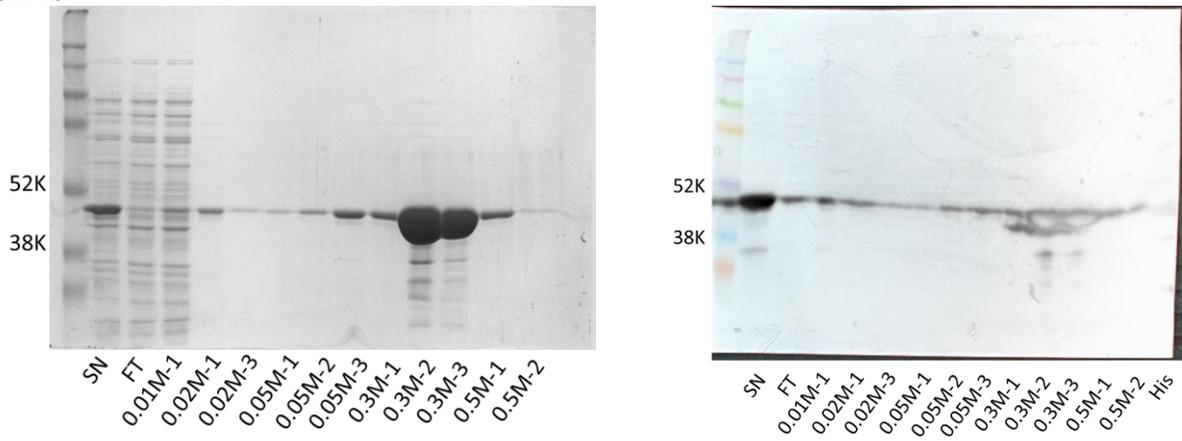

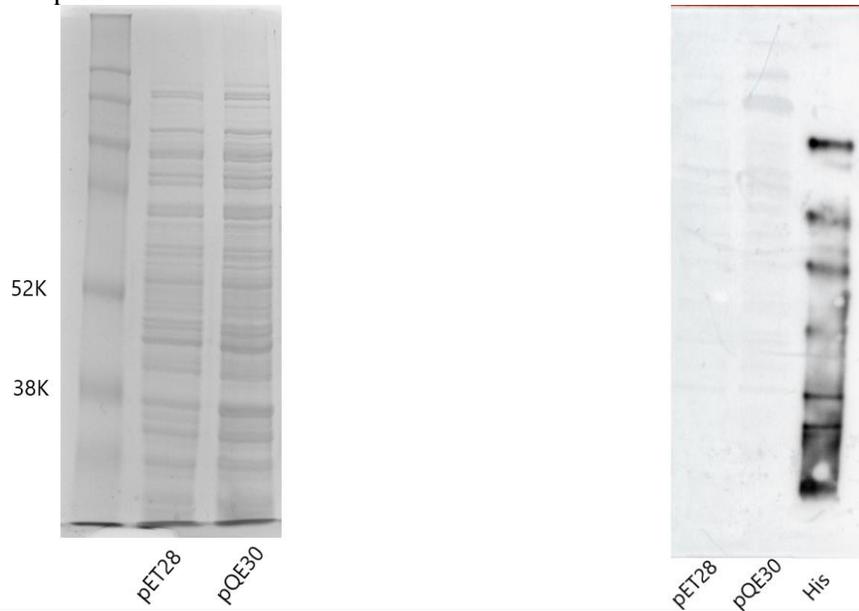**ASAT3 VI030462-w**

ASAT3 VI030462-c
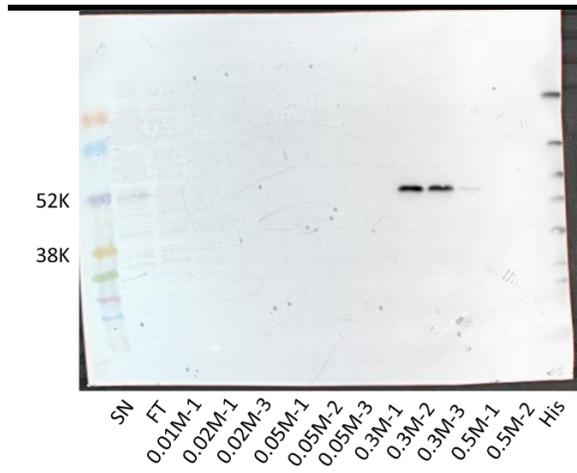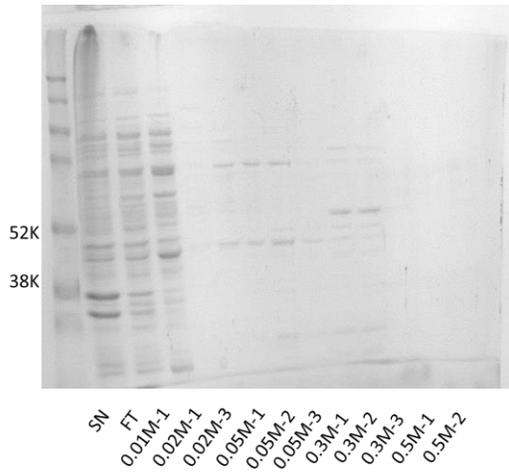
ASAT3 LA1777

AMD5

SAMS12

Empty vectors - supernatant

Appendix 4. MRM

A) HPLC

Q1 Mass(Da)	Q3 Mass (Da)	Time (min)	ID	DP(volts)	CE(volts)
621.2	250.1	3.4	SAM-Fmoc-1	25.393	30
621.2	146	3.4	SAM-Fmoc-2	25.393	30
621.2	298.1	3.4	SAM-Fmoc-3	25.393	30
563.2	136	5	dcSAM-Fmoc-1	23.881	30
563.2	206	5	dcSAM-Fmoc-2	23.881	30
563.2	428.1	5	dcSAM-Fmoc-3	23.881	20
393.4	171	6.5	F-D5-1	20	15
393.4	125	6.5	F-D5-2	20	30
393.4	179	6.5	F-D5-3	20	30

B) QTRAP

ID	Q1 Mass (Da)	Q3 Mass (Da)	Time (msec)	DP (volts)	CE (volts)	CXP (volts)
ADP	426.2	78.85	10.8	-75	-66	-3
ATP	506.2	158.78	11.1	-80	-38	-9
CoA	766.1	78.9	11.6	-40	-122	-1
C2-coA	808.12	808.12	-	-240	-38	-13
C4-coA	836.15	836.15	11.91	-240	-38	-13
C5-coA	850.17	850.17	12.22	-240	-38	-13
C6-coA	864.19	864.18	12.65	-240	-38	-13
C8-coA	892.21	892.21	13.3	-240	-38	-13
C10-coA	920.24	920.24	13.96	-240	-38	-13
C12-coA	948.27	948.27	14.61	-240	-38	-13
C2-FAs	59.01	59.01	-	-28	-10	-10
C4-FAs	87.05	87.05	4.55	-28	-10	-10
C5-FAs	101.06	101.06	6.97	-28	-10	-10
iC5-FAs	101.06	101.06	6.56	-28	-10	-10
aiC5-FAs	101.06	101.06	6.48	-28	-10	-10
C6-FAs	115.08	115.08	8.8	-28	-10	-10
C8-FAs	143.21	143.21	13.23	-28	-10	-10
C10-FAs	171.14	171.14	15.56	-28	-10	-10
C12-FAs	199.17	199.17	17.5	-28	-10	-10

C) QTOF

MS source parameters

Curtain gas	35 arbitrary
Gas 1/2	60/70 psi
Voltage	-4500V
Source temperature	600 °C
Source housing	Duo spray, ion source

SWATH MS and MS/MS parameters

MS1 Exp. Index:	Start mass (Da)	Stop mass (Da)	Declustering potential (V)	Entrance (V)	Collision energy (V)	Collision energy spread	Dwell time (ms)
0	65	1250	-35	-10	-10	-	150
1	65	90	-35	-10	-45	35	20
2	89	115	-35	-10	-45	35	20
3	114	140	-35	-10	-45	35	20
4	139	165	-35	-10	-45	35	20
5	164	190	-35	-10	-45	35	20
6	189	215	-35	-10	-45	35	20
7	214	240	-35	-10	-45	35	20
8	239	265	-35	-10	-45	35	20
9	264	290	-35	-10	-45	35	20
10	289	315	-35	-10	-45	35	20
11	314	340	-35	-10	-45	35	20
12	339	365	-35	-10	-45	35	20
13	364	390	-35	-10	-45	35	20
14	389	415	-35	-10	-45	35	20
15	414	440	-35	-10	-45	35	20
16	439	465	-35	-10	-45	35	20
17	464	490	-35	-10	-45	35	20
18	489	515	-35	-10	-45	35	20
19	514	540	-35	-10	-45	35	20
20	539	565	-35	-10	-45	35	20
21	564	590	-35	-10	-45	35	20
22	589	615	-35	-10	-45	35	20
23	614	640	-35	-10	-45	35	20
24	639	665	-35	-10	-45	35	20
25	664	690	-35	-10	-45	35	20
26	689	715	-35	-10	-45	35	20
27	714	740	-35	-10	-45	35	20
28	739	765	-35	-10	-45	35	20
29	764	790	-35	-10	-45	35	20
30	789	815	-35	-10	-45	35	20
31	814	840	-35	-10	-45	35	20
32	839	865	-35	-10	-45	35	20
33	864	890	-35	-10	-45	35	20
34	889	915	-35	-10	-45	35	20
35	914	940	-35	-10	-45	35	20
36	939	965	-35	-10	-45	35	20
37	964	990	-35	-10	-45	35	20
38	989	1015	-35	-10	-45	35	20
39	1014	1040	-35	-10	-45	35	20
40	1039	1065	-35	-10	-45	35	20
41	1064	1090	-35	-10	-45	35	20
42	1089	1115	-35	-10	-45	35	20
43	1114	1140	-35	-10	-45	35	20
44	1139	1165	-35	-10	-45	35	20
45	1164	1190	-35	-10	-45	35	20
46	1189	1215	-35	-10	-45	35	20
47	1214	1240	-35	-10	-45	35	20
48	1239	1250	-35	-1	-45	35	20

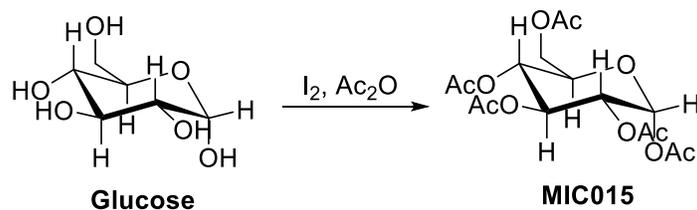## Appendix 5. Synthesis of AS

Glucose ring

Figure 51. Synthesis of MIC015 from glucose.

Glucose (15 g, 83.3 mmol) was suspended in acetic anhydride (75.41 mL, 799 mmol). Diiodide (0.74 g, 2.916 mmol) was added dropwise and the reaction was stirred at room temperature for 15 minutes. The reaction was diluted with dichloromethane (DCM) (100 mL) and washed a 5 % aqueous solution of sodium-thiosulphate ($Na_2S_2O_3$) (20 mL). The organic phase was collected and dried under reduced pressure. Ice water (250 mL) was added to the product and the mixture was stirred for 2 h and after neutralized with saturated sodium bicarbonate ($NaHCO_3$). The organic phase was extracted three times with DCM (3 x 100 mL), dried over sodium sulphate (Na_2SO_4) and concentrated under reduced pressure to afford glucose penta-acetate (**MIC015**), as a thick syrup (30.3 g, 96 %). For the major diastereoisomer (α-anomer): ^{1}HNMR (400 MHz, $CDCl_3$): δ = 1.99 (s, 3H, CH_3); 2.00 (s, 3H, CH_3); 2.01 (s, 3H, CH_3); 2.01 (s, 3H, CH_3); 2.06 (s, 3H, CH_3); 2.15 (s, 3H, CH_3); 4.0 – 4.12 (m, 2H); 4.22 – 4.27 (m, 1H); 5.07 (dd, J = 10.3/3.8 Hz, 1H, CH-2); 5.09 – 5.13 (m, 1H); 5.44 (dd, J = 10.3/9.5, 1H, CH); 6.3 (d, J = 3.8 Hz, CH-1). ^{13}CNMR (100 MHz, $CDCl_3$): δ = 20.5, 20.6, 20.7, 20.8, 20.9 (CH_3); 61.5 (CH_2); 68.0, 69.3, 69.9, 89.1 (CH); 168.8, 169.5, 169.7, 170.3, 170.7 (C=O). ESI-MS: m/z calculated for $C_{16}H_{22}NaO_{11}$ [M+Na]$^+$: 413.1; found: 413.0.


Figure 52. Synthesis of MIC016 from MIC015.

To a solution of **MIC015** (30.3 g, 77.7 mmol) in dried DCM (10 mL) was added thiophenol (PhSH) (7.13 mL, 69.9 mmol) and freshly distilled boron trifluoride etherate ($BF_3.Et_2O$) (17.26 mL, 139.8 mmol) under argon atmosphere and the reaction mixture was stirred for 24 h. The reaction mixture was then diluted with DCM (100 mL) and washed successively with H_2O (50 mL), saturated $NaHCO_3$ aqueous solution (3 x 20 mL) and H_2O (30 mL). The organic phase was dried over Na_2SO_4 and concentrated under reduced pressure. The dried crude product was purified by column chromatography (*n*-hexane/EtOAc 5:1 to 3:1 (*v/v*), Retention factor (Rf) = 0.51 (*n*-hexane/EtOAc 2:1 (*v/v*)) to afford **MIC016** as a white solid (13.4 g, 39%). ^{1}HNMR (400 MHz, $CDCl_3$): δ = 1.92 (s, 3H, CH_3); 1.95 (s, 3H, CH_3); 2.01 (s, 3H, CH_3); 2.02 (s, 3H, CH_3); 3.66 (ddd, J = 10.1/5.0/2.6 Hz, 1H, CH-5); 4.13 (dd, J = 12.3/2.8 Hz, 1H, CH_2); 4.14 (dd, J = 12.3/5.0 Hz, 1H, CH_2); 4.65 (d, J = 10.0 Hz, 1H, CH-1); 4.91 (dd, J =10.0/9.2 Hz, 1H, CH); 4.97 (t, J = 10.0 Hz, 1H, CH); 7.23 – 7.27 (m, 3H, SPh); 7.40 – 7.45 (m, 2H, SPh). ^{13}CNMR (100 MHz, $CDCl_3$): δ = 20.7, 20.8 (CH_3); 62.2 (CH_2); 68.3, 70.0, 74.0, 75.9, 85.8, 128.5, 129.0 (CH); 131.7 (C); 133.2 (CH); 169.3, 169.5, 170.2, 170.6 (C=O). ESI-MS: m/z calculated for $C_{20}H_{24}NaO_9S$ [M+Na]$^+$: 463.1; found: 462.8.
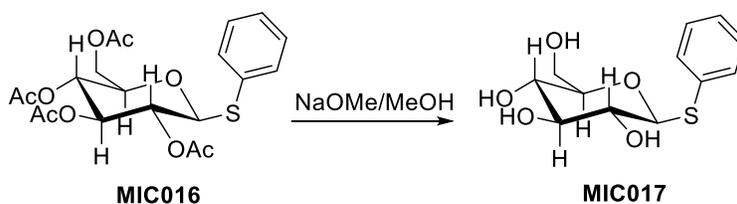
Figure 53. Synthesis of MIC017 from MIC016.

To a solution of **MIC016** (13.4 g; 30,4 mmol) in dried methanol (MeOH) (80 mL), a fresh solution of sodium methoxide (NaOMe) (0.3 mol/L, 7.17 mL) were added. The reaction mixture was stirred for 2 h at room temperature under argon atmosphere and then neutralized with Amberlite® IR120 (H+) resin. The salt was filtered out by filtered paper and the solution was evaporated to afford **MIC017** as a pale yellow solid (8.3 g, ~100%). ^{1}HNMR (400 MHz, CD$_3$OD): δ = 3.22 (dd, J = 9.8/8.6 Hz, 1H, CH); 3.28 – 3.34 (m, 2H); 3.36 – 3.42 (m, 1H); 3.64 – 3.69 (m, 1H); 3.84 – 3.88 (m, 1H); 4.59 (d, J = 9.8 Hz, 1H, CH-1); 7.21 – 7.32 (m, 3H, SPh); 7.53 – 7.57 (m, 2H, SPh). ^{13}CNMR (100 MHz, CD$_3$OD): δ = 62.8 (CH_2); 71.3, 73.7, 79.6, 81.9, 89.3, 128.3, 129.8, 132.6 (CH); 135.2 (C). ESI-MS: m/z calculated for C$_{12}$H$_{16}$NaO$_5$S [M+Na]$^+$: 295.1; found: 295.0.
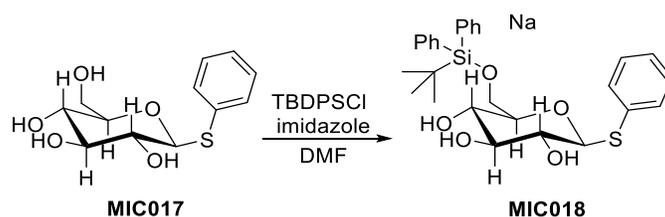
Figure 54. Synthesis of MIC017 to MIC018.

To a stirred solution of **MIC017** (8.3 g, 30.2 mmol) and imidazole (4.6 g, 67.1 mmol) in dimethylformamide (DMF) (15 mL) at 0 °C was added drop-wise *tert*-butylchlorodiphenylsilane (TBDPSCl) (10.15 mL, 39.669 mmol). The reaction mixture was allowed to reach room temperature and was stirred for two days. The reaction was stopped with the addition of H$_2$O (20 mL). The aqueous phase was extracted with diethyl ether (Et$_2$O) (3 x 50 mL) and the organic phase was dried over Na$_2$SO$_4$ and concentrated under reduced pressure. The crude product was purified by column chromatography (*n*-hexane/EtOAc 2:1 to 1:1 (*v/v*)) to afford **MIC018** (14 g, 90%), Rf = 0.5 (*n*-hexane/EtOAc 1:2 (*v/v*)). ^{1}HNMR (400 MHz, CDCl$_3$): δ = 1.12 (s, 9H, 3 x CH_3, Si*t*Bu); 3.39 – 3.46 (m, 1H); 3.48 – 3.55 (m, 1H); 3.59 – 3.68 (m, 2H); 3.95 (dd, J = 11.0/5.1 Hz, 1H, CH_2); 4.04 (dd, J = 11.0/3.3 Hz, 1H, CH_2); 4.62 (d, J = 9.7 Hz, 1H, CH-1); 7.21 – 7.28 (m, 3H, Ar); 7.38 – 7.49 (m, 6H, Ar); 7.57 – 7.62 (m, 2H, Ar); 7.76 – 7.81 (m, 4H, Ar). [3]CNMR (100 MHz, CDCl$_3$): δ = 19.3 (C); 26.9 (CH_3); 64.2 (CH_2); 70.7, 71.9, 78.1, 79.7, 87.9, 127.8, 129.0, 129.8, 132.1(CH); 132.8, 133.1, 133.3 (C); 135.7, 135.8 (CH). ESI-MS: m/z calculated for C$_{28}$H$_{34}$NaO$_5$SSi [M+Na]$^+$: 533.2; found: 533.1.

Figure 55. Synthesis of MIC019 from MIC018.
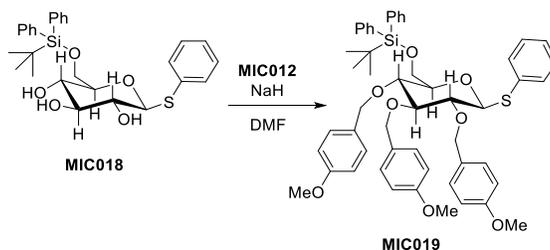
To a suspension of **MIC018** (14 g, 27.5 mmol) in DMF (50 mL) cooled to 0 °C was added sodium hydride (NaH) (2.8 g, 115.3 mmol, from 60% dispersion in mineral oil) and the reaction mixture was stirred over a period of 30 minutes before the addition of 4-methoxybenzyl bromide (PMBBr) (20.9 g, 104.3 mmol). The reaction mixture was allowed to reach room temperature and was stirred overnight. The reaction mixture was cooled in an ice bath and quenched with MeOH. The organic phase was extracted with DCM (4 x 100 mL) and then washed with brine (3 x 100 mL) and H_2O (3 x 100 mL). The product was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by column chromatography (*n*-hexane/EtOAc 9:1 (*v/v*)) to afford **MIC019** as a transparent syrup (13.6 g, 57%). Rf = 0.33 (*n*-hexane/EtOAc 9:1 (*v/v*)). ^{1}HNMR (400 MHz, CDCl$_3$): δ = 1.08 (s, 9H, 3 x C*H*$_3$, Si*t*Bu); 3.37 (ddd, *J* = 9.4/3.9/1.8 Hz, 1H, C*H*-5); 3.64 – 3.70 (m, 1H); 3.70 – 3.76 (m, 1H); 3.78 (s, 3H, OC*H*$_3$); 3.80 (s, 3H, OC*H*$_3$); 3.81 (s, 3H, OC*H*$_3$); 3.90 (dd, *J* = 11.4/3.9 Hz, 1H, C*H*$_2$); 3.97 (dd, *J* = 11.4/1.8 Hz, 1H, C*H*$_2$); 4.57 – 4.63 (m, 2H); 4.64 – 4.69 (m, 2H); 4.77 – 4.86 (m, 4H); 6.75 – 6.80 (m, 2H, Ar); 6.84 – 6.92 (m, 4H, Ar); 7.02 – 7.07 (m, 2H, Ar); 7.18 – 7.23 (m, 3H, Ar); 7.24 – 7.45 (m, 10H, Ar); 7.56 – 7.63 (m, 2H, Ar); 7.69 – 7.78 (m, 4H, Ar). ^{13}CNMR (100 MHz, CDCl$_3$): δ = 19.4 (*C*); 27.0, 31.0, 55.4 (*C*H$_3$); 62.8, 74.9, 75.1, 75.7 (*C*H$_2$); 77.3, 80.1, 80.7, 86.7, 87.6, 114.0, 114.1, 127.3, 127.8, 127.9, 128.8, 129.0, 129.7, 129.8, 129.9, (*C*H); 130.4, 130.6, 130.8 (*C*); 131.7 (*C*H); 133.1, 133.7, 134.4 (*C*); 135.8, 136.0 (*C*H); 159.4, 159.5 (*C*). ESI-MS: m/z calculated for C$_{52}$H$_{58}$NaO$_8$SSi [M+Na]$^+$: 893.4; found: 893.3.
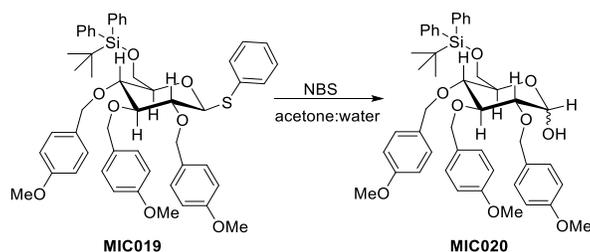

Figure 56. Synthesis of MIC020 from MIC019.

MIC019 (11.6 g, 13.0 mmol) was dissolved in 50 mL of a mixture of acetone: H_2O 24:1 (*v/v*) and stirred with external EtOH/ice bath. N-bromosuccinimide (NBS) (7.37 g, 41.4 mmol) was added to the reaction mixture at room temperature during a period of 5 minutes and stirred for 20 min The reaction mixture was diluted with H_2O (50 mL), extracted with Et$_2$O (3 x 50 mL) and washed with Na$_2$CO$_3$ (50 mL). The organic phase was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by column chromatography (*n*-hexane/EtOAc 3:1 (*v/v*)) to afford **MIC020** (7.4 g, 71%). Rf = 0.3 (*n*-hexane/EtOAc 3:1 (*v/v*)). For the major rotamer: ^{1}HNMR (400 MHz, CDCl$_3$): δ = 1.05 (s, 9H, 3 x C*H*$_3$, Si*t*Bu); 3.52 – 3.63 (m, 1H); 3.66 – 3.77 (m, 1H); 3.79 (s, 3H, OC*H*$_3$); 3.80 (s, 3H, OC*H*$_3$); 3.82 (s, 3H, OC*H*$_3$); 3.84 – 3.95 (m, 3H); 4.55 – 4.90 (m, 8H); 6.77 – 6.82 (m, 2H, Ar); 6.84 – 6.91 (m, 4H, Ar); 7.06 – 7.12 (m, 2H, Ar); 7.23 – 7.44 (m, 11H, Ar); 7.65 – 7.70 (m, 3H, Ar). ^{13}CNMR (100 MHz, CDCl$_3$): δ = 19.5 (*C*); 27.0, 31.1, 55.4 (*C*H$_3$); 62.8 (*C*H$_2$); 71.8 (*C*H); 73.1, 74.9, 75.7 (*C*H$_2$); 77.2, 80.4, 81.7, 91.4, 113.9, 114.0, 114.1, 127.7, 127.8, 128.8, 129.6, 129.7, 129.9 (*C*H); 130.3, 130.8, 131.1, 133.3, 133.4, 133.9 (*C*); 135.8, 136.0 (*C*H); 159.3, 159.6 (*C*). ESI-MS: m/z calculated for C$_{46}$H$_{54}$NaO$_9$Si [M+Na]$^+$: 801.3; found: 801.2.
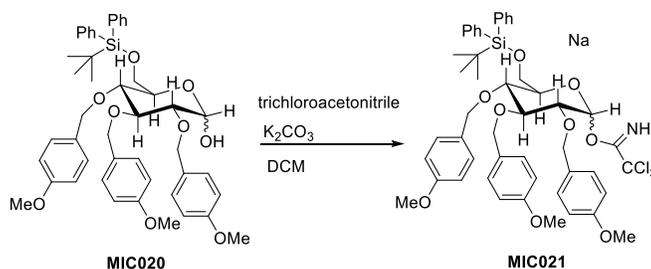

Figure 57. Synthesis of MIC021 from MIC020.

To a suspension of **MIC020** (4.7 g, 6.0 mmol) in dried DCM (50 mL) was added 4.7 g potassium carbonate (K_2CO_3) (4.7 g, 34.1 mmol) and trichloroacetonitrile (3.224 mL, 32.0 mmol) and the reaction mixture was strongly stirred for 5 h under nitrogen atmosphere. The mixture was filtered over celite and the pad was washed several times with DCM. The filtrate was concentrated under reduced pressure and purified by column chromatography (*n*-hexane/EtOAc 7:1 (*v/v*)) to afford **MIC021** (3.5 g, 63%). Rf = 0.21 (*n*-hexane/EtOAc 7:1 (*v/v*)). ^{1}HNMR (400 MHz, CDCl$_3$): δ =1.04 (s, 9H, 3 x C*H*$_3$, Si*t*Bu); 3.69 – 3.73 (m, 1H); 3.78 (s, 3H, OC*H*$_3$); 3.80 (s, 3H, OC*H*$_3$); 3.81 (s, 3H, OC*H*$_3$); 3.89 (brs, 2H); 4.02 (t, *J* = 9.3 Hz, 1H); 4.12 (q, *J* = 7.1 Hz, 1H); 4.55 – 4.65 (m, 2H); 4.69 – 4.78 (m, 2H); 4.81 – 4.88 (m, 2H); 6.55 (d, *J* = 3.5 Hz, 1H, Ar); 6.74 – 6.80 (m, 2H, Ar); 6.82 – 6.90 (m, 4H); 7.04 – 7.10 (m, 2H); 7.23 – 7.30 (m, 4H); 7.30 – 7.44 (m, 6H); 7.65 – 7.70 (m, 3H); 8.55 (s, 1H, N*H*). ^{13}CNMR (100 MHz, CDCl$_3$): δ = 19.4 (*C*); 27.0, 31.1, 55.4 (*C*H$_3$); 62.6, 72.5 (*C*H$_2$); 74.5 (*C*H); 75.1, 75.6 (*C*H$_2$); 76.7, 79.8, 81.3 (*C*H); 91.6 (*C*Cl$_3$); 94.5, 113.9, 114.0, 127.7, 127.8, 129.3, 129.8, 130.0 (*C*H); 130.4, 130.5, 131.0, 133.3, 133.7 (*C*); 135.8, 136.0 (*C*H); 159.3, 159.4 (*C*); 161.5 (*C*=NH). ESI-MS: m/z calculated for $C_{48}H_{54}Cl_3NNaO_9Si$ [M+Na]$^+$: 944.3; found: 946.0.
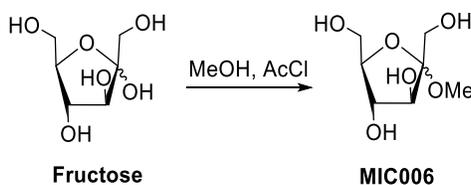
Fructose ring

Figure 58. Synthesis of MIC006 from fructose.

Acetyl chloride (17.835 mL, 249.9 mmol) was added dropwise to a suspension of fructose (20 g, 111.1 mmol) in MeOH (500 mL). The reaction mixture was stirred for 2 h until all the fructose was consumed (checked by thin layer chromatorgraphy: Rf (fructose) = 0.30; Rf (methyl-fructofuranoside) = 0.46 (chloroform/AcOH/H$_2$O 3.0:3.5:0.5 (*v/v*)). The reaction was quenched by the addition of triethylamine (34.9 mL, 249.9 mmol) and the solvent was evaporated under reduced pressure to afford **MIC006**. ESI-MS: m/z calculated for $C_7H_{14}ClO_6$ [M+Cl]$^-$: 229.0; found: 228.8. Note: the triethyl ammonium chloride was filtered off before the next step.
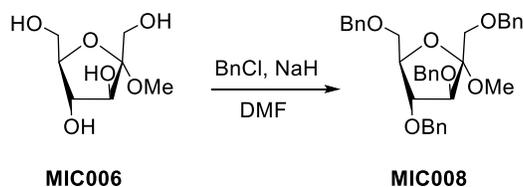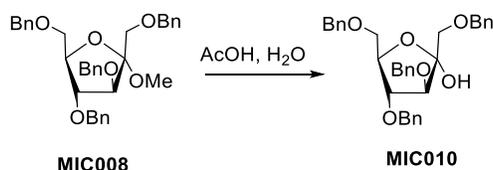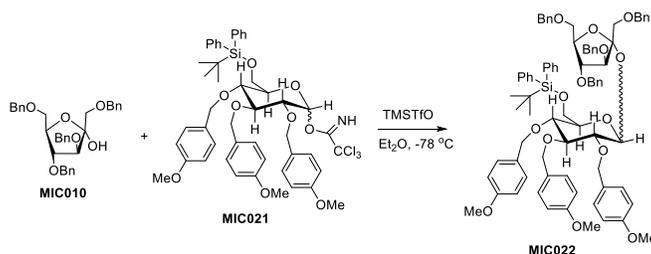

Figure 59. Synthesis of MIC008 from MIC006.

MIC006 (21.6 g, 119.6 mmol) was suspended in dried DMF at 0 °C. NaH from 60 % dispersion in mineral oil (23.0 g, 959.6 mmol) was added over a period of 40 minutes followed by the addition of benzyl chloride (110.4 mL, 959.7 mmol). The reaction mixture was allowed to reach room temperature and stirred overnight. After the reaction was completed, the mixture was cooled in a water bath and quenched with MeOH. The organic layer was extracted four times with DCM and washed with brine (3 x 100 mL) and H$_2$O (3 x 100 mL). The organic phase was dried over Na$_2$SO$_4$, concentrated under reduced pressure and purified by column chromatography (*n*-hexane/EtOAc 20:1 to 9:1 (*v/v*)) to afford **MIC008** (8.5 g, 14%). Rf = 0.22 (*n*-hexane/EtOAc 9:1 (*v/v*)). For the major diastereoisomer: ^{1}HNMR (400 MHz, CDCl$_3$): δ = 3.28 (s, 3H, OC*H*$_3$); 3.50 – 3.55 (m, 2H); 3.62 (m, 1H); 4.01 – 4.15 (m, 2H); 4.27 – 4.81 (m, 10H); 7.14 – 7.41 (m, 20H, Ar). ^{13}CNMR (100 MHz, CDCl$_3$): δ = 48.7 (*C*H$_3$); 65.9, 70.4, 72.0, 72.6, 73.4, 73.7 (*C*H$_2$); 80.9, 84.7, 87.1 (*C*H); 108.3 (*C*); 127.7, 127.8, 127.9, 128.0, 128.1, 128.3, 128.4, 128.5 (*C*H); 138.0, 138.1, 138.2, 138.3 (*C*). ESI-MS: m/z calculated for $C_{35}H_{38}NaO_6$ [M+Na]$^+$: 577.3; found: 576.8.
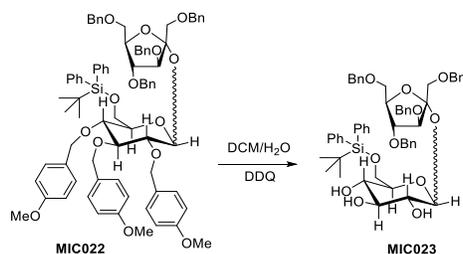
Figure 60. Synthesis of MIC010 from MIC008.

MIC008 (1.5 g, 2.7 mmol) was suspended on a mixture of acetic acid (AcOH)/H_2O 6:1 (*v/v*) and heated on a steam bath for 2 h. After cooled, the mixture was diluted with H_2O (80 mL) and extracted with DCM (4 x 15 mL). The organic phase was washed with $NaHCO_3$ (10 mL) and H_2O (10 mL). The organic phase was dried over Na_2SO_4 and concentrated under reduced pressure to afford **MIC010** as a thick yellow syrup (1.4 g, 96%). For the major diastereoisomer: ^{1}HNMR (400 MHz, $CDCl_3$): δ = 3.50 – 3.63 (m, 2H); 4.04 (q, *J* = 4.9 Hz, 1H, C*H*-5); 4.12 (t, *J* = 4.9 Hz, 1H; C*H*-4); 4.17 (d, *J* = 4.9 Hz, 1H, C*H*-3); 4.39 – 4.66 (m, 10H), 6.99 – 7.43 (m, 20H, Ar). ^{13}CNMR (100 MHz, $CDCl_3$): δ = 70.7, 72.0, 72.2, 72.8, 73.6, 73.7 (*C*H$_2$); 80.1, 83.5, 83.8, (*C*H), 102.6 (*C*); 127.8, 127.9, 128.0, 128.2, 128.5, 128.6 (*C*H); 137.7, 138.0, 138.1 (*C*). ESI-MS: m/z calculated for $C_{34}H_{36}NaO_6$ [M+Na]$^+$: 563.2; found: 563.9.

Acylated sucrose

Figure 61. Synthesis of MIC022 from MIC021.

MIC021 (1.9 g, 2.4 mmol) and **MIC010** (1.4 g, 2.7 mmol) were dissolved in dry Et_2O (30 mL) in an acetone/dry ice bath at – 78 °C. Trimethylsilyltriflate (TMSTfO) (0.2 mL, 1.0) mmol was added and the reaction mixture was stirred for 3 h and after allowed to reach room temperature. The reaction mixture was washed with saturated Na_2CO_3 (10 mL) and H_2O (10 mL). The organic phase was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by column chromatography (*n*-hexane/EtOAc 10:1 to 6:1 (*v/v*)) to afford **MIC022** (1.4 g, 84%). Rf = 0.29 (*n*-hexane/EtOAc 7:1 (*v/v*)). For the major diastereoisomer: ^{1}HNMR (400 MHz, $CDCl_3$): δ = 1.05 (s, 9H, 3 x C*H*$_3$, Si*t*Bu); 3.52 – 3.59 (m, 3H); 3.68 – 3.73 (m, 3H); 3.76 (s, 3H, OC*H*$_3$); 3.81 (s, 3H, OC*H*$_3$); 3.84 (s, 3H, OC*H*$_3$); 3.88 – 3.99 (m, 3H); 4.04 – 4.25 (m, 4H); 4.29 – 4.89 (m, 15H); 6.80 – 6.93 (m, 8H, Ar); 7.19 – 7.47 (m, 31H, Ar); 7.70 – 7.76 (m, 3H, Ar). ^{13}CNMR (100 MHz, $CDCl_3$): δ = 19.4 (*C*); 27.0 (*C*H$_3$); 55.4 (*C*H$_3$); 62.3, 69.3, 70.4, 71.5 (*C*H$_2$); 72.3 (*C*H); 72.8, 72.9, 73.0, 73.4, 75.0, 75.5 (*C*H$_2$); 77.4, 80.3, 80.5, 82.0, 84.8, 87.7, 90.0 (CH); 108.5 (*C*); 113.7, 113.9, 127.0, 127.5, 127.6, 127.8, 128.1, 128.3, 128.4, 129.2, 129.4, 129.7, 129.9 (*C*H); 131.0, 131.4, 133.5, 134.0 (*C*); 135.8, 136.2 (*C*H); 138.2, 138.3, 138.4, 138.5, 159.0, 159.3 (*C*). ESI-MS: m/z calculated for $C_{80}H_{88}NaO_{14}Si$ [M+Na]$^+$: 1323.6; found: 1323.8.

Figure 62. Synthesis of MIC023 from MIC022.

MIC022 (1.2 g, 1.0 mmol) was suspended in a mixture DCM/H_2O 20:1 (*v/v*) (12 mL) and 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) (0.2 g, 0.7 mmol) was added at room temperature. After the reaction was completed, a saturated aqueous solution of $NaHCO_3$ (5 mL) was added and the mixture was extracted with DCM (3 x 50 mL). The organic phase was washed with brine (20 mL), dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified by column chromatography (*n*-hexane/ EtOAc 2:1 to 1:1 (*v/v*)) to yield two different deprotected diastereomers: **MIC023-F1** with Rf_1 = 0.22 (240 mg) and **MIC023-F2** with Rf_2 = 0.11 (200 mg) (*n*-hexane/EtOAc 2:1 (*v/v*)). Total yield = 61 %. For **MIC023-F1**: ^{1}HNMR (400 MHz, CD_3OD): δ = 1.05 (s, 9H, 3 x C*H*$_3$, Si*t*Bu); 3.34 – 3.45 (m, 2H); 3.53 – 3.62 (m, 3H); 3.65 (s, 2H, C*H*$_2$); 3.73 (dd, *J* = 10.0/9.0 Hz, 1H); 3.81 – 3.87 (m, 2H); 3.92 (d, *J* = 2.3 Hz, 1H); 4.01 – 4.08 (m, 2H); 4.16 (d, *J* = 11.7, 1H, C*H*$_2$); 4.35 (td, *J* = 6.6/2.3 Hz, 1H); 4.40 (s, 2H, C*H*$_2$); 4.45 – 4.58 (m, 4H); 5.39 (d, *J* = 4.0 Hz, 1H); 6.94 – 6.98 (m, 2H, Ar); 7.07 – 7.15 (m, 5H, Ar); 7.20 – 7.26 (m, 3H, Ar); 7.26 – 7.37 (m, 9H, Ar); 7.37 – 7.48 (m, 7H, Ar); 7.68 – 7.77 (m, 4H, Ar). ^{13}CNMR (100 MHz, CD_3OD): δ = 20.2 (*C*); 27.4 (*C*H$_3$); 63.5, 69.8 (*C*H$_2$); 70.4 (*C*H); 71.3, 72.8, 73.1 (*C*H$_2$); 74.0, 74.2 (*C*H); 74.2 (*C*H$_2$); 76.5, 84.3, 84.8, 85.4, 92.4 (*C*H); 110.5 (*C*); 128.3, 128.4, 128.6, 128.7, 128.8, 128.9, 129.0, 129.2, 129.4, 129.7, 130.7, 130.8 (*C*H); 134.9, 135.1 (*C*); 136.8, 137.3 (*C*H); 138.4, 139.0, 139.1, 139.4 (*C*). For **MIC023-F2**: ^{1}HNMR (400 MHz, CD_3OD): δ = 1.03 (s, 9H, 3 x C*H*$_3$, Si*t*Bu); 3.30 (dd, *J* = 9.3/3.8 Hz, 1H); 3.51 – 3.76 (m, 8H); 3.80 – 3.87 (m, 2H); 3.96 – 4.01 (m, 1H); 4.10 (t, *J* = 7.2 Hz, 1H); 4.28 – 4.62 (m, 9H); 4.68 (d, *J* = 11.6 Hz, 1H, C*H*$_2$); 5.51 (d, *J* = 3.8 Hz, 1H); 7.08 – 7.12 (m, 2H, Ar); 7.15 – 7.27 (m, 13H, Ar); 7.31 – 7.45 (m, 11H, Ar); 7.69 – 7.77 (m, 4H, Ar). ^{13}CNMR (100 MHz, CD_3OD): δ = δ = 20.2 (*C*); 27.4 (*C*H$_3$); 64.1 (*C*H$_2$); 70.9 (*C*H); 72.0, 72.8, 73.4 (*C*H$_2$); 73.5, 74.0 (*C*H); 74.1, 74.6 (*C*H$_2$); 75.7, 80.8, 83.5, 85.3, 93.5 (*C*H); 105.5 (*C*); 128.6, 128.7, 128.8, 128.9, 129.2, 129.3, 129.4, 129.5, 130.7, 130.8 (*C*H); 134.9, 135.0 (*C*); 136.8, 137.0 (*C*H); 139.2, 139.3, 139.4 (*C*). ESI-MS: m/z calculated for $C_{56}H_{64}NaO_{11}Si$ [M+Na]$^+$: 963.4; found: 963.3.
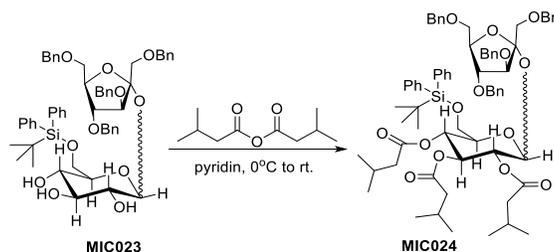
Figure 63. Synthesis of MIC024 from MIC023.

MIC023-F1 (0.24 g, 0.25 mmol) was suspended in pyridine (6 mL) at 0 °C and isovaleric anhydride (0.30 mL, 1.5 mmol) was added. The reaction mixture was stirred for 5 h at 70 °C. Then, the reaction mixture was diluted with H_2O (10 mL) and extracted with EtOAc (3 x 20 mL). The organic phase was washed with brine (2 x 20 mL), dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by column chromatography (*n*-hexane/EtOAc 30:1 (*v/v*)) to afford **MIC024-F1** (200 mg, 65%). Rf = 0.29 (*n*-hexane/EtOAc 30:1 (*v/v*)). ^{1}HNMR (400 MHz, $CDCl_3$): δ = 0.72 (d, *J* = 6.2 Hz, 3H, C*H*$_3$); 0.76 (d, *J* = 6.3 Hz, 3H, C*H*$_3$); 0.85 (d, *J* = 6.6 Hz, 3H, C*H*$_3$); 0.87 (d, *J* = 6.6 Hz, 3H, C*H*$_3$); 0.92 (d, *J* = 6.6 Hz, 6H, 2 x C*H*$_3$); 1.02 (s, 9H, 3 x C*H*$_3$, Si*t*Bu); 1.86 – 2.10 (m, 9H); 3.30 (dd, *J* = 11.8/2.5 Hz, 1H, C*H*$_2$); 3.42 (dd, *J* = 11.8/2.0 Hz, 1H, C*H*$_2$); 3.51 – 3.60 (m, 2H); 3.65 – 3.72 (m, 2H); 3.89 (dd, *J* = 5.3/2.0 Hz; 1H); 3.99 (d, *J* = 11.8 Hz, 1H, C*H*$_2$); 4.05 (q, *J* = 5.2 Hz, 1H, C*H*); 4.13 – 4.22 (m, 3H); 4.34 (d, *J* = 11.5 Hz, 1H, C*H*$_2$); 4.51 – 4.55 (m, 3H); 4.58 (d, *J* = 11.8 Hz, 1H, C*H*$_2$); 4.64 (d, *J* = 11.8 Hz, 1H, C*H*$_2$); 4.96 (dd, *J* = 10.2/3.8 Hz, 1H, C*H*-2); 5.42 (t, *J* = 10.2 Hz, 1H, C*H*-4); 5.62 (t, *J* = 10.2 Hz, 1H, C*H*-3); 5.65 (d, *J* = 3.8 Hz, 1H, C*H*-1) ; 6.87 – 6.93 (m, 2H, Ar); 7.02 – 7.11 (m, 3H, Ar); 7.21 – 7.31 (m, 13H, Ar); 7.32 – 7.41 (m, 8H, Ar); 7.53 – 7.60 (m, 2H, Ar); 7.65 – 7.71 (m, 2H, Ar). ^{13}CNMR (100 MHz, $CDCl_3$): δ = 19.4 (*C*); 22.4, 22.5, 22.6 (*C*H$_3$); 25.3, 25.4, 25.5, 26.8 (*C*H); 31.1 (*C*H$_3$); 42.9, 43.3, 61.5 (*C*H$_2$); 68.1 (*C*H); 69.4, 70.3 (*C*H$_2$); 70.5, 70.6 (*C*H); 72.1, 72.8, 73.0, 73.4 (*C*H$_2$); 82.1, 85.0, 86.7, 89.2 (*C*H); 109.2 (*C*); 127.1, 127.2, 127.7, 127.8, 128.0, 128.1, 128.4, 128.5, 129.6, 129.7 (*C*H); 133.3, 133.7 (*C*); 135.8,

136.2 (C*H*); 137.9, 138.0, 138.1, 138.3 (*C*); 171.3, 172.4, 172.6 (*C*=O). ESI-MS: m/z calculated for $C_{71}H_{88}NaO_{14}Si$ [M+Na]$^+$: 1215.6; found: 1215.5.

Table 8. Complete MIC024 characterization.

MIC024_F1 (apparently α,α- configuration)	**MIC024_F2 (apparently α,β configuration)**
MIC024_F1 NOE contacts:	**MIC024_F2 NOE contacts:**

MIC023-F2 (0.20 g, 0.21 mmol) was suspended in pyridine (6 mL) at 0 °C and isovaleric anhydride (0.25 mL, 1.3 mmol) was added. The reaction mixture was stirred for 5 h at 70 °C. Then, the reaction mixture was diluted with H_2O (10 mL) and extracted with EtOAc (3 x 20 mL). The organic phase was washed with brine (2 x 20 mL), dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by column chromatography (*n*-hexane/EtOAc 30:1 (*v/v*)) to afford **MIC024-F2** (155 mg, 61%). Rf = 0.32 (*n*-hexane/EtOAc 30:1 (*v/v*)). ^{1}HNMR (400 MHz, CDCl$_3$): δ = 0.84 – 0.89 (m, 12H, 4 x C*H*$_3$); 0.92 (d, *J* = 6.6 Hz, 6H, 2 x C*H*$_3$); 1.01 (s, 9H, 3 x C*H*$_3$, Si*t*Bu); 1.86 – 2.10 (m, 9H); 3.26 (dd, *J* = 11.8/3.0 Hz, 1H, C*H*$_2$); 3.40 – 3.46 (m, 2H); 3.62 (d, *J* = 10.8 Hz, 1H, C*H*$_2$); 3.67 (d, *J* = 4.8 Hz, 2H); 4.01 – 4.12 (m, 3H); 4.28 (d, *J* = 11.5 Hz, 1H, C*H*$_2$); 4.37 (d, *J* = 7.4 Hz, 1H); 4.41 – 4.47 (m, 3H); 4.51 – 4.60 (m, 5H); 4.92 (dd, *J* = 10.1/3.7 Hz, 1H, C*H*-2); 5.37 (t, *J* = 10.1 Hz, 1H, C*H*-4); 5.54 (t, *J* = 10.1 Hz, 1H, C*H*-3); 5.78 (d, *J* = 3.7 Hz, 1H, C*H*-1) 7.08 – 7.13 (m, 2H, Ar); 7.14 – 7.20 (m, 5H, Ar); 7.20 – 7.35 (m, 19H, Ar); 7.56 – 7.60 (m, 2H, Ar); 7.64 – 7.67 (m, 2H, Ar). ^{13}CNMR (100 MHz, CDCl$_3$): δ = 19.4 (*C*); 22.5, 22.6 (*C*H$_3$); 25.5, 26.9 (*C*H); 43.2, 43.3, 43.4, 61.9 (*C*H$_2$); 68.0, 70.2, 70.4, 70.7 (*C*H); 70.9, 71.8, 72.8, 73.3, 73.5, 73.7 (*C*H$_2$); 79.5, 82.4, 84.0, 89.2 (*C*H); 104.7 (*C*); 127.6, 127.7, 127.8, 127.9, 128.0, 128.3, 128.4, 128.5, 129.6, 129.7 (*C*H); 133.4, 133.6 (*C*); 135.8, 136.0 (*C*H); 137.9, 138.0, 138.2, 138.3 (*C*); 171.2, 172.1, 172.4 (*C*=O). ESI-MS: m/z calculated for $C_{71}H_{88}NaO_{14}Si$ [M+Na]$^+$: 1215.6; found: 1215.5.
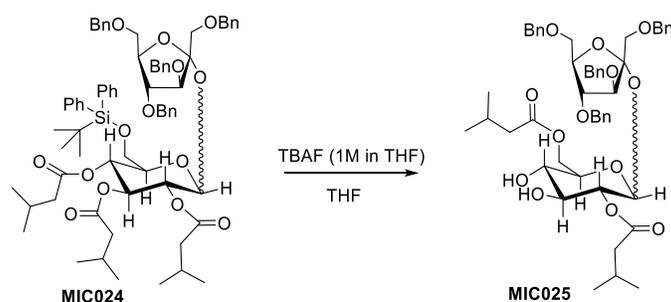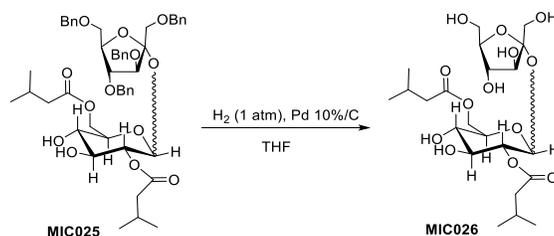
Figure 64.Synthesis of MIC025 from MIC024.

MIC024-F1 (200 mg, 0.17 mmol) was dissolved in 3 mL of THF, tetrabutylammonium fluoride (TBAF) (25 mg, 0.09 mmol _90 µL from a 1 M stock solution in THF) was added and the mixture was stirred for 3 h at room temperature. Then, an excess of $CaCO_3$, DOWEX 50WX8-400 and MeOH and the suspension was stirred at rt for 1 h. All insoluble materials were removed by filtration through a pad of celite, and the filter cake was washed with MeOH thoroughly. Combined filtrates were evaporated under vacuum to afford the crude product that was further purified by column chromatography (*n*-hexane/EtOAc) to afford **MIC025-F1** (40 mg, 28%). ^{1}HNMR (400 MHz, CDCl$_3$): δ = 0.83 (d, *J* = 7.1 Hz, 2H, C*H*$_3$); 0.85 (d, *J* = 7.0 Hz, 3H, C*H*$_3$); 0.95 (d, *J* = 6.6 Hz, 3H, C*H*$_3$); 0.96 (d, *J* = 6.5 Hz, 3H, C*H*$_3$); 0.99 (d, *J* = 6.7 Hz, 1H, C*H*$_3$); 2.04 (m, 2H); 2.13 (d, *J* = 7.7 Hz, 2H, C*H*$_2$); 2.24 (d, *J* = 7.3 Hz, 2H, C*H*$_2$); 3.36 (t, *J* = 9.5 Hz, 1H, C*H*-4); 3.55 (t, *J* = 5.8 Hz, 1H, C*H*-4´); 3.70 – 3.81 (m, 2H); 3.84 – 3.93 (m 2H); 4.05 (d, *J* = 9.5 Hz, 1H, C*H*$_2$); 4.07 – 4.11 (m, 1H); 4.13 (d, *J* = 2.2 Hz, 1H, C*H*-3´); 4.19 (q, *J* = 5.3 Hz, 1H, C*H*-5´), ; 4.41 (dd, *J* = 12.1/ 2.4 Hz, 1H, C*H*$_2$-6); 4.46 – 4.54 (m, 4H); 4.56 (dd, J = 12.1 Hz, 1H, C*H*$_2$); 4.62 (dd, *J* = 11.8 Hz, 1H, C*H*$_2$) ; 4.76 (dd, *J* = 10.1/3.7 Hz, 1H, C*H*-2); 5.51 (d, *J* = 3.7 Hz, 1H, C*H*-1); 7.20 – 7.34 (m, 20H, Ar). ^{13}CNMR (100 MHz, CDCl$_3$): δ = 22.4, 22.6 (C*H*$_3$); 25.6, 25.8 (*C*H); 43.1, 43.3, 62.8, 69.3, 70.3 (*C*H$_2$); 70.6, 71.3 (*C*H); 71.9 (*C*H$_2$); 72.5 (*C*H); 73.5 (*C*H$_2$); 82.2, 84.8, 87.1, 89.7 (*C*H); 109.1 (C); 127.3, 127.7, 127.8, 128.4, 128.5 (*C*H); 137.8, 138.0 (*C*)*; 172.5, 174.3 (C=O)*. * from HMBC spectrum. ESI-MS: m/z calculated for C$_{50}$H$_{62}$FO$_{13}$ [M+F]$^-$: 889.4; found: 889.7.

MIC024-F2 (155 mg, 0.13 mmol) was dissolved in 3 mL of THF, tetrabutylammonium fluoride (TBAF) (19 mg, 0.07 mmol _70 µL from a 1 M stock solution in THF) was added and the mixture was stirred for 3 h at room temperature. Then, an excess of $CaCO_3$, DOWEX 50WX8-400 and MeOH and the suspension was stirred at rt for 1 h. All insoluble materials were removed by filtration through a pad of celite, and the filter cake was washed with MeOH thoroughly. Combined filtrates were evaporated under vacuum to afford the crude product that was further purified by column chromatography (*n*-hexane/EtOAc) to afford **MIC025-F2** (90 mg, 79%). ^{1}HNMR (400 MHz, CDCl$_3$): δ = 0.95 (d, *J* = 6.6 Hz, 6H, 2 x C*H*$_3$); 0.99 (d, *J* = 6.6 Hz, 6H, 2 x C*H*$_3$); 2.06 – 2.13 (m, 2H); 2.21 – 2.25 (m, 4H); 3.29 (dd, *J* = 10.1/8.9 Hz, 1H, C*H*-4); 3.41 (d, *J* = 10.8 Hz, 1H, C*H*$_2$); 3.57 – 3.67 (m, 4H); 3.83 (dd, *J* = 12.5/2.2 Hz, 1H, C*H*$_2$-6); 3.93 (dd, *J* = 10.1/9.0 Hz, 1H, C*H*$_2$); 3.98 – 4.07 (m, 2H); 4.15 (t, *J* = 7.4 Hz, 1H, C*H*-4´); 4.28 (dd, *J* = 12.5/3 Hz, 1H, C*H*$_2$-6); 4.38 (d, *J* = 7.4 Hz, 1H, CH-3´); 4.46 (d, *J* = 12.0 Hz, 1H, C*H*$_2$); 4.49 – 4.61 (m, 6H); 4.64 (m, 1H, C*H*-2); 5.62 (d, *J* = 3.7 Hz, 1H, C*H*-1); 7.20 – 7.36 (m, 20H, Ar). ^{13}CNMR (100 MHz, CDCl$_3$): δ = 22.5, 22.6 (C*H*$_3$); 25.7, 25.8 (*C*H); 43.3, 42.7, 62.7 (*C*H$_2$); 70.2, 70.5 (CH); 70.7(*C*H$_2$); 71.3 (*C*H); 71.3, 72.6, 72.8, 73.4, 73.7 (*C*H$_2$); 77.4, 79.5, 82.1, 83.9, 89.8 (*C*H); 104.7 (*C*); 127.7, 127.9, 128.0, 128.2, 128.5, 128.6, 129.8 (*C*H); 134.9, 137.9, 138.1, 138.3 (*C*); 172.8, 174.5 (*C*=O). ESI-MS: m/z calculated for C$_{50}$H$_{62}$FO$_{13}$ [M+F]$^-$: 889.4; found: 889.5.

Figure 65. Synthesis of MIC026 from MIC025.

  MIC025-F1 (40 mg, 0.05 mmol) was dissolved in THF (2 mL) and placed in a three necked flask, then catalytic amounts of Pd 10%/C were added and the reaction was carried out under hydrogen atmosphere (1 atm) for 7 days. Then, the Pd/C was filtered off and the solvent concentrated under reduced pressure to afford **MIC026-F1** (15mg) ^{1}HNMR (400 MHz, CDCl$_3$): δ = 0.96 (d, J = 6.6 Hz, 12H, 4 x CH_3); 2.03 – 2.15 (m, 2H); 2.18 – 2.35 (m, 4H); 3.42 (t, J = 9.6 Hz, 1H, CH-4); 3.63 – 3.71 (m, 2H); 3.73 – 3.82 (m, 3H); 3.83 – 3.94 (m, 4H); 3.96 – 4.09 (m, 3H); 4.10 – 4.15 (m, 1H, CH-5); 4.30 (dd, J = 12.2/5.4 Hz, 1H, CH_2-6); 4.39 (d, J = 11.7 Hz, 1H); 4.87 (dd, J = 10.1/3.8 Hz, CH-2); 5.45 (d, J = 3.8 Hz, 1H, CH-1); ^{13}CNMR (100 MHz, CDCl$_3$): δ = 22.5, 22.8 (CH_3); 25.7, 25.8 (CH); 43.0, 43.2, 61.3, 61.8, 62.9, 63.5 (CH_2); 70.7, 70.9, 71.6, 72.1, 78.5, 80.4, 87.3, 89.6 (CH); 111.0 (C); 172.5, 174.0 (C=O). ESI-MS: m/z calculated for C$_{22}$H$_{38}$NaO$_{13}$ [M+Na]$^+$: 533.2; found: 533.3

MIC025-F2 (90 mg, 0.1 mmol) was dissolved in THF (2 mL) and placed in a three necked flask, then catalytic amounts of Pd 10%/C were added and the reaction was carried out under hydrogen atmosphere (1 atm) for 7 days. Then, the Pd/C was filtered off and the solvent concentrated under reduced pressure to afford
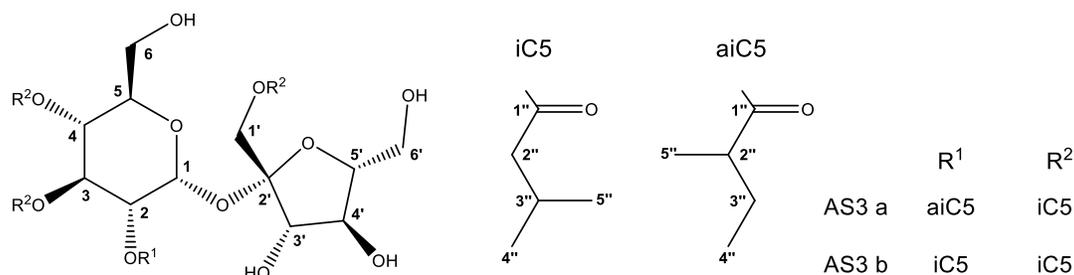
MIC026-F2 ^{1}HNMR (61mg, 400 MHz, CDCl$_3$): δ = could not be determined due to high impurity. ESI-MS: m/z calculated for C$_{22}$H$_{38}$NaO$_{13}$ [M+Na]$^+$: 533.2; found: 533.3

Appendix 6. RNAseq result based on average TPM using three replicates.

Gene ID	ITAG2.4 description	AVT014 24 Leaf	VI03046 2 Leaf	AVT01424 Trichome	VI030462 Trichome
AACS					
Solyc01g066310.2	Acyl-CoA dehydrogenase	30.463	7.705	38.582	11.255
Solyc02g081370.2	Acyl-CoA synthetase	0.031	2.494	0.028	1.402
Solyc02g082870.2	Acyl-CoA synthetase	3.369	0.512	217.495	51.987
Solyc02g082880.2	Acyl-CoA synthetase	0	0.986	0.008	33.316
Solyc06g073560.2	Isovaleryl-CoA dehydrogenase	9.788	1.575	102.998	66.944
Solyc07g043630.1	Acyl-CoA synthetase	0.023	0.6	0.151	18.573
Solyc07g043640.2	Acyl-CoA synthetase	0.014	0.027	0.184	1.487
Solyc07g043660.2	Acyl-CoA synthetase	0.139	0.393	1.779	10.892
Solyc08g075810.2	Acyl-CoA synthetase	4.542	9.269	13.5	25.164
Solyc08g078400.2	Acyl-CoA oxidase	2.45	0.329	119.03	10.966
Solyc10g076600.1	Acyl-CoA dehydrogenase	16.545	42.805	24.703	100.682
Solyc10g085200.1	Acyl-CoA dehydrogenase	27.548	86.056	71.967	130.555
Solyc12g044300.1	Acyl-CoA synthetase	76.527	81.992	167.216	304.202
Solyc12g099360.1	Acyl-CoA synthetase	0.199	0.893	0.044	1.227
Genes in QTL Chr.5					
Solyc01g101060.2	S-adenosylmethionine synthase	242.537	364.259	230.509	302.032
Solyc01g080380.2	S-adenosylmethionine decarboxylase proenzyme	0.551	0.714	0.06	0.049
Solyc02g089610.1	S-adenosylmethionine decarboxylase proenzyme	118.558	230.501	104.157	91.927
Solyc05g010420.1	S-adenosylmethionine decarboxylase proenzyme	1817.029	418.773	4383.6	169.842
Solyc09g008280.1	S-adenosylmethionine synthase	342.063	400.236	341.237	529.371
Solyc10g083970.1	S-adenosylmethionine synthase	113.671	78.344	95.404	109.311
Solyc12g099000.1	S-adenosylmethionine synthase	39.19	15.953	43.644	12.467
Genes in QTL Chr.11					
Solyc11g010420.1	Ovarian cancer-associated gene 2 protein homolog	6.525	22.547	23.564	199.375
Solyc11g010430.1	Ovarian cancer-associated gene 2 protein homolog	10.087	37.103	78.06	127.039
Solyc11g067270.1	Acyltransferase-like protein	0.188	14.017	4.762	431.48
Solyc11g067290.1	Acyltransferase-like protein	0	0.195	0	7.756
Solyc11g071610.1	Xanthine dehydrogenase	1.606	2.501	6.249	102.545
Solyc11g071620.1	Aldehyde oxidase	10.397	7.628	36.172	292.109
Solyc11g071800.1	Strictosidine synthase family protein	33.855	45.959	111.268	205.299
Flavonoid					
Solyc05g053170.2	Chalcone synthase-A	0	0.169	0	0.8
Solyc05g053550.2	*SlCHS1*	69.228	90.358	838.423	590.327
Solyc05g010320.2	*SlCHI1*	68.828	67.324	462.492	405.124
Solyc05g052240.2	*SlCHIL*	76.756	18.096	423.373	686.261
Lipoxygenase					
Solyc01g006540.2	Tomlox C	40.639	27.12	3251.463	4070.217
Solyc01g006560.2	Tomlox F	136.427	157.867	47.068	11.267
Solyc03g122340.2	Tomlox D	9.503	15.704	4.892	5.739
Solyc12g011040.1	Lipoxygenase	0.807	0.868	0.251	13.813
Solyc03g093360.2	Wound stress protein	10.072	12.543	28.146	144.579
Solyc03g096460.2	Lipoxygenase homology domain-containing protein	44.769	17.453	376.703	367.44
Solyc01g009680.2	Lipoxygenase	0.085	1.641	0.034	0.503
Solyc01g099150.2	Lipoxygenase	4.987	4.216	101.8	130.853

Solyc01g099160.2	Lipoxygenase	0.626	6.443	12.675	307.636
Solyc01g099170.2	Lipoxygenase	0	0.112	0.013	5.73
Solyc01g099180.2	Lipoxygenase	6.785	24.031	1.161	18.217
Solyc01g099190.2	Tomlox B	0	0.306	0.013	15.171
Solyc01g099200.2	Lipoxygenase	0.005	0.351	0.384	0.628
Solyc09g055890.2	Lipoxygenase	1.453	1.385	0.126	0.092
Solyc09g055900.2	Lipoxygenase	2.915	2.892	0.222	0.239
Fatty acid desaturase					
Solyc01g009960.2	Stearoyl-acyl-carrier protein desaturase	0.162	32.443	7.791	2284.779
Solyc06g054670.1	Stearoyl-acyl carrier protein desaturase	0.223	1.075	0.741	148.5
Solyc06g059710.2	Stearoyl-acyl carrier protein desaturase	45.906	0.845	2506.525	65.539
Solyc06g059720.2	Stearoyl-acyl carrier protein desaturase	0	0.013	0.064	0.989
Solyc11g008680.1	Acyl- Fatty acid desaturase	5.593	5.849	11.251	33.138
Solyc01g006430.2	SlFAD2-1	293.181	195.935	3834.545	524.609
Solyc03g058430.1	SlFAD2-2	34.697	28.436	13.567	7.701
Solyc06g007140.2	Omega-3 fatty acid desaturase	24.339	1.329	194.527	2.8
Solyc05g050090.2	Delta-6 fatty acid desaturase	12.409	30.535	1.991	1.598
Solyc07g005510.2	SlFAD6	294.97	184.92	157.234	29.684
Terpene synthase					
Solyc01g101170.2	TPS 31 - Viridiflorene	0.706	0.11	47.938	0
Solyc01g105850.2	TPS1 - Pseudo gene	3.947	0.061	313.235	5.318
Solyc01g105870.2	TPS3 - Camphene, tricyclene	0	0.238	0.025	20.2
Solyc01g105880.2	TPS4 (MTS2) - βPhellandrene	0.006	0.756	0.072	33.714
Solyc01g105890.2	TPS5 (MTS1) - Linalool	6.201	12.976	160.934	909.29
Solyc01g105900.2	TPS6 - Pseudo gene	0.023	37.732	0.03	43.448
Solyc05g026590.1	TPS43 Pseudo gene	0.623	0.161	88.273	13.191
Solyc06g059910.2	TPS10 - αBisabolol	0.204	1.517	17.131	137.289
Solyc06g059930.2	TPS11 - Pseudo gene	13.862	14.952	919.703	1651.02
Solyc06g060180.1	TPS36 - cis-Muurola-3,5-diene	0	2.215	0.156	206.542
Solyc07g008690.2	TPS16 - δCadinene	4.54	12.425	131.004	1623.66
Solyc07g051940.2	TPS15 - Pseudo gene	6.474	23.384	130.383	2333.036
Solyc08g005640.2	TPS21 - Lycosantalene	11.529	6.217	756.635	981.755
Solyc09g092470.2	TPS14 - Alpha-humulene,E-beta-caryophyllene synthase	0.015	1.489	0.225	262.748
Solyc10g005390.2	TPS39 - Linalool, E-nerolidol	0.064	0.762	3.115	63.984
Solyc10g005410.2	TPs37- Linalool, E-nerolidol	3.99	1.022	233.974	123.312
Solyc12g006570.1	TPS17 - Valencene	12.802	240.935	847.733	217.815
Solyc12g019240.1	TPS29 - Pseudo gene	1.125	0.019	15.516	0.225
Solyc08g005710.2	TPS41 - Copalyl diphosphate	9.38	0.012	634.451	0.041
Solyc07g066670.2	TPS24 (KS) - ent-kaurene	7.121	2.57	6.472	1.52
Solyc08g005720.2	TPS18	6.033	3.16	417.065	433.054
Solyc08g005670.2	TPS19 - βMyrcene, βOcimene	11.013	4.813	663.479	649.116
Solyc07g052150.2	TPS52 - E-Nerolidol, α-Bisabolol	22.793	1.029	838.992	100.97
Solyc01g105860.2	TPS1 - Pseudo gene	4.117	0.019	362.008	0.222
Solyc01g105950.2	TPS22 - Pseudo gene	0	0.46	0.018	42.345
Solyc02g079910.1	TPS27 - αFarnesene	0.026	0	3.26	0.052
Solyc05g026600.2	TPS43 - Pseudo gene	0.654	0.045	23.567	0.362
Solyc07g008680.2	TPS16 - δCadinene	0.769	12.931	42.678	1231.715
Solyc07g052120.2	TPS51- E-Nerolidol, α-Bisabolol	0.057	0.722	0.504	71.426
Solyc07g052140.2	TPS52- E-Nerolidol, α-Bisabolol	69.362	0.07	1278.782	1.66
Solyc10g005420.1	TPS42 - Pseudo gene	0	0	0	7.839

AVT01424 vs VI030462 trichomes

Solyc09g007910.2	Phenylalanine ammonia-lyase	7.18	16.267	23.31	16.996
Solyc03g045020.2	Tryptophan decarboxylase 5	0.123	8.236	1.54	833.216
Solyc06g053400.2	2-isopropylmalate synthase 1	25.157	41.589	51.369	25.523
Solyc02g063490.2	Malate dehydrogenase	22.576	25.158	50.285	49.285
Solyc12g014180.1	Malate dehydrogenase	9.364	5.347	25.023	21.857
Solyc08g066260.2	Histidine decarboxylase	4.57	0.044	205.937	1.846
Solyc08g016770.2	Decarboxylase family protein	0.434	0	11.66	1.304
Solyc08g014230.2	2-isopropylmalate synthase 3	0.71	3.225	63.15	189.408
Solyc07g054280.1	Tryptophan decarboxylase 2	9.287	7.665	317.685	909.3
Solyc08g068670.2	Histidine decarboxylase 18	9.801	1.637	450.622	169.884
Solyc08g066220.2	Histidine decarboxylase 9	7.527	14.706	708.004	1813.929
Solyc04g049130.2	At5g03900 (Fragment)	24.831	43.539	16.313	26.77
Solyc05g051700.2	Iron-sulfur cluster insertion protein erpA	146.297	321.37	104.23	197.079
Solyc02g087360.2	Paired amphipathic helix protein SIN3	12.924	19.739	18.522	19.216
Solyc11g067020.1	Histone deacetylase 4	14.694	26.244	20.895	24.997
Solyc02g067570.2	Paired amphipathic helix protein SIN3	13.002	19.414	19.362	22.835
Solyc02g087390.1	Paired amphipathic helix protein SIN3	0	59.25	0.04	69.691
Solyc09g091440.2	Histone deacetylase 1	43.56	40.597	67.969	53.962
Solyc05g006540.2	Paired amphipathic helix protein SIN3	18.362	22.943	35.351	30.666
Solyc02g014470.2	Lipase-like protein	4.298	3.696	0.331	0.41
Solyc02g077100.2	Lipase-like protein	25.296	37.462	10.931	4.047
Solyc12g010910.1	Lipase (Fragment)	37.618	25.073	13.811	6.19
Solyc02g077030.2	Lipase-like	8.206	1.714	2.22	0.132
Solyc07g005430.2	Unknown Protein	34.864	33.912	32.725	31.395
Solyc12g036790.1	NADH-ubiquinone oxidoreductase subunit	261.011	219.394	241.582	323.187
Solyc03g025910.2	Unknown Protein	83.925	55.127	84.273	104.334
Solyc08g078860.2	Homology to unknown gene	100.421	110.122	101.362	184.281
Solyc07g065280.2	Os08g0431500 protein (Fragment)	86.01	96.838	89.907	140.011
Solyc04g080570.2	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase	45.643	38.307	50.695	65.052
Solyc11g005880.1	NADH-ubiquinone oxidoreductase kD subunit	112.918	84.474	124.45	186.262
Solyc08g080240.2	NADH ubiquinone dehydrogenase	81.273	65.936	90.429	134.084
Solyc03g096940.2	NADH ubiquinone oxidoreductase subunit	72.247	84.171	82.6	132.738
Solyc09g065830.2	NADH ubiquinone oxidoreductase subunit	77.059	88.069	88.628	111.272
Solyc11g068510.1	F1F0-ATPase inhibitor protein	151.979	147.793	176.613	203.655
Solyc06g075810.2	NADH dehydrogenase	102.187	94.339	121.254	145.484
Solyc05g013910.2	Unknown Protein	93.309	118.879	116.815	207.579
Solyc10g005230.2	Unknown Protein	76.626	62.301	98.196	139.116
Solyc01g110390.2	NADH dehydrogenase	84.354	78.433	108.898	148.859
Solyc01g102830.2	Unknown Protein	38.321	53.784	55.598	140.947

Appendix 7. NMR data of AS3a and AS3b (molar ratio ca. 1: 1 in CD3OD)

Pos.	^{1}H^a [ppm] m (J [Hz])		^{13}C^b [ppm]	
	AS3a	AS3b	AS3a	AS3b
1	5.702 d (3.8)	5.669 d (3.8)	90.5	90.5
2	4.888 dd (10.4; 3.8)	4.923 dd (10.4; 3.8)	71.9	71.7
3	5.558 dd (10.4; 10.3)	5.542 dd (10.4; 10.3)	71.4	71.4
4	5.102 dd (10.3; 9.5)	5.106 dd (10.3; 9.5)	69.9	69.9
5	4.206 m		71.8	
6	3.656 br d (12.2) / 3.535 dd (12.2; 4.6)		61.5	
1'	4.11 /4.05		64.2	
2'	---		104.2	104.2
3'	4.10		78.3	
4'	4.00		74.8	
5'	3.73		84.1	
6'	3.76 / 3.74		63.5	
iC5				
1'' @ 2	---		---	173.6
1'' @ 3	---		173.5	
1'' @ 4	---		173.2	
1'' @ 1'	---		173.9	
2''	2.30 – 2.10		44.1	
3''	2.10 – 1.96		26.9 – 26.5	
4''	0.98 – 0.91		22.8	
5''				
aiC5				
1''	---	---	177.4	---
2''	2.445 m	---	42.0	---
3''	1.651 m / 1.460 m	---	27.8	---
4''	0.884 t (7.5)	---	11.9	---
5''	1.141 d (6.7)	---	16.3	---
^{a} ^{1}H chemical shifts with only two decimal places are chemical shifts of HSQC correlation peaks; ^{b} ^{13}C chemical shifts are chemical shifts of HSQC or HMBC correlation peaks.				

^{1}H, ^{13}C, 2D (^{1}H,^{1}H gDQCOSY; ^{1}H,^{1}H zTOCSY;^{1}H,^{1}H ROESYAD; ^{1}H,^{13}C gHSQCAD; ^{1}H,^{13}C gHMBCAD) NMR spectra were measured with an Agilent VNMRS 600 instrument at 599.831 MHz (^{1}H) using standard CHEMPACK 8.1 pulse sequences implemented in the VNMRJ 4.2A spectrometer software. TOCSY mixing time: 80 ms; ROESY mixing time: 300 ms; HSQC optimized for ^{1}J$_{CH}$ = 146 Hz; HMBC optimized for ^{n}J$_{CH}$ = 8 Hz. All spectra were obtained with CD$_3$OD + 0.03% TMS as solvent at +25°C. Chemical shifts were referenced to internal TMS (^{1}H = 0 ppm; ^{13}C = 0 ppm).

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende wissenschaftliche Arbeit selbständig und ohne fremde Hilfe angefertigt habe. Ich erkläre, dass ich keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe und die den Werken wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Ich versichere weiterhin, dass ich mich erstmals mit dieser Arbeit um die Erlangung des Doktorgrades bewerbe. Diese Arbeit wurde an keiner anderen Fakultät oder Universität zur Begutachtung eingereicht.

Micha Gracianna Devi

Halle (Saale)

Acknowledgement

Curriculum vitae

First and last name	:	Micha Gracianna Devi
Email address	:	michagrcnndv7@gmail.com
Place, date of birth	:	Bogor, July 2nd 1991
Nationality	:	Indonesian

Academic education

2021	:	**Postdoctoral researcher at the Department of Plant Breeding Wageningen University and Research**
2017-2021	:	**Martin Luther University, Halle-Wittenberg** **Doctoral researcher at the Leibniz Institute for Plant Biochemistry (IPB), Halle** Department of cell and metabolic Biology GT and isoprenoid biosynthesis research group **Doctoral thesis:** "The genetic of white fly resistance in tomato"
2015-2017	:	**Wageningen University and Research** Master's in plant biotechnology Specialization molecular plant breeding **Major thesis:** "Functional characterization of candidate susceptibility genes for downy mildew in cucumbers (*Cucumis sativus L.*)" Supervisor: Henk Schouten & Jeroen Berg **Minor thesis:** "Exploring allelic variation of candidate genes in sugar metabolism and phenolic derived volatiles biosynthesis of tomato (*Solanum lycopersicum*)" Supervisor: Arnold Bovy
2009-2014	:	**Bogor Agriculture University (University IPB)** Faculty of agriculture Department of agronomy and horticulture **Thesis:** "Globular embryo induction of sugar palm (*Arenga pinnata* (Wurmb))" Supervisor: Prof. Agus Purwito & Dr. Ali Husni
2014 (1 year exchange)	:	**University of Santiago de Compostela** Department of Crop Production and Projects of Engineering **Thesis:** "Analysis of northwestern Spain apples using microsatellite markers, Universidad Santiago de Compostela" Superviosor: Prof. Santiago Pereira-Lorenzo & Dr. Ana Ramos-Cabrer **Internship:** Micropropagation Laboratory Supervisor: Fernández Lorenzo Juan Luis
2011 (6 months exchange)	:	**Kasetsart University** Faculty of Tropical Agriculture

Conferences and symposium

2019	:	Terpenoid Conference 2019, IPB Halle
2019	:	Plant Science Student Symposium (organizer)
2018	:	German Plant Breeding Conference, IPK
2018	:	Solanceae Conference 2018, Thailand
2017, 2018	:	Plant Science Student Symposium, IPB/IPK

Courses and workshops during doctoral research

2020	:	Legal Basics for Plant Scientists (Plant 2030 Academy)
2019	:	Communication & presentation in academic context (Golin Wissenschafts)
2018	:	Teamwork & leadership in academia (Golin Wissenschafts)
	:	Time- and selfmanagement for doctoral candidate (INGRA MLU)
	:	Conflict management in academi (Golin Wissenschafts)
	:	Multivariat statistics with R (INGRA MLU)
	:	Introduction to genome-wide association studies (deNBI)
	:	Introduction to RNA-seq data analysis with Galaxy
	:	Science 2.0 – How to accelerate your research (INGRA MLU)
	:	Statistics for natural scientis (Sum of Square)
2017	:	Processing and analysis of scientific images (INGRA MLU)
	:	ImageJ/Fiji Workshop (MLU)
	:	Joint CIBI user meeting on OpenMS, MetFrag and SeqAn (deNBI)
	:	Current challenges in ecometabolomics (iDiv)

Language

Bahasa Indonesia	:	Mother tongue
English	:	C1
Spanish	:	B1
German	:	B1
Dutch	:	A2