Sucrase-transglucosidases and the interplay between detoxification and osmoregulation in phloem feeding insects

Michael Lachlan Alan Einar Easson

MAX PLANCK INSTITUTE FOR CHEMICAL ECOLOGY

seit 1558

MAX-PLANCK-GESELLSCHAFT

Sucrase-transglucosidases and the interplay between detoxification and osmoregulation in phloem feeding insects

Dissertation

To Fulfill the
Requirements for the Degree of
„doctor rerum naturalium“ (Dr. rer. nat.)

Submitted to the Council of the Faculty
of Biological Sciences
of the Friedrich Schiller University Jena

by Michael Lachlan Alan Einar Easson (B.Sc. Honours Biochemistry)

born on April 25th, 1993 in St. Catharines (Canada)

List of Publications

This dissertation is based on the following research articles:

1. Malka, O*., Easson, M. L*., Paetz, C., Götz, M., Reichelt, M., Stein, B., ... & Mondaca, L. L. (2020). Glucosylation prevents plant defense activation in phloem-feeding insects. Nature Chemical Biology, 1-7.

2. Easson, M.L., Malka, O., Paetz, C., Hojná, A., Reichelt, M., Stein, B., Colvin, J., Winter, S., Morin, S., Gershenzon, J., Vassão, D.G. (2020). Cassava cyanogenic glucosides are detoxified via multiple pathways in the pest Bemisia tabaci. (In preparation)

3. Easson, M.L., Gershenzon, J., Vassão, D.G. (2021). Isomerization and oligomerization of dietary disaccharides by Bemisia tabaci transglucosidases. (In preparation).

Contents

Introduction

"Bugs are not going to inherit the earth. They own it now. So we might as well make peace with the landlord."

-Thomas Eisner

"If I have seen further than others, it is by standing on the shoulders of giants."

-Isaac Newton

General

Detoxification, or interactions with toxic chemical compounds as a process is something only whispered about in the case of our daily human lives. Part of the reason for this is that throughout history, humans have even gone to extreme lengths to reduce our interactions with these compounds. In terms of crop plants, humans have endeavoured painstaking breeding efforts and more recently targeted approaches in order to reduce the toxicity of the food that we eat. The motivation for these actions are in actuality obvious, as foods with lowered toxicity will result in a reduction of health complications or even death associated with individuals who consume them. However, these efforts are not without consequence, as the very chemical constituents that we often seek to reduce in these plants also serves to protect them from other organisms that humans frequently term "pests".

The study of plant-pest interactions is part of the discipline known as chemical ecology, which in a broad sense seeks to understand all the chemical interactions among organisms in an ecosystem. Much of the research in chemical ecology has focused on the interactions of herbivorous insects with the plants that they feed on. Due to the antagonistic nature of this relationship, production of defense chemicals by plants and detoxification by herbivores are two pillars of this field of study, with the dominating hypothesis being that these interactions are a product of an evolutionary arms race. Although this concept was originally coined in 1940 (Cott, 1940) it was hinted at by many before (Darwin, 1859), and finally became popularized by usage in co-evolutionary theory (Ehrlich & Raven, 1964). The evolutionary arms-race as a whole anthropomorphises the cyclical development of

adaptations and counter-adaptations between an organism and its antagonists. With reference to plant-herbivore interactions, it most often pertains to the selection pressures put on plants to produce defensive metabolites, and in parallel, the mechanisms that herbivores utilize to eliminate these pressures through detoxification or other adaptations.

Among animals, insects represent the largest and most diverse group of organisms on the planet, with many of them being herbivores (Grimaldi & Engel, 2005). These insects are studied not only in a purely scientific manner, but motivations also stem from the impact that they have on agricultural around the world, where annual crop loss from the activities of herbivorous insects amounts to an estimated 470 billion US dollars annually (Sharma, Kooner, & Arora, 2017). This is exacerbated by the increase of international trade, which has increased the introduction of pest species to new places. Introduced pest species may be able to feed with absolute impunity on plants in their new environment (Bradshaw et al., 2016) as plant species and predators in the new range did not co-evolve with them. Therefore, studies of these insects in their natural habitats including interactions with predatory insects, and importantly plants that exhibit resistance to herbivory contribute to the understanding and development of control measures for these insects in other parts of the world.

In parallel to their enormous phylogenetic diversity, insect herbivores also display an enormous range of feeding preferences, from insects which are narrowly specialized to feed on a handful of plant species such as the parsnip webworm *Depressaria pastinacella* (Ali & Agrawal, 2012), to those that can feed on several hundreds of plant species. The broad or narrow tuning of host preference may also be reflected in the respective insect's detoxification repertoire, where insects that feed on few plants are very successful at the detoxification of only a few metabolites, while insects that feed on many plants can detoxify a plethora of metabolites, with perhaps a lower efficiency (Ali & Agrawal, 2012). A prolific example of a polyphagous insect is the phloem-feeder *Bemisia tabaci*, capable of feeding on over 1000 observed plant species (Abd-Rabou & Simmons, 2010; Oliveira, Henneberry, & Anderson, 2001). Not only can this insect cause direct damage to plants through feeding behaviours, but it also happens to be an excellent vector for many plant disease-causing viruses. Through the combination of its broad host range and vector capabilities (Navas-Castillo, Fiallo-Olivé, & Sánchez-Campos, 2011), this insect is alone responsible for substantial crop losses worldwide (W. Chen et al., 2016; Oliveira et al., 2001). As a consequence of its extremely large host range, *Bemisia tabaci* represents an excellent model

species for gaining insight into the detoxification of plant chemical defense compounds, due to the diversity of compounds in its host plants.

The following sections of the introduction give a more detailed overview of (I) the diversity of plant-chemical defense compounds, (II) detoxification with specific emphasis on phase II detoxification in insects, (III) phloem feeding insects and *Bemisia tabaci*, and finally (IV) plant responses and defenses to phloem feeders. The introduction concludes with (V) the motivations of the thesis, highlighting the major contributions to the field.

I. Secondary Metabolites as Defense Compounds in Plants

As sessile organisms, plants are unable to simply uproot and leave when their biological environment is dire, and thus produce chemical defenses to protect against herbivores and pathogens. In the realm of natural product chemistry, especially when referring to plants, there exist two major classifications of metabolites, primary metabolites and secondary metabolites, the latter often also referred to as specialized metabolites. Primary metabolites are ubiquitously produced across all or a majority of kingdoms of life. Representative members of this class include nucleotides and amino acids used in DNA and protein synthesis, lipids, sterols, sugars and other substances essential to the normal growth, development and reproduction of an organism (Croteau, Kutchan, & Lewis, 2000). Secondary metabolites in contrast are much more narrowly distributed taxonomically (Bourgaud, Gravot, Milesi, & Gontier, 2001) and are produced to enable persistence in a specific environment. These two classes of metabolites can be seen as a product of trade-offs or compromises, being that an organism must find a balance between growth and development (associated with primary metabolism) and defense against enemies (associated with secondary metabolism), as stated in several plant defense theories (Bazzaz & Grace, 1997; Coley, Bryant, & Chapin, 1985). In terms of plant insect-interactions, especially involving phloem feeders, the distributions of both primary (mainly sugars and free amino acids) and secondary metabolites are of great importance from an insect perspective feeding on a host plant. Secondary metabolites have been studied for over 100 years, with the term first described in 1891 (Kossel, 1891), and expanded 30 years later (Czapek, 1921). These metabolites have since been major factors spurring the development of multiple facets of analytical science with the goal of isolation and characterization of these bioactive molecules (Croteau et al., 2000). Indeed, many secondary metabolites are

currently used in medicine such as vinblastine from *Catharanthus roseus* for the treatment of certain cancers (Hudes et al., 1992; Samuels & Howe, 1970); or at the very least these metabolites represent important starting points for the development of synthetic pharmaceutical analogues (Sparks, Hahn, & Garizi, 2017; Winter & Tang, 2012). For convenience, secondary metabolites are often broken into broad sub-categories, such as nitrogen-containing metabolites (the major focus of this thesis), phenylpropanoids, terpenoids, and fatty acid-derived compounds (Croteau et al., 2000). The following subsection will focus on metabolites known as activated two-component defenses that are highlighted in this thesis and important for phloem-feeding herbivore-plant interactions.

I.I Activated two-component defense metabolites

As a subclass of metabolites not restricted to any of the aforementioned major metabolite classes, two-component defense instead span a majority of specialized metabolite categories. As the name implies, these defense compounds require the presence of a second class of molecules in order to exert full toxicity in their role as defense metabolites. Many of these compounds are glycosides (Jones & Vogt, 2001; Mithöfer & Boland, 2012) that are associated with an activating glycosidase. The glycoside and glycosidase are typically spatially isolated until tissue disruption through actions, such as herbivore damage, cause their mixture and release of a toxic product in a localized area of destruction (Mithöfer & Boland, 2012). These metabolites are well-known as defenses against chewing herbivores (Cooper-Driver & Swain, 1976; Pentzold, Zagrobelny, Rook, & Bak, 2014; Ratzka, Vogel, Kliebenstein, Mitchell-Olds, & Kroymann, 2002; Wittstock et al., 2004), since maceration of plant tissue results in maximal compound activation. However, two-component defenses have usually not been considered to be activated by phloem-feeding insects since tissue damage is minimal (Pentzold et al., 2014).

I.I.I Glucosinolates

Glucosinolates (GSLs) are nitrogen containing, and sulfur rich natural products which are ubiquitously produced across the order Brassicales, including the family Brassicaceae (Tripathi & Mishra, 2007), and importantly present in the model plant *Arabidopsis thaliana,* thus contributing to their being very well studied (Blažević et al.,

2020; Winde & Wittstock, 2011). There are approximately 88-137 individual glucosinolate structures (Agerbirk & Olsen, 2012; Blažević et al., 2020) of which 88 have been satisfactorily characterized, and 49 remain only partially characterized in literature (Blažević et al., 2020). The compounds share common structural features of a thio-linked β-D-glucose linked to a (*Z*)-N-hydroxyamino sulfate ester (Halkier & Gershenzon, 2006). The variability in these compounds arises from the various amino acids from which the "R" side chain is derived and further modified, being alanine, leucine, isoleucine, methionine, and valine in the case of aliphatic glucosinolates, phenylalanine or tyrosine for benzenic glucosinolates, and tryptophan for indolic glucosinolates (Halkier & Gershenzon, 2006). Modification of side chains is extensive and variable, and includes chain elongation, *O*-methylation, desaturation, hydroxylation, and acylation (Halkier & Gershenzon, 2006). Biosynthesis of glucosinolates starts in many cases with the elongation of the amino acid side chain, and then proceeds with modification of the amino acid core structure via the sequential action of two cytochrome p450 enzymes (Du, Lykkesfeldt, Olsen, & Halkier, 1995) with the first (from CYP79 family) (Bak, Nielsen, & Halkier, 1998) resulting in the formation of an aldoxime (Figure 1) and the second (from the CYP83 family) (Bak & Feyereisen, 2001; Bak, Tax, Feldmann, Galbraith, & Feyereisen, 2001) resulting in the formation of an unstable nitrile oxide. The responsible enzyme for the synthesis of the *S*-alkyl-thiohydroximate (Figure 1) has not been isolated to date but is hypothesized to be a glutathione-*S*-transferase (GST)-like enzyme (Halkier & Du, 1997) with the cysteine conjugate then formed by cleavage of the remainder of the glutathione moiety. This product is also relatively unstable, and therefore a tightly associated enzyme, CS-lyase (Geu-Flores et al., 2009; Mikkelsen, Naur, & Halkier, 2004) produces the thiohydroximic acid. The action of subsequent enzymes *S*-glucosyl transferase (S-GT) (Gachon, Langlois-Meurinne, Henry, & Saindrenan, 2005) and sulphur transferase (ST) (Piotrowski et al., 2004) yield the final glucosinolate product (Figure 1).

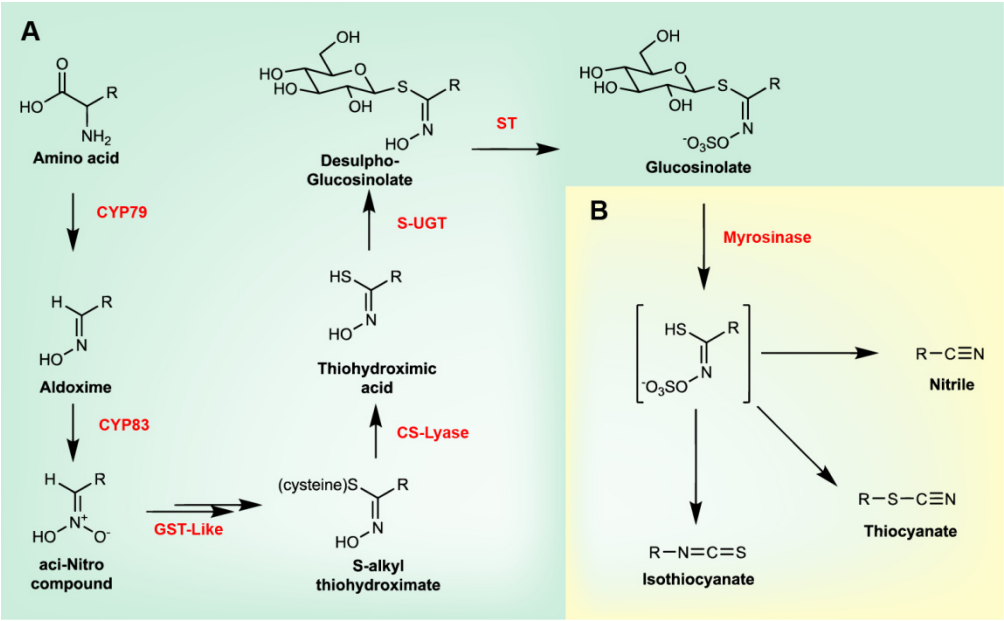


Figure 1: Glucosinolate biosynthesis and typical hydrolysis products. (**A**) Glucosinolate biosynthesis begins with an amino acid that is oxidized by sequentially acting P450 enzymes. Next, reaction of glutathione is followed by cleavage of the glycine and γ-glutamyl moieties of the resulting conjugate to leave a cysteine conjugate. The sulphur of cysteine is then reduced by CS-lyase to form the thiohydroximic acid. The last two steps involve a glucosyltransferase followed by a sulphotransferase to form the final glucosinolate. (**B**) The glucosinolate may be activated by contact with the enzyme myrosinase forming the depicted unstable intermediate in brackets, which can rearrange to some typical products, the most toxic of which is the isothiocyanate. Abbreviations: CYP (**CY**tochrome **P**450), GST (**G**lutathione-**S**-**T**ransferase), CS (**C**ysteine-**S**-conjugate), S-UGT (**S**ulphur-**U**ridine diphosphate **G**lucosyl **T**ransferase), ST (**S**ulphate **T**ransferase). Adapted from (Halkier & Gershenzon, 2006).

The activating enzyme associated with glucosinolates is a β-thioglucosidase known as myrosinase (Björkman & Janson, 1972; Bussy, 1840) of the glycoside hydrolase (GH) family 1 (Rask et al., 2000) and taken together with the glucosinolates themselves is affectionately referred to as the mustard oil bomb (Lüthy & Matile, 1984). Myrosinase is responsible for the hydrolysis of the β-thioglucose moiety of the glucosinolate, resulting in the release of a highly unstable intermediate (Figure 1) which can then undergo a series of rearrangements to form compounds of variable toxicity (Benn, 1977; Stauber et al., 2012), the most notable of which being the electrophilic isothiocyanates (Figure 1), which if left unchecked are hypothesized to react with proteins and other macromolecules important for

normal homeostasis (Borek, Elberson, McCaffrey, & Morra, 1998; Duus, 1979). Even without myrosinase, some glucosinolates, such as the indolic type, are reported to be chemically unstable resulting in spontaneous activation and degradation (Kim, Lee, Schroeder, & Jander, 2008). The distribution of these two components within *Arabidopsis thaliana* has also been extensively studied, with specific sulfur rich or "S"-cells (Koroleva et al., 2000) containing large quantities of glucosinolates and dedicated idioblasts or "M"-cells (Bones & Iversen, 1985; Guinard, 1890; Peche, 1913) serving in myrosinase storage. It is important to note that indolic glucosinolates seem to be localized in the plant surrounding the phloem tissue with aliphatic glucosinolates being more ubiquitously distributed (Nintemann et al., 2018). In the Col-0 ecotype of the model plant species *Arabidopsis thaliana*, the most abundant glucosinolate is the methionine-derived aliphatic glucosinolate glucoraphanin or 4-methylsulfinylbutyl glucosinolate (4msob-GSL) (P. D. Brown, Tokuhisa, Reichelt, & Gershenzon, 2003).

I.I.II Cyanogenic Glycosides

Cyanogenic glycosides are another class of two-component chemical defenses believed to have originated over 300 million years ago (Bak et al., 2006). It is present in over 2500 species of plants (Bak et al., 2006; Conn, 1980), as well as being produced by some insects as defense compounds (Moore, 1967; Nahrstedt & Davis, 1983). Chemically cyanogenic glycosides are perhaps more simple than their cousins, the glucosinolates, and unsurprisingly share many chemical features and their biosynthetic origin (Bak et al., 1998). These compounds are *O*-β-glycosides of α-hydroxynitriles that are typically classified as aliphatic or aromatic, which is again dependent on the "R" group or the amino acid from which their biosynthesis begins (Figure 2). The amino acids that contribute especially to aliphatic cyanogenic glycosides are valine, leucine and isoleucine with aromatic cyanogens derived from phenylalanine and tyrosine (Zagrobelny, Bak, & Møller, 2008). Interestingly, these compounds are also known to exist as diglycosides in addition to typical monoglycosides, and are hypothesized to have further roles in nitrogen storage (Busk & Møller, 2002), transport (Dirk Selmar, Reinhard Lieberei, & Böle Biehl, 1988; Selmar et al., 1987) and protection from selenium toxicity in animals (Palmer, El Olson, Halverson, Miller, & Smith, 1980; Smith Jr, Weisleder, Miller, Palmer, & Olson, 1980), even having their own dedicated disaccharidases (Fan & Conn, 1985).

A well-known crop plant that contains large amounts of cyanogenic glycosides as a defense is the tropical carbohydrate crop *Manihot esculenta* (McMahon, White, & Sayre, 1995) or more commonly cassava. The most abundant cyanogenic glycoside present is the valine derived aliphatic cyanogenic monoglycoside linamarin (Nartey, 1968), with contributions from the isoleucine derived lotaustralin (Nartey, 1968). Biosynthesis of linamarin like that of glucosinolates begins with the consecutive action of two cytochrome p450s. The first, from the CYP79 family forms the corresponding oxime (Andersen, Busk, Svendsen, & Møller, 2000) (Figure 2), and the second P450 from the CYP71 family catalyzes the loss of water and subsequent hydroxylation to form an acetone cyanohydrin (Jørgensen et al., 2011). Following this is the dedicated glucosyl-transfer reaction to form the final cyanogenic glycoside product (Kannangara et al., 2011). Due to the ancient nature of cyanogenic glycosides, and the conservation of the beginning CYP79 oxidation steps to oximes, cyanogenic glycosides are hypothesized to be the predecessor of glucosinolates (Bak et al., 1998; Bak et al., 2006).

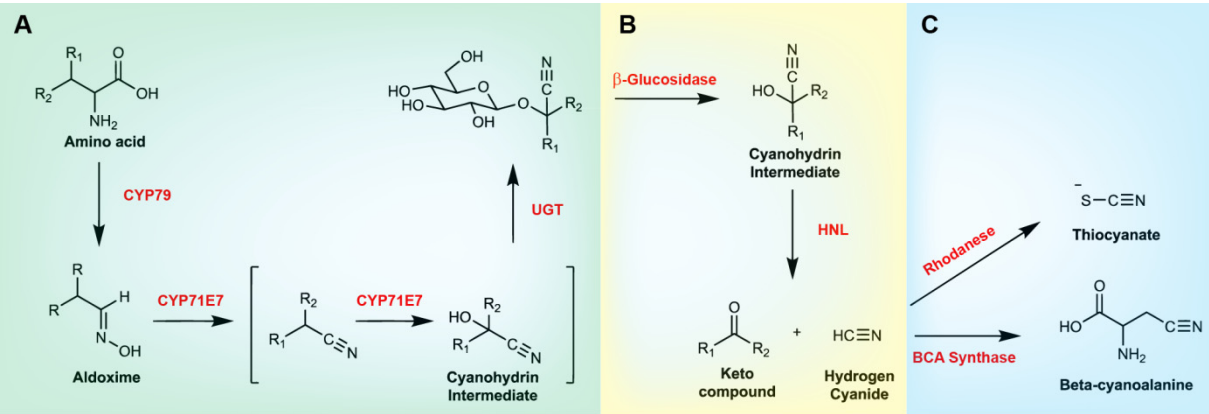


Figure 2: General biosynthesis, hydrolysis and detoxification of cyanogenic glycosides. (**A**) Similar to glucosinolate biosynthesis, cyanogenic glycoside biosynthesis begins with an amino acid which is oxidized by a CYP79 enzyme to form an aldoxime. Following this, an NADPH-dependent CYP71 enzyme forms an unstable cyanohydrin intermediate succeeded by immediate glucosylation by a dedicated glucosyltransferase. (**B**) Hydrolysis occurs via the action of dedicated β-glucosidase, which again forms the unstable cyanohydrin which can either spontaneously degrade or has its degradation enzymatically facilitated forming a keto-compound and the toxin hydrogen cyanide. (**C**) General detoxification reactions of cyanogenic glycosides typically found in plants include the cysteine transferring enzyme beta-cyanoalanine synthase and the thiocyanate forming

rhodanese. Abbreviations CYP (**CY**tochrome **P**450), UGT (**U**ridine diphosphate **G**lucosyl **T**ransferase), HNL (**H**ydroxy**N**itrile **L**yase), BCA (**B**eta-**C**yano**A**lanine).

The dedicated enzyme for the activation of linamarin in cassava and other linamarin-accumulating species is known as linamarase (Mkpong, Yan, Chism, & Sayre, 1990), which is responsible for the hydrolysis of the β-glucose of linamarin releasing an unstable cyanohydrin intermediate (Figure 2). The cyanohydrin produced can quickly degrade at pH values above 6 (Cooke, 1978), to form acetone and the toxin hydrogen cyanide. To make the release of hydrogen cyanide occur even more rapidly, many plants possess an enzyme known as hydroxynitrile lyase (HNL, Figure 2), which speeds the formation of hydrogen cyanide approximately six-fold faster in crude enzyme preparations than with linamarase alone (White & Sayre, 1992). Hydrogen cyanide is a respiratory toxin because of its inhibition of the electron transport chain. Since this substance is also toxic to the plant itself as well as to herbivore enemies, mechanisms have been developed by plants to reduce its toxicity. Beta-cyanoalanine synthase and rhodanese (Figure 2) are enzymes dedicated to the detoxification of hydrogen cyanide (Conn, 1980) with the former being utilized by cyanogenic and non-cyanogenic plants alike, since hydrogen cyanide is formed in all plants during the biosynthesis of the hormone ethylene (Peiser et al., 1984).

II. Detoxification

Detoxification as a broad definition should refer to the avoidance or mitigation of the potential effects of harmful substances, and with herbivorous insects and plants having together evolved for over 350 million years (Gatehouse, 2002) they have developed multiple adaptations and counter-adaptations for one another. Although activated two-component defenses represent one of the more sophisticated adaptations of plants to deal with their pest insects, some insects are able to feed on plants that contain these toxic metabolites with little to no mortality. Strategies employed by insect herbivores include behavioral modulations such as leaf mining and leaf cutting to avoid chemical defense activation (Pentzold et al., 2014). This is mirrored in phloem feeders, having a specialized feeding locale (the phloem) where it is hypothesized that they can subvert two-component defenses by avoiding causing enough tissue damage to activate them (Heil, 2009). In addition, insects may possess digestive systems that do not present favorable conditions for the activation of two-component defenses (Pentzold et al., 2014) and some specialized

insects may even sequester plant defenses in certain organs or tissues before they can exert their toxicity (Opitz, Jensen, & Müller, 2010). If an insect is not so incredibly adapted to avoid such defenses as is most commonly associated with insects having a broader host range; then metabolites will most likely become activated or exert their toxicity during regular feeding. Other adaptations involve various types of detoxification reactions commonly referred to as phase I and phase II detoxification strategies (Gibson & Skett, 2013), the major enzymatic pathways of which are discussed in the following subsections. Phase III detoxification, which typically involves the efflux and excretion of metabolites (Ioannides, 2002), is also known but not covered in this section.

II.I. Phase I detoxification

Phase I detoxification is known as the functionalization stage of detoxification (Ioannides, 2002), being largely catalyzed by enzymes involved in oxidation (cytochrome P450s), reduction and hydrolysis such as carboxylesterases (Liska, 1998). Cytochrome P450s are known to play an extensive role in the detoxification of many xenobiotics and can have a very large substrate affinity. The oxidized products can be more readily excreted due to an increase in polarity or may become substrates in phase II detoxification. In humans, famous examples can be found in everyday life in the form of alcohol oxidases (Grant, 1991), which are often targets for the pharmaceutical industry to reduce the metabolism of certain drugs and so increase their efficacy (Purnapatre, Khattar, & Saini, 2008) as treatments. In insects, these enzymes are believed to be perhaps one of the most important for insecticide resistance (Isaac Ishaaya, 1993; Lee & Scott, 1989). Numerous examples of monooxygenase activity are observed in the metabolism of insecticides (Bergé, Feyereisen, & Amichot, 1998; Hodgson, 1985; C. Wilkinson & Brattsten, 1972), as well as observations that perturbations to this pathway is met with susceptibility in otherwise resistant species (Tang, Zhao, Feng, Liu, & Qiu, 2012; Zhang et al., 2019).

Additionally in insects, hydrolases can play very important roles in the detoxification of various insecticides (Isaac Ishaaya, 1993), such as organophosphates (Dauterman, 1983) pyrethroids (I. Ishaaya, Ascher, & Casida, 1983) and benzoylphenyl ureas (Eto, Kishimot.K, Matsumura, Ohshita, & Oshima, 1966). Indeed, the pest species *C. carnea* displays a uniquely high tolerance to pyrethroids, attributed to the highly active and specific pyrethroid esterase (I. Ishaaya & Casida, 1981). Further examples of

hydrolysis are known in insects that pre-emptively metabolize glucosinolates by hydrolysis of the sulphate residue by sulphatases (Falk & Gershenzon, 2007; Ratzka et al., 2002), thus rendering these toxins inert to activation by myrosinase, a reaction interestingly also present in the whitefly *B. tabaci* (Malka et al., 2016). As mentioned previously, enzymatic action of phase I detoxification enzymes may be sufficient for the inactivation of a given toxin, however the functionalization reactions performed allow the further utility of phase II detoxification processes.

II.II. Phase II detoxification

Phase II detoxification is often referred to as conjugation detoxification, where toxic compounds may be made inert by conjugation to a molecule that also serves to increase its excretion via increasing polarity. Phase II enzymes may work after phase I detoxification processes or independently when conjugation is possible without prior functionalization. Enzymes which are typical members of this family are glutathione-*S*-transferases (GSTs), sulphotransferases, methyltransferases, glycine acyltransferases and UDP-glycosyltransferases (UDPGTs) (Meyer, 1996).

Glutathione conjugation and cysteine transferring enzymes are important members of the phase II detoxification machinery that are inducible in a tissue dependent manner following exposure to xenobiotics (Bhagwat, Mullick, Avadhani, & Raza, 1998). In the case of glucosinolate hydrolysis products, detoxification involves the mercapturic acid pathway, a well-characterized sequence of enzymatic steps utilized in the formation and processing of stable conjugates of isothiocyanates and other dangerous electrophiles (Figure 3). In this case, glutathione (GSH) is reacted with the electrophilic center of the isothiocyanate, reducing the toxicity and the potential of the toxins to react with proteins and other nucleophilic sites (Borek et al., 1998; Duus, 1979). The pathway does not simply end after the initial conjugation, but the non-cysteine amino acid residues of the tripeptide GSH are scavenged (Figure 3) so as to reduce the loss of nitrogen upon excretion. In virtually all studied plants and select insects, cysteine transfer enzymes mimicking GST action are also utilized in the detoxification of hydrogen cyanide via the action of beta-cyanoalanine synthase (Conn, 1980; van Ohlen, Herfurth, Kerbstadt, & Wittstock, 2016).

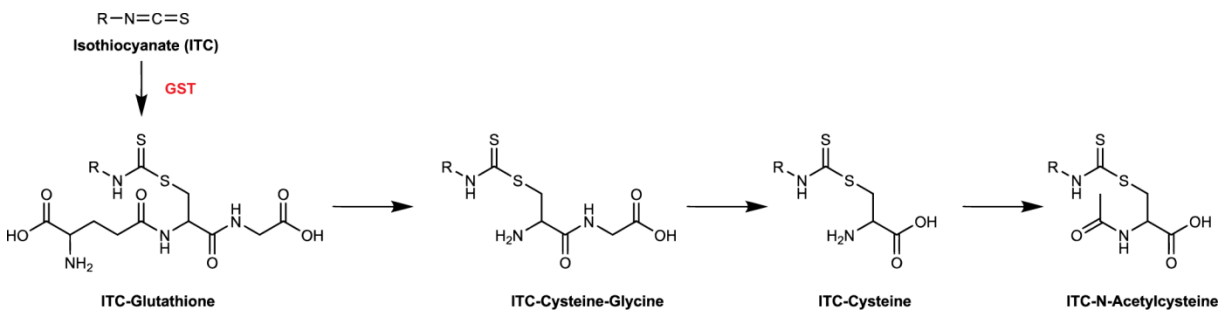


Figure 3: Meracapturic acid pathway for a general isothiocyanate. Isothiocyanates produced from glucosinolate hydrolysis can be detoxified via phase II conjugation with the tripeptide glutathione. This tripeptide is further broken down usually prior to excretion with the hydrolysis of glutamate, followed by glycine and the product may be further N-acetylated to aid in excretion.

Sulphotransferase detoxification is a multi-purpose transformation of xenobiotics (B.-H. Chen, Wang, Hou, Mao, & Yang, 2015), neurotransmitters (Roth, 1986) and hormones (Visser et al., 1998) alike. Flavanols in plants are well-known examples of secondary metabolites that can undergo multiple degrees of sulphation in response to UV stress (Gidda & Varin, 2006). In detoxification, this conjugation serves to increase the solubility of the products and is usually associated with a decrease in biological toxicity; however, it has been utilized in the past as a mechanism to activate certain pharmaceuticals. Indeed, in some instances, the sulphation of certain chemicals has been shown to further toxify them into dangerous carcinogens (Y. Wang et al., 2002). In plants especially, and in some insects, the activity of a thiosulphate sulphotransferase known as rhodanese is responsible for the conversion of the toxin hydrogen cyanide to the corresponding thiocyanate (Figure 2) (Antony et al., 2006).

Glycosyltransfer reactions are a major detoxification step in all vertebrates (Karl W Bock et al., 2012) and also extensively utilized in insects (Karl Walter Bock, 2016). This closely related family of enzymes are typically membrane bound in animals and are responsible for the transfer of a glycosyl moiety to molecules containing a nucleophilic nitrogen, sulphur, carbon or oxygen atom (Gibson & Skett, 2013). Importantly, they transfer an activated form of the glycosyl unit in the form of a diphosphate, and as a consequence the name uridine diphosphate glycosyltransferases (UDPGT) is employed (Karl Walter Bock, 2016). This conjugation is typically associated with a loss of activity of a given toxin and a major increase to the solubility of resulting glycoside. In humans,

drugs from virtually all classes are subject to the action of these enzymes, demonstrating the catalytic flexibility of UDPGT enzymes. A unique example of this transformation in insects is the detoxification of the plant toxin DIMBOA-glucose, in various lepidopteran herbivores (Wouters et al., 2014). In this case the two-component defense metabolite DIMBOA-glucose exists in the plant as β-linked glycoside, and is hydrolyzed within the insect gut when feeding (Vassão et al., 2018). However, dedicated UDPGTs from this insect can re-glucosylate the compound with an inversion in stereochemistry, resulting in a compound inert to the activating β-glycosidases (Wouters et al., 2014).

II.III Unique mechanisms in insect detoxification

Much of the work and research pertaining to detoxification in general has its roots in mammalian research where pharmaceutical metabolism and its potential effects in humans holds a bulk of all literature entries since its original categorization (Williams, 1959). However, the growing body of literature on insect detoxification processes shows deviations from classical mammalian detoxification strategies, especially in regards to phase II conjugation chemistry. The first of these differences pertains to the utilization of sugars and sugar derivatives in transfer reactions. In vertebrates, the unit which is almost always transferred is the glucose-derived sugar acid, glucuronic acid, which has the C-6 oxidized to a carboxylic acid rather than a hydroxyl group. In insects, plants and invertebrates, on the other hand, the transferred unit is exclusively glucose (Karl Walter Bock, 2016).

Further deviations in phase II detoxification in insects involve the transfer of phosphate moieties to xenobiotics, a process completely absent in the mammalian detoxification repertoire. Phosphorylation is not an uncommon modification of proteins, where its presence or absence may change substrate specificity or activate an enzyme (Mitra et al., 2011). However, in vertebrates or mammals the phosphorylation of metabolites is not reported as a detoxification strategy, possibly due to metabolic cost (Boeckler, Paetz, Feibicke, Gershenzon, & Unsicker, 2016), with phosphate supply limiting basic energy metabolism (Acin-Perez, Gatti, Bai, & Manfredi, 2011). In insects however, some of the first reports of this modification were seen as far back as the 1960s (Darby, Heenan, & Smith, 1966), where phenolic metabolites were seen to be phosphorylated in the excreta of *Costelytra zealandica*. Furthermore, the polyphagous pest *S. gregaria* has

been known to modify numerous insecticides via this conjugation mechanism, rendering them non-toxic (Olsen, Gabel-Jensen, Nielsen, Hansen, & Badolo, 2014; Olsen et al., 2016). Even more exciting is the clear link between classical phase I detoxification activities of monooxygenases (Olsen et al., 2014) in this insect, followed by phosphorylation of the hydroxylated metabolite, or even phosphorylation of conjugated glucose units transferred in a phase II detoxification manner (Olsen et al., 2016). With the growing evidence for insect phosphorylation of dietary toxins and other xenobiotics by insects (Scanlan, Gledhill-Smith, Battlay, & Robin, 2020), this should be recognized as a subcategory of phase II detoxification.

III. Phloem feeding insects

III.I General characteristics of phloem feeding insects

Phloem-feeding insects, members of the order Hemiptera or "true bugs" (Raven, 1983) are piercing-sucking feeders that consume the contents of the phloem specifically rather than the entirety of leaves, stems or other organs. They perform this task with a uniquely adapted stylet which can navigate apoplastically between cells in search of the sugar-laden phloem, which is pierced mechanically (Tjallingii & Esch, 1993). Examples of phloem-feeding insects include aphids, leafhoppers, plant hoppers and whiteflies, with varying ranges of host utilization or preference (Raven, 1983). Many of the studies on feeding behavior in this guild have been performed on aphids, where it has been found through electrical penetration graph (EPGs) (Tjallingii & Esch, 1993) that the insect first probes the plant superficially (Esch & Tjallingii, 1990) to determine if it will continue in search of the phloem. If the insect determines that a plant is suitable for consumption, further probing begins with eventual searching for the phloem or sieve element (Tjallingii & Esch, 1993). Once the insect has pierced the phloem, the turgor pressure of the sieve element allows for a steady stream of phloem content to exude and be consumed by the insect, facilitated by sucking (Raven, 1983). The plant, through calcium-dependent signaling cascades, naturally repairs damage to the phloem. Thus it is proposed that aphids and other phloem feeders may have mechanisms to block calcium-dependent signaling such as the use of calcium-chelating proteins (Will, Tjallingii, Thönnessen, & van Bel, 2007).

Phloem feeders excrete sugary frass known as honeydew, which can serve as an excellent substrate for the growth of fungi on the leaves of plants (Raven, 1983).

III.I.I The nitrogen barrier

It is generally accepted that phloem feeders (and most insects) are limited by the supply of nitrogen, which is usually present as free amino acids (Douglas, 2006). Interestingly, in some cases total nitrogen content may be adequate but there may be a lack of essential amino acids (Douglas, 1993). In the aphid host plant *Vicia faba*, the dominant amino acid within the phloem is asparagine (a non-essential amino acid), while the ratio of essential amino acids to non-essential amino acids is extremely low when compared to that present in aphid proteins (Douglas, 1993, 2006). In order to overcome this barrier, aphids and other phloem feeders utilize bacterial endosymbionts that are capable of producing essential amino acids lacking in their diets (Douglas, 1998; Hansen & Moran, 2011). Nitrogen quality in various species of plants also seems to have a distinct effect on the population of aphid symbionts, showing the importance of nitrogen or amino acid content in the utilization of a plant by a phloem feeder (TL Wilkinson, Koga, & Fukatsu, 2007).

III.I.II The sugar barrier and the importance of osmoregulation

The most dominant metabolites in the phloem are sugars, with a majority being the disaccharide sucrose, formed directly from the photosynthetic fixation of carbon (Byrne & Miller, 1990; Hayashi & Chino, 1990). However, phloem sap composition differs in some plants, such as in the case of cucurbits where raffinose and higher order saccharides make a large contribution (Haritatos, Keller, & Turgeon, 1996). The problem with this extremely high concentration of sugar in the form of sucrose is that it can cause the dehydration of feeding insects as sugars represent important osmolytes that influence the influx and efflux of water (Douglas, 2006). Aphids, whiteflies and other phloem feeders, however, still manage to feed on a sucrose-rich diet without ill effects through the use of sugar modifying enzymes, and tight regulation of aquaporins (Douglas, 2006).

The most important class of osmoregulatory enzymes, critical for all phloem feeders are the sugar modifying enzymes commonly referred to as sucrase-transglucosidases. They

are part of the glucohydrolase (GH) family of enzymes within the GH13 subfamily. The mechanism of these enzymes (Figure 4) begins with the formation of a glucosyl-enzyme intermediate (André, Potocki-Véronese, Morel, Monsan, & Remaud-Siméon, 2010; Moulis et al., 2006) and the release of fructose, followed by a concentration dependent fate of this intermediate (Cristofoletti, Ribeiro, Deraison, Rahbé, & Terra, 2003). In the case of low sucrose or nucleophile concentrations, water may enter the catalytic pocket and result in a net hydrolysis of sucrose to its substituent monosaccharides glucose and fructose (Figure 4 C). However, when sucrose or nucleophile concentrations are high enough, then these nucleophiles may enter the catalytic pocket instead (Figure 4 C'), resulting in transglucosidation to the nucleophilic center, which in the case of sucrose itself results in a trisaccahride. Fructose is typically absorbed and transported selectively out of the insect gut and utilized for energy purposes (Ashford, Smith, & Douglas, 2000). Thus, the overall conversion is two molecules of sucrose to one molecule of a trisaccharide, which halves the osmotic strength. It is important to note that this process may repeat with the same metabolite, producing larger and larger saccharides (Figure 4 D) until an equilibrium is reached as seen in the measurement of the iso-osmotic pressure of honeydew excreted from aphids (T Wilkinson, Ashford, Pritchard, & Douglas, 1997). Although protein fractions with such a transglucosidase activity have been isolated from aphids (Cristofoletti et al., 2003), and attempts have been made to characterize enzymes with such activities in other insects (Jing et al., 2016; Price et al., 2007), the only genes encoding enzymes with these activities are from bacteria (Robyt, Yoon, & Mukerjea, 2008).

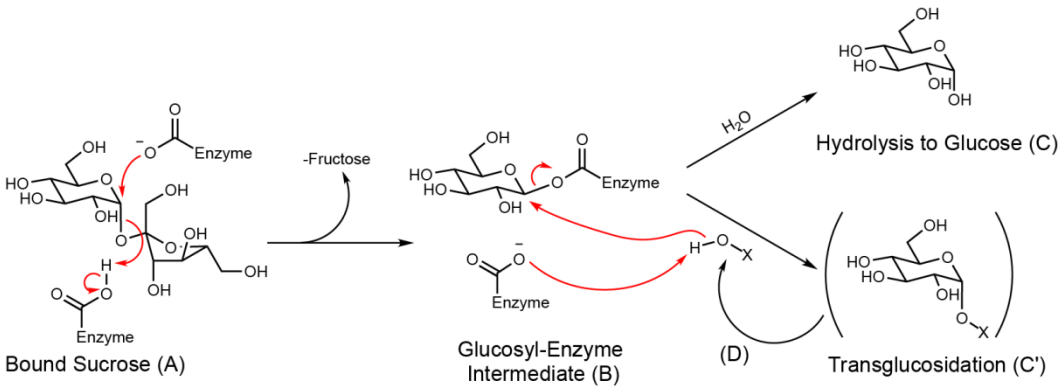


Figure 4: Basic mechanism of a retaining sucrase-transglucosidase. The binding of sucrose (**A**) causes an initial inversion of stereochemistry at the anomeric carbon with the release of fructose to form a glucosyl-enzyme intermediate (**B**). From here the intermediate can have two fates which are typically hypothesized to be concentration dependent. When

nucleophile (depicted as general H-O-X) concentrations are low, water enters the catalytic pockets resulting in net hydrolysis (**C**). However, when nucleophile concentrations are sufficient the result is a net transglucosidation (**C'**) which may happen numerous times (**D**) to the same molecule following the binding of another sucrose molecule.

III.II The whitefly *Bemisia tabaci*

III.II.I General

Bemisia tabaci was originally described in 1889 in Greece (Gennadius, 1889), where it was noted to persist with tobacco plantations, originally named *Aleyrodes tabaci* or the tobacco whitefly. Some of the first collections of *B. tabaci* in North America were made from sweet potatoes in the United States, at the time described as *Aleyrodes inconspicua* and commonly referred to as the sweet potato whitefly (Quaintance & Banks, 1900). Whitefly populations of the genus *Bemisia* have expanded recently from subtropical and tropical to temperate regions of the world, and are currently distributed on every continent, excluding Antarctica (Martin, Mifsud, & Rapisarda, 2000). As a result in part of the individual documentation history of whiteflies, *B. tabaci* represents a number of herbivorous phloem-feeding whitefly species, that until recently (Xu, De Barro, & Liu, 2010) was regarded as a complex species. Now being represented as a cryptic species complex, composed of at least thirty six (Firdaus et al., 2013) morphologically indistinguishable species which were commonly referred to as biotypes (Boykin et al., 2007). Typically species have been identified by means of cytochrome oxidase I (COI) sequencing with a divergence threshold of 3.5% (Dinsdale, Cook, Riginos, Buckley, & De Barro, 2010). The life cycle of *B. tabaci* is broken into six developmental stages that last anywhere from 14-28 days (Byrne & Bellows Jr, 1991), including egg deposition, hatching of the first mobile nymphal instar, the second-fourth sessile nymphal instars, the fourth nymphal instar with red eyes, and finally emergence of the adult from red eyed nymphs. *B. tabaci* is an arrhenotokous parthogenic species in which unfertilized eggs develop into haploid (male), or diploid (female) offspring (Byrne & Bellows Jr, 1991).

III.II.II Economic impact

Along with the observed expansion in global distribution, *B. tabaci* has now been observed to have a massive host range being able to feed and reproduce on over 1000 plant species to date (Abd-Rabou & Simmons, 2010). This whitefly causes extensive crop damage globally and has been named as one of the top 100 worst crop pests in the world (W. Chen et al., 2016; Zubair et al., 2020). The whitefly causes damage to plants by phloem feeding and by depositing honeydew. This sugary frass represents a perfect medium for growth of sooty molds, causing a decrease in crop yields (V. Srivastava & Thakre, 2000). In certain crops such as cotton, the deposition of honeydew represents a large problem with regards to crop harvests with processing costs being raised in order to eliminate contaminating sugar (Cheung, Roberts, & Perkins Jr, 1980). Whiteflies also serve as ideal vectors for plant disease-causing viruses (J. Brown, 1994). *B. tabaci* is capable of vectoring over 100 different viruses (W. Chen et al., 2016) which cause various forms of mosaic phenotypes (Basu, 2019; Jacobson, Duffy, & Sseruwagi, 2018), crippling crop yields, and even loss of harvestable material as is the case in cassava brown-streak disease (Jacobson et al., 2018), where tubers have observed tissue necrosis (figure 5) rendering them useless for consumption.

Figure 5: Necrotic tissue caused by cassava brown streak disease (CBSD). The whitefly in addition to causing damage via feeding and the deposition of sugar honeydew on the

leaves of plant is also an excellent vector for plant disease causing viruses such as CBSD depicted above, which results in tubers with necrotic tissue, seen as brown circular veins.

III.III.III Sugar metabolism

Sugar metabolism in whiteflies, specifically *B. tabaci*, is unique in comparison to other phloem feeders because of the activity of sucrose-isomerizing enzymes known as trehalulose synthases. Trehalulose synthase is an enzyme responsible for the conversion of sucrose [α-D-fructofuranosyl-(2↔1)-D-glucose] to trehalulose [α-D-fructofuranosyl-(1→1)-D-glucose], with the majority of reports from bacteria (Rhimi, Haser, & Aghajari, 2008). Trehalulose was first detected in the honeydew of Aleyrodoidea in the early 1990s and seems to be particularly present in whitefly metabolism (Byrne, Hendrix, & Williams III, 2003; Byrne & Miller, 1990; Hendrix, Wei, & Leggett, 1992). Following this, and in part due to the lack of reports of any insect performing this conversion and the numerous reports in bacteria (Rhimi et al., 2008), research was conducted to determine whether this isomerization was the product of whitefly symbionts or a distinct whitefly enzyme (Davidson, Segura, Steele, & Hendrix, 1994), with the latter found to be true (Salvucci, 2003). The metabolic advantages of sucrose to trehalulose conversion are not entirely obvious, however this reaction is hypothesized to represent a relatively fast conversion of sucrose to a less hydrolysable form, thus lowering the risk of sucrose inversion, which would exacerbate osmoregulatory challenges (Salvucci, Wolfe, & Hendrix, 1997). Whiteflies, particularly *B. tabaci,* have also been shown to produce a host of other unique sugar polymers such as the trisaccharides bemisiose (Hendrix & Wei, 1994) and isobemisiose (Hendrix & Salvucci, 2001), as well as other larger order saccharides (Y.-a. Wei, Hendrix, & Nieman, 1996; Y. A. Wei, Hendrix, & Nieman, 1997).

IV. Plant defenses to phloem-feeders and *B. tabaci*

Phloem feeders, being inconspicuous herbivores in their feeding mode, do not cause significant tissue disruption (Heil, 2009), especially when compared to chewing herbivores, where maceration of leaf plant tissue triggers large defensive responses. The plant however, is still able to recognize phloem-feeders and can mount several, fairly specific responses upon detection of phloem feeder damage. Hormonally, these responses seem to be mediated

via salicylic acid, rather than the jasmonic acid response typically associated with tissue damage and chewing herbivory (Soler et al., 2012). One of the most characteristic defenses against phloem feeders is the induction of phloem-blocking proteins that in essence cut off the supply of food for the feeding herbivore. Typically, these phloem disruption responses are triggered by an influx of calcium due to loss of local tissue integrity (Luna et al., 2011; V. K. Srivastava & Tuteja, 2014). Two major phloem-blocking mechanisms are known. One is the deposition of callose proteins in the sieve plates (Hao et al., 2008), resulting in a permanent disruption to phloem bulk flow. This phloem associated mechanism was originally discovered in the late 19th century (Naegeli, 1861), and later was found to be a defensive strategy for the deterrence of phloem feeding aphids, since callose deposits were found to be significantly increased in plants plagued by aphids (Saheed et al., 2007). A second example, restricted to the Fabaceae, is the more transient forisomes, which are protein aggregates within the sieve element that upon calcium influx, disperse to clog the vasculature current (Giordanengo et al., 2010). Once again, studies with aphids have found an increase in the number of forisomes when generalist aphids feed on the broad bean *Vicia faba* in comparison to more specialist aphids (Medina-Ortega & Walker, 2015).

Another prominent response to phloem-feeders is the induction of plant defense metabolites or secondary metabolites that serve as toxins and deterrents to phloem-feeder herbivory. In certain aphid species feeding on *Pisum sativum* the induction of phenylpropanoids was highlighted by the large accumulation in leaves where aphids had previously fed (Morkunas et al., 2016), correlating directly with the number of aphids feeding. This induction is hypothesized to have defensive roles against aphids from other studies (Lattanzio, Arpaia, Cardinali, Di Venere, & Linsalata, 2000) including one in which certain flavonoids were shown to inhibit aphid feeding (Goławska & Łukasik, 2012). Interestingly, it has been shown that specific two-component defenses, the glucosinolates can have major deterrent effects on aphids (Kim et al., 2008; Pfalz, Vogel, & Kroymann, 2009) feeding on *Brassica* plants, and are also induced upon herbivory by the same aphids (Kim & Jander, 2007). Indolic glucosinolates are hypothesized to be particularly active against phloem feeders, due to their inherent instability, resulting in hydrolysis without the presence of myrosinase (Kim et al., 2008). Furthermore, it is important to note that in *Arabidopsis thaliana*, “S”-cells containing high glucosinolate concentrations seem to be localized around the phloem, with a particular abundance of indolic glucosinolates being found in the surrounding tissues (Nintemann et al., 2018). In a study outlining the effects

of various glucosinolates on whitefly performance (Markovich et al., 2013), it was observed that aliphatic and indolic glucosinolates had variable effects on the fecundity of *B. tabaci* with performance effects associated with the overexpression of certain transcription factors associated with glucosinolate biosynthesis. With regards to a related two component plant defense, cyanogenic glycosides, the induction of detoxification enzyme activities in *B. tabaci* such as rhodanese and beta-cyanoalanine synthase, represents indirect evidence of the toxic effects of cyanogenic glycosides present in cassava (Antony et al., 2006).

V. Goals and scope of the thesis

Bemisia tabaci is capable of feeding on over 1000 species of plants (Abd-Rabou & Simmons, 2010), and represents a major agricultural and ornamental pest species. Through its inconspicuous mode of feeding and ability to vector a multitude of plant viruses, it is responsible for major economic losses across the globe. This loss becomes even more important when considering the impact it has on small shareholder farmers in places such as Sub Saharan Africa where whitefly populations have recently exploded. Despite the economic impact of whiteflies, few studies have been published that investigate the mechanisms that allow this insect to feed on so many different plants, some known to contain high concentrations of defenses, without apparent ill effects.

The subject of *B. tabaci* speciation remains very controversial. However, regardless of possible biotypes and cryptic sub-species, many questions can be raised about the mechanisms that have allowed whiteflies to become such a successful group of herbivores. Do these insects have interactions with classical plant defenses, such as two-component defenses? If so, what mechanisms exist to detoxify them, and are they unique to the phloem feeding guild? Finally, what are the enzymes that perform these reactions in the prolific phloem feeder *B. tabaci*? These unanswered questions were the motivation for my thesis, where I aimed to expand our knowledge of chemical-based interactions between whiteflies and plants.

Among the more specific interactions I explored were the interactions of the whitefly *B. tabaci* and two-component defenses from plants via analysis of the honeydew produced after feeding on phloem tissues. Even though whiteflies were not previously believed to interact with two-component defenses, in **Manuscript I** my colleagues and I

were able to show that the two-component defense 4msob-GSL was activated during regular feeding behaviours on *Arabidopsis thaliana* with classical phase II mercapturic acid pathway detoxification mitigating the toxicity of this activation. Furthermore, we were able to show a unique pre-emptive detoxification pathway utilized by whiteflies that resulted in repeated glucosylation of glucosinolates. This pathway was demonstrated to be exclusive to phloem feeding insects and seems to have been derived from osmoregulatory enzymes of the glucohydrolase (GH) family. Phylogenetic analysis showed a large number of GH enzymes within this insect. Utilizing sequence homology analysis to other known GH13 enzymes and complementary expression data, we characterized two potential GH enzymes from *B.tabaci* Middle-East-Asia-Minor 1 biotype (MEAM1) with transglucosidation activity with 4msob-GSL.

Following the discovery of this unique form of detoxification and the potential threat that two-component defenses can have towards the whitefly and other phloem feeders; the ubiquity of this detoxification mechanism was tested with whitefly populations feeding on cassava (**Manuscript II**). We were curious to find out if the same detoxification pathway would be utilized for other two-component defense systems in plants, and the related cyanogenic glycosides was an excellent candidate to investigate. We once again found evidence that cyanogenic glycosides were activated during whitefly feeding through the accumulation of a cyanide detoxification product in the insect body. We also discovered that the *B. tabaci* Sub-Saharan-Africa 1 biotype (SSA1) was capable of producing α-glycoside derivatives of linamarin similar to those produced with 4msob-GSL. However, the relative abundance of derivatives with glucosylation at different positions varied between these two defense compounds. Finally, we were able to detect yet another pre-emptive detoxification/modification not detected with glucosinolates, being the phosphorylation of the cyanogenic glycoside linamarin. The same two GH enzymes were tested again for potential transglucosidation activity with linamarin and a range of other secondary metabolites.

Whiteflies and other phloem feeders are perhaps most often associated with primary metabolites such as sugars (i.e sucrose) which are abundant in the phloem. Interestingly, the mechanism of secondary metabolite detoxification discovered here seems to have its origins in osmoregulation, or the modification of sugars. In order to gain further insight into the important ecological role of sugar isomerization in these insects, we searched for a potential trehalulose synthase in *B. tabaci* MEAM1. Using TIMS-ToF analysis in

Manuscript III, we were able to characterize three genes encoding proteins from *B. tabaci* with trehalulose synthase activity *in vitro*. Additionally, these enzymes also functioned as active transglucosidases of sucrose and other related sugars, highlighting their role in an osmoregulatory context. These enzymes were also tested for their ability to conjugate sugars to the secondary metabolite 4msob-GSL for which transglucosidation activity was previously observed. We found that none of the enzymes were capable of modifying 4msob-GSL and propose a division between osmoregulatory functions and secondary metabolite modification, perhaps providing a reason for the large GH gene expansion observed in these insects as compared to other herbivores.

Overview of Manuscripts

Manuscript I

Glucosylation prevents plant defense activation in phloem-feeding insects

Osnat Malka [1,7], Michael L. A. E. Easson [2,7], Christian Paetz [2], Monika Götz [3], Michael Reichelt [2], Beate Stein [3], Katrin Luck [2], Aleksa Stanisic [2], Ksenia Juravel [1,4], Diego Santos-Garcia [1], Lilach L. Mondaca [5], Simon Springate [6], John Colvin [6], Stephan Winter [3], Jonathan Gershenzon [2], Shai Morin [1] and Daniel G. Vassão [2]

[1] The Hebrew University of Jerusalem, Rehovot, Israel.

[2] Max Planck Institute for Chemical Ecology, Jena, Germany.

[3] Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany.

[4] Ludwig Maximilian University, Munich, Germany.

[5] Sapir Academic Collage, D.N. Hof, Askelon, Israel.

[6] Natural Resources Institute, University of Greenwich, Chatham Maritime, Kent, UK.

[7] These authors contributed equally: Osnat Malka, Michael L.A.E. Easson.

Corresponding author: Daniel G. Vassão

Brief Summary

In this chapter, we examined the fate of ingested two-component defenses, the glucosinolates in phloem feeding insects, specifically the whitefly *B. tabaci*. Here we discovered that during feeding by *B. tabaci* on *Arabidopsis thaliana*, the most abundant aliphatic defense compound 4msob-GSL is hydrolyzed in the insect and detoxified using the mercapturic acid pathway. We further investigated the metabolism of this glucosinolate through metabolic analysis of the honeydew from whiteflies feeding on glucosinolate containing plants and discovered that a majority of 4msob-GSL was being converted to myrosinase-inert poly-glucosylated derivatives and excreted, in line with a phase II detoxification process. Through labeling studies we showed that the responsible enzyme was a sucrase-transglucosidase and not a UDPGT, due to the exclusive utilization of sucrose for the transfer of glucose to 4msob-GSL. We further characterized two enzymes from *B. tabaci*, being SUC2 and SUC5 which were capable of utilizing sucrose to catalyze the transfer of glucose to 4msob-GSL. The phenomenon of poly-glucosylation was determined to be a unique detoxification mechanism to the phloem feeding-guild, with these compounds being inherently absent in insect excreta from other feeding guilds.

Author contributions:

Project conception: SM, JG, DGV, OM, MLAEE (20%)

Direction and supervision: DGV, JG, SM, OM

Funding acquisition: JG, SM, JC, DGV

Experimental design: MLAEE (80%), OM, DGV

Provision of insect and genetic material: OM, MG, BS, SW, LLM, SS, KL

Chemical analysis of honeydew: MLAEE (75%), DGV, MR

Purification of glycosides: DGV, MLAEE (33%)

NMR analysis: CP

Artificial diet feedings: MLAEE (60%), OM

Phylogenetic analysis: DSG

Molecular and enzymatic work: MLAEE (100%)

Wrote the manuscript: MLAEE (75%), DGV, JG, SM, OM

Manuscript II

Cassava cyanogenic glucosides are detoxified via multiple pathways in the pest *Bemisia tabaci*

Michael L.A.E. Easson[1], Osnat Malka[2], Christian Paetz[1], Anna Hojná[1], Michael Reichelt[1], Beate Stein[3], John Colvin[4], Stephan Winter[3], Shai Morin[2], Jonathan Gershenzon[1], Daniel G. Vassão[1]

[1]Max Planck Institute for Chemical Ecology, Jena 07745, Germany

[2]The Hebrew University of Jerusalem, Rehovot 7610001, Israel

[3]Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig 38104, Germany

[4]Natural Resources Institute, University of Greenwich, Chatham Maritime, Kent ME4 4TB, UK

Corresponding author: Daniel G. Vassão

Brief Summary

Manuscript II further explored the metabolism of two-component defenses in *B. tabaci* with cyanogenic glycosides from cassava, specifically linamarin. In order to find evidence of cyanogenic glycoside hydrolysis during whitefly feeding, a cyanide detoxification product, beta-cyanoalanine was measured and found to be enriched in whitefly bodies that fed on cyanogenic cassava versus non-cyanogenic eggplant. Previous observations in **Manuscript I** showed that *B. tabaci* may metabolize cyanogenic glycosides in a similar fashion to glucosinolates and indeed, poly-glucosylated linamarin derivatives were purified from honeydew produced by whiteflies feeding on cassava. The enzymes which catalyzed the transglucosidation of 4msob-GSL (BtSUC2 and BtSUC5) were utilized in enzyme assays with linamarin and found to also transglucosidate this metabolite with variable efficiency. Interestingly, an additional metabolism of linamarin was detected in the honeydew of cassava fed whiteflies which upon purification was revealed to be the addition of a phosphate to both native linamarin and the insect modified glycosides. The addition of phosphate also rendered linamarin inert to the activating enzyme linamarase, adding to this insect's detoxification toolset. This variation in the metabolism of linamarin illustrates the plethora of unique detoxification processes associated with *B. tabaci*, and further highlights the roles of two-component defenses in phloem-feeder and plant interactions.

Author contributions:

Project conception: MLAEE (50%), SM, JG, DGV, OM,

Direction and supervision: DGV, JG,

Funding acquisition: JG, SM, JC, DGV

Experimental design: MLAEE (90%), DGV

Provision of insect and genetic material: OM, BS, SW

Chemical analysis of honeydew: MLAEE (90%) MR

Purification of insect metabolites: MLAEE (90%), AH

NMR analysis: CP

Artificial diet feedings: MLAEE (75%), OM

Expression analysis: OM

Molecular and enzymatic work: MLAEE (100%)

Wrote the manuscript: MLAEE (75%), DGV, JG, SM, OM

Manuscript III

Isomerization and oligomerization of dietary disaccharides by *Bemisia tabaci* transglucosidases

Michael L.A.E. Easson[1], Jonathan Gershenzon[1], Daniel G. Vassão[1]

[1]Max Planck Institute for Chemical Ecology, Jena 07745, Germany

Brief Summary

Manuscript III sought to investigate the enzymes responsible for the unique activity observed in whitefly sugar metabolism, being the isomerization of sucrose to trehalulose. Here we were able to show for the first time in insects, three *B. tabaci* enzymes of the glycoside hydrolase 13 family which showed activity in the isomerization of sucrose to form trehalulose *in vitro*, aided by TIMS-ToF analysis for the identification of this product. Further analysis of these enzymes activities showed the utilization of sucrose, as well as other disaccharides in the formation of sugar oligomers of variable composition with some products identified by their respective ion-mobilities and corresponding momentum collision cross-sections (CCSs). Incubations with various sugar isotopomers provided mechanistic details to the observed *in vitro* activities, as well as the composition of unknown transglucosidation products. The characterization of these enzyme activities marks the first trehalulose synthases identified in insects, as well as the first saccharide transgluosidating enzymes identified in *B. tabaci* which are most likely important for osmoregulatory functions. Taken together with the other manuscripts in this thesis, it further highlights the utility of GH13 enzymes in *B. tabaci* metabolism of ingested plant substrates.

Author contributions:

Project conception: MLAEE (90%), JG, DGV

Direction and supervision: DGV, JG,

Funding acquisition: JG, DGV

Experimental design: MLAEE (90%), DGV

TIMS-TOF analysis of sugar products: MLAEE (100%)

Molecular and enzymatic work: MLAEE (100%)

Wrote the manuscript: MLAEE (90%), DGV, JG

Manuscript I

Glucosylation prevents plant defense activation in phloem-feeding insects

nature chemical biology

ARTICLES

https://doi.org/10.1038/s41589-020-00658-6

Check for updates

Glucosylation prevents plant defense activation in phloem-feeding insects

Osnat Malka[1,7]✉, Michael L. A. E. Easson[2,7], Christian Paetz[2], Monika Götz[3], Michael Reichelt[2], Beate Stein[3], Katrin Luck[2], Aleksa Stanišić[2], Ksenia Juravel[1,4], Diego Santos-Garcia[1], Lilach L. Mondaca[5], Simon Springate[6], John Colvin[6], Stephan Winter[3], Jonathan Gershenzon[2], Shai Morin[1] and Daniel G. Vassão[2]✉

The metabolic adaptations by which phloem-feeding insects counteract plant defense compounds are poorly known. Two-component plant defenses, such as glucosinolates, consist of a glucosylated protoxin that is activated by a glycoside hydrolase upon plant damage. Phloem-feeding herbivores are not generally believed to be negatively impacted by two-component defenses due to their slender piercing-sucking mouthparts, which minimize plant damage. However, here we document that glucosinolates are indeed activated during feeding by the whitefly *Bemisia tabaci*. This phloem feeder was also found to detoxify the majority of the glucosinolates it ingests by the stereoselective addition of glucose moieties, which prevents hydrolytic activation of these defense compounds. Glucosylation of glucosinolates in *B. tabaci* was accomplished via a transglucosidation mechanism, and two glycoside hydrolase family 13 (GH13) enzymes were shown to catalyze these reactions. This detoxification reaction was also found in a range of other phloem-feeding herbivores.

Two-component chemical defenses of plants commonly consist of a glucosylated protoxin that is cleaved by a glycoside hydrolase (GH), yielding an unstable aglycone that rearranges to form toxic products. The protoxin and glycoside hydrolase are stored in separate compartments that mix upon plant damage, activating the toxin. Two-component defenses, which include glucosinolates (GSLs), cyanogenic glycosides, benzoxazinoids and iridoids, have long been known to play decisive roles in interactions between plants and chewing herbivores[1,2]. However, how such compounds affect other herbivorous guilds, especially piercing-sucking, phloem-feeding herbivores such as aphids and whiteflies, is less well established. Piercing-sucking herbivores reach the sugary phloem sap by maneuvering their flexible mouthparts (stylets) between plant cells[3,4]. Therefore, in comparison to chewing insect herbivores such as caterpillars, phloem-feeding insects cause considerably less tissue damage and may not activate two-component defenses[5,6]. Nevertheless, infestations of phloem-feeding insects can still be very harmful to plants, as these pests vector a multitude of plant viruses and excrete large amounts of a sugary honeydew that supports fungal colonization, thus promoting disease and reducing yields of crop plants[7,8].

GSLs are among the best-studied examples of two-component plant chemical defenses. These amino acid-derived β-thioglucosides are produced by species of the order Brassicales, such as cabbages, turnips, radishes and mustard plants, as well as the model plant *Arabidopsis thaliana*. The hydrolytic enzymes required for GSL activation are the β-thioglucose hydrolases referred to as myrosinases[9]. Following tissue damage, myrosinase-catalyzed GSL cleavage results in unstable *O*-sulfated thiohydroximate aglycones, which rearrange to toxic isothiocyanates (ITCs) and other hydrolysis products (Fig. 1a). Some herbivorous insects are able to feed on GSL-containing plants without apparent negative effects. They may metabolize ITCs to glutathione derivatives (reviewed in ref. [10]) or convert GSLs to desulfated products before myrosinase action[10–12], which produces derivatives that cannot be cleaved by myrosinase. This latter reaction has also been observed in a phloem feeder, the whitefly *Bemisia tabaci* (species MED-Q2)[13]. The presence of such a preemptive detoxification mechanism in a phloem-feeding insect suggests the threat of exposure to GSL hydrolysis products, as does the fact that plant GSLs have negative effects on the performance of phloem feeders[14,15].

Besides plant chemical defenses, herbivores that live on phloem sap, such as aphids and whiteflies, must also overcome other hurdles. The sugar-rich phloem sap has very high osmolarity, which may cause dehydration and cell damage. Sucrose is often the main saccharide in the phloem vasculature, with concentrations ranging from 0.4 M to 0.8 M (refs. [16,17]), and the phloem sap gains further osmotic contributions from monosaccharides and larger oligosaccharides[18,19]. To prevent dehydration, phloem feeders tightly regulate water transport and employ sugar-modifying enzymes known as sucrase-transglucosidases[20,21]. These enzymes catalyze sucrose hydrolysis, followed by transglucosidation of the bound glucose moiety onto a growing sugar chain. After fructose absorption for energy metabolism, this effectively lowers the osmolarity of the insect gut content, reducing the threat of dehydration. The honeydew excreted by whiteflies therefore contains an oligosaccharide composition very different from that of the incoming plant sap, with a majority of di- and higher-order ($n \geq 3$) saccharides containing differing regio- and stereo-chemistries than those of the plant-derived saccharides[22].

Researchers have long wondered how plant defenses affect phloem-feeding insects, and whether these insects possess mechanisms to minimize toxicity after exposure. Here, we offer evidence

[1]The Hebrew University of Jerusalem, Rehovot, Israel. [2]Max Planck Institute for Chemical Ecology, Jena, Germany. [3]Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany. [4]Ludwig Maximilian University, Munich, Germany. [5]Sapir Academic Collage, D.N. Hof, Askelon, Israel. [6]Natural Resources Institute, University of Greenwich, Chatham Maritime, Kent, UK. [7]These authors contributed equally: Osnat Malka, Michael L.A.E. Easson. ✉e-mail: osnat226@gmail.com; vassao@ice.mpg.de

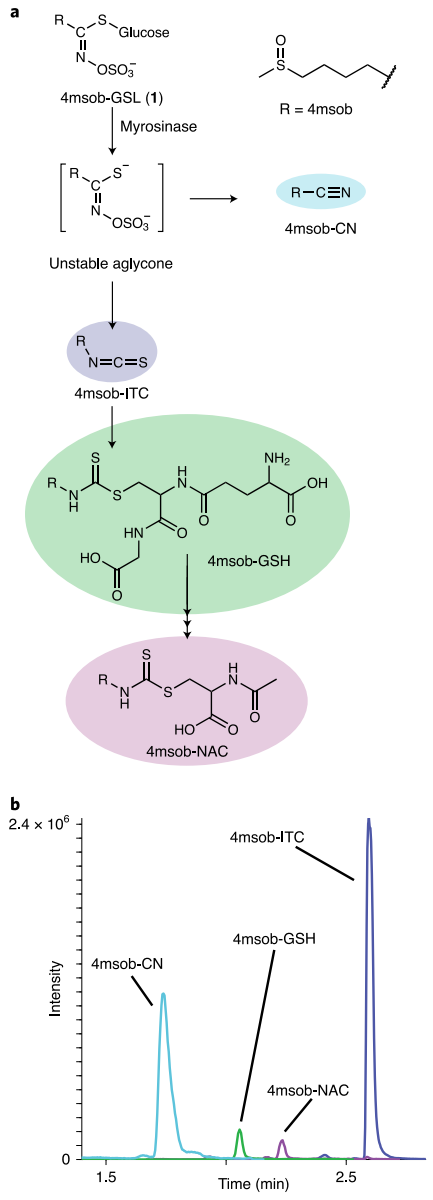


Fig. 1 | Activation and metabolism of 4-methylsulfinylbutyl glucosinolate (4msob-GSL) in the whitefly *Bemisia tabaci*. a, Myrosinase-catalyzed hydrolysis activates this GSL (**1**), forming the corresponding nitrile (light blue oval background) and isothiocyanate (dark blue oval background). The resulting isothiocyanate is then conjugated to glutathione (green oval background) and metabolized via the previously described mercapturic acid pathway (purple oval background). **b**, Representative LC-MS extracted ion chromatogram of concentrated honeydew from whiteflies fed on GSL-containing plants, showing the activation products of 4msob-GSL (4msob, 4-methylsulfinylbutyl; 4msob-CN, 4msob-nitrile; 4msob-ITC, 4msob-isothiocyanate) and known isothiocyanate detoxification products (4msob-GSH, glutathione conjugate of 4msob-ITC; 4msob-NAC, *N*-acetylcysteine conjugate of 4msob-ITC).

demonstrating the risks of two-component defenses for phloem feeders. We describe the detection of GSL-derived hydrolysis products in the honeydew of the phloem feeder *B. tabaci* MEAM1 and a novel preemptive GSL detoxification pathway based on the stereoselective addition of glucose moieties. Feeding of isotopically labeled sugars implicated the involvement of transglucosidase activity in these reactions, and we identified two different *B. tabaci* enzymes capable of catalyzing these transformations. Furthermore, we found this same detoxification reaction in a range of other phloem-feeding herbivores, but not in insect herbivores from other feeding guilds.

Results

Glucosylation by phloem feeders blocks toxin activation. To determine if *B. tabaci* whiteflies encounter GSLs while feeding and metabolize these compounds, we chemically analyzed the honeydew excreted from whiteflies fed on *A. thaliana*. Targeted HPLC-MS analyses identified intact and desulfated GSLs[13], as well as known GSL hydrolysis products and GSH conjugates (Fig. 1b). Whiteflies therefore not only encounter GSLs during feeding, but are also exposed to their toxic hydrolysis products. Untargeted HPLC-MS analyses of whitefly honeydew revealed large additional peaks of unknown metabolites with MS fragmentation patterns similar to those of 4-methylsulfinylbutyl (4msob)-GSL (**1**), the most abundant GSL in the leaves of *A. thaliana* accession Col-0. These unidentified compounds were also excreted in honeydew after whiteflies fed on other crucifer species containing 4msob-GSL, such as Brussels sprouts and broccoli, and on sucrose-based artificial diets containing 4msob-GSL. However, they were not found in any plant extracts, confirming that they were insect-derived GSL metabolites.

The mass spectra of these unknown compounds suggested a serial addition of one to five hexose units to the basic GSL skeleton. Because these additions did not affect MS fragmentation of the methylsulfinyl and sulfate groups, they appeared to be attached to the pre-existing GSL glucose residue. To elucidate further aspects of their structures, HPLC-purified 4msob-GSL derivatives with one and two hexose additions were investigated by NMR. The spectra obtained (Supplementary Table 1 and Supplementary Note) established that the hexose units were glucose moieties that had been added to the existing GSL β-thioglucose via α-(1→6) (**2** and **4**) or α-(1→4) (**3**) bonds (Fig. 2a). Such GSL metabolites have not been reported previously from any other organism. To determine if this glucosylation is a major metabolic fate of GSLs in whiteflies, NMR and targeted HPLC-MS/MS were used for the quantification of these derivatives. The combined mono- and di-glucosides of 4msob-GSL were present in a 14.2:1 ratio (standard error (s.e.) = 1.3, $N = 3$) relative to the unmodified GSL in honeydew from *A. thaliana*-fed *B. tabaci* MEAM1, while the previously reported desulfo-4msob-GSL was produced in a 2:1 ratio relative to unmodified 4msob-GSL in this species (s.e. = 0.4, $N = 3$). These results support glucosylation being a major route of 4msob-GSL metabolism in *B. tabaci*.

We found analogous derivatives of other GSLs in honeydew from *B. tabaci* feeding on *A. thaliana*, and on other GSL-containing plants (Extended Data Fig. 1). Therefore, *B. tabaci* can glucosylate other GSLs with short aliphatic side chains, although glucosylated indolic GSLs were low in abundance and these GSLs were primarily metabolized to desulfated GSLs[13]. In addition to whiteflies, HPLC-MS-based screening of extracts from the honeydew and feces of other herbivores indicated that glucosylated 4msob-GSL derivatives were also produced by other piercing-sucking insects feeding on *A. thaliana*, including aphids, mealybugs and other whiteflies, but not by chewing herbivores, such as lepidopterans, beetles, grasshoppers and single-cell feeders such as mites (Supplementary Table 2). Furthermore, we also detected analogous glucosylated derivatives of a cyanogenic glucoside in the honeydew of cassava-fed *B. tabaci* (species SSA1-SG3, which can feed on cassava, while *B. tabaci* MEAM1 cannot), but not when other

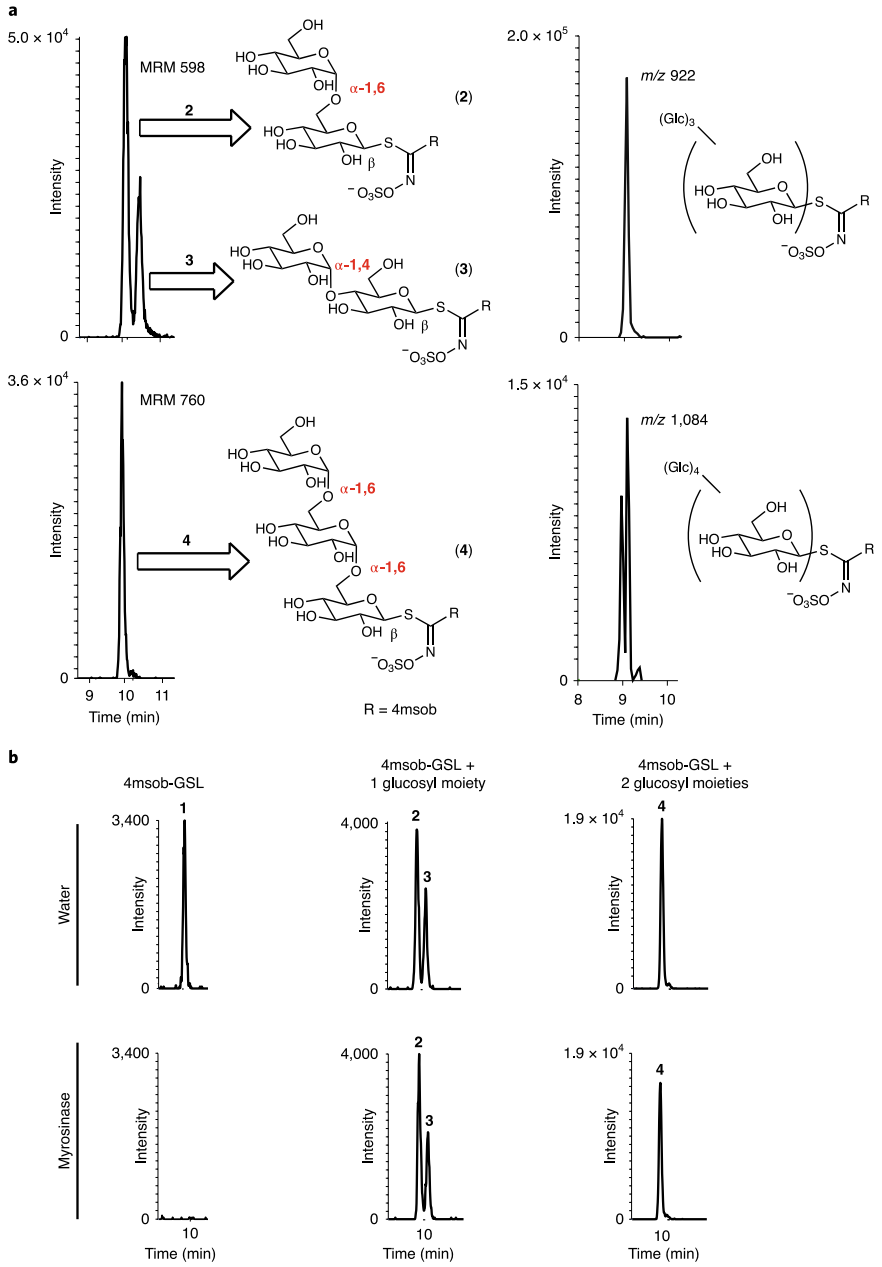


Fig. 2 | 4msob-GSL metabolites in *B. tabaci* whitefly honeydew. a, 4msob-GSL metabolites in honeydew contain additional α-linked glucose moieties, with up to four further glucose residues. Compounds **2**, **3** and **4** were purified and their structures determined by MS and NMR. **b**, Glucosylated 4msob-GSLs are resistant to activation by plant myrosinase, while the parent 4msob-GSL is readily hydrolyzed. Samples were incubated for 0.5 h with myrosinase in phosphate buffer at pH 7.0.

glucosylated defense compounds (benzoxazinoid, iridoid and coumarin glucosides) were fed in artificial diets to *B. tabaci* MEAM1 (Supplementary Table 3). These results suggest that the capacity to glucosylate GSLs is a broad characteristic of piercing-sucking insects, and this reaction may occur on other activated plant defenses present in their diets.

To determine whether serial α-glucosylation of GSLs might serve as a potential detoxification strategy, insect-produced glucosylated GSLs were incubated in vitro with a GSL-activating myrosinase (β-thioglucosidase) of plant origin. Myrosinase readily hydrolyzed 4msob-GSL, but the insect-derived glucosylated GSLs resisted hydrolysis (Fig. 2b). Taken together, these observations provide

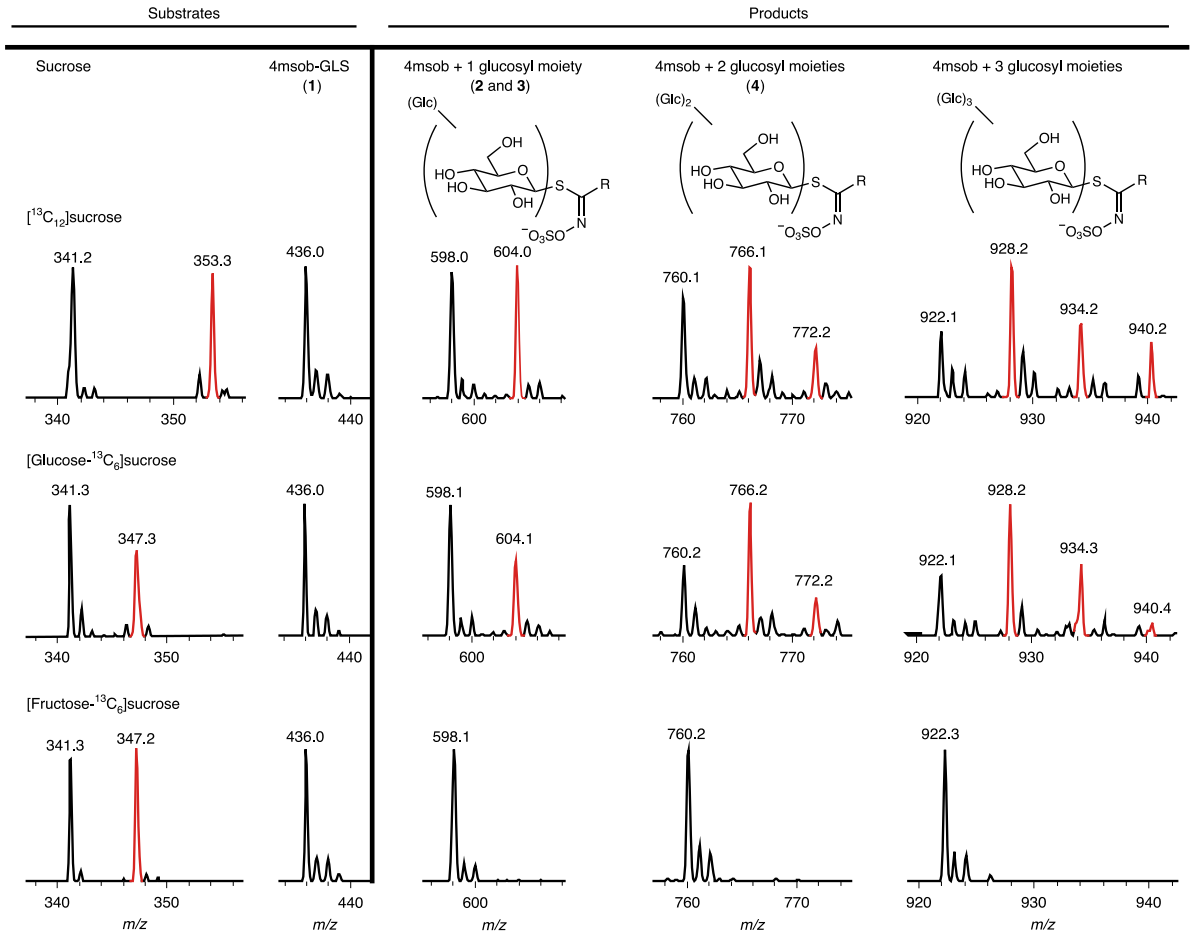


Fig. 3 | Elucidation of the biochemical mechanism for *B. tabaci* whitefly glucosylation of GSLs. Glucosylation of GSLs in *B. tabaci* is catalyzed by a transglucosidase activity, based on feeding of [^{13}C]sucrose isotopologues and the glucosinolate 4msob-GSL to insects on an artificial diet. Shown are regions of mass spectra from LC-MS analyses of substrates and products of feeding experiments on three different diets. Feeding fully ^{13}C-labeled sucrose and sucrose ^{13}C-labeled in the glucose moiety gave labeling in the newly added glucose of glucosylated 4msob-GSL (**1**) metabolites (**2**, **3**, **4** and others). However, feeding sucrose ^{13}C-labeled in the fructose moiety gave no label in the products, demonstrating initial hydrolysis of sucrose and reaction of only the glucose portion with the glucosinolate. Feeding of ^{13}C-labeled glucose and fructose with 4msob-GSL gave no incorporation into glucosylated derivatives of 4msob-GSL (Extended Data Fig. 2b). Peaks in red are isotopically enriched.

strong evidence that serial α-glucosylation serves as an important preemptive detoxification mechanism of certain glucosylated plant defenses in piercing-sucking herbivorous insects, generating derivatives that cannot be activated by plant glycoside hydrolases.

Glucosylation is carried out by GH13 transglucosidases. Glucosylation of dietary constituents by herbivores typically results from the action of UDP-glucosyltransferases, but transglucosidases (Extended Data Fig. 2a) might also be involved. To obtain information on the enzyme class and mechanism involved in serial GSL glucosylation, we fed *B. tabaci* MEAM1 with 4msob-GSL in diets containing ^{13}C isotopologues of sucrose, glucose or fructose. ^{13}C-labeled 4msob-GSL derivatives were only formed in vivo when [$^{13}C_{12}$]sucrose and [glucose-$^{13}C_6$]sucrose were added to diets, but not when [fructose-$^{13}C_6$]sucrose, [$^{13}C_6$]glucose or [$^{13}C_6$]fructose were offered (Fig. 3 and Extended Data Fig. 2b). Therefore the additional glucose moieties in whitefly-produced glucosylated GSLs originated from the glucose residue of sucrose, but not from its fructose group or from free monosaccharides, consistent with α-transglucosidase activity. This type of enzyme, belonging to the glycoside hydrolase class, converts glucose-containing saccharides to enzyme-bound glucose intermediates, which can then either be hydrolyzed to give free glucose, or the enzyme-bound glucose residue can be transglucosidated to an acceptor molecule. Transglucosidation is isoenergetic and often serves in vivo to add glucose residues to a growing chain, for example for carbohydrate storage[23] or to reduce osmolarity[24], but has not been previously implicated in detoxifying plant defense compounds.

A glycoside hydrolase carrying out transglucosidation reactions has been previously characterized in the gut of a phloem-feeding insect, the aphid *Acyrthosiphon pisum* (where it was named sucrase 1 (SUC1)[20,25]) and some of these enzymes have been identified in honeybees[26]. They belong to the large GH family 13 (GH13, http://www.cazy.org/GH13.html), whose members catalyze

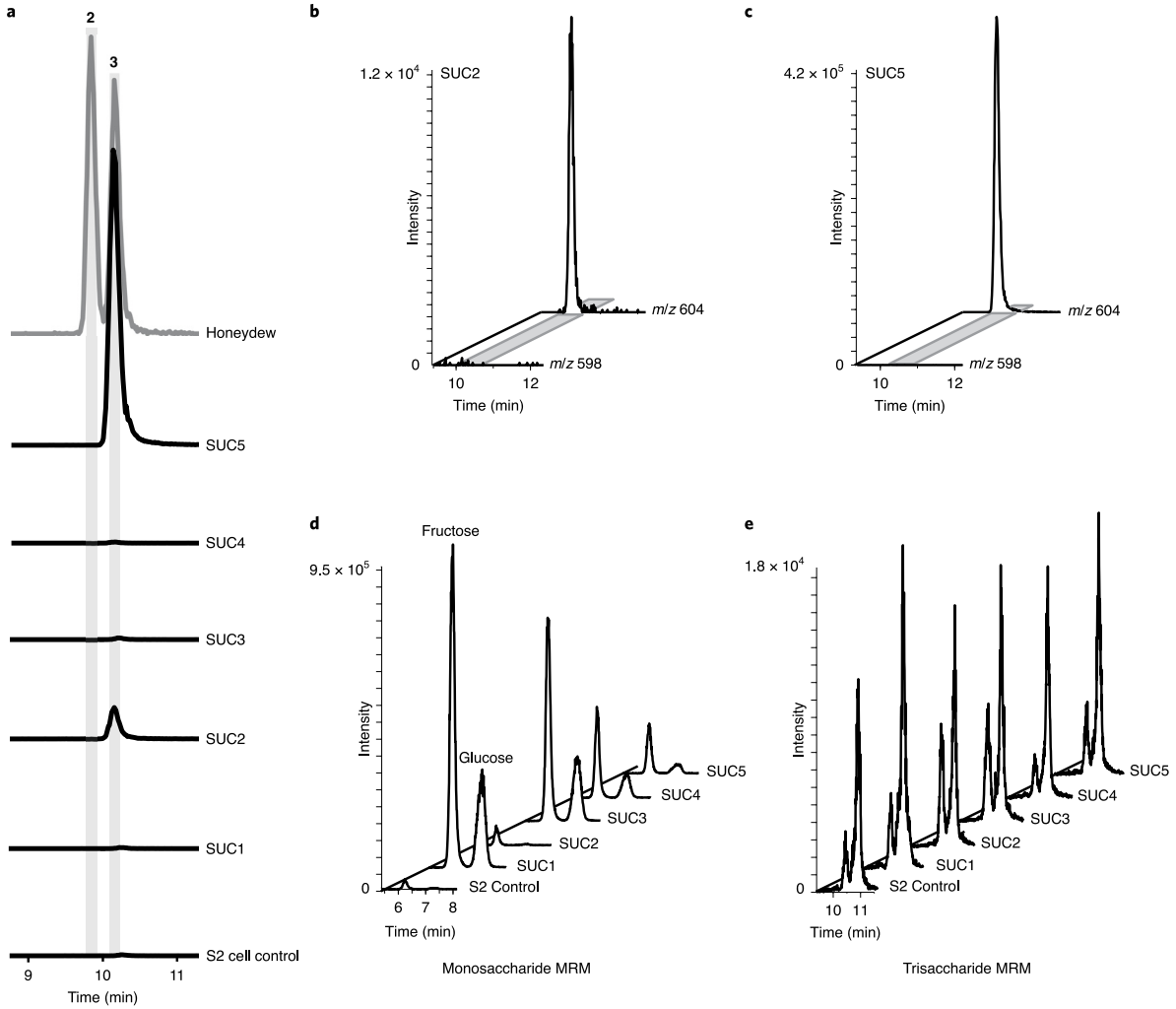


Fig. 4 | Chromatographic analyses of products from BtMEAM1 SUC1–5 enzymes heterologously produced in *D. melanogaster* S2 cells. Cell medium was assayed with 4msob-GSL and sucrose. **a**, Extracted multiple reaction monitoring (MRM) LC-MS chromatograms for mono-glucosylated 4msob-GSL derivatives. The enzymes SUC2 and SUC5 showed transglucosidation activity, producing α-(1→4)-linked glucose derivatives of 4msob-GSL (**3**), which are present in *B. tabaci* honeydew along with α-(1→6)-linked derivatives (**2**). **b**,**c**, When incubated with 4msob-GSL and sucrose ^{13}C-labeled in the glucose moiety, the enzymes SUC2 (**b**) and SUC5 (**c**) produced the ^{13}C-labeled, α-(1→4)-linked glucosylated derivative of 4msob-GSL. Depicted are the extracted MRM LC-MS chromatograms for unlabeled transglucosidase product (*m/z* 598) and the ^{13}C-labeled product (*m/z* 604). **d**,**e**, LC-MS analyses of sucrose metabolism by S2 control and S2 cells producing the SUC1–5 enzymes. Depicted are the monosaccharide MRM results showing hydrolysis of sucrose to glucose and fructose (**d**) and the trisaccharide MRM results showing that there was no significant activity above control levels for the transglucosidation of sucrose (**e**). The S2 control assay used extracts of the cell medium of untransformed cells.

stereochemistry-retaining reactions on α-glucosidic substrates, and which contains thousands of enzymes in a wide range of organisms from all kingdoms of life. To identify the transglucosidase enzymes responsible for the glucosylation of GSLs, we carried out a comparative phylogenetic analysis of the GH13 families of *B. tabaci* MEAM1 and eight other herbivorous arthropod species, including phloem feeders and representatives of other feeding guilds. The resulting tree showed a large expansion of the GH13 family in *B. tabaci* and other phloem-feeding insects (Extended Data Fig. 3).

For heterologous expression, we selected the most likely homolog of the *A. pisum* transglucosidase-encoding gene *SUC1* (ref. [27]) in *B. tabaci* MEAM1 (BtMEAM1 *SUC1*). We also selected four additional candidate GH13 genes (BtMEAM1 *SUC2–5*) to represent all three major *B. tabaci* GH13 sub-groups and to include genes encoding enzymes with high similarity to the characterized honeybee transglucosidase enzymes[26] (Extended Data Fig. 3). Most of the selected genes (except for *SUC2*) were more highly expressed in *B. tabaci* MEAM1 guts than in other tissues[28] (Supplementary Table 4), and all the encoded proteins possessed predicted signal peptides. The candidate genes were expressed in *Drosophila melanogaster* S2 cells and the medium was assayed for activity using 4msob-GSL and sucrose. Both SUC2 and SUC5 enzymes generated the α-(1→4)

mono-glucosylated derivative of 4msob-GSL (Fig. 4a–c) and only low amounts of free glucose and fructose (Fig. 4d). The enzymes SUC1, 3 and 4 did not exhibit transglucosidase activity with 4msob-GSL, but readily hydrolyzed sucrose (Fig. 4d). None of the enzymes was able to produce detectable transglucosidase products from sucrose itself (higher-order glucose oligomers of sucrose) (Fig. 4e), as had been previously described for the activity of *A. pisum* midgut enzymes[21]. SUC5 displayed apparent Michaelis–Menten kinetics with an estimated K_M of ~0.2 M for sucrose and 460 µM for 4msob-GSL. At 1 M sucrose, SUC2 reached 1/2 v_{max} at ~1,000 µM 4msob-GSL, but we were unable to saturate SUC2 with sucrose and therefore could not estimate its K_M towards this substrate, which must be greater than 0.5 M (Supplementary Fig. 1). Trehalose, a major disaccharide in hemolymph[29], was only poorly used as a glucose donor (<1% relative to sucrose, Supplementary Fig. 2). For a phloem-feeding insect, such as *B. tabaci*, the concentration of sucrose in its gut originating from the plant is expected to be several hundred millimolar[17,30,31].

Discussion

Activated defenses are well known to defend plants against chewing herbivores, but piercing-sucking insects, such as aphids and whiteflies, have not been thought to cause sufficient plant tissue damage to trigger defense activation. In this research, we found that *B. tabaci* MEAM1 whiteflies feeding upon glucosinolate-containing plants excrete toxic hydrolysis products such as ITCs and known ITC detoxification products in their honeydew, indicating activation (Fig. 1b). Hydrolysis might occur from GSLs present on the leaf surface[32] or found in the mesophyll along the path to the phloem[3,4]. The activating enzyme myrosinase is present in cells surrounding the phloem[32,33], and after stylet penetration might come into contact with phloem GSLs, or GSLs could be hydrolyzed by insect enzymes[14]. Previous evidence of GSL breakdown by phloem-feeding insects comes from work on aphids[14,34]. The activation of GSLs by phloem feeders indicates that these herbivores might be well served by preemptively metabolizing GSLs and other protoxins to non-hydrolyzable derivatives to prevent their activation.

Here we discovered that the phloem-feeding *B. tabaci* MEAM1 converts ingested GSLs into α-(1→6) and α-(1→4) glucosylated conjugates (Fig. 2a). These represent the major GSL metabolites detected in whitefly honeydew when feeding on *A. thaliana*, outnumbering the parent GSLs 14:1 and the previously described[13] desulfated GSL metabolites by 7:1. Glucosylated 4msob-GSL is not susceptible to cleavage by plant myrosinases (Fig. 2b) and thus can no longer be activated. Similar glucosylated GSLs were found in all other investigated phloem-feeding insects (Supplementary Table 2). However, glucosylated 4msob-GSL derivatives were absent in excreta from chewing insects such as grasshoppers, caterpillars and beetles and also from cell content feeders such as mites. Additionally, glucosylation was used by *B. tabaci* (species SSA1-SG3) to modify a related group[35] of activated defenses, the cyanogenic glycosides (Supplementary Table 3). Representatives of other classes of activated glucoside defense metabolites fed to *B. tabaci* (the benzoxazinoid DIMBOA-Glc, the iridoid catalpol and the hydroxycoumarin esculin) were not converted to their corresponding glucosylated derivatives, suggesting that not all glucosidic defense compounds can be glucosylated by this process.

The *B. tabaci* glucosylated GSLs were shown to be produced by a transglucosidase activity. Transglucosidation has already been described in phloem-feeding insects to form higher-order sugar oligomers from sucrose, helping to lower the high osmotic pressures encountered when feeding on phloem sap[20,25]. When supplied with sucrose and 4msob-GSL, two of the five enzymes we tested formed glucosylated derivatives of 4msob-GSL with α-(1→4) linkages (Fig. 4a–c), while the other three only catalyzed sucrose hydrolysis (Fig. 4d). The enzymes catalyzing the α-(1→6) modification remain elusive. None of the enzymes tested could produce higher-order oligomers of sucrose, as previously described for aphid gut transglucosidase(s)[20] (Fig. 4e). These data suggest that *B. tabaci* transglucosidases use a variety of different substrates, and the enzymes responsible for osmoregulation may be different from those capable of detoxifying plant defenses.

Glucose conjugation can render plant defenses less harmful, not only by preventing enzymatic activation, but also by increasing their polarity (to enhance excretion) or by otherwise altering their structures to prevent toxic or deterrent activity. Moreover, transglucosidases move glycosidic linkages between substrates without extra inputs of chemical energy, as opposed to glucosylation via UDP-glucosyltransferases, which utilizes phosphodiester bonds. From the phylogenetic analysis performed (Extended Data Fig. 3), the GH13 family expanded substantially in phloem feeders, with *B. tabaci* possessing the greatest number of genes among the species surveyed[27] (Supplementary Table 7). Given that GH13 proteins are involved in sugar metabolism, it is perhaps not surprising that phloem-feeding insects possess a large diversity of these enzymes. They may be responsible for various functions, including osmoregulation, furnishing energy, glycosylation of defenses for detoxification and others to be determined. The *B. tabaci* species complex is an ideal group for future studies on the function, evolution and structure–activity relationships of GH13 enzymes, because different taxa in the group have different feeding preferences[6], which might be mediated by the spectrum of GH13 activities present.

The identification of new traits important in the adaptation of a pest insect to its host plant can provide new targets for control measures. Thus, chemical inhibitors of GH13 detoxification enzymes or RNA interference (RNAi) directed against the corresponding genes could reduce whitefly damage to agricultural crops by decreasing insect growth and survival. RNAi, in particular supplied by plants or by direct application, is considered to have great promise in the control of hemipteran pests[36,37].

In summary, the modification of activated plant defenses via transglucosidase-catalyzed glucose addition represents a novel route for avoiding the release of toxic hydrolysis products in phloem feeders. The GH13 enzymes responsible are abundant in *B. tabaci* and other phloem-feeding insects, and might help to explain the evolutionary radiation of hemipterans and their success as plant feeders.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41589-020-00658-6.

Received: 3 April 2020; Accepted: 21 August 2020;
Published online: 28 September 2020

References

1. Morant, A. V. et al. ß-glucosidases as detonators of plant chemical defense. *Phytochemistry* **69**, 1795–1813 (2008).
2. Pentzold, S., Zagrobelny, M., Rook, F. & Bak, S. How insects overcome two-component plant chemical defence: plant ß-glucosidases as the main target for herbivore adaptation. *Biol. Rev.* **89**, 531–551 (2014).
3. Tjallingii, W. F. & Esch, T. H. Fine structure of aphid stylet routes in plant tissues in correlation with EPG signals. *Physiol. Entomol.* **18**, 317–328 (1993).
4. Walker, G. P. & Perring, T. M. Feeding and oviposition behavior of whiteflies (Homoptera, Aleyrodidae) interpreted from AC electronic feeding monitor wave forms. *Ann. Entomol. Soc. Am.* **87**, 363–374 (1994).
5. Walling, L. L. Avoiding effective defenses: strategies employed by phloem-feeding insects. *Plant Physiol.* **146**, 859–866 (2008).
6. Wang, X. W., Li, P. & Liu, S. S. Whitefly interactions with plants. *Curr. Opin. Insect Sci.* **19**, 70–75 (2017).
7. De Barro, P. J., Liu, S. S., Boykin, L. M. & Dinsdale, A. B. *Bemisia tabaci*: a statement of species status. *Annu. Rev. Entomol.* **56**, 1–19 (2011).

8. Oliveira, M. R. V., Henneberry, T. J. & Anderson, P. History, current status and collaborative research projects for *Bemisia tabaci*. *Crop Prot.* **20**, 709–723 (2001).
9. Bones, A. M. & Rossiter, J. T. The myrosinase-glucosinolate system, its organisation and biochemistry. *Physiol. Plant.* **97**, 194–208 (1996).
10. Jeschke, V., Gershenzon, J. & Vassao, D. G. Insect detoxification of glucosinolates and their hydrolysis products. *Adv. Bot. Res.* **80**, 199–245 (2016).
11. Falk, K. L. & Gershenzon, J. The desert locust, *Schistocerca gregaria*, detoxifies the glucosinolates of *Schouwia purpurea* by desulfation. *J. Chem. Ecol.* **33**, 1542–1555 (2007).
12. Ratzka, A., Vogel, H., Kliebenstein, D. J., Mitchell-Olds, T. & Kroymann, J. Disarming the mustard oil bomb. *Proc. Natl Acad. Sci. USA* **99**, 11223–11228 (2002).
13. Malka, O. et al. Glucosinolate desulfation by the phloem-feeding insect *Bemisia tabaci*. *J. Chem. Ecol.* **42**, 230–235 (2016).
14. Kim, J. H., Lee, B. W., Schroeder, F. C. & Jander, G. Identification of indole glucosinolate breakdown products with antifeedant effects on *Myzus persicae* (green peach aphid). *Plant J.* **54**, 1015–1026 (2008).
15. Markovich, O. et al. *Arabidopsis thaliana* plants with different levels of aliphatic and indolyl-glucosinolates affect host selection and performance of *Bemisia tabaci*. *J. Chem. Ecol.* **39**, 1361–1372 (2013).
16. Hayashi, H. & Chino, M. Chemical composition of phloem sap from the uppermost internode of the rice plant. *Plant Cell Physiol.* **31**, 247–251 (1990).
17. Lohaus, G. et al. Solute balance of a maize (*Zea mays L.*) source leaf as affected by salt treatment with special emphasis on phloem retranslocation and ion leaching. *J. Exp. Bot.* **51**, 1721–1732 (2000).
18. Haritatos, E., Keller, F. & Turgeon, R. Raffinose oligosaccharide concentrations measured in individual cell and tissue types in *Cucumis melo* L. leaves: implications for phloem loading. *Planta* **198**, 614–622 (1996).
19. Rennie, E. A. & Turgeon, R. A comprehensive picture of phloem loading strategies. *Proc. Natl Acad. Sci. USA* **106**, 14162–14167 (2009).
20. Cristofoletti, P. T., Ribeiro, A. F., Deraison, C., Rahbe, Y. & Terra, W. R. Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid *Acyrthosiphon pisum*. *J. Insect Physiol.* **49**, 11–24 (2003).
21. Douglas, A. E. Phloem-sap feeding by animals: problems and solutions. *J. Exp. Bot.* **57**, 747–754 (2006).
22. Byrne, D. N. & Miller, W. B. Carbohydrate and amino acid composition of phloem sap and honeydew produced by *Bemisia tabaci*. *J. Insect Physiol.* **36**, 433–439 (1990).
23. Monsan, P., Remaud-Simeon, M. & Andre, I. Transglucosidases as efficient tools for oligosaccharide and glucoconjugate synthesis. *Curr. Opin. Microbiol.* **13**, 293–300 (2010).
24. Fisher, D. B., Wright, J. P. & Mittler, T. E. Osmoregulation by the aphid *Myzus persicae*: a physiological role for honeydew oligosaccharides. *J. Insect Physiol.* **30**, 387–393 (1984).
25. Price, D. R. G. et al. Molecular characterisation of a candidate gut sucrase in the pea aphid, *Acyrthosiphon pisum*. *Insect Biochem. Mol. Biol.* **37**, 307–317 (2007).
26. Ngiwsara, L. et al. Amino acids in conserved region II are crucial to substrate specificity, reaction velocity, and regioselectivity in the transglucosylation of honeybee GH-13 α-glucosidases. *Biosci. Biotechnol. Biochem.* **76**, 1967–1974 (2012).
27. Jing, X. et al. Evolutionary conservation of candidate osmoregulation genes in plant phloem sap-feeding insects. *Insect Mol. Biol.* **25**, 251–258 (2016).
28. Wang, X. W. et al. Analysis of a native whitefly transcriptome and its sequence divergence with two invasive whitefly species. *BMC Genomics* **13**, 529 (2012).
29. Hendrix, D. L. & Salvucci, M. E. Isobemisiose: an unusual trisaccharide abundant in the silverleaf whitefly, *Bemisia argentifolii*. *J. Insect Physiol.* **47**, 423–432 (2001).
30. Riens, B., Lohaus, G., Heineke, D. & Heldt, H. W. Amino acid and sucrose content determined in the cytosolic, chloroplastic, and vacuolar compartments and in the phloem sap of spinach leaves. *Plant Physiol.* **97**, 227–233 (1991).
31. Merritt, S. Z. Within-plant variation in concentrations of amino acids, sugar, and sinigrin in phloem sap of black mustard, *Brassica nigra* (L) Koch (Cruciferae). *J. Chem. Ecol.* **22**, 1133–1145 (1996).
32. Nintemann, S. J. et al. Localization of the glucosinolate biosynthetic enzymes reveals distinct spatial patterns for the biosynthesis of indole and aliphatic glucosinolates. *Physiol. Plant.* **163**, 138–154 (2018).
33. Andreasson, E., Jorgensen, L. B., Hoglund, A. S., Rask, L. & Meijer, J. Different myrosinase and idioblast distribution in *Arabidopsis* and *Brassica napus*. *Plant Physiol.* **127**, 1750–1763 (2001).
34. Danner, H., Desurmont, G. A., Cristescu, S. M. & van Dam, N. M. Herbivore-induced plant volatiles accurately predict history of coexistence, diet breadth and feeding mode of herbivores. *New Phytol.* **220**, 726–738 (2018).
35. Bak, S., Nielsen, H. L. & Halkier, B. A. The presence of CYP79 homologues in glucosinolate-producing plants shows evolutionary conservation of the enzymes in the conversion of amino acid to aldoxime in the biosynthesis of cyanogenic glucosides and glucosinolates. *Plant Mol. Biol.* **38**, 725–734 (1998).
36. Eakteiman, G. et al. Targeting detoxification genes by phloem-mediated RNAi: a new approach for controlling phloem-feeding insect pests. *Insect Biochem. Mol.* **100**, 10–21 (2018).
37. Luo, Y. A. et al. Towards an understanding of the molecular basis of effective RNAi against a global insect pest, the whitefly *Bemisia tabaci*. *Insect Biochem. Mol.* **88**, 21–29 (2017).

Methods

Plants. *Arabidopsis thaliana* accession Col-0 was cultivated in a growth chamber under short-day conditions (10:14 h, light:dark) at 21 °C and 50–60% relative humidity. Brussels sprout (*Brassica oleracea* var. *gemmifera*, cv. Franklin), broccoli (*Brassica oleracea* var. *italica*, cv. Autumn Spear), kale (*Brassica oleracea* var. *sabellica*, cv. Dwarf Green Curled), eggplant (*Solanum melongena*, cv. Black Beauty), potato (*Solanum tuberosum*, cv. Marabel), cotton (*Gossypium hirsutum*, cv. Acala) and cassava (*Manihot esculenta*, cv. MCol22) plants were grown under standard greenhouse conditions at 26 ± 2 °C, with supplemental lighting at a photoperiod of 14:10 h (light:dark).

Insects. *Bemisia tabaci* (Hemiptera: Aleyrodidae) species MEAM1 were collected in southern Israel in 2003 and from Sudan in the late 1990s, then reared continuously on cotton. *B. tabaci* of species SSA1-SG3 (sub-Saharan Africa species group) were collected on Bagamoyo Road, Tanzania in 2013 and reared on cassava plants. *Aleyrodes proletella* (Hemiptera: Aleyrodidae) were collected in Kent (UK) and reared on kale. *Myzus persicae* (Hemiptera: Aphididae) were from a culture initiated from individuals collected in Hannover, Germany in 2005 and reared on *A. thaliana* Col-0. *Planococcus citri* (Hemiptera: Pseudococcidae) were collected in Israel in 2016 from cotton plants and reared on potato tubers. *Schistocerca gregaria* (Orthoptera: Acrididae) were provided by A. Ayali and maintained as previously described[38]. *Tetranychus cinnabarinus* (Trombidiformes: Tetranychidae) were obtained from the Israel Cohen Institute for Biological Control (Rehovot, Israel) and reared on bean plants. These insects were reared under standard greenhouse conditions of 26 ± 2 °C, with supplemental lighting at a photoperiod of 14:10 h (light:dark). *Spodoptera littoralis* (Lepidoptera: Noctuidae), *Helicoverpa armigera* (Lepidoptera: Noctuidae), *Mamestra brassicae* (Lepidoptera: Noctuidae), *Trichoplusia ni* (Lepidoptera: Noctuidae), *Pieris rapae* (Lepidoptera: Pieridae), *Plutella xylostella* (Lepidoptera: Plutellidae), *Diabrotica balteata* (Coleoptera: Chrysomelidae) and *Phyllotreta striolata* (Coleoptera: Chrysomelidae) were collected and maintained as described previously[39–41].

Insect feeding on *A. thaliana* and other species. Groups of adult whiteflies (50–100 individuals), mealybugs (20 individuals), aphids (>100 individuals) and locusts (10 individuals) were collected from their host plants and switched to *A. thaliana* Col-0 plants. *A. thaliana* plants were used before bolting when the rosette had 5–7 leaves. Insects were fed on leaves enclosed within Petri dishes with aluminum foil beneath the plant leaf. After 72–96 h, the honeydew and feces deposited on the aluminum foil were washed with water:methanol (20:80, vol/vol). *B. tabaci* whiteflies were also raised on Brussels sprouts (MEAM1 species) or cassava (SSA1-SG3 species). Nymph honeydew was collected under binoculars into a glass vial containing water:methanol (20:80, vol/vol). Groups of 200 adult mites were transferred from beans to *A. thaliana* and their feces were collected under binoculars into a glass vial containing water:methanol (20:80, vol/vol). Collection and extraction of feces from lepidopteran (three larvae, at third or fourth larval stages) and coleopteran (three adult insects) herbivores were performed as previously described[39–41].

***B. tabaci* feeding on artificial diets with different glucosylated protoxins.** Groups of 150 *B. tabaci* (MEAM1 species) adults were collected from cotton plants and switched to artificial diet feeders (consisting of a glass tube, 3 cm height × 2 cm diameter, with a liquid diet covered with a double layer of Parafilm). Insects were allowed to feed through the Parafilm on a 10% sucrose solution containing no additives (control) or 4-methylsulfinylbutyl GSL (4msob, from *Brassica oleracea* seeds) at a concentration of 5 mM. Other glucosides, including linamarin, DIMBOA-Glc, quercitrin, catalpol or esculin were also tested using this concentration. After 96 h, the honeydew deposited on the glass tubes and the aluminum foil was washed with water:methanol (20:80, vol/vol).

***B. tabaci* feeding on artificial diets containing isotopically labeled sugars.** Four different sucrose isotopologues were added to artificial diets: [$^{12}C_{12}$]sucrose, [$^{13}C_{12}$]sucrose, [glucose-$^{13}C_6$]sucrose and [fructose-$^{13}C_6$]sucrose. The monosaccharides [$^{13}C_6$]fructose and [$^{13}C_6$]glucose were also fed. The artificial feeding devices consisted of a glass tube (5 cm high × 2.5 cm diameter) with the liquid diet (50 µl) held within a double layer of Parafilm. About 50 *B. tabaci* MEAM1 adults were placed in each tube. Feeding assays were performed for 72 h on diets that contained 5 mM 4msob-GSL and 0.29 M labeled sugars. After feeding, the vials were held at −80 °C to kill the whiteflies, and the bodies were transferred and stored for analysis. The honeydew deposited on the glass tubes was washed with water:methanol (20:80, vol/vol) and stored at −20 °C until analysis. A full summary of the artificial diet constituents is provided in Supplementary Table 5.

Purification and LC-MS analysis of glucosylated glucosinolates, glucosinolate hydrolysis products and sugars. Purification of glucosylated glucosinolates was performed via fractionation on a Nucleodur Sphinx RP column (250 × 4.6 mm, 5 µm, Macherey–Nagel) using an HP 1200 HPLC system (Agilent Technologies) coupled to a fraction collector (Advantec). Chromatographic separation was achieved using a gradient of 0.05% aqueous formic acid (solvent A) and acetonitrile (solvent B) at a flow rate of 1 ml min^{-1} at 25 °C as follows: 0.5% B (9 min), 0.5–6% B (8 min), 6–7% B (2 min), 7–100% B (0.1 min), a 3.8-min hold at 100% B, 100–0.5% B (0.1 min) and a 6-min hold at 0.5% B.

Qualitative analysis of glucosylated glucosinolates in feces and honeydew extracts was also performed on an HP 1100 series HPLC system. Separation was achieved on a Nucleodur Sphinx RP column (250 × 4.6 mm, 5 µm; Macherey–Nagel) with a gradient of 0.2% aqueous formic acid (solvent A) and acetonitrile (solvent B) with a flow rate of 1 ml min^{-1} at 25 °C as follows: 0.5% B (9 min), 0.5–6% B (8 min), 6–10% B (23 min), 10–50% B (14 min), 50–70% B (3.9 min), 70–100% B (0.1 min), a 3-min hold at 100% B, 100–0.5% B (0.1 min) and a 3.9-min hold at 0.5% B. The HPLC was coupled to an Esquire 6000 ESI-Ion Trap mass spectrometer (Bruker Daltonics) operated in both positive and negative modes in the range of *m/z* 50–1,500, with a skimmer voltage of −40 V, capillary exit voltage of −146.7 V, capillary voltage of 4,000 V, nebulizer pressure of 35 p.s.i., drying gas rate of 11 l min^{-1} and gas temperature of 330 °C. DataAnalysis software V4 (Bruker Daltonics) was used for chromatogram analysis.

Qualitative analysis of isotopically labeled glucosylated glucosinolates in feces and honeydew extracts was performed on an HP 1100 series HPLC system. Separation was achieved on a Nucleodur Sphinx RP column (250 × 4.6 mm, 5 µm, Macherey–Nagel) with a gradient of 0.2% aqueous formic acid (solvent A) and acetonitrile (solvent B) with a flow rate of 1 ml min^{-1} at 25 °C as follows: 5–55% B (25 min), 55–100% B (0.1 min), 100% B 0.9-min hold, 100–5% B (0.1 min), 5% B 3.9-min hold. The HPLC device was coupled to an Esquire 6000 ESI-ion trap mass spectrometer (Bruker Daltonics) operated in both positive and negative modes in the range of *m/z* 60–1,400, with skimmer voltage of −40 V, capillary exit voltage of −128.5 V, capillary voltage of 4,000 V, nebulizer pressure of 35 p.s.i., drying gas rate of 11 l min^{-1} and gas temperature of 330 °C. DataAnalysis software V4 (Bruker Daltonics) was used for chromatogram analysis.

Quantification of the glucosylated glucosinolates in transglucosidation and myrosinase assays and in honeydew was accomplished using an HP 1260 series HPLC system coupled to an AB Sciex API 5000 mass spectrometer (Applied Biosystems). The column was a Nucleodur Sphinx RP column (250 × 4.6 mm, 5 µm, Macherey–Nagel) and used a chromatographic gradient of 0.05% aqueous formic acid (solvent A) and acetonitrile (solvent B) with a flow rate of 1 ml min^{-1} at 25 °C as follows: 1.5% B (2 min), 1.5–10% B (2.5 min), 10–40% B (7.5 min), 40–70% B (5 min), 70–100% B (0.1 min), hold at 100% B (2.4 min), 100–1.5% B (0.1 min) and hold at 1.5% B (3.9 min). The mass spectrometer was operated in negative mode with a collision gas value of 2, curtain gas pressure of 35 p.s.i., spray gas pressures of 70 p.s.i., ion spray voltage of −4,500 V and turbogas temperature of 700 °C. Compounds were detected using MRM detection with the parameters outlined in Supplementary Table 6. Quantification was achieved using external calibration curves constructed from solutions of purified glucosylated glucosinolates of known concentrations (determined in solution via NMR as described below). Analyst 1.5 software (Applied Biosystems) was used for data acquisition and processing. Averages and standard errors were calculated from three independent biological replicates. No other statistical tests were performed.

Analysis of 4msob-GSL hydrolysis products in concentrated honeydew from *B. tabaci* MEAM1 fed on broccoli and Brussels sprouts was performed on an HP 1260 HPLC system coupled to an AB Sciex API 5000 mass spectrometer. The column utilized was a Agilent XDB-C18 column (50 × 4.6 mm, 1.8 µm, Agilent Technologies) and used a chromatographic gradient of 0.05% aqueous formic acid (solvent A) and acetonitrile (solvent B) with a flow rate of 1.1 ml min^{-1} at 25 °C as follows: 3–15% B (0.5 min), 15–85% B (2.0 min), 85–100% B (0.1 min), hold of 100% B for 0.9 min, 100–3% B (0.1 min) and a hold at 3% B (2.4 min). The mass spectrometer was operated in positive mode with collision gas value of 4, curtain gas pressure of 35 p.s.i., spray gas pressures of 60 p.s.i., ion spray voltage of 5,500 V and turbogas temperature of 700 °C. Compounds were detected using scheduled MRM detection with the parameters outlined in Supplementary Table 6. Analyst 1.5 software was used for data acquisition and processing.

Sugar products from the transglucosidase enzyme assays were analyzed on an HP 1260 system coupled to an AB Sciex API 5000 mass spectrometer with an apHeraNH2 polymer column (150 × 4.6 mm, 5 µm, Supelco Analytical) with a chromatographic gradient of water (solvent A) and acetonitrile (solvent B) at a flow rate of 1 ml min^{-1} at 20 °C as follows: 80% B (0.5 min), 80–55% B (12.5 min), 55–80% B (1 min), hold at 80% for 4 min. The mass spectrometer was operated in negative mode with a collision gas value of 2, curtain gas pressure of 35 p.s.i., spray gas pressure of 70 p.s.i., ion spray voltage of −4,500 V and turbogas temperature of 700 °C. Compounds were detected using scheduled MRM detection with the parameters outlined in Supplementary Table 6. Analyst 1.5 software was used for data acquisition and processing.

NMR spectroscopy. NMR spectra (^{1}H, ^{1}H-^{1}H COSY, ^{1}H-^{13}C HSQC, ^{1}H-^{13}C HMBC and ^{1}H-^{1}H SELTOCSY) were acquired on a 700-MHz Avance III HD spectrometer equipped with a 1.7-mm cryoprobe (Bruker Biospin). Further information is provided in the Supplementary Information. Data acquisition and processing were accomplished using TopSpin version 3.2 (Bruker Biospin). Samples were measured in MeOH-d_3 or D_2O, as indicated, at 293 K. For quantification via ^{1}H NMR (10-s delay between scans), purified compounds were dried under a N_2 flow and resuspended in D_2O and a sucrose solution (3.13 mM in D_2O) was used as an external quantification standard.

Phylogenetic analysis of glycoside hydrolase family 13 proteins. Protein sequences of species representing different orders and families of arthropod herbivores, including *A. pisum* (Hemiptera: Aphididae), *B. tabaci* (MEAM1) (Hemiptera: Aleyrodidae), *Bombyx mori* (Lepidoptera: Bombycidae), *Frankliniella occidentalis* (Thysanoptera: Thripidae), *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae), *Manduca sexta* (Lepidoptera: Sphingidae), *M. persicae* (Hemiptera: Aphididae), *P. rapae* (Lepidoptera: Pieridae) and *T. urticae* (Trombidiformes: Tetranychidae), were downloaded from open databases (Supplementary Table 7). Putative GH13 protein sequences were identified in the downloaded proteomes using the pfam alpha-amylase domain (PF00128)[42] as a query for hmmsearch from the HMMER 3.1b1 software package (http://hmmer.org/ (2013)). The relevant amino acid sequences were then aligned using MUSCLE with default parameters[43].

The alignment was quality-trimmed, obtaining a final alignment of 456 positions, with trimal v1.4 automated1 option[44]. Model Finder selected VT + R7 (heterogenity free rate with seven categories)[45]. Maximum likelihood tree inference was conducted with IQ-TREE multicore version 1.5.5 with 1,000 ultrafast bootstrap[46] and 5,000 SH-aLRT[47] (Shimodaira–Hasegawa-like approximate likelihood ratio test) as node support values[46,48]. Tree representation was performed in Figtree version 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

After identifying putative GH13 domain-containing proteins, two *B. tabaci* MEAM1 RNA-Seq paired libraries were downloaded from the SRA-NCBI: adult midguts (SRX272314)[49] and adult whole bodies (SRX022878)[49]. Trimmomatic v0.33[50] (TruSeq2-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:35) was used to clip/trim bad sequences and generate clean libraries. Both clean libraries and the *B. tabaci* MEAM1annotation v1.1 (http://www.whiteflygenomics.org, last accessed 19 March 2020) were used as queries and reference, respectively. RSEM[51] was used to extract all coding genes from the *B. tabaci* MEAM1 GFF annotation file and non-GH13-encoding putative genes were filtered out. Gene expression quantification as fragments per kilobase million (FPKM) was performed with RSEM in paired mode (--paired-end --bowtie2). Finally, genes of interest (*SUC1–5*) were selected and their expression in the midgut[52] was compared against the whole body (Supplementary Table 4).

The list of proteins from each species was analyzed to identify the corresponding GH13 subfamilies. The search was done using the dbCAN2 meta server[53] and the dbCAN CAZyme HMM database[53,54] for identifying CAZyme domain boundaries[55] with HMMER v3.3. A default threshold of iE-value $<1\times10^{-15}$ and coverage >0.35 was used for protein classification. For some of the proteins that did not reach the above-mentioned threshold, a more relaxed threshold was applied using the coverage parameter >0.3 (ref. [56]). Even after the relaxed threshold was applied, the coverage parameter of some proteins was below the threshold. These proteins presented open reading frames that were shorter than the HMM GH13 subfamily profile. Therefore, these proteins were assigned to a subfamily based on their phylogenetic location (Extended Data Fig. 3). The HMMER classification of each protein was compared to the phylogenetic assignment in the tree (Extended Data Fig. 3). Cases in which conflicts were identified between the phylogenetic assignment and the HMMER classification are indicated in the comments of Supplementary Table 8 and also in some cases in the tree (Extended Data Fig. 3). The different enzyme activities and substrate specificities[57–60] are described in Supplementary Table 9.

Cloning and expression in *Drosophila* S2 cells. Vectors, cells and cell media were obtained from ThermoFisher Scientific. Candidate GH13-encoding genes were amplified from *B. tabaci* MEAM1 whitefly guts and whole-body cDNA using primer sets outlined in Supplementary Table 10, cloned into PCR4 Blunt TOPO and released using corresponding restriction enzymes according to each primer. The digested fragment was further ligated into the pAc5.1/V5-His A vector for expression in *Drosophila* S2 cells using Schneider's medium. Insect cells were co-transformed with pCOBLAST vector for selection via blasticidin. Transformation was achieved via $CaCl_2$ incubation following the manufacturer's standard protocols and selected with 50 μg ml^{-1} blasticidin in six-well plates. After one week of selection, cells were maintained in T-flasks at 27 °C with 10 μg ml^{-1} blasticidin. Cells were regularly maintained by splitting 1:20 when cells had achieved a confluency of 90% (approximately every 4–5 days). Cells and media were collected for enzyme activity studies when cells reached 90% confluency.

Enzyme assays. *Drosophila* S2 cells producing candidate GH13 enzymes as well as non-transfected control cells were centrifuged at 100*g* for 5 min. The resulting supernatant was utilized as the secreted protein fraction for enzyme assays. For determination of 4msob-GSL transglucosidation activity, 5 μl of each supernatant containing an equivalent total protein content (verified by Bradford assays) was mixed with 5 μl 50 mM phosphate buffer at pH 7.0 containing 2 M sucrose and 5 mM 4msob-GSL, and reacted for 6 h at 25 °C with no stirring. Assay conditions were chosen after screening a range of substrate concentrations and pH values that reflect previous work on glycoside hydrolases and phloem-feeding insects[20,25]. The reaction was stopped using 15 μl of methanol and immediately stored at −20 °C. Before analysis, enzyme assays were centrifuged at 5,200*g* for 5 min and the supernatant obtained was analyzed by LC-MS. A dilution of 1:10 in water was used for analysis of glucosylated 4msob-GSL and a dilution of 1:1,000 in water for analyzing sugars such as glucose, fructose and trisaccharides formed from these incubations.

For K_M estimation, crude enzyme preparations were assumed to catalyze no other reactions on the substrates supplied except transglucosidation and sucrose hydrolysis. Enzyme assays were performed in the same manner as described above, except incubation times were reduced to 30 min. Peak areas were integrated and compared based on duplicate analysis. Sucrose K_M determination was performed at a constant 4msob-GSL concentration of 2.5 mM with sucrose concentrations ranging from 0.05 to 1 M. The 4msob-GSL K_M determination utilized sucrose at a final concentration of 1 M and a range of 4msob-GSL concentrations from 25 μM to 20 mM. All assays were carried out under linear reaction conditions with respect to time and protein concentration. Substrate concentration was never reduced below 95% during the 30-min assay period. K_M calculations were performed in SigmaPlot 12.0.

For tests on the reactivity of glucosylated GSL with plant myrosinase, purified 4msob-GSL (10 μl of a 5 mM solution) and honeydew from *A. thaliana* Col-0-reared adult MEAM1 whiteflies containing insect-derived glucosylated 4msob-GSL glycosides were mixed with a 20 mM phosphate buffer solution (10 μl), pH 7.0, containing one unit of *S. alba* myrosinase (Sigma Aldrich). Negative control reactions were supplemented with 5 μl of water instead of the myrosinase solution. Reactions were incubated at room temperature with no stirring, stopped after 1 h with 20 μl acetic acid and stored frozen until LC-MS analysis.

Statistics and reproducibility. All experimental repetitions (≥2) and biological replicates analyzed showed consistent results.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this Article.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding authors upon reasonable request. Public databases used in the construction of the protein phylogenetic tree are provided in Supplementary Table 7 and are available from the following websites: HGSC (https://www.hgsc.bcm.edu/arthropods/colorado-potato-beetle-genome-project), BIPAA (https://bipaa.genouest.org/sp/acyrthosiphon_pisum/), Whitefly Genome Database (http://www.whiteflygenomics.org/cgi-bin/bta/index.cgi), BIPAA (https://bipaa.genouest.org/sp/myzus_persicae/), NCBI (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bombyx_mori/101/), HGSC (https://www.hgsc.bcm.edu/arthropods/tobacco-hornworm-genome-project), NCBI (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Pieris_rapae/100/), NCBI (https://www.ncbi.nlm.nih.gov/assembly/GCA_000697945.4), Ensembl (http://metazoa.ensembl.org/Tetranychus_urticae/Info/Index) and dbCAN (http://bcb.unl.edu/dbCAN/). All other data supporting this work, if not already indicated, are available in the Supplementary Information.

References

38. Guershon, M. & Ayali, A. Innate phase behavior in the desert locust, *Schistocerca gregaria*. *Insect Sci.* **19**, 649–656 (2012).
39. Beran, F. et al. *Phyllotreta striolata* flea beetles use host plant defense compounds to create their own glucosinolate-myrosinase system. *Proc. Natl Acad. Sci. USA* **111**, 7349–7354 (2014).
40. Jeschke, V. et al. How glucosinolates affect generalist Lepidopteran larvae: growth, development and glucosinolate metabolism. *Front. Plant Sci.* **8**, 1995 (2017).
41. Robert, C. A. M. et al. A specialist root herbivore exploits defensive metabolites to locate nutritious tissues. *Ecol. Lett.* **15**, 55–64 (2012).
42. Finn, R. D. et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285 (2016).
43. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
44. Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
45. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
46. Minh, B. Q., Nguyen, M. A. T. & von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* **30**, 1188–1195 (2013).
47. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
48. Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
49. Wang, X. W. et al. Transcriptome analysis and comparison reveal divergence between two invasive whitefly cryptic species. *BMC Genomics* **12**, 458 (2011).

50. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
51. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
52. Ye, X. D. Transcriptomic analyses reveal the adaptive features and biological differences of guts from two invasive whitefly species. *BMC Genomics* **15**, 370 (2014).
53. Zhang, H. et al. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **46**, W95–W101 (2018).
54. Yin, Y. B. et al. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **40**, W445–W451 (2012).
55. Cantarel, B. L. et al. The carbohydrate-active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* **37**, D233–D238 (2009).
56. Sablok, G., Kumar, S., Ueno, S., Kuo, J. & Varotto, C. (eds) *Advances in the Understanding of Biological Sciences Using Next Generation Sequencing* (*NGS*) *Approaches* (Springer, 2015).
57. Franceus, J. & Desmet, T. Sucrose phosphorylase and related enzymes in glycoside hydrolase family 13: discovery, application and engineering. *Int. J. Mol. Sci.* **21**, 2526 (2020).
58. Majzlova, K., Pukajova, Z. & Janecek, S. Tracing the evolution of the α-amylase subfamily GH13_36 covering the amylolytic enzymes intermediate between oligo-1,6-glucosidases and neopullulanases. *Carbohydr. Res.* **367**, 48–57 (2013).
59. Oslancova, A. & Janecek, S. Oligo-1,6-glucosidase and neopullulanase enzyme subfamilies from the α-amylase family defined by the fifth conserved sequence region. *Cell. Mol. Life Sci.* **59**, 1945–1959 (2002).
60. Stam, M. R., Danchin, E. G. J., Rancurel, C., Coutinho, P. M. & Henrissat, B. Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of α-amylase-related proteins. *Protein Eng. Des. Sel.* **19**, 555–562 (2006).

Acknowledgements

We thank A. Douglas (Cornell University) for the SUC1 sequence, K. Falk for assistance with graphics, the MPI-CE, DSMZ and HUJI greenhouse teams for plant and insect maintenance, and other members of the African Cassava Whitefly Project (cassavawhitefly.org) for helpful discussions. This work was supported financially by the Max Planck Society, the Deutsche Forschungsgemeinschaft (DFG Collaborative Research Center 1127 ChemBioSys) and the Natural Resources Institute, University of Greenwich from a grant provided by the Bill and Melinda Gates Foundation (OPP1058938).

Author contributions

The authors contributed in the following manner: conceptualization (O.M., M.L.A.E.E., J.G., S.M. and D.G.V.); direction and supervision (D.G.V., J.G., O.M. and S.M.); funding acquisition (J.G., J.C., S.M. and D.G.V.); provision of insects and genetic materials (O.M., M.G., B.S., S.W., K.L., A.S., L.L.M., J.C. and S.S.); chemical analysis of honeydew by LC-MS (M.L.A.E.E., D.G.V., M.R., A.S. and O.M.); purification of insect-derived products (M.L.A.E.E. and D.G.V.); execution and interpretation of the NMR analyses of the purified products (C.P.); carrying out labeled artificial diet feedings (M.L.A.E.E., D.G.V., M.G., B.S. and S.W.); performance of phylogenetic analysis (K.J. and O.M.); estimation of gene expression (D.S.-G. and O.M.); cloning of candidate GH13 genes, generation of recombinant proteins and performance of enzyme assays (M.L.A.E.E., with contributions from K.L.); development of methods (M.L.A.E.E., D.G.V., M.R., C.P. and O.M.); drafting of the manuscript (M.L.A.E.E. and D.G.V., with input from J.G., O.M. and S.M.).

Competing interests

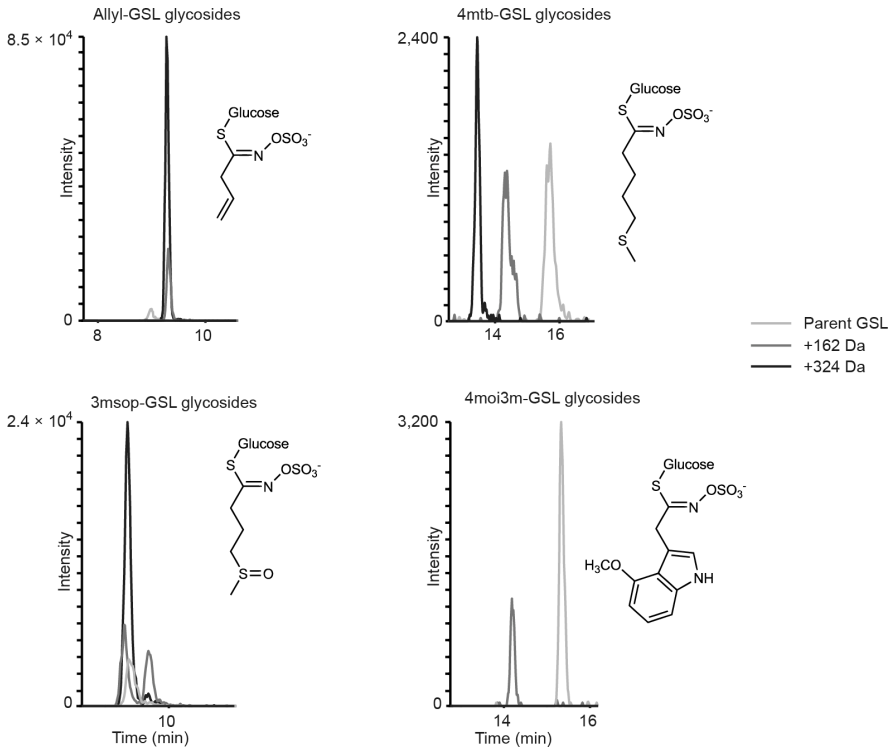
The authors declare no competing interests.

Additional information

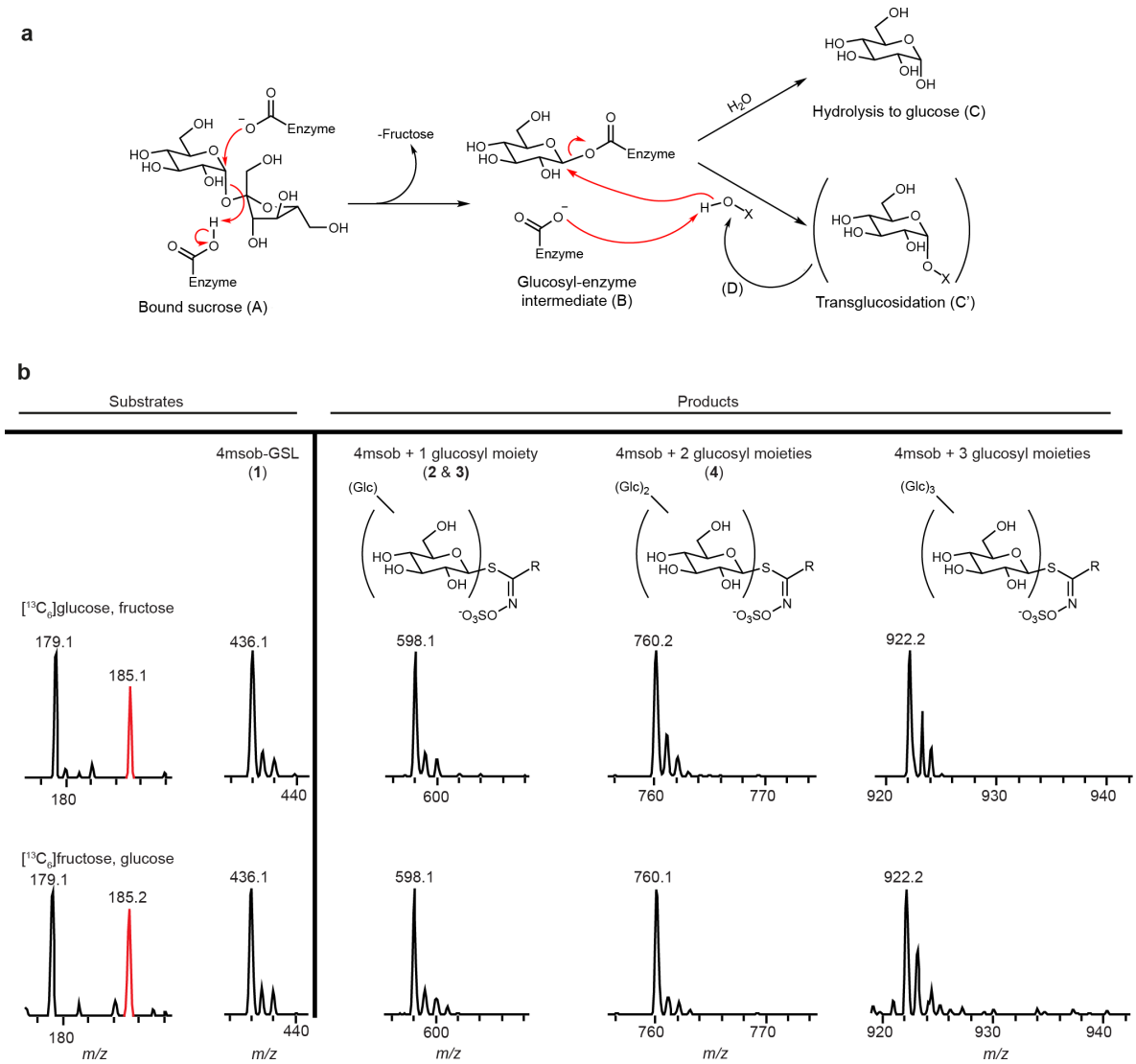
Extended data is available for this paper at https://doi.org/10.1038/s41589-020-00658-6.

Supplementary information is available for this paper at https://doi.org/10.1038/s41589-020-00658-6.

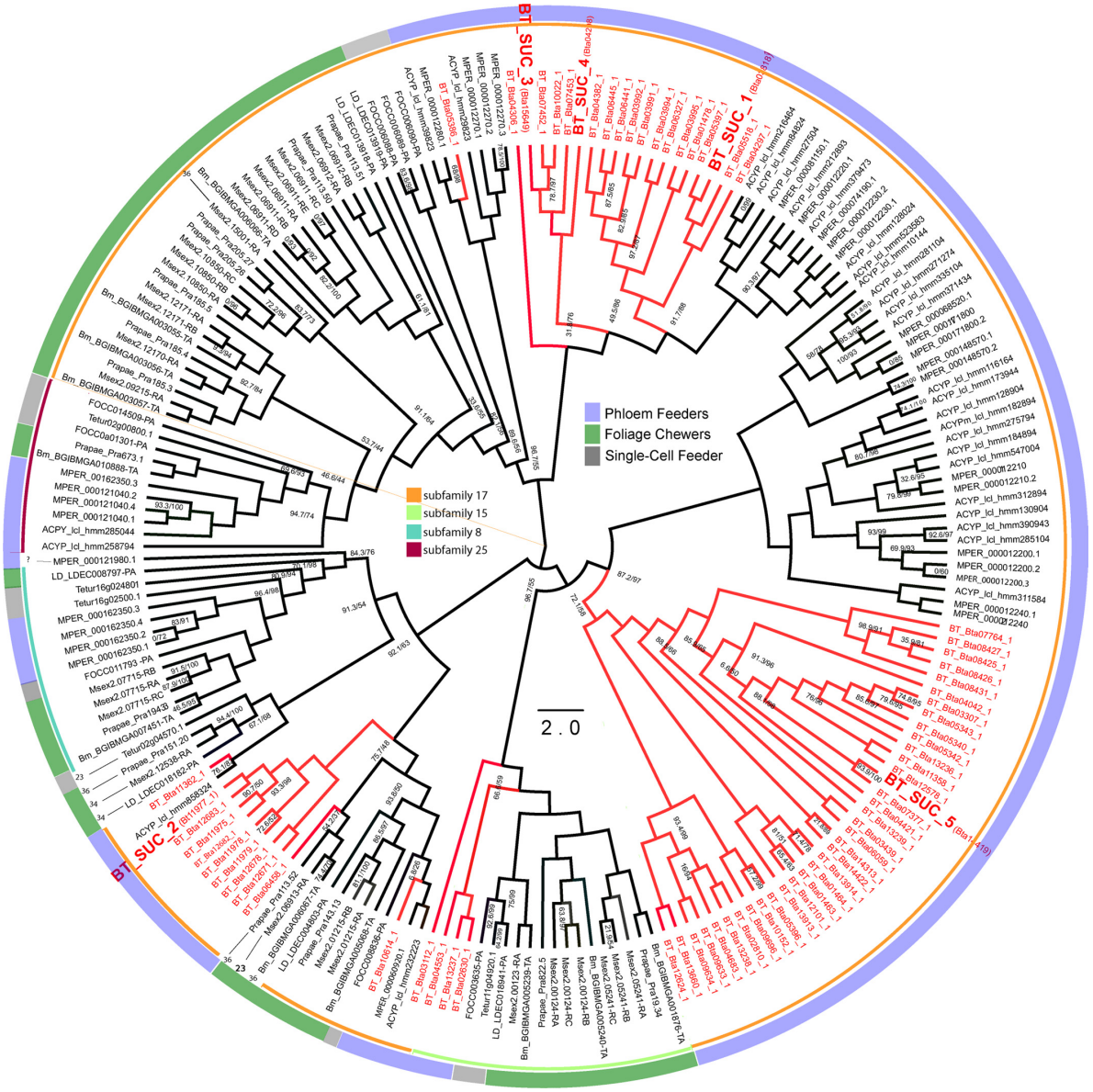
Correspondence and requests for materials should be addressed to O.M. or D.G.V.



Extended Data Fig. 1 | *Bemisia tabaci* glucosylates GSLs with various side chains. LC-MS traces of allyl-GSL, 4mtb [4-methylthiobutyl]-GSL, 3msop [3-methylsulfinylpropyl]-GSL and 4moi3m [4-methoxyindolyl-3-methyl]-GSL and their glycosides detected in the combined honeydew of 50-100 adult *Bemisia tabaci* MEAM1 whiteflies feeding on GSL-containing plants (kale or *A. thaliana* Col-0). The detected parent GSL is indicated with structure and representative color for the mass spectral trace (light gray), and the dectected subsequent glycosides represented as +162 Da and +324 Da (dark gray and black respectively).



Extended Data Fig. 2 | GSL glucosylation in the whitefly *Bemisia tabaci* is catalyzed by a transglucosidase activity. a, Simplified reaction mechanism of a sucrase-transglucosidase showing the two competing reaction paths: After binding of sucrose to the enzyme (A), hydrolysis of the fructose residue occurs with retention of bound glucose (B). Glucose is released (C) when sucrose concentrations are low, while transglucosidation to an acceptor (C') occurs when acceptor concentrations are sufficiently high. This product may undergo further transglucosidation (D). **b**, Depiction of the results from two of the five diets not shown in Fig. 3, those diets with the ^{13}C-labeled monosaccharides glucose and fructose. None gave labeled glycosylated GSLs, unlike feeding with sucrose labeled in the glucose portion. The results are consistent with a transglucosidase activity that initially hydrolyzes sucrose and links the resulting glucose moiety to the plant GSL.



Extended Data Fig. 3 | Maximum likelihood circular cladogram showing the relationship of glycoside hydrolase family 13 enzymes from nine chosen herbivore species. The tree was inferred using a total of 205 sequences. Ultrafast bootstrap[45] and Shimodaira–Hasegawa approximate likelihood ratio test (SH-aLRT)[46] validation values lower than 95 are presented close to the corresponding nodes. GH13 members from *Bemisia tabaci* are highlighted by red nodes. SUC1–5 are indicated by bold text in red. Colors surrounding the cladogram indicate the feeding guild of the corresponding species. Thin colored inner circles specify the subfamily of the corresponding GH13 enzymes. Subfamilies with less than four proteins are marked only by the subfamily number (for more details see Supplementary Table 8). The protein sequences are named according to their GenBank accession numbers, or their names in the released proteome. Species name abbreviations are as indicated: Ld, *Leptinotarsa decemlineata*; ACYP, *Acyrthosiphon pisum*; BT, *Bemisia tabaci*; MPER, *Myzus persicae*; Bm, *Bombyx mori*; Msex, *Manduca sexta*; Prapae, *Pieris rapae*; FOCC, *Frankliniella occidentalis*; Tetur, *Tetranychus urticae*.

nature research

Corresponding author(s): Daniel Giddings Vassão, Osnat Malka

Last updated by author(s): Jul 23, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Mass Spectrometry: Esquire 6000 ESI-Ion Trap mass spectrometer (Bruker Daltonics), AB Sciex API 5000 mass spectrometer (Applied Biosystems) NMR: 700 MHz Avance III HD spectrometer equipped with a 1.7 mm cryoprobe (Bruker Biospin)
Data analysis	Mass Spectrometry: DataAnalysis software V4 (Bruker Daltonics), Analyst 1.5 software (Applied Biosystems) NMR: TopSpin ver. 3.2 (Bruker Biospin) Phylogeny: HMMER 3.1b1 software package, MUSCLE with default parameters, trimal v1.4 automated1 option, Model Finder selected VT +R7 (heterogenity free rate with 7 categories), IQ-TREE multicore version 1.5.5 with 1000 ultrafast bootstrap, Figtree version 1.4.3, RSEM v1.3.1, Trimmomaticv0.33, HMMER v3.3 Enzyme Kinetics: SigmaPlot 12.0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding authors upon reasonable request. Public databases used in the construction of the protein phylogenetic tree are provided in Supplementary table 7 and available from the following websites: HGSC (https://www.hgsc.bcm.edu/arthropods/colorado-potato-beetle-genome-project), BIPAA (https://bipaa.genouest.org/sp/acyrthosiphon_pisum/), Whitefly Genome

Database (http://www.whiteflygenomics.org/cgi-bin/bta/index.cgi), BIPAA (https://bipaa.genouest.org/sp/myzus_persicae/), NCBI (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Bombyx_mori/101/), HGSC (https://www.hgsc.bcm.edu/arthropods/tobacco-hornworm-genome-project), NCBI (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Pieris_rapae/100/), NCBI (https://www.ncbi.nlm.nih.gov/assembly/GCA_000697945.4), Ensembl (http://metazoa.ensembl.org/Tetranychus_urticae/Info/Index) dbCAN (http://bcb.unl.edu/dbCAN/). All other data supporting this work, if not already indicated, are available in the Supplementary Materials.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes for frass or honeydew analysis were chosen based on the amount of material needed to produce consistent and detectable MS signals. Numbers of replicates used (>=3) were sufficient, as no statistical analyses were necessary.
Data exclusions	No data were excluded.
Replication	All experimental repetitions (>=2) and biological replicates (>=3) analyzed showed consistent results.
Randomization	Insects were allocated randomly among replicates and treatments.
Blinding	Blinding not used; due to the randomization of insects among groups, and quantification of chemical compounds based on data directly extracted from chromatograms, this was not deemed necessary.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Drosophila S2 cells: Invitrogen Gibco, derived from a culture of late stage Drosophila melanogaster embryos.
Authentication	Cells lines were not authenticated.
Mycoplasma contamination	The cell line was tested for contamination by the manufacturer using a Mycoplasma qPCR Detection Assay.
Commonly misidentified lines (See ICLAC register)	None of the commonly misidentified cell lines listed were used in this study.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Insects were used in this study.

Wild animals	The study did not involve wild animals.
Field-collected samples	No samples directly collected from the field were chemically analyzed in this study.
Ethics oversight	No ethical approval was necessary for the use of the insects in this study.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

nature research

https://doi.org/10.1038/s41589-020-00658-6

Supplementary information

Glucosylation prevents plant defense activation in phloem-feeding insects

In the format provided by the authors and unedited

SUPPLEMENTARY INFORMATION

Glucosylation prevents plant defense activation in phloem-feeding insects

Osnat Malka[§*], Michael L.A.E. Easson[§], Christian Paetz, Monika Götz, Michael Reichelt, Beate Stein, Katrin Luck, Aleksa Stanisic, Ksenia Juravel, Diego Santos-Garcia, Lilach L. Mondaca, Simon Springate, John Colvin, Stephan Winter, Jonathan Gershenzon, Shai Morin, Daniel G. Vassão[*]

[§]These authors contributed equally to this work

*Corresponding authors: DGV (vassao@ice.mpg.de), OM (osnat226@gmail.com)

Supplementary Table 1: **^{1}H and ^{13}C NMR shifts and coupling constants for glucosylated glucosinolates.**

	Monoglucosylated α-(1,6)-4msob-GSL (2)[a]			Diglucosylated α-(1,6), α-(1,6)-4msob-GSL (4)[a]		
	δH	mult., *J* [Hz]	δC	δH	mult., *J* [Hz]	δC
Aglycone						
0	-	-	163.0	-	-	163.0
1	2.72	dd, 6.5/6.5	31.6	2.72	dd, 6.5/6.5	31.6
2	1.83	m	25.3	1.83	m	25.3
3	1.77	m	21.1	1.77	m	21.1
4	2.88	m	51.9	2.88	m	51.9
-S(O)CH3	2.62	s	36.2	2.62	s	36.2
Original glucose of GSL						
1'	5.00	d, 10.0	81.8	5.00	d, 10.0	81.8
2'	3.38	dd, 10.0/9.5	71,6	3.38	dd, 10.0/9.5	71,7
3'	3.49	dd, 9.5/9.5	77.1	3.49	m	77.1
4'	3.50	dd, 9.5/9.5	68.7	3.48	m	68.9
5'	3.70	m	78.3	3.72	m	78.3
6a'	3.88	dd, 12.0/5.0	65.5	3.90	bd, 12.0	65.5
6b'	3.69	bd, 12.0	65.5	3.70	bd, 12.0	65.5
First added glucose moiety						
1''	4.86	d, 3.5	97.8	4.87	d, 3.5	97.7
2''	3.46	dd, 9.6/3.5	71.3	3.48	dd, 9.6/3.5	71.3
3''	3.62	dd, 9.6/9.0	73.2	3.61	dd, 9.6/9.0	73.2
4''	3.35	dd, 9.5/9.0	69.4	3.43	dd, 9.5/9.0	69.4
5''	3.60	m	70.2	3.78	m	70.2
6a''	3.74	bd, 12.0	65.4	3.87	bd, 12.0	65.4
6b''	3.68	dd, 12.0/4.5	65.4	3.66	bd, 12.0	65.4
Second added glucose moiety						
1'''	-	-	-	4.88	d, 3.5	97.6
2'''	-	-	-	3.46	dd, 9.5/3.6	71,4
3'''	-	-	-	3.64	dd, 9.5/9.5	73.2
4'''	-	-	-	3.34	dd, 9.5/9.5	69.4
5'''	-	-	-	3.63	m	71.8
6a'''	-	-	-	3.77	bd, 12.0	60.4
6b'''	-	-	-	3.68	bd, 12.0	60.4

	Monoglucosylated α-(1,4)-4msob-GSL (3)[b]			4msob-GSL (1)[b]		
	δH	mult., *J* [Hz]	δC	δH	mult., *J* [Hz]	δC
Aglycone						
0	-	-	160.6	-	-	161.6
1a	2.80	ddd, 15.0/7.5/7.5	32.9	2.94	ddd, 13.0/6.7/9.2	33.0
1b	2.74	ddd, 15.0/7.5/7.5	32.9	2.83	ddd, 13.0/6.7/9.2	33.0
2	1.93	m	26.7	1.93	m	27.0
3	1.86	m	22.6	1.87	m	22.9
4a	2.88	m	54.1	2.82	ddd 15.5/7.5/7.5	54.2
4b				2.76	ddd 15.5/7.5/7.5	54.2
-S(O)CH3	2.64	s	37.7	2.65	s	38.1
Original glucose of GSL						
1'	4.89	d, 10.0	83.4	4.86	ovlp.	83.6
2'	3.31	dd, 9.5/9.0	73.7	3.25	dd, 10.0/9.5	74.3
3'	3.67	dd, 9.5/9.5	79.2	3.40	dd, 9.5/9.5	79.6
4'	3.55	dd, 9.5/9.5	80.2	3.29	dd, 9.5/9.5	71.4
5'	3.49	m	80.6	3.36	ddd, 9.5/6.5/2.0	82.3
6a'	3.90	dd, 12.5/2.5	62.2	3.86	dd, 12.3/2.0	62.9
6b'	3.76	dd, 12.5/5.0	62.2	3.62	dd, 12.3/6.5	62.9
First added glucose moiety						
1''	5.18	d, 3.6	102.6	-	-	-
2''	3.43	dd, 9.5/3.6	74.2	-	-	-
3''	3.61	dd, 9.5/9.5	75.1	-	-	-
4''	3.27	dd, 9.5/9.5	71.4	-	-	-
5''	3.67	m	74.7	-	-	-
6a''	3.82	m	62.5	-	-	-
6b''	3.66	m	62.5	-	-	-
Second added glucose moiety						
1'''	-	-	-	-	-	-
2'''	-	-	-	-	-	-
3'''	-	-	-	-	-	-
4'''	-	-	-	-	-	-
5'''	-	-	-	-	-	-
6a'''	-	-	-	-	-	-
6b'''	-	-	-	-	-	-

[a] Measured in D_2O

[b] Measured in MeOH-d_3

Supplementary Table 2: **Glucosylated GSLs are produced by phloem-feeding hemipterans, but not by representatives of other insect groups.** Ability of various selected insect species to produce glucosylated 4msob-GSLs when fed on *A. thaliana* Col-0 leaves containing 4msob-GSL as the major GSL. Honeydew or feces extracts were analyzed by LC-MS.

Class	Order		Species	Glucosylated 4msob-GSLs
Insecta	Hemiptera	Whiteflies	*Bemisia tabaci*	+
			Aleyrodes proletella	+
		Aphid	*Myzus persicae*	+
		Mealybug	*Planococcus citri*	+
	Coleoptera		*Diabrotica balteata*	–
			Phyllotreta striolata	–
	Lepidoptera		Six species (1)	–
	Orthoptera		*Schistocerca gregaria*	–
Arachnida	Trombidiformes		*Tetranychus cinnabarinus*	–

(1): *Helicoverpa armigera* , *Mamestra brassicae* , *Pieris rapae* , *Plutella xylostella* , *Spodoptera littoralis* , *Trichoplusia ni*

Supplementary Table 3: ***Bemisia tabaci* can glucosylate glucosinolates and cyanogenic glucosides, but not other plant glycosides tested.** Ability of *B. tabaci* to produce glucosylated derivatives of various plant glycosides fed on artificial diets. Resulting honeydews were screened by LC-MS.

	4msob-GSL	Linamarin	DIMBOA-Glc	Catalpol	Esculin	Quercitrin
	Glucosinolates	Cyanogenic glucoside	Benzoxazinoid glucoside	Iridoid glucoside	Hydroxycoumarin glucoside	Flavonoid glucoside
Glucosylated derivatives in honeydew	+	+	–	–	–	–

Supplementary Table 4: **Ratios of normalized *SUC1-5* gene expression.** Values are based on RNA-Seq SRX272314 (midgut tissue) and SRX022878 (whole body).

SUC Name	Bta Number	FPKM Ratio Midgut/All body
SUC1	Bta03818	2.74
SUC2	Bta11977	No detection in midgut
SUC3	Bta15649	5.54
SUC4	Bta04298	5.54
SUC5	Bta14419	4.63

FPKM (fragments per kilobase of exon model per million reads mapped)

Supplementary Table 5: **Constituents of artificial diets fed to *Bemisia tabaci* MEAM1.**

Diet number	Isotopically-labeled compound (concentration)	Unlabeled constituents (concentration)
1	[$^{13}C_{12}$] sucrose (0.145 M)	Sucrose (0.145 M) 4msob-GSL (5 mM)
2	[glucose-$^{13}C_6$] sucrose (0.145 M)	Sucrose (0.145 M) 4msob-GSL (5 mM)
3	[fructose-$^{13}C_6$] sucrose (0.145 M)	Sucrose (0.145 M) 4msob-GSL (5 mM)
4	[$^{13}C_6$] glucose (0.0725 M)	Sucrose (0.145 M) 4msob-GSL (5 mM) Fructose (0.0725 M)
5	[$^{13}C_6$] fructose (0.0725 M)	Sucrose (0.145 M) 4msob-GSL (5 mM) Glucose (0.0725 M)

Supplementary Table 6: **List of multiple reaction monitoring (MRM) parameters for individual compounds analyzed by LC-MS.**

Compound	Quadrupole 1 (precursor mass Da)	Quadrupole 2 (product mass Da)	Declustering potential (V)	Entrance potential (V)	Collision energy (V)	Collision cell exit potential (V)
4msob-ITC	178.11	114	51	5	13	4
4msob-GSH	485.11	179.1	76	5.5	29	6
4msob-CG	356.07	136.1	46	11	15	4
4msob-Cys	299.06	136.1	51	3	15	4
4msob-NAC	341.07	178.1	51	3	17	6
4msob-CN	146	129	63	10	13	4
3msop-GSL	421.8	95.9	-100	-4.5	-60	0
4msob-GSL	435.9	95.9	-100	-5	-60	0
4mtb-GSL	420	95.9	-100	-5	-58	0
4moi3m-GSL	477	95.9	-100	-12	-50	0
Allyl-GSL	358	95.9	-100	-5	-60	0
4msob-GSL + Glucose	598	95.9	-100	-5	-60	0
4msob-GSL + 2 Glucoses	760	95.9	-100	-5	-60	0
4msob-GSL + 3 Glucoses	922	95.9	-100	-5	-60	0
4moi3m-GSL + Glucose	639	95.9	-100	-12	-50	0
4moi3m-GSL + 2 Glucoses	801	95.9	-100	-12	-50	0
4mtb-GSL+ Glucose	582	95.9	-100	-5	-60	0
4mtb-GSL + 2 Glucoses	744	95.9	-100	-5	-60	0
Ally-GSL + Glucose	520	95.9	-100	-5	-60	0
Allyl-GSL + 2 Glucoses	682	95.9	-100	-5	-60	0
3msop-GSL + Glucose	584	95.9	-100	-5	-60	0
3msop-GSL+ 2 Glucoses	746	95.9	-100	-5	-60	0
Monosaccharide	178.8	89	-50	-9.5	-10	0
Disaccharide	340.9	59	-65	-10	-46	0
Trisaccahride	503.1	179	-95	-10	-28	-4
Tetrasaccharide	665.2	179	-100	-10	-48	-4

Supplementary Table 7: **List of insect protein databases searched for GH13 sequences used in the production of the phylogenetic tree.**

Order	Species	Number of GH13	Genome version	Database
Coleoptera	*Leptinotarsa decemlineata*	6	Maker annotation v0.5.3	HGSC
Hemiptera	*Acyrthosiphon pisum*	30	Version v1.0	BIPAA
Hemiptera	*Bemisia tabaci* (MEAM1)	71	Version v1.1	Whitefly genome database
Hemiptera	*Myzus persicae*	31	Version v1.0	BIPAA
Lepidoptera	*Bombyx mori*	11	ASM15162v1	NCBI
Lepidoptera	*Manduca sexta*	29	OGS2	HGSC
Lepidoptera	*Pieris rapae*	14	GCF_001856805.1	NCBI
Thysanoptera	*Frankliniella occidentalis*	8	Focc v1.0	NCBI
Trombidiformes	*Tetranychus urticae*	5	ASM23943v1	Ensembl

Supplementary Table 8: **List of GH13 proteins and their subfamilies used to produce the Maximum likelihood circular cladogram** (Extended Data Figure 3).

Species	[a]Protein ID	[ab]Subfamily	Hit location	[b]Comments
Bemisia tabaci (MEAM1)	Bta02630_1	GH13_15	58-309	
Bemisia tabaci (MEAM1)	Bta04553_1	GH13_15	56-316	
Bemisia tabaci (MEAM1)	Bta13237_1	GH13_15	5-219	
Bombyx mori	BGIBMGA001876-TA	GH13_15	108-398	
Bombyx mori	BGIBMGA005239-TA	GH13_15	56-341	
Bombyx mori	BGIBMGA005240-TA	GH13_15	52-342	
Frankliniella occidentalis	FOCC003635-PA	GH13_15	67-305	
Leptinotarsa decemlineata	LDEC018941-PA	GH13_15	55-335	
Manduca sexta	Msex2.00123-RA	GH13_15	54-339	
Manduca sexta	Msex2.00124-RA	GH13_15	54-344	
Manduca sexta	Msex2.00124-RB	GH13_15	54-344	
Manduca sexta	Msex2.00124-RC	GH13_15	54-344	
Manduca sexta	Msex2.05241-RA	GH13_15	53-343	
Manduca sexta	Msex2.05241-RB	GH13_15	1-262	
Manduca sexta	Msex2.05241-RC	GH13_15	159-449	
Pieris rapae	Pra19.34	GH13_15	55-342	
Pieris rapae	Pra822.5	GH13_15	52-339	

Tetranychus urticae	tetur11g04920.1	GH13_24 (15)	60-357	Based on tree location (HMMER prediction is for subfamily 15 coverage: 0.99, iE-value: 3.30E-106)
Acyrthosiphon pisum	lcl|hmm10144	GH13_17	40-386	
Acyrthosiphon pisum	lcl|hmm116164	GH13_17	57-219	
Acyrthosiphon pisum	lcl|hmm128024	GH13_17	1-336	
Acyrthosiphon pisum	lcl|hmm128904	GH13_17	2-233	
Acyrthosiphon pisum	lcl|hmm173944	GH13_17	1-145	
Acyrthosiphon pisum	lcl|hmm182894	GH13_17	56-339	
Acyrthosiphon pisum	lcl|hmm184894	GH13_17	202-557	
Acyrthosiphon pisum	lcl|hmm212893	GH13_17	60-409	
Acyrthosiphon pisum	lcl|hmm216464	GH13_17	192-531	
Acyrthosiphon pisum	lcl|hmm232223	GH13_17	119-341	
Acyrthosiphon pisum	lcl|hmm271274	GH13_17	43-391	
Acyrthosiphon pisum	lcl|hmm275794	GH13_17	57-412	
Acyrthosiphon pisum	lcl|hmm281104	GH13_17	71-416	
Acyrthosiphon pisum	lcl|hmm285104	GH13_17	1-231	
Acyrthosiphon pisum	lcl|hmm29823	GH13_17	57-407	
Acyrthosiphon pisum	lcl|hmm311584	GH13_17	73-389	
Acyrthosiphon pisum	lcl|hmm312894	GH13_17	-	Based on tree location (HMMER prediction is also for subfamily 17, but coverage: 0.17, iE-value: 1.60E-21)

Acyrthosiphon pisum	lcl|hmm335104	GH13_17	31-380	
Acyrthosiphon pisum	lcl|hmm371434	GH13_17	33-309	
Acyrthosiphon pisum	lcl|hmm379473	GH13_17	57-298	
Acyrthosiphon pisum	lcl|hmm390943	GH13_17	57-412	
Acyrthosiphon pisum	lcl|hmm39823	GH13_17	49-397	
Acyrthosiphon pisum	lcl|hmm523583	GH13_17	57-406	
Acyrthosiphon pisum	lcl|hmm84824	GH13_17	58-352	
Acyrthosiphon pisum	lcl|hmm858324	GH13_17	-	Based on tree location (HMMER prediction is also for subfamily 17, but coverage: 0.2, iE-value: 9.60E-11)
Acyrthosiphon pisum	lcl|hmm130904	GH13_17	1-108	
Acyrthosiphon pisum	lcl|hmm27504	GH13_17		Based on tree location (HMMER prediction is also for subfamily 17, but coverage: 0.12, iE-value: 3.30E-12)
Acyrthosiphon pisum	lcl|hmm547004	GH13_17	-	Based on tree location (HMMER prediction is also for subfamily 17, but coverage: 0.12, iE-value: 2.40E-09)
Bemisia tabaci (MEAM1)	Bta01464_1	GH13_17	9-167	
Bemisia tabaci (MEAM1)	Bta01478_1	GH13_17	107-474	
Bemisia tabaci (MEAM1)	Bta02810_1	GH13_17	25-371	
Bemisia tabaci (MEAM1)	Bta03112_1	GH13_17	1-337	
Bemisia tabaci (MEAM1)	Bta03307_1	GH13_17	89-434	
Bemisia tabaci (MEAM1)	Bta03439_1	GH13_17	52-395	
Bemisia tabaci (MEAM1)	Bta03818_1	GH13_17	53-401	

Bemisia tabaci (MEAM1)	Bta03991_1	GH13_17	51-401	
Bemisia tabaci (MEAM1)	Bta03992_1	GH13_17	52-402	
Bemisia tabaci (MEAM1)	Bta03994_1	GH13_17	62-430	
Bemisia tabaci (MEAM1)	Bta03995_1	GH13_17	619-856	
Bemisia tabaci (MEAM1)	Bta04042_1	GH13_17	58-217+17(245-367)	
Bemisia tabaci (MEAM1)	Bta04297_1	GH13_17	106-453	
Bemisia tabaci (MEAM1)	Bta04298_1	GH13_17	67-413	
Bemisia tabaci (MEAM1)	Bta04306_1	GH13_17	53-407	
Bemisia tabaci (MEAM1)	Bta04382_1	GH13_17	48-363	
Bemisia tabaci (MEAM1)	Bta04421_1	GH13_17	110-445	
Bemisia tabaci (MEAM1)	Bta04683_1	GH13_17	56-217	
Bemisia tabaci (MEAM1)	Bta05340_1	GH13_17	47-358	
Bemisia tabaci (MEAM1)	Bta05342	GH13_17	-	Base on tree location (HMMER prediction is also for subfamily 17, but coverage: 0.19, iE-value: 1.60E-25)
Bemisia tabaci (MEAM1)	Bta05343_1	GH13_17	1-186	
Bemisia tabaci (MEAM1)	Bta05386_1	GH13_17	92-442	
Bemisia tabaci (MEAM1)	Bta05396_1	GH13_17	67-416	
Bemisia tabaci (MEAM1)	Bta05397_1	GH13_17	79-425	
Bemisia tabaci (MEAM1)	Bta05518_1	GH13_17	56-403	
Bemisia tabaci (MEAM1)	Bta06059_1	GH13_17	1-316	
Bemisia tabaci (MEAM1)	Bta06327_1	GH13_17	55-289	
Bemisia tabaci (MEAM1)	Bta06441_1	GH13_17	52-237	

Bemisia tabaci (MEAM1)	Bta06445_1	GH13_17	56-201	
Bemisia tabaci (MEAM1)	Bta06458_1	GH13_17	54-393	
Bemisia tabaci (MEAM1)	Bta07452_1	GH13_17	53-344	
Bemisia tabaci (MEAM1)	Bta07453_1	GH13_17	68-415	
Bemisia tabaci (MEAM1)	Bta07764_1	GH13_17	52-392	
Bemisia tabaci (MEAM1)	Bta08425_1	GH13_17	56-393	
Bemisia tabaci (MEAM1)	Bta08426_1	GH13_17	56-393	
Bemisia tabaci (MEAM1)	Bta08427_1	GH13_17	52-390+17(427-650)	
Bemisia tabaci (MEAM1)	Bta08431_1	GH13_17	51-347	
Bemisia tabaci (MEAM1)	Bta09633_1	GH13_17	51-399+GH13(578-882)	
Bemisia tabaci (MEAM1)	Bta09634_1	GH13_17	50-211	
Bemisia tabaci (MEAM1)	Bta09696_1	GH13_17	55-422	
Bemisia tabaci (MEAM1)	Bta10022_1	GH13_17	59-405	
Bemisia tabaci (MEAM1)	Bta10152_1	GH13_17	39-389	
Bemisia tabaci (MEAM1)	Bta10614_1	GH13_17	122-317	
Bemisia tabaci (MEAM1)	Bta11358_1	GH13_17	52-398	
Bemisia tabaci (MEAM1)	Bta11977_1	GH13_17	50-389	
Bemisia tabaci (MEAM1)	Bta11978_1	GH13_17	54-382	
Bemisia tabaci (MEAM1)	Bta11979_1	GH13_17	77-420	
Bemisia tabaci (MEAM1)	Bta12101_1	GH13_17	49-396	
Bemisia tabaci (MEAM1)	Bta12578_1	GH13_17	59-397	
Bemisia tabaci (MEAM1)	Bta12624_1	GH13_17	139-397	

Bemisia tabaci (MEAM1)	Bta12671_1	GH13_17	46-384	
Bemisia tabaci (MEAM1)	Bta12678_1	GH13_17	49-385	
Bemisia tabaci (MEAM1)	Bta12682_1	GH13_17	102-449	
Bemisia tabaci (MEAM1)	Bta12683_1	GH13_17	90-436	
Bemisia tabaci (MEAM1)	Bta13236_1	GH13_17	49-395	
Bemisia tabaci (MEAM1)	Bta13238_1	GH13_17	41-405	
Bemisia tabaci (MEAM1)	Bta13239_1	GH13_17	1-325	
Bemisia tabaci (MEAM1)	Bta13660_1	GH13_17	50-398	
Bemisia tabaci (MEAM1)	Bta13913_1	GH13_17	1-202	
Bemisia tabaci (MEAM1)	Bta13914_1	GH13_17	45-392	
Bemisia tabaci (MEAM1)	Bta14313_1	GH13_17	1-123	
Bemisia tabaci (MEAM1)	Bta14419_1	GH13_17	55-399	
Bemisia tabaci (MEAM1)	Bta14422_1	GH13_17	53-400	
Bemisia tabaci (MEAM1)	Bta15649_1	GH13_17	59-405	
Bemisia tabaci (MEAM1)	Bta01463_1	GH13_17	-	Based on tree location (HMMER prediction is also for subfamily 17, but coverage: 0.19, iE-value: 1.30E-16)
Bemisia tabaci (MEAM1)	Bta07377_1	GH13_17	1-119	
Bemisia tabaci (MEAM1)	Bta11362	GH13_17	-	Based on tree location (HMMER prediction is also for subfamily 17, but coverage: 0.14, iE-value: 7.40E-07)
Bemisia tabaci (MEAM1)	Bta11975	GH13_17	-	Based on tree location (HMMER prediction is also for subfamily 17, but coverage: 0.27, iE-value: 3.90E-28)

Bombyx mori	BGIBMGA003055-TA	GH13_17	50-404	
Bombyx mori	BGIBMGA003056-TA	GH13_17	50-404	
Bombyx mori	BGIBMGA003057-TA	GH13_17	50-404	
Bombyx mori	BGIBMGA005068-TA	GH13_17	193-374	
Bombyx mori	BGIBMGA006066-TA	GH13_17	70-428	
Frankliniella occidentalis	FOCC006088-PA	GH13_17	68-303	
Frankliniella occidentalis	FOCC006089-PA	GH13_17	109-457	
Frankliniella occidentalis	FOCC008836-PA	GH13_17	152-504	
Frankliniella occidentalis	FOCC006090-PA	GH13_17	241-349	
Leptinotarsa decemlineata	LDEC004803-PA	GH13_17	157-342	
Leptinotarsa decemlineata	LDEC013918-PA	GH13_17	1-193	
Leptinotarsa decemlineata	LDEC013919-PA	GH13_17	57-180	
Manduca sexta	Msex2.01215-RA	GH13_17	132-278	
Manduca sexta	Msex2.01215-RB	GH13_17	204-350	
Manduca sexta	Msex2.06911-RA	GH13_17	70-428	
Manduca sexta	Msex2.06911-RB	GH13_17	70-428	
Manduca sexta	Msex2.06911-RC	GH13_17	55-413	
Manduca sexta	Msex2.06911-RD	GH13_17	70-428	
Manduca sexta	Msex2.06911-RE	GH13_17	1-337	
Manduca sexta	Msex2.06912-RA	GH13_17	68-427	
Manduca sexta	Msex2.06912-RB	GH13_17	1-338	
Manduca sexta	Msex2.09215-RA	GH13_17	70-424	

Manduca sexta	Msex2.10850-RA	GH13_17	48-234	
Manduca sexta	Msex2.10850-RB	GH13_17	48-234	
Manduca sexta	Msex2.10850-RC	GH13_17	48-234	
Manduca sexta	Msex2.12170-RA	GH13_17	143-496	
Manduca sexta	Msex2.12171-RA	GH13_17	49-403	
Manduca sexta	Msex2.12171-RB	GH13_17	67-421	
Myzus persicae	MYZPE13164_G006_v1.0_000012200.1	GH13_17	57-412	
Myzus persicae	MYZPE13164_G006_v1.0_000012200.2	GH13_17	57-412	
Myzus persicae	MYZPE13164_G006_v1.0_000012200.3	GH13_17	57-412	
Myzus persicae	MYZPE13164_G006_v1.0_000012210.1	GH13_17	58-413+17(673-1028)	
Myzus persicae	MYZPE13164_G006_v1.0_000012210.2	GH13_17	58-413+17(673-1028)	
Myzus persicae	MYZPE13164_G006_v1.0_000012220.1	GH13_17	58-407	
Myzus persicae	MYZPE13164_G006_v1.0_000012230.1	GH13_17	57-406	
Myzus persicae	MYZPE13164_G006_v1.0_000012230.2	GH13_17	91-440	
Myzus persicae	MYZPE13164_G006_v1.0_000012240.1	GH13_17	59-414	
Myzus persicae	MYZPE13164_G006_v1.0_000012240.2	GH13_17	105-460	
Myzus persicae	MYZPE13164_G006_v1.0_000012270.1	GH13_17	57-407	
Myzus persicae	MYZPE13164_G006_v1.0_000012270.2	GH13_17	52-402	
Myzus persicae	MYZPE13164_G006_v1.0_000012270.3	GH13_17	57-407	
Myzus persicae	MYZPE13164_G006_v1.0_000012280.1	GH13_17	50-399	
Myzus persicae	MYZPE13164_G006_v1.0_000060920.1	GH13_17	119-313	
Myzus persicae	MYZPE13164_G006_v1.0_000068520.1	GH13_17	52-400	

Myzus persicae	MYZPE13164_G006_v1.0_000074190.1	GH13_17	57-406	
Myzus persicae	MYZPE13164_G006_v1.0_000081150.1	GH13_17	135-293+17(298-440)	
Myzus persicae	MYZPE13164_G006_v1.0_000148570.1	GH13_17	54-400	
Myzus persicae	MYZPE13164_G006_v1.0_000148570.2	GH13_17	69-415	
Myzus persicae	MYZPE13164_G006_v1.0_000171800.1	GH13_17	52-400	
Myzus persicae	MYZPE13164_G006_v1.0_000171800.2	GH13_17	52-400	
Pieris rapae	Pra113.50	GH13_17	70-428	
Pieris rapae	Pra113.51	GH13_17	67-422	
Pieris rapae	Pra143.13	GH13_17	204-394	
Pieris rapae	Pra185.3	GH13_17	47-401	
Pieris rapae	Pra185.4	GH13_17	49-403	
Pieris rapae	Pra185.5	GH13_17	177-531	
Pieris rapae	Pra205.26	GH13_17	43-397	
Pieris rapae	Pra205.27	GH13_17	42-396	
Acyrthosiphon pisum	lcl|hmm258794	GH13_18	85-427	Based on tree location in subfamily 25, but no support in HMMER for placement in subfamily 25
Manduca sexta	Msex2.06913-RA	GH13_23	44-208	
Tetranychus urticae	tetur02g04570.1	GH13_23	311-677	
Acyrthosiphon pisum	lcl|hmm285044	GH13_25	1097-1536	
Bombyx mori	BGIBMGA010888-TA	GH13_25	-	Based on tree location. No significant support in HMMER for any subfamily
Frankliniella occidentalis	FOCC001301-PA	GH13_25	1145-1590	

Leptinotarsa decemlineata	LDEC008797-PA	GH13_25	-	Based on tree location. No significant support in HMMER for any subfamily
Myzus persicae	MYZPE13164_G006_v1.0_000121040.1	GH13_25	841-1280	
Myzus persicae	MYZPE13164_G006_v1.0_000121040.2	GH13_25	841-1280	
Myzus persicae	MYZPE13164_G006_v1.0_000121040.3	GH13_25	1015-1454	
Myzus persicae	MYZPE13164_G006_v1.0_000121040.4	GH13_25	841-1280	
Pieris rapae	Pra673.1	GH13_25	1104-1540	
Tetranychus urticae	tetur02g00800.1	GH13_25	1093-1523	
Leptinotarsa decemlineata	LDEC018182-PA	GH13_34	2-243	Based on tree location and support values in HMMER can be also placed in subfamily 23 (iE-value-: 4.20E-13, coverage: 0.21) or subfamily 36 (iE-value: 4.50E-09, coverage: 0.57)
Manduca sexta	Msex2.12538-RA	GH13_34	2-242	Based on tree location and support values in HMMER can be also placed in subfamily 36 (iE-value: 1.00E-16, coverage: 0.23)
Bombyx mori	BGIBMGA006067-TA	GH13_36	44-347	
Manduca sexta	Msex2.15001-RA	GH13_36	33-144	Based on tree location can be placed in subfamily 17 (HMMER prediction is also for subfamily 17, coverage: 0.35, iE-value: 5.10E-58)
Pieris rapae	Pra113.52	GH13_36	35-241	
Pieris rapae	Pra151.20	GH13_36	203-456	Based on tree location (HMMER prediction is for subfamily 36, but coverage: 0.23, iE-value: 2.4E-17)
Bombyx mori	BGIBMGA007451-TA	GH13_8	310-495	
Frankliniella occidentalis	FOCC011793-PA	GH13_8	1-293	
Manduca sexta	Msex2.07715-RA	GH13_8	226-519	

Manduca sexta	Msex2.07715-RB	GH13_8	233-526	
Manduca sexta	Msex2.07715-RC	GH13_8	233-526	
Myzus persicae	MYZPE13164_G006_v1.0_000162350.1	GH13_8	240-533	
Myzus persicae	MYZPE13164_G006_v1.0_000162350.2	GH13_8	240-533	
Myzus persicae	MYZPE13164_G006_v1.0_000162350.3	GH13_8	240-533	
Myzus persicae	MYZPE13164_G006_v1.0_000162350.4	GH13_8	98-391	
Myzus persicae	MYZPE13164_G006_v1.0_000121980.1	GH13_8	-	Based on tree location (no significant support from HMMER)
Pieris rapae	Pra1943.3	GH13_8	233-526	
Tetranychus urticae	tetur16g02500.1	GH13_8	238-531	
Tetranychus urticae	tetur16g02480.1	GH13_8	-	Based on tree location (HMMER prediction is also for subfamily 8, but coverage: 0.28, E-value: 3.40E-34)

[a]The names of the proteins are according to their GenBank accession numbers or their name in the released proteome. The individual CAZy GH13 subfamilies were established by Stam et al. [60]

[b]The predicted classification to a subfamily in HMMER was verified by checking the corresponding branch assignment in the cladogram (Supplementary Figure S3). In cases where conflicting classifications occurred, the final assignment is justified in the "Comments" column.

Supplementary Table 9: Enzyme and Substrate specificities of the predicted GH13 subfamily.

CAZy GH13 subfamilies[a]	Enzyme specificity[b]	Substrate[b]	Origin[b]	References
GH13_8	1,4-α-Glucan branching enzyme	1,4-α-Glucan chain	Eukaryota	60
GH13_15	α-Amylase	α-1,4 glycosidic bonds inside a maltopolysaccharide linear chain, mainly in starch and glycogen	Insects	60
GH13_17	α-Glucosidase	Sucrose and p-nitrophenyl α-d-gluco-pyranoside, maltooligosaccharides	Insects	60
GH13_18	Sucrose phosphorylase	Sucrose 6F-phosphate, glucosylglycerol or glucosylglycerate	Bacteria	57,60
GH13_23	α-Glucosidase	Maltose (glucosyl donor substrate)	Bacteria	60
GH13_25	Bifunctional glycogen debranching enzyme	Glycogen	Eukaryota	60
[c]GH13_34	[d]NA	Amino acid transporter	Eukaryota	60
GH13_36	α-amylase	Mixed enzyme specificity of α-amylase and an additional GH13 specificity	Bacteria	58,59

[a]The individual CAZy GH13 subfamilies were established by Stam et al [60].

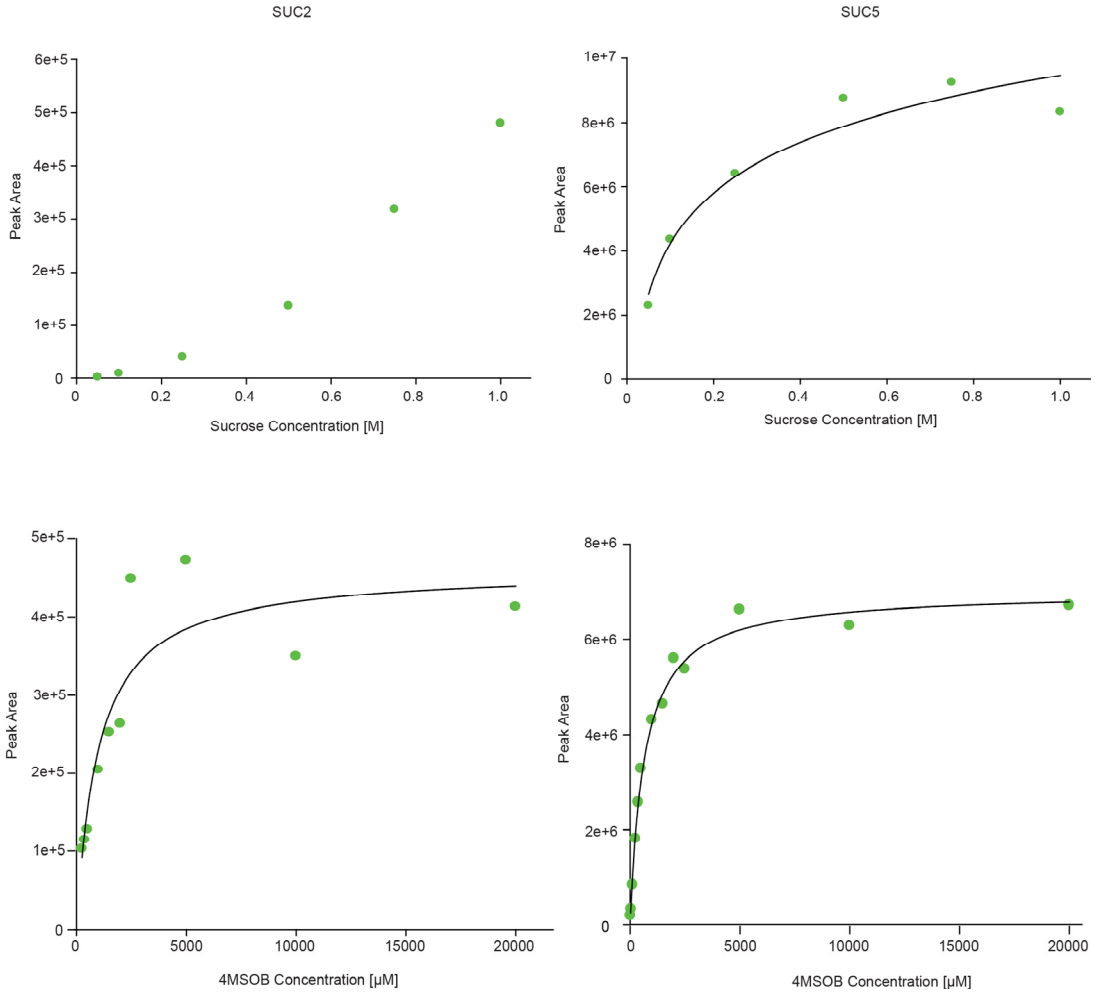
[b]The enzyme and substrate specificities have been described in the relevant reference.

[c]GH13 members that have lost their catalytic machinery and evolved to a novel function.

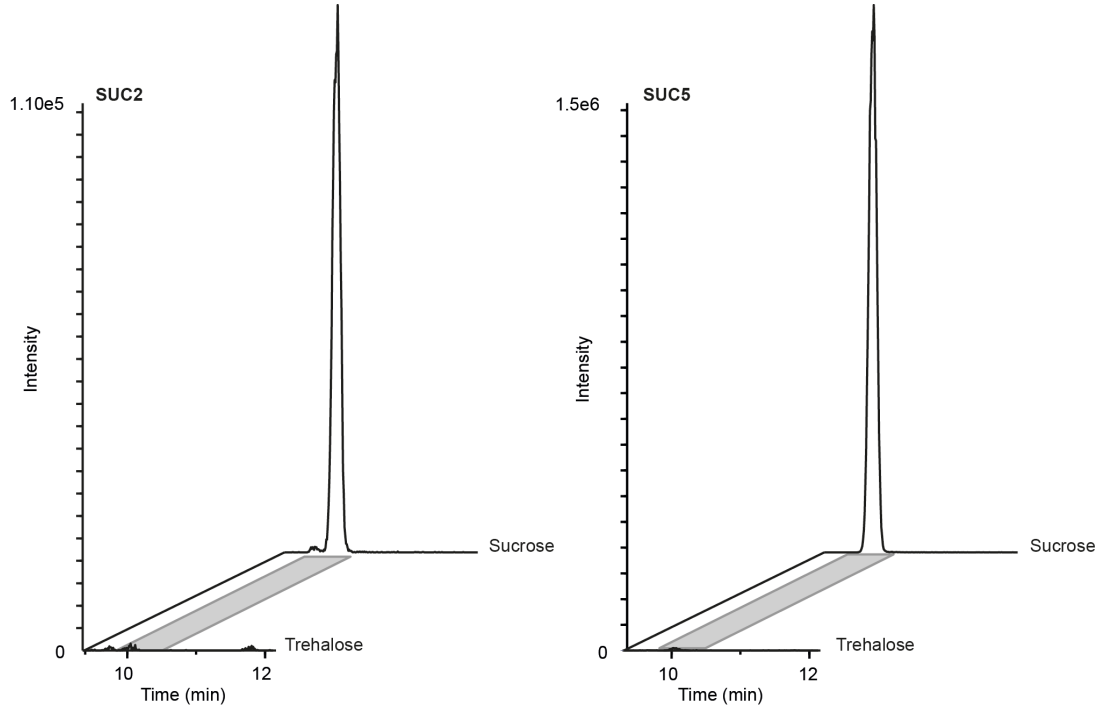
[d]NA - not applicable

Supplementary Table 10: **Primer sets used for cloning *B. tabaci* GH13 family genes in pAc5.1/V5-HIS A for expression in *D. melanogaster* S2 cells.**

Primer name	Sequence
SUC1 pACV5-HIS Full EcoR1 Forward S2	CGTACGAATTCGAGATGAAAATAGCAGTGCTTTCATTT
SUC1 pACV5-HIS Full XhoI Reverse S2	TACCTCGAGCAGAGGTTTTTAGGCCACCCC
SUC2 pACV5-HIS Full KpnI Forward S2	GCACGTGGTACCATGAGTCGGAATTTGACAATACTGC
SUC2 pACV5-HIS Full XbaI Reverse S2	CACCGGCTCTAGAATTGGATTTACCGGTTAATCTACTGC
SUC3 pACV5-HIS Full KpnI Forward S2	GCACGTGGTACCATGCTGTTAGAGATAATGTGCAAATTTACG
SUC3 pACV5-HIS Full XbaI Reverse S2	CACCGGCTCTAGATAGGAGTAACCTATGAGTGAAAAGAAATACA
SUC4 pACV5-HIS Full KpnI Forward S2	GCACGGGGTACCATGATAGTTTTAAATTATTTCATAATCGATCTCT ATTACAA
SUC4 pACV5-HIS Full XbaI Reverse S2	CACCGGCTCTAGACAGGGAATATTTACAGTACAGGAACAG
SUC5 pACV5-HIS Full EcoRI Forward S2	GCACGTGAATTCACCATGGAGAGACTACTCTACTTTGTTGTG
SUC5 pACV5-HIS Full XbaI Reverse S2	CACCGGCTCTAGAGCTGATGAAGATTTTTAAATAAATTGATAGTAGAAG



Supplementary Figure 1: **Kinetic characterization of SUC2 and SUC5 shows activity at high sucrose concentrations.** Depicted are the relations between substrate concentration and reaction rate (product peak area) for both sucrose and 4msob-GSL. Assays were carried out as described in the methods section. When the sucrose concentration was varied from 0.05 to 1.0 M, 4msob-GSL concentration was held at 2.5 mM. When the 4msob-GSL concentration was varied from 25 µM to 20 mM, sucrose concentration was held at 1 M.



Supplementary Figure 2: **SUC2 and SUC5 enzymes utilize sucrose to glucosylate GSLs as a preferential substrate to trehalose.** The enzymes SUC2 and SUC5 were incubated *in vitro* with 4msob-GSL and either sucrose or trehalose as potential glucose donors for 6 h at 25°C and pH 7.0. Depicted are extracted multiple reaction monitoring LC-MS chromatograms for the glucosylated α-(1→4) 4msob-GSL products of SUC2 and SUC5. Both enzymes utilized trehalose as a glucose donor only poorly, at <1 % efficiency when compared to sucrose.

Supplementary Note 1

Supporting information NMR

NMR spectra were recorded on a Bruker Avance III HD 700 MHz spectrometer, equipped with a cryoplatform and a 1.7 mm cryoprobe (Bruker Biospin GmbH, Rheinstetten, Germany). Spectrometer control and data processing was accomplished using Bruker TopSpin ver. 3.2. NMR spectra for compounds **1** and **3** were recorded in D_2O and the chemical shifts were left uncorrected. ^{13}C chemical shifts were determined indirectly by means of ^{1}H-^{13}C HSQC (heteronuclear single quantum coherence) and ^{1}H-^{13}C HMBC (heteronuclear multiple bond correlation) experiments. The glucosidic substitutions were elucidated with the help of 1D ^{1}H-^{1}H SELTOCSY (selective total correlation spectroscopy) experiments. The resulting spectra served as projection spectra for the 2D homo- and heteronuclear experiments. The positions of the chemical shifts in the molecules were furthermore determined based on $^{3}J_{HH}$ coupling constants and ^{13}C chemical shifts. Important spectral details have been assembled for deeper understanding of the structure elucidation as follows.

Figure SN1-1. Structure of glucoraphanin (4MSOB) with numbering

Figure SN1-2. Structure of glucoraphanin (4MSOB) with chemical shifts in D_2O

Figure SN1-3. ^{1}H NMR spectrum of Glucoraphanin (4MSOB) in D_2O

Figure SN1-4. ^{1}H-^{13}C HSQC spectrum of Glucoraphanin (4MSOB) in D_2O

Figure SN1-5. Structure of **1** with numbering

Figure SN1-6. ^{1}H NMR spectrum of **1** in D_2O with water suppression

Figure SN1-7. ^{1}H NMR spectra of the S-glucosidic moieties of **1** and 4MSOB

Figure SN1-8. ^{1}H-^{1}H SELTOCSY spectrum of **1,** transmitter on 1‘‘, in D_2O

Figure SN1-9. ^{1}H-^{13}C HSQC spectrum of **1** in D_2O, glucosidic range

Figure SN1-10. Structure of **3** with numbering

Figure SN1-11. ^{1}H NMR spectrum of **3** in D_2O with water suppression

Figure SN1-12. ^{1}H NMR spectra of the S-glucosidic moieties of **3** and 4MSOB

Figure SN1-13. ^{1}H-^{1}H SELTOCSY spectra of the second and third glucosidic moiety of **3**

Figure SN1-14. ^{1}H-^{13}C HSQC spectrum of **3** in D_2O, glucosidic range

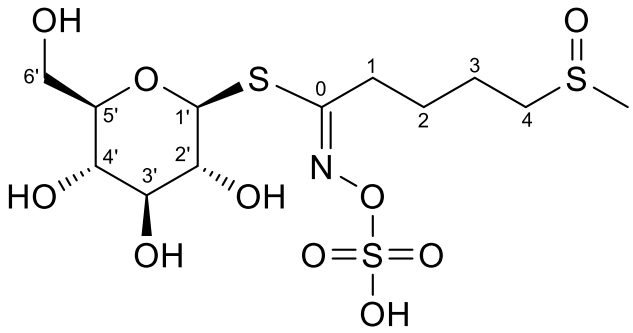


Figure SN1-1. Structure of glucoraphanin (4MSOB) with numbering.

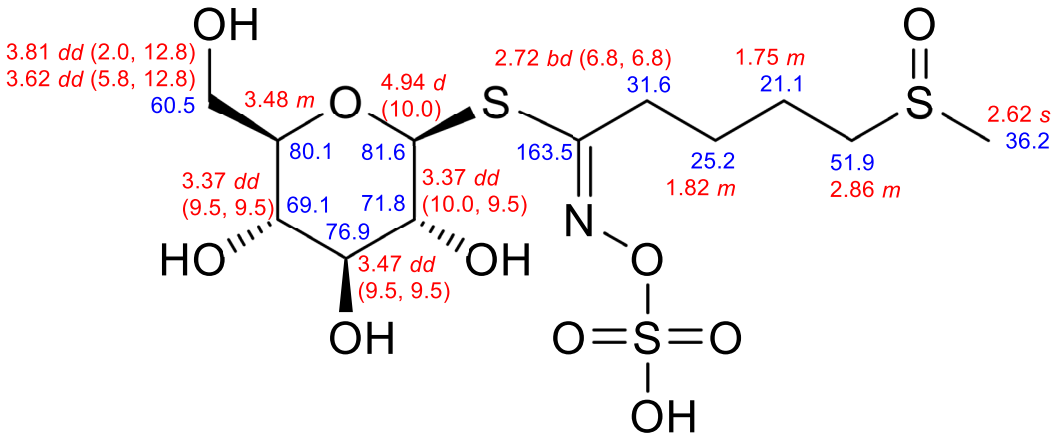


Figure SN1-2. Structure of glucoraphanin (4MSOB) with chemical shifts (δ ppm; δ_C blue, δ_H red) and coupling constants (*mult.*, $^3J_{HH}$ in Hz) from NMR measurements in D_2O.

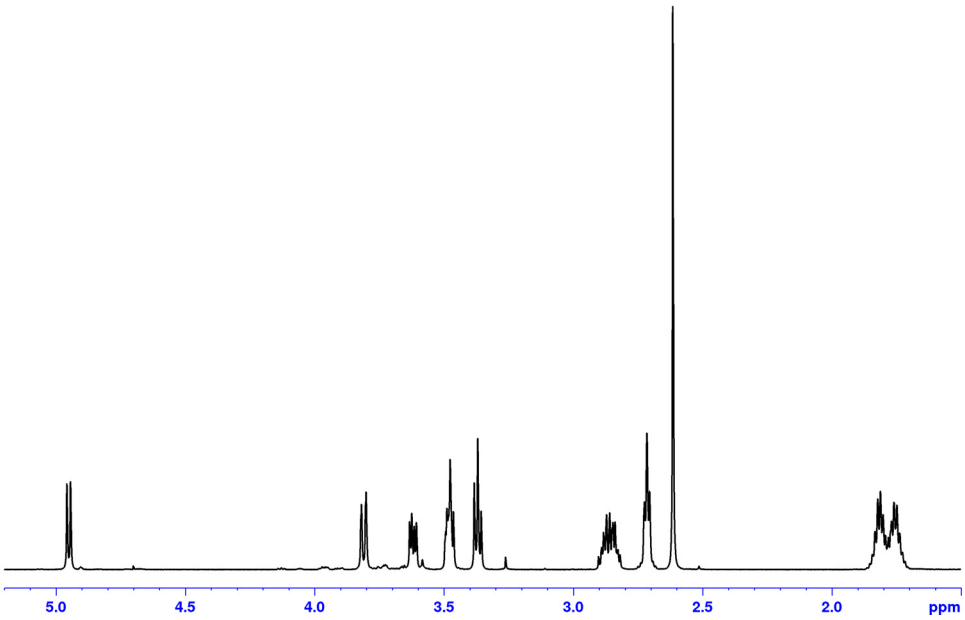


Figure SN1-3. ^{1}H NMR spectrum of Glucoraphanin (4MSOB) in D_2O with water suppression.

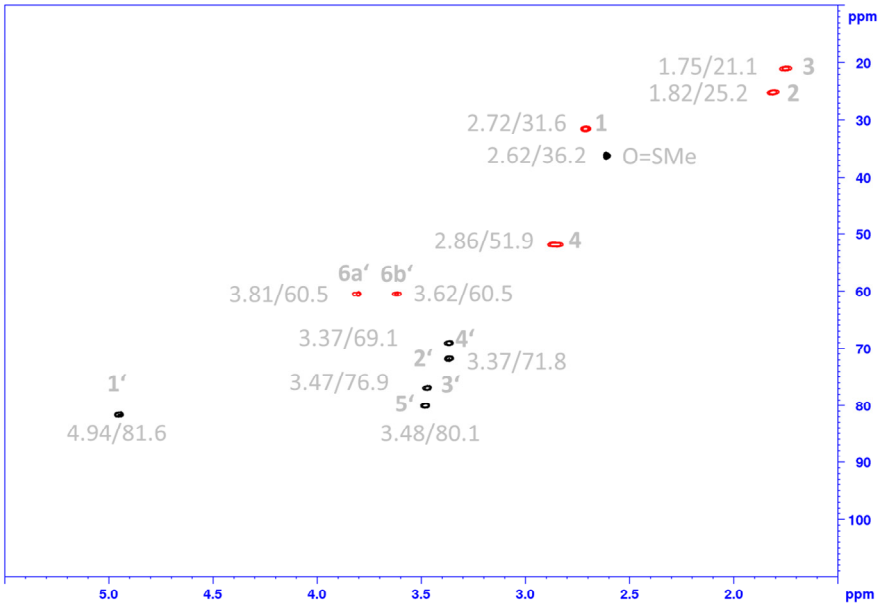


Figure SN1-4. ^{1}H-^{13}C HSQC spectrum of Glucoraphanin (4MSOB) in D_2O.

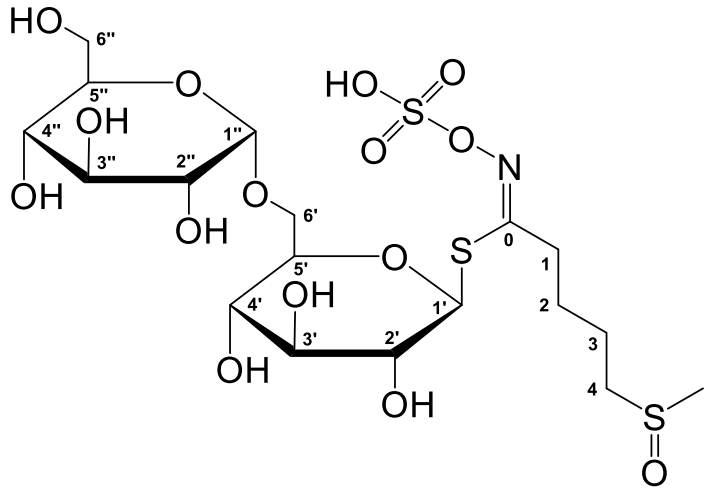


Figure SN1-5 Structure of **1** with numbering.

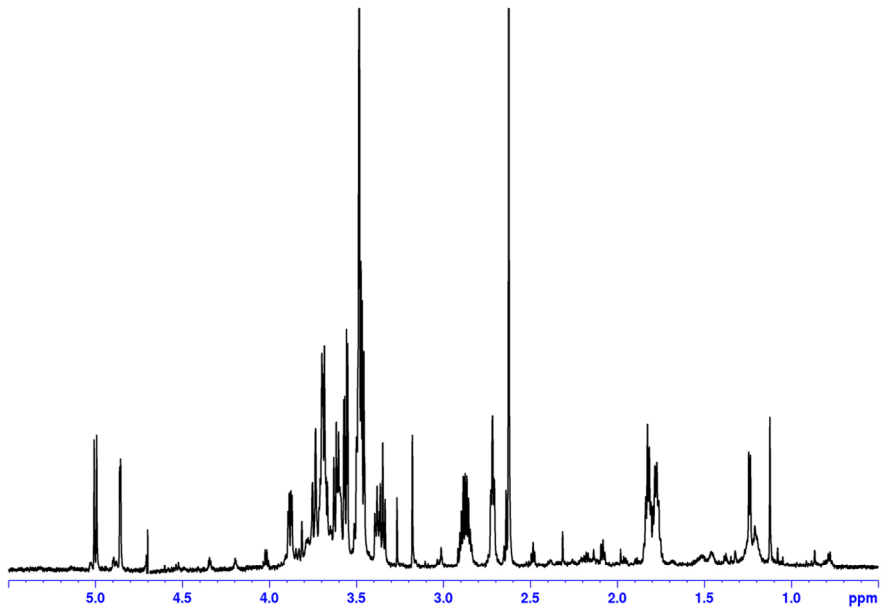


Figure SN1-6. ^{1}H NMR spectrum of **1** in D_2O with water suppression.

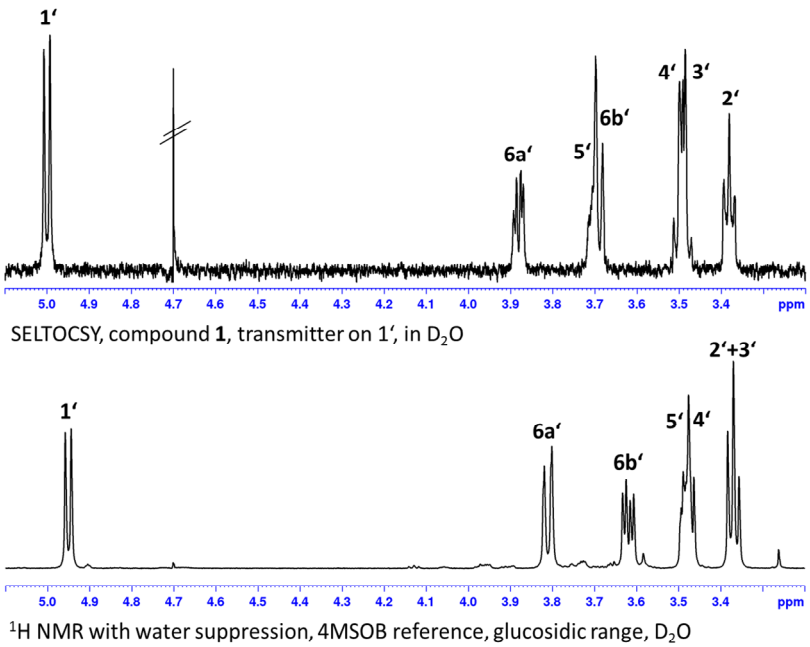


Figure SN1-7. ^{1}H NMR spectra of the S-glucosidic moieties of **1** and 4MSOB.

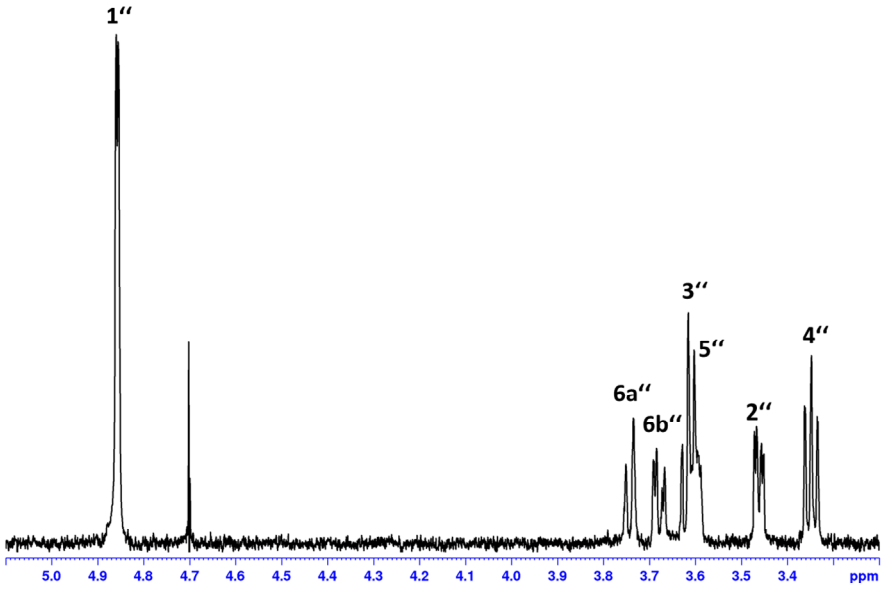


Figure SN1-8. ^{1}H-^{1}H SELTOCSY spectrum of **1,** transmitter on 1‘‘, in D_2O.

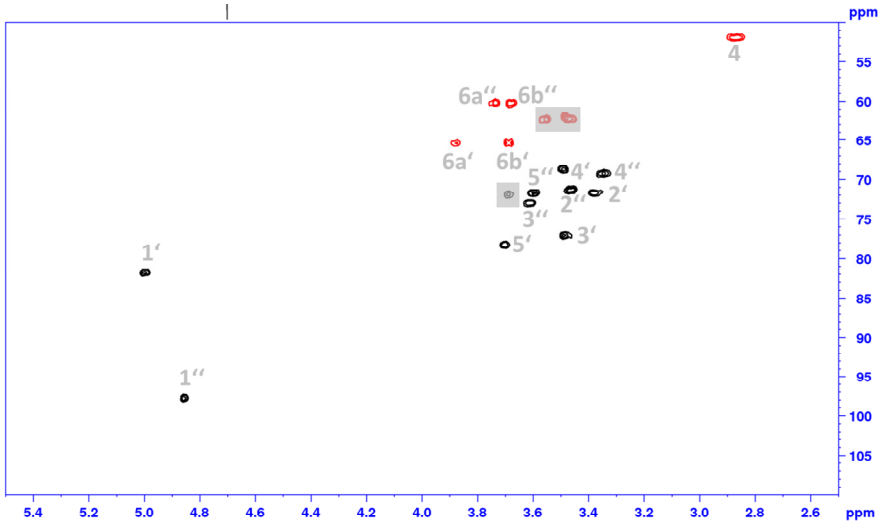


Figure SN1-9. ^{1}H-^{13}C HSQC spectrum of **1** in D_2O, glucosidic range, with numbering. Grey fields cover impurities.

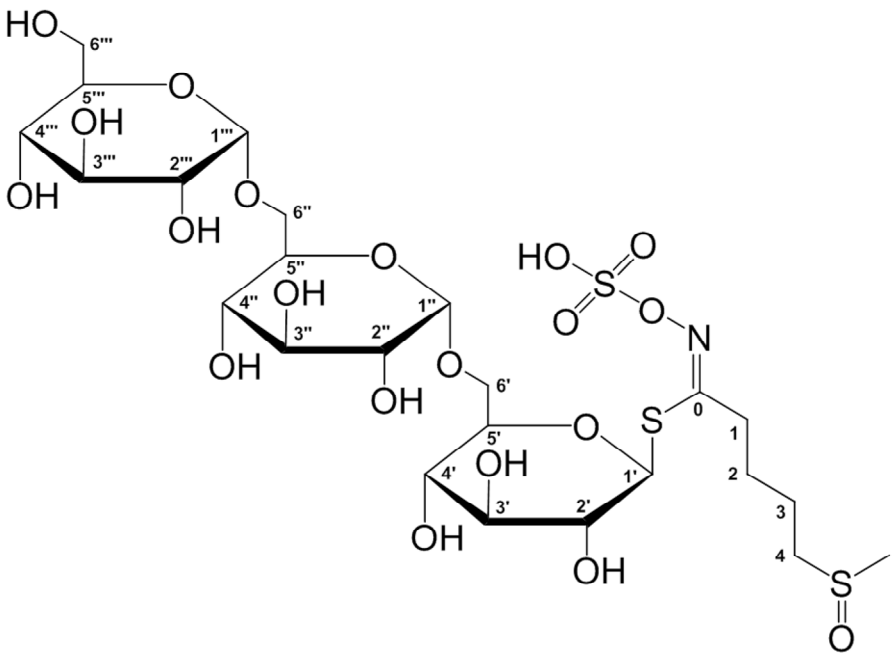


Figure SN1-10. Structure of **3** with numbering.

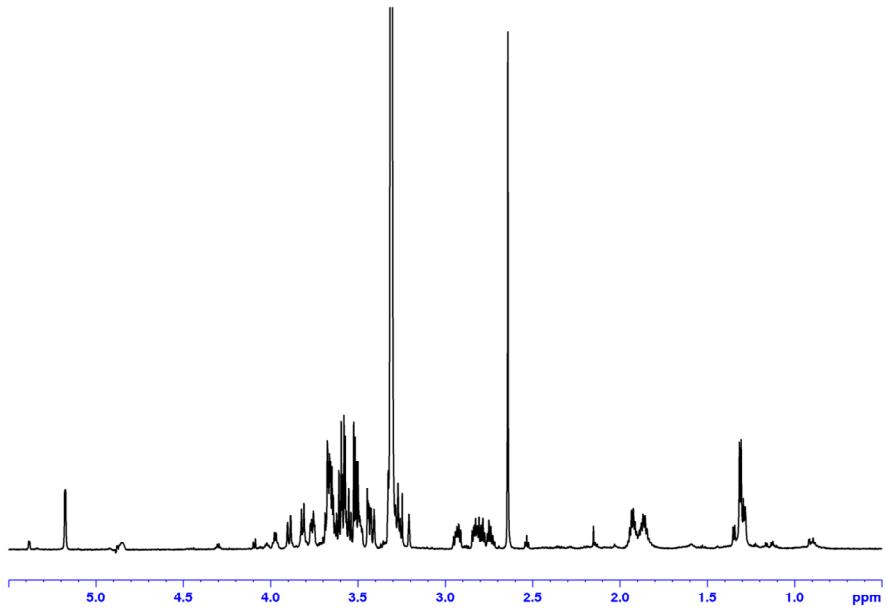


Figure SN1-11. ^{1}H NMR spectrum of **3** in D_2O with water suppression.

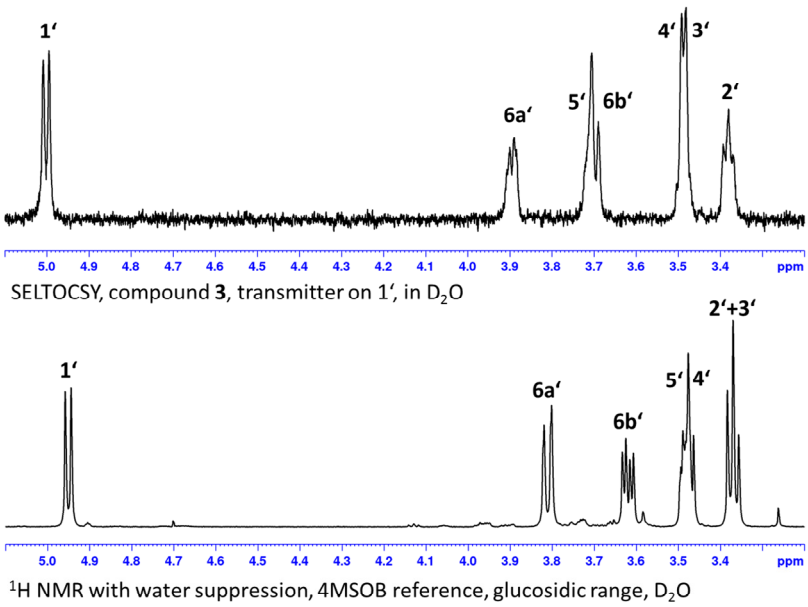


Figure SN1-12. ^{1}H NMR spectra of the S-glucosidic moieties of **3** and 4MSOB.

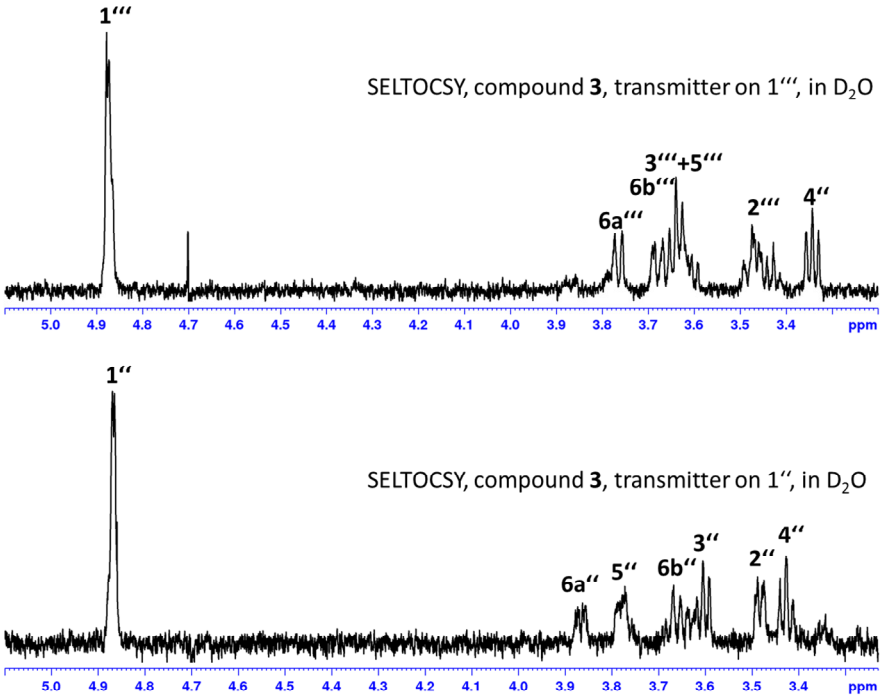


Figure SN1-13. ^{1}H-^{1}H SELTOCSY spectra of the second and third glucosidic moiety of **3**.

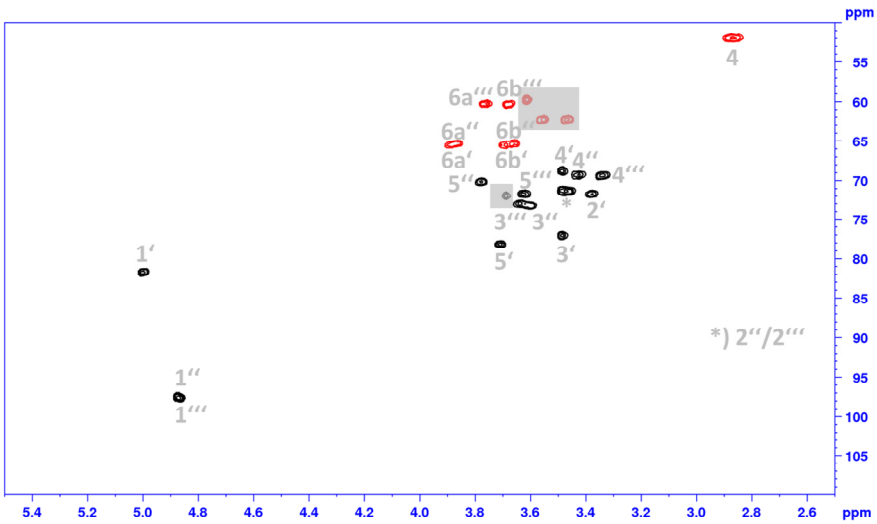


Figure SN1-14. ^{1}H-^{13}C HSQC spectrum of **3** in D_2O, glucosidic range, with numbering. Grey fields cover impurities.

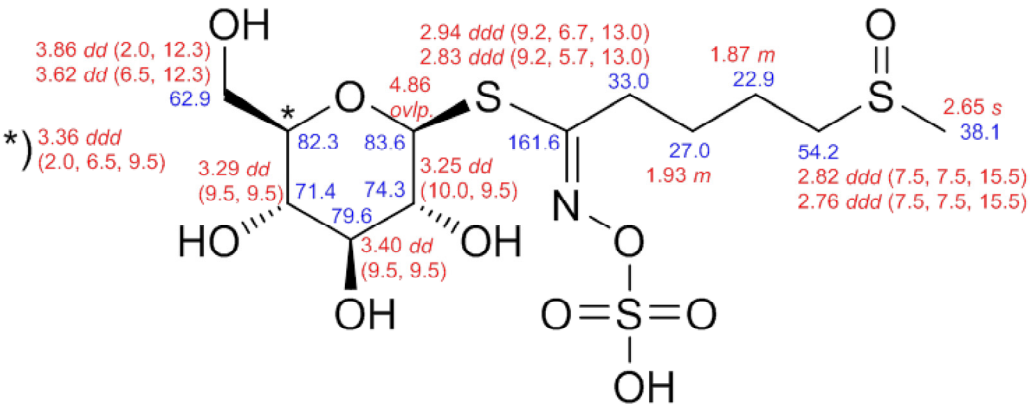


Figure SN1-15. Structure of glucoraphanin (4MSOB) with chemical shifts (δ ppm; δ_C blue, δ_H red) and coupling constants (*mult.*, $^3J_{HH}$ in Hz) from NMR measurements in MeOH-d_3.

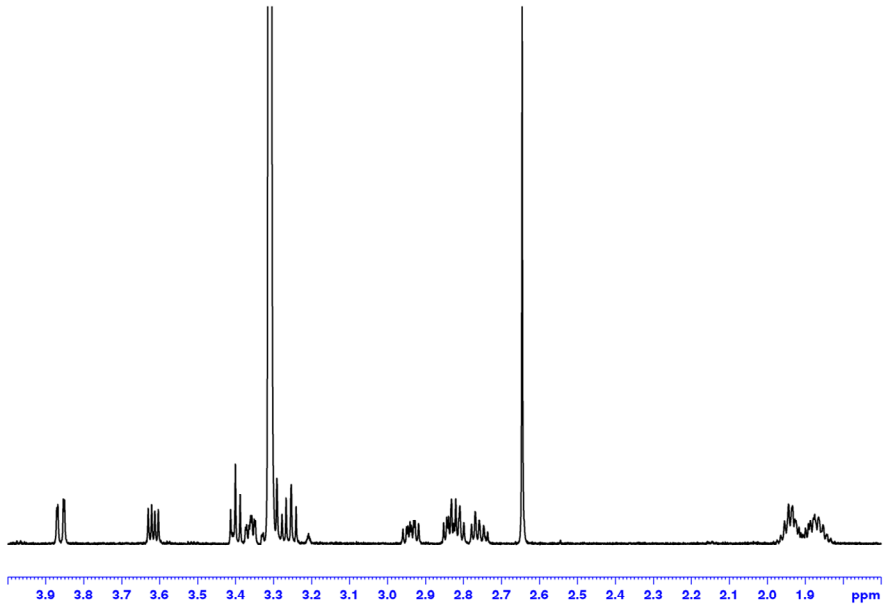


Figure SN1-16. ^{1}H NMR spectrum of glucoraphanin (4MSOB) in MeOH-d_3.

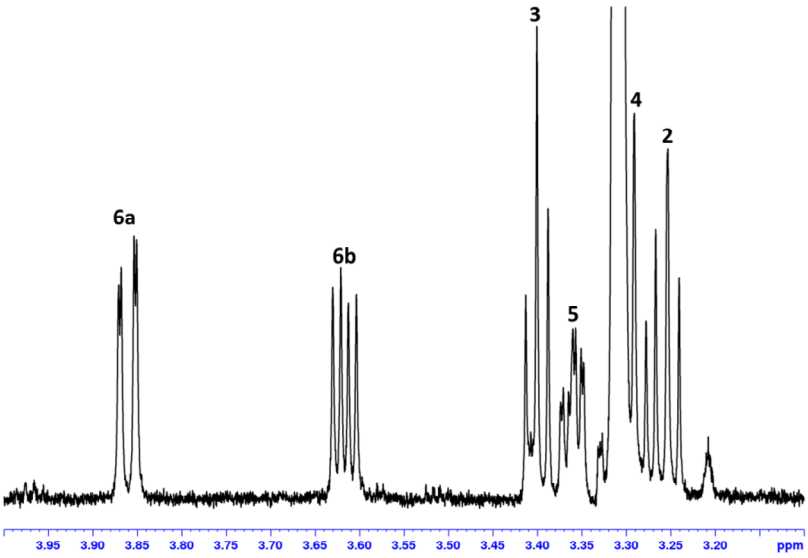


Figure SN1-17. Detail of the ^{1}H NMR spectrum of glucoraphanin (4MSOB) in MeOH-d_3.

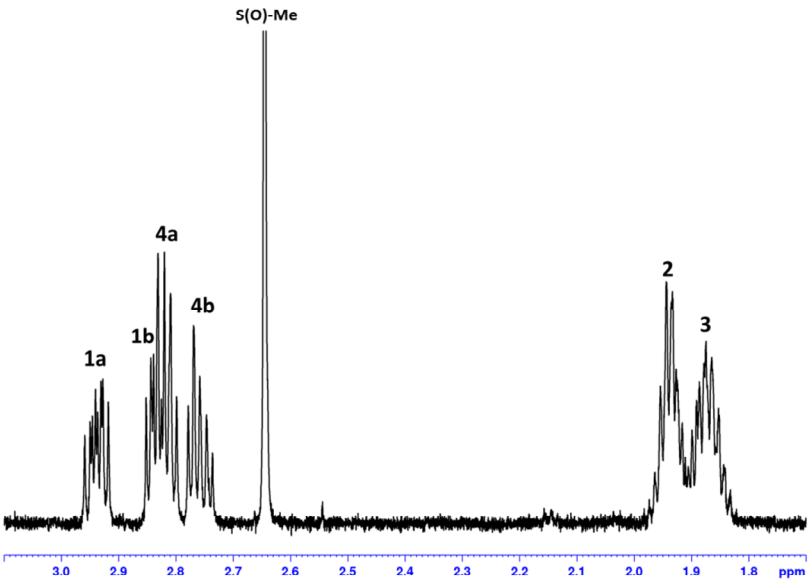


Figure SN1-18. Detail of the ^{1}H NMR spectrum of glucoraphanin (4MSOB) in MeOH-d_3.

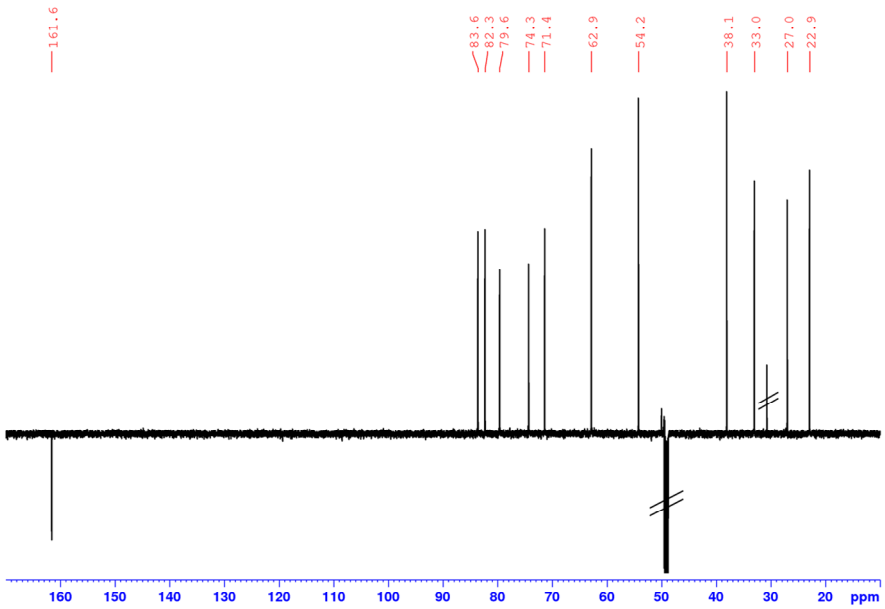


Figure SN1-19. ^{13}C NMR (DEPTQ) spectrum of glucoraphanin (4MSOB) in MeOH-d_3.

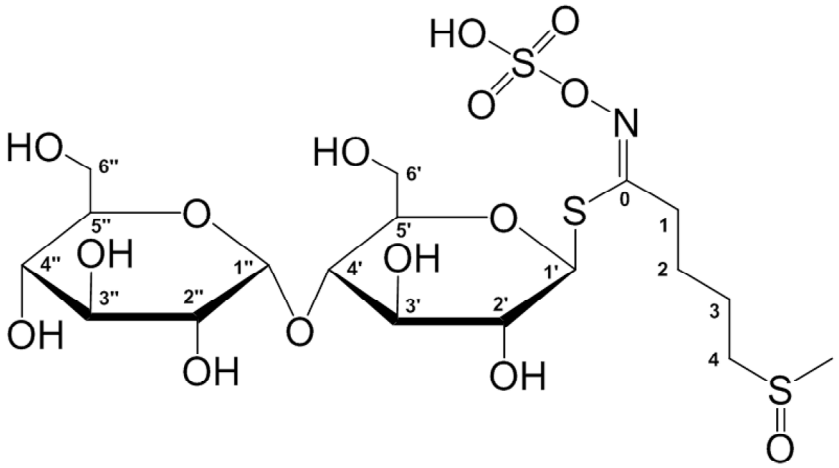


Figure SN1-20. Structure of **2** with numbering.

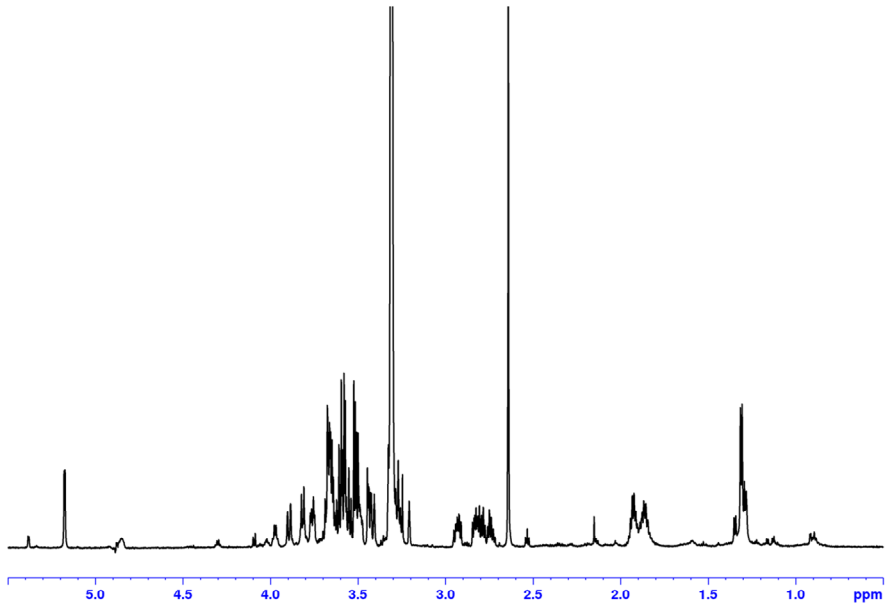


Figure SN1-21. ^{1}H NMR spectrum of **2** in MeOH-d_3 with water suppression.

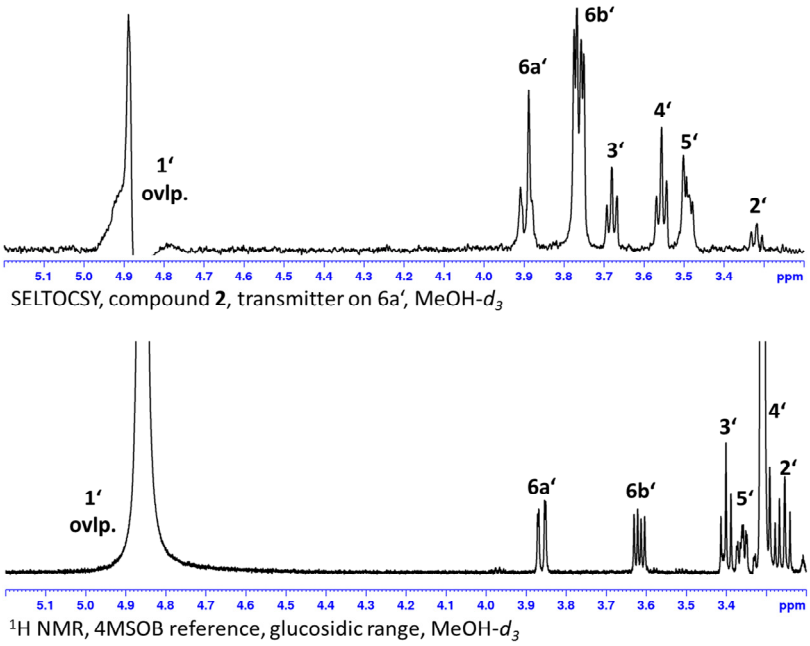


Figure SN1-22. ^{1}H NMR spectra of the S-glucosidic moieties of **2** and 4MSOB.

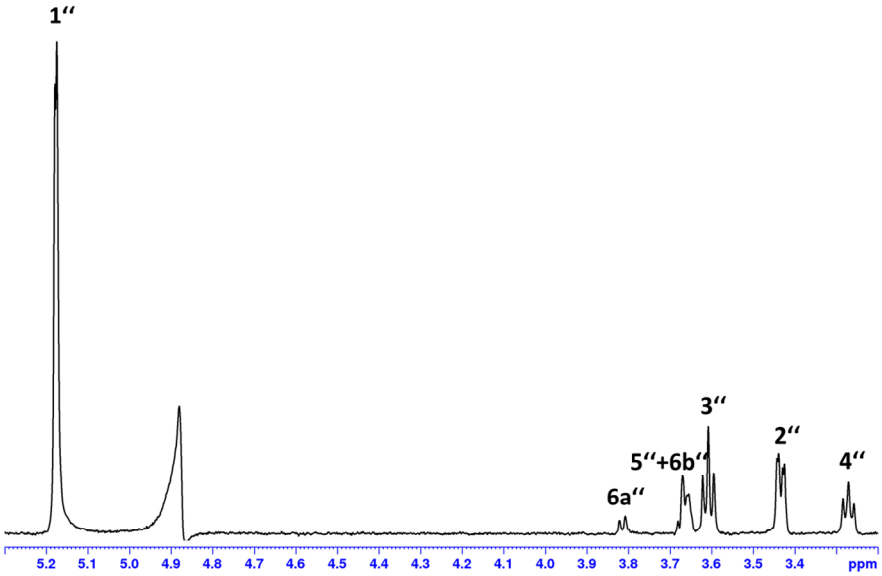


Figure SN1-23. ^{1}H-^{1}H SELTOCSY spectrum of **2,** transmitter on 1'', in MeOH-d_3.

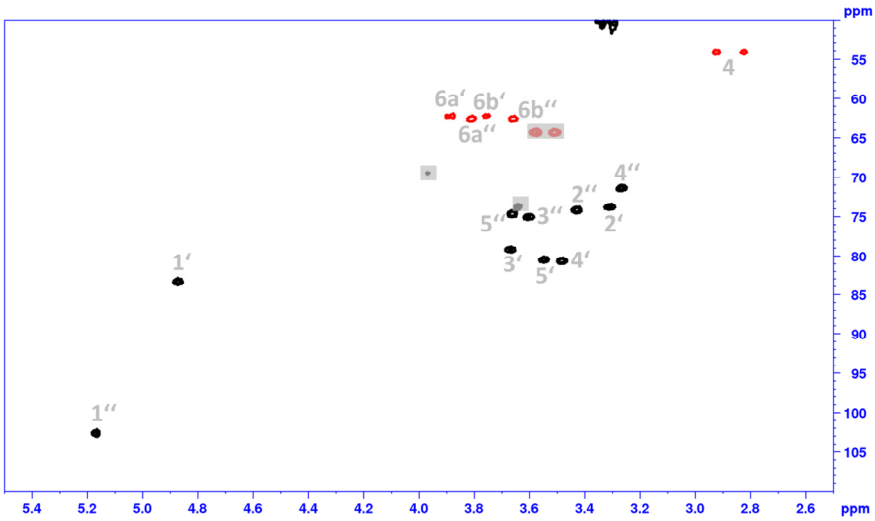


Figure SN1-24. ^{1}H-^{13}C HSQC spectrum of **2** in MeOH-d_3, glucosidic range, with numbering. Grey fields cover impurities.

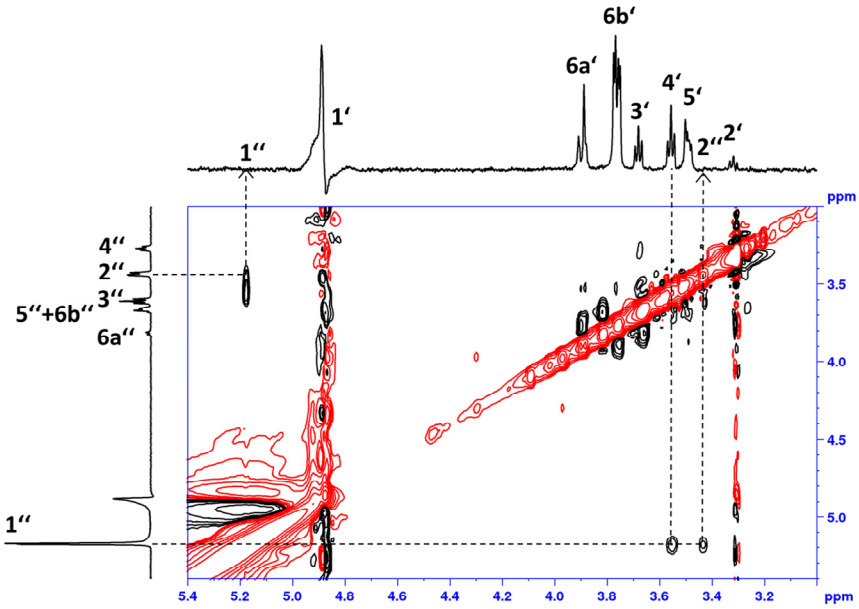


Figure SN1-25. ^{1}H-^{1}H ROESY spectrum of **2** in MeOH-d_3, glucosidic range. SELTOCSY spectra of the glucosidic moieties were used as projection spectra.

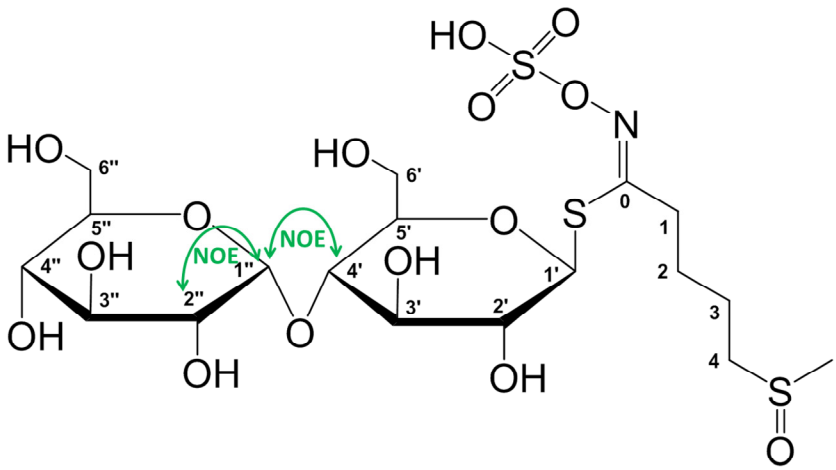


Figure SN1-26. Structure of **2** with prominent NOE correlations

Manuscript II

Cassava cyanogenic glucosides are detoxified via multiple pathways in the pest *Bemisia tabaci*

Article Preprint (ISSN 2470-8224)

Cassava cyanogenic glucosides are detoxified via multiple pathways in the pest Bemisia tabaci

Michael L.A.E. Easson[1], Osnat Malka[2], Christian Paetz[1], Anna Hojná[1], Michael Reichelt[1], Beate Stein[3], John Colvin[4], Stephan Winter[3], Shai Morin[2], Jonathan Gershenzon[1], Daniel G. Vassão[1]

[1]*Max Planck Institute for Chemical Ecology, Jena 07745, Germany*

[2]*The Hebrew University of Jerusalem, Rehovot 7610001, Israel*

[3]*Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig 38104, Germany*

[4]*Natural Resources Institute, University of Greenwich, Chatham Maritime, Kent ME4 4TB, UK*

Corresponding author: Daniel G. Vassão

Email: (vassao@ice.mpg.de)

Abstract

Two-component activated defenses such as cyanogenic glucosides are produced by many plant species, but phloem feeding herbivores have long been thought to cause little to no activation of these compounds due to their mode of feeding. Here, however, we report that cyanogenic glycoside defenses from cassava (*Manihot esculenta*), a major staple crop in Africa, are activated during feeding by a pest insect, the tobacco whitefly *Bemisia tabaci*, and the resulting hydrogen cyanide is detoxified by conversion to beta-cyanoalanine. Additionally, the whitefly was found to utilize two detoxification pathways to prevent the activation of cyanogenic glucosides. First, a transglucosidation reaction converts the cyanogenic glycoside linamarin to glucosides that are excreted in the honeydew. Two glycoside hydrolase family 13 enzymes were shown to glucosylate linamarin in vitro utilizing sucrose as a co-substrate. Second, phosphorylation of both linamarin and the insect-derived linamarin glucosides is described. Both phosphorylation and glucosidation of linamarin render this plant pro-toxin inert to the activating plant enzyme linamarase, thus representing pre-emptive detoxification strategies that avoid cyanogenesis.

Keywords: *Bemisia tabaci*, Transglucosidase, Phosphorylation, Cyanogenic glycosides, Detoxification

Introduction

Many plants produce two-component chemical defenses as protection against attacks from herbivores and pathogens. In these systems, protoxins that are often chemically protected by a glucosyl group are activated by an enzyme such as a glycoside hydrolase, yielding an unstable aglycone that is toxic or rearranges to form toxic products [1]. The glycoside and the hydrolase are stored in separate compartments that mix upon plant damage, activating the toxin only when necessary

for defense. Two-component defenses include cyanogenic, benzoxazinoid and iridoid glycosides and glucosinolates, and have long been known to play decisive roles in interactions between plants and herbivores, especially when extensive plant tissue disruption happens during feeding, such as during attack by chewing herbivores[1,2]. However, the activation of such defenses by piercing-sucking phloem-feeding herbivores such as aphids and whiteflies is poorly understood [3] in spite of the agricultural importance of these insects.

Cyanogenic glycosides are well-studied two-component plant chemical defenses believed to have arisen over 300 million years ago, and found in many diverse plant species [4-6]. Cyanogenic glycosides are O-β-glycosides of α-hydroxynitriles which are typically classified as aliphatic or aromatic, depending on the amino acid from which they are derived [7]. They also occur as disaccharides in some plant species, believed to be important as more stable transport forms [8,9]. Crops that produce cyanogenic glycosides include several legumes and fruits, as well as the tropical staple crop cassava (manioc, yuca; Manihot esculenta). Cassava originated in the Amazon basin [10] and was introduced to Africa in the 16th century [10], where it has become an extremely important crop, being the currently most produced carbohydrate staple crop in Central and Western Africa [11,12] with production expected to grow further to 100 Mt in Sub-Saharan Africa by 2025 [12]. The main cyanogenic glycoside present in cassava is a valine-derived cyanogenic mono-glycoside called linamarin, highly abundant in both aerial and root tissues of this plant [13], reaching up to 2 mg cyanide per g dry weight [14]. The hydrolytic enzyme required for the activation of linamarin and other cyanogenic glycosides in cassava is a β-glucosidase commonly referred to as linamarase [13,15] that, upon tissue damage, cleaves linamarin to give an unstable hydroxynitrile that rearranges to produce the respiratory toxin cyanide and acetone. This reaction happens spontaneously but can be accelerated by an enzyme known as hydroxynitrile lyase [16]. As a crop, cassava displays resistance to drought, however it is challenged by several viral diseases (e.g. cassava mosaic, cassava brown streak, and cassava vein mosaic virus to name a few) [17], many of which are vectored by biotypes of the tobacco whitefly (Bemisia tabaci Sub-Saharan Africa, BtSSA) that feed on cassava phloem tissue [18]. This problem of severe viral epidemics is gaining increased relevance in Africa in recent years [11] and is often associated with large whitefly outbreaks [11] that contribute as vectors for viral complexes.

The whitefly B. tabaci is a complex of cryptic species or biotypes that are morphologically indistinguishable [19]. Collectively they are polyphagous phloem feeders able to feed on over 600 species of plants [20], and thus an extremely important as a crop pest named in the top 100 of the world's worst [21]. This is in part due to its ability to vector over 300 plant viruses [21]. As a consequence, discovery of the mechanisms by which whiteflies feed so successfully on crops represents an important area of research for the protection of African small share-holder farmers, and could lead to new control measures in conjunction with other whitefly research [11,21-23].

There is evidence for the involvement of two-component defenses in the interactions of phloem feeders and their host plants [24-30]. Glucosinolates have been shown to affect aphid feeding with certain indolic glucosinolate hydrolysis products detected after Myzus persicae feeding [31], In B. tabaci, glucosinolates were observed to alter performance and host selection [32]. Furthermore these latter insects were shown to be able to detoxify glucosinolates via the formation of desulpho-glucosinolates [33], and more recently via novel sucrase-transglucosidases and the transfer of multiple additional glucose residues [30]. Finally, in the case of cassava cyanogenic glycosides, the activities of the cyanide detoxification enzymes rhodanese and beta-cyanoalanine synthase in whiteflies are increased when feeding on cassava in comparison to sweet potato (cyanogenic versus non-cyanogenic plants) [24].

Here we provide evidence supporting the role of cassava cyanogenic glycosides in defense against the phloem feeder B. tabaci, and elucidate multiple pathways for the detoxification of these

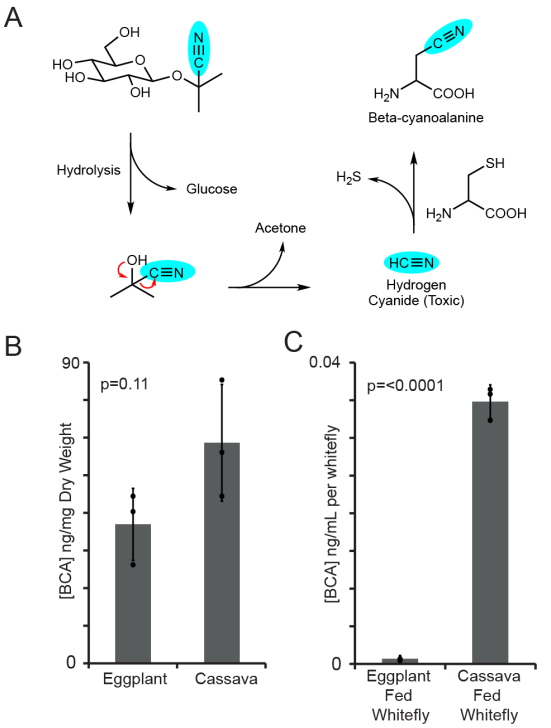


Figure 1. Cyanogenic glycosides are activated during whitefly feeding. (A) Scheme of linamarin hydrolysis and the known detoxification pathway of beta-cyanoalanine formation. **(B)** Native beta-cyanoalanine concentrations in both eggplant and cassava leaves were not different (P-value from unpaired t-test with N=3). (C) Whitefly extracts from insects feeding on cassava showed elevated levels of beta-cyanoalanine (BCA) in comparison to eggplant-fed insects, supporting the hydrolysis of cyanogenic glycosides during feeding (P-value from unpaired t-test with N=3).

compounds in this cosmopolitan pest. We first describe the formation of beta-cyanoalanine as a strategy to mitigate the toxicity of the cyanogenic glycoside hydrolysis product hydrogen cyanide [34] in the bodies of B. tabaci SSA1 (BtSSA1). Then we elucidate the detoxification of cyanogenic glycosides via both transglucosidation and phosphorylation of sugar moieties. Since the glucoslylated and phosphorylated derivatives are resistant to hydrolysis by the plant activating enzymes, they therefore serve as pre-emptive detoxification pathways.

Results

Beta-cyanoalanine is increased in cassava-fed whiteflies

The potential activation of cyanogenic glycosides during whitefly feeding would lead to the release of hydrogen cyanide. We therefore investigated the levels of a known cyanide detoxification metabolite, beta-cyanoalanine (Figure 1 A), in whiteflies feeding on leaves of the cyanogenic plant cassava (M. esculenta) in comparison to beta-cyanoalanine levels in a non-cyanogenic plant, eggplant (Solanum melongena). Since all plants produce beta-cyanoalanine as a natural by-product of ethylene biosynthesis [34,35] the concentrations of this amino acid in the tissues of both of these plant species were also measured to determine the native levels of this compound prior to ingestion. The levels of beta-cyanoalanine in the respective plants were not different (p=0.11 N=3) (Figure 1 B); however, cassava-fed whiteflies produced much higher levels of this compound than whiteflies feeding on eggplant (p=<0.0001, N=3) (Figure 1 C). The accumulation of beta-cyanoalanine provides evidence for activation of cyanogenic glycosides during feeding by the whitefly B. tabaci.

Cassava cyanogenic glycosides are metabolized by glucosylation by whiteflies

Honeydew from whiteflies (BtSSA1-SG3) feeding on cassava was collected and chemically analyzed via an untargeted LC-MS approach for the presence of the cyanogenic glycoside linamarin and possible derivatives. A peak corresponding to the native glycoside was easily observed, confirming that this insect was exposed to linamarin while

feeding on cassava phloem sap. Additionally, MS signals consistent with glycosylated linamarin derivatives were also detected. The molecular masses of these putative products presented a characteristic serial mass addition of +162Da units up to +648Da, suggestive of glucosylation, with elution times being slightly shorter with each addition. These peaks were also detected in honeydew from whiteflies (BtMEAM1) fed artificial diets consisting of sucrose and linamarin, but not when fed sucrose alone, confirming they were linamarin derivatives produced by the insects.

In order to elucidate the structure of these whitefly linamarin derivatives, the compounds corresponding to the addition of 1 and 2 apparent glucose moieties, which also corresponded to the most intense MS signals, were purified using HPLC fractionation. NMR analyses revealed that the structures of these compounds were consistent with metabolism via glucose conjugation. The derivatives with additional glucose moieties showed sugar addition to the original glucose of linamarin in either an α-(1→6) or α-(1→4)

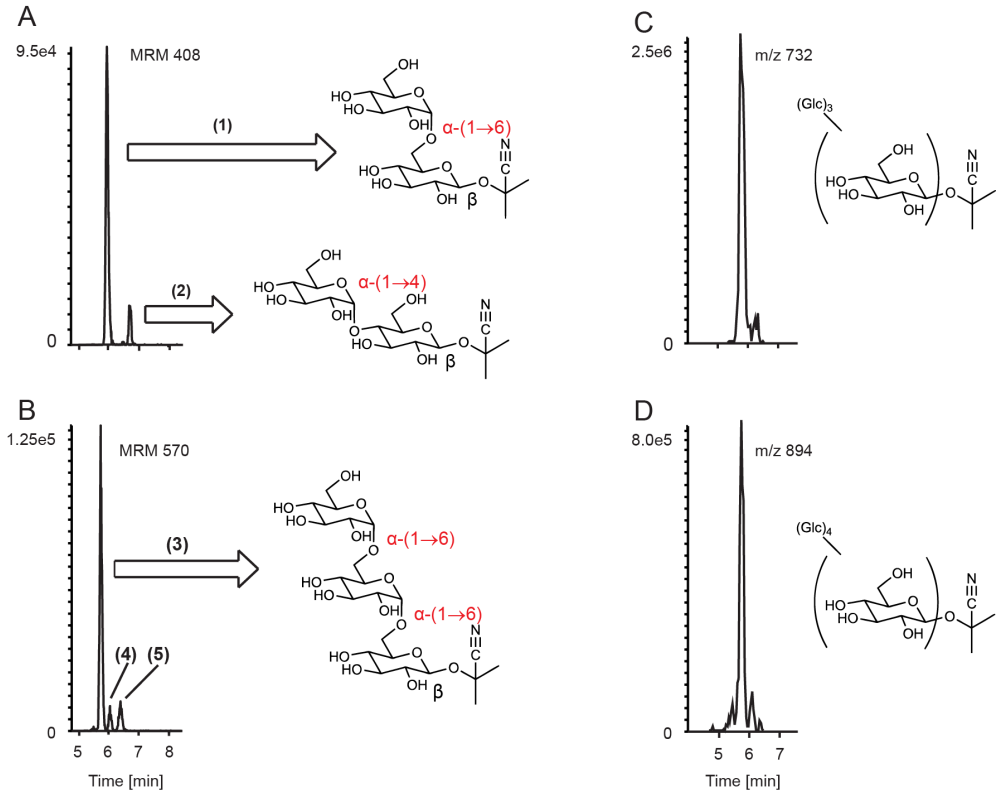


Figure 2. Novel linamarin-derived metabolites in *B. tabaci* whitefly honeydew. Linamarin-derived metabolites in the honeydew contain additional alpha-linked glucose moieties. Panel (**A**) **s**hows metabolites corresponding to 1 glucose addition (compounds **1** and **2**) and Panel (**B**) shows metabolite peaks corresponding to 2 glucose additions (Compounds **3, 4** and **5**). Metabolites with up to 4 further hexose additions were detected via untargeted analysis (Panels **C** and **D**). Compounds **1, 2** and **3** were purified from honeydew and their structures determined by MS and NMR.

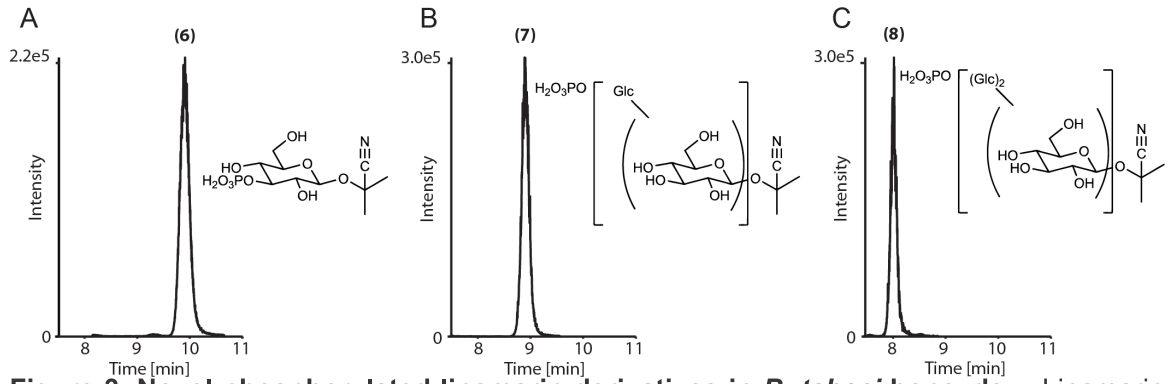


Figure 3. Novel phosphorylated linamarin derivatives in *B. tabaci* honeydew. Linamarin and its insect-produced glycosides were further modified by *B. tabaci* via phosphorylation. (**A**) The position for phosphorylation of linamarin was elucidated as C3 **6**. The structures of both the monoglucosylated (**7**, panel **B**) and diglucosylated (**8**, panel **C**) phosphorylated derivatives could not be determined by NMR analysis due to low abundance, but the mass spectral data are completely consistent with the structures given

orientation (compounds **1** and **2,** respectively Figure 2 A and Supplemental Note 1 Table SN1-1) with the latter having slightly greater retention. The diglucose derivative showed the serial addition of two α-(1→6) linked glucose moieties to the previously existing β-linked glucose (Compound **3**, Figure 2 B and Supplemental Note 1 Table SN1-1). Quantitative ^{1}H-NMR techniques using sucrose as an external standard were utilized to estimate the amount of the monoglucose derivative **1** and diglucose derivative **3** purified (5.88 µg and 3.38 µg respectively). LC-MS standard curves for these purified metabolites were then constructed, allowing for the calculation of molar ratios for these compounds in comparison to intact linamarin in the honeydew of BtSSA whiteflies feeding on cassava. It was found that glycosylated linamarin derivatives corresponding to 1 and 2 glucose additions were present in the honeydew in a combined 5.62:1 (SE=0.89, N=3) ratio to intact linamarin. Additional peaks with mass spectra consistent with two glucose additions were also observed (Compounds **4** and **5**); however the quantities and purities of the fractions obtained after chromatographic separation were not sufficient for adequate structure elucidation. Glycosides with masses corresponding to three and four glucose additions to the original β-linked glucose were also detected (Figure 2 C and D).

<u>The cyanogenic glycoside linamarin is metabolized by phosphorylation</u>

In addition to glucosides, further unknown metabolites were detected in the honeydew of both cassava-fed whiteflies (BtSSA1) and those that fed on linamarin-containing artificial diets (BtMEAM1), but were absent in diets not containing linamarin. These metabolites again showed a pattern suggesting serial glucose additions ($+162_n$), with characteristic progressively earlier eluting peaks. The smallest of these metabolites displayed a mass of 326 Da (Figure 3 A) **6**, which is 80 mass units greater than linamarin, but eluted much later. There were also two earlier eluting metabolites consistent with +162 and +324 Da additions to **6** (Figure 3 B and C, compounds **7** and **8** respectively). We hypothesized that the addition of 80 Da could correspond to either a sulphate or phosphate group linked to linamarin, with the earlier eluting peaks being a result of subsequent glycosylation. The addition of a phosphate was further supported by accurate mass data (Supplemental Figure S1 and Supplementary Table S2) and the disappearance

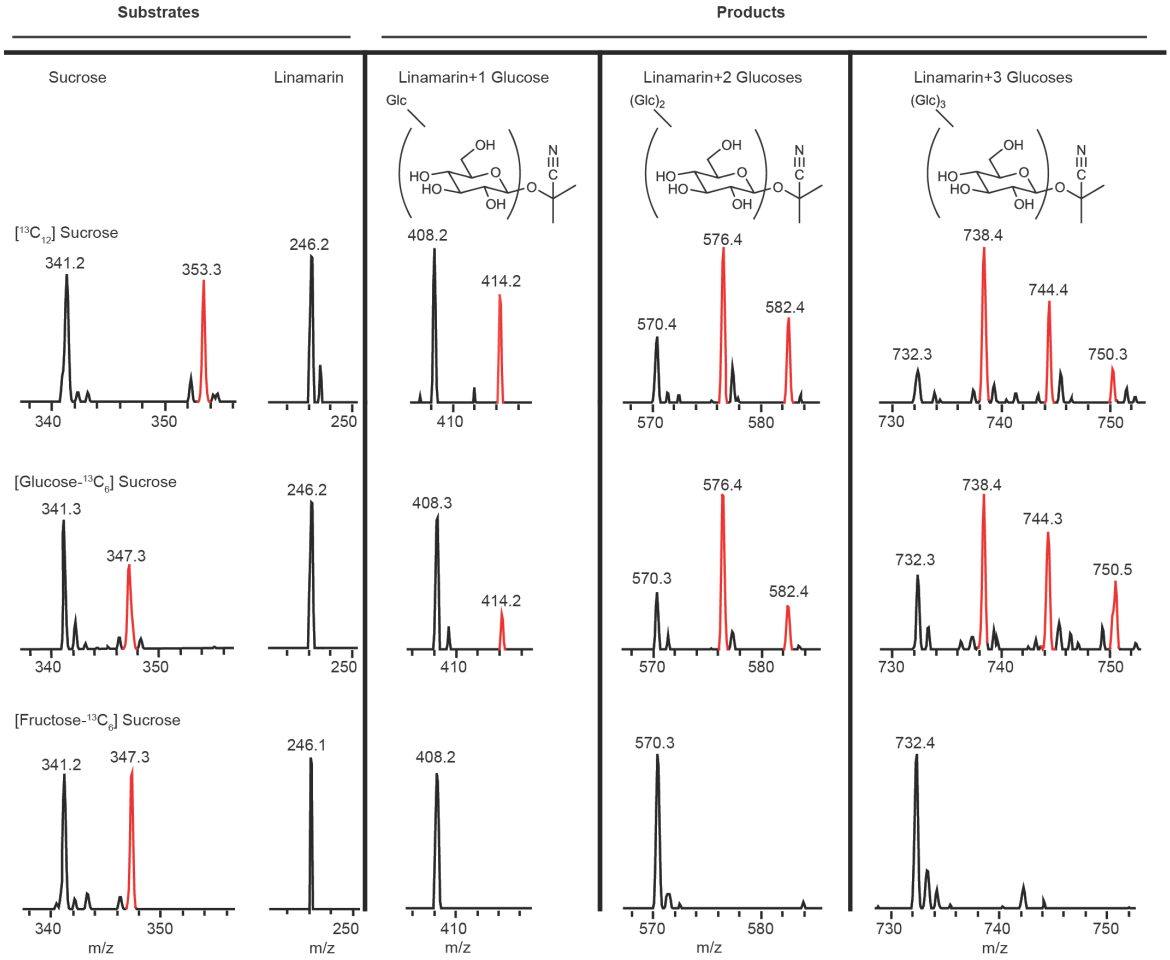


Figure 4. Incorporation of glucose during *B. tabaci* glucosylation of linamarin. Glucosylation of cyanogenic glycosides in *B. tabaci* is catalyzed by a transglucosidase activity based on feeding of [^{13}C] sucrose isotopologues and the CN-Glc linamarin to insects in artificial diet. Shown are MS regions from LC-MS analyses of substrates and products of feeding experiments on five different diets. Feeding fully ^{13}C-labeled sucrose and sucrose ^{13}C-labeled in the glucose moiety gave labeling in the newly added glucose of glucosylated linamarin metabolites. However, feeding sucrose ^{13}C-labeled in the fructose moiety gave no label in the products, demonstrating incorporation of only the glucose of sucrose with the CN-Glc. Feeding of free ^{13}C-labeled glucose and fructose with linamarin also gave no incorporation into glucosylated derivatives of linamarin (Supplementary Figure S3), showing that the glucose must originate from sucrose.

of these metabolites upon incubation with alkaline phosphatase (Supplemental Figure S2). Purification via HPLC fractionation followed by NMR analysis revealed the addition of a phosphate to linamarin at position 3 of the β-linked sugar **(6)** based on the deshielding of the ^{1}H and ^{13}C signals at this position (Supplemental Note 1 Figure SN1-5 A). Due to the low abundance of the purified products,

no NMR spectroscopic evidence could be obtained for the glucosylated phosphate derivatives.

Mechanism of glucose additions to linamarin

Glycosylation reactions are most often catalyzed by either of two known enzyme classes, UDP-glucosyl-transferases (UGTs) and transglucosidases The latter of these two enzymes utilizes a mechanism which transfers a glucose unit from a donor disaccharide directly to an acceptor molecule, while the former utilizes an activated form of glucose (UDP-glucose). In order to elucidate the mechanism by which glucose units are being added to linamarin by whiteflies, a simple artificial diet experiment using ^{13}C sucrose isotopologues was performed. Upon feeding isotopically labeled sugars in diets containing linamarin, labeled glucose was incorporated into the glycosides in the cases of diets containing [$^{13}C_{12}$] sucrose and [glucose-$^{13}C_6$]sucrose, but not in diets containing [fructose-$^{13}C_6$]sucrose, [$^{13}C_6$]glucose or [$^{13}C_6$]fructose (Figure 4 and Supplemental Figure S3). The labels were also incorporated into phosphorylated glycosides in the same manner as described for the unphosphorylated glycosides (Supplemental Figure S4). This mechanism is consistent with a transglucosidase activity, which is typically carried out by enzymes of the glucohydrolase (GH) family [36].

Insect-derived linamarin products resist enzymatic activation

In order to determine whether the linamarin glycosides and further phosphorylated metabolites can be activated by the enzymes present within the plant to form cyanide, extracts of cassava leaves containing linamarase activity, as well as linseed extracts containing linustatinase [8,37] (a disaccharidase) activity were incubated with pure linamarin and with the honeydew of cassava-fed *B. tabaci*. Linamarin was degraded in the presence of both cassava enzyme extracts and linseed enzyme

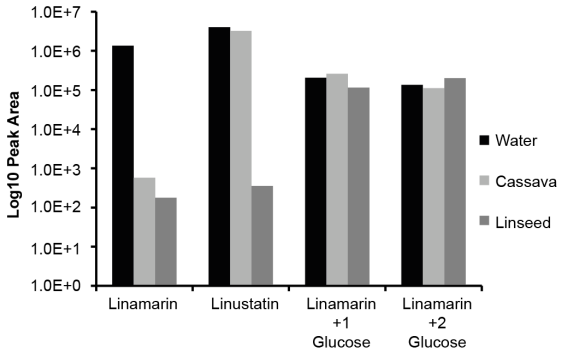


Figure 5. Stability of linamarin and glycosylated derivatives to activation by plant enzymes. Linamarin and linustatin as well as glycosides in the honeydew from whiteflies fed on cassava were tested. Honeydew and standards were incubated with crude enzyme extracts from cassava and linseed or with water alone. Linamarin was readily hydrolyzed by both cassava and linseed extracts, and linustatin was hydrolyzed by linseed enzymes. Conversely, the insect-derived glycosides were not substrates for both of the enzyme mixtures.

extracts, while the disaccharide linustatin was only hydrolyzed in the presence of linseed enzyme extract. The insect-derived glycosides, however, remained stable in the presence of both enzyme extracts, resisting both plant monosaccharidase and disaccharidase activities (Figure 5). Phosphorylated linamarin and phosphorylated linamarin glycosides also showed similar stability against enzymatic activation in the presence of the cassava extracts (Supplemental Figure S5). Therefore, formation of these derivatives likely serves as a true pre-emptive detoxification of linamarin rendering products that can no longer by hydrolyzed with the release of hydrogen cyanide.

Expression profiles of BtSUC2 and BtSUC5 in *B. tabaci* feeding on various host plant

BtSUC2 and 5 were previously identified

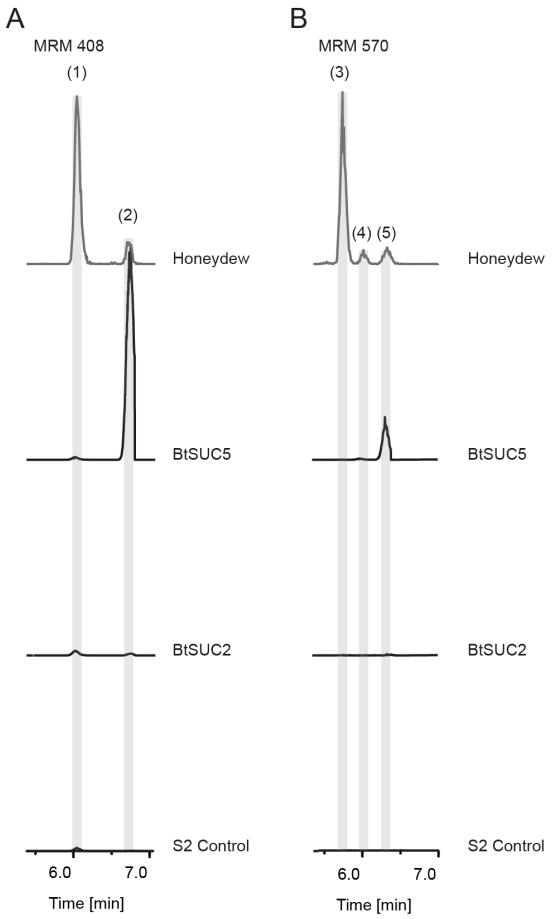


Figure 6. Chromatographic analyses of products from *B. tabaci* BtSUC2 and 5. Medium of *Drosophila* S2 cells expressing these enzymes was assayed with linamarin and sucrose. Depicted are extracted multiple reaction monitoring (MRM) LC-MS chromatograms for mono- and di-glucosylated linamarin derivatives. (**A**) BtSUC2 and 5 showed transglucosidation activity, producing α-(1→4)-linked glucose derivative **2** of linamarin with BtSUC2 also producing small amounts of the α-(1→6)-linked derivative **1** above control levels. (**B**) BtSUC5 additionally produced the diglycosylated product **5**. S2 cell control assay was of cell medium extracts of untransformed cells. (**1**: α-(1→6)-linked glucose derivative of linamarin; **2**: α-(1→4)-linked glucose derivative of linamarin; **3**: α-(1→6),α-(1→6)-linked diglucosylated derivative of linamarin; **4**, **5**: unknown diglucosylated derivatives of linamarin).

as whitefly transglucosidases that glucosylate glucosinolates to prevent their hydrolysis and release of toxic products [30]. Thus their potential to participate in the glucosylation of cyanogenic gluycosides was investigated. First the genomic sequences of *BtSUC2* and *BtSUC5* in the cassava-feeding *B. tabaci* biotype SSA1-SG3 were found to be approximately 95% identical to their counterparts in the Brassica-feeding *B. tabaci* biotype MEAM1 (Supplemental Figure 6A). Next, the expression of the *BtSUC2* and *BtSUC5* genes was compared in the two whitefly biotypes feeding on two different plant species. It was found that *BtSUC2* was expressed to a greater extent on cassava in MEAM1 whiteflies with no significant increase in expression on eggplant in comparison to SSA1-SG3 whiteflies, and *BtSUC5* was overall expressed to a much larger extent in the cassava associated whiteflies SSA1-SG3(supplemental figure 6B), both on cassava and eggplant as food sources. The inducibility of these genes was also assessed for each whitefly species and it was observed that both genes were expressed to a greater extent in MEAM1 when feeding on cassava in comparison to eggplant. Only *BtSUC2* was seemingly induced in SSA1-SG3 whiteflies when feeding on cassava, with *BtSUC5* showing signs of constitutive expression of this gene (Supplemental Figure S6C).

In vitro activities of previously characterized GH enzymes BtSUC2 and BtSUC5

Drosophila S2 cells expressing BtSUC2 and BtSUC5 were utilized in enzyme assays with cyanogenic glycosides and other substrates with the donor disaccharide sucrose and the acceptor molecule linamarin. BtSUC2 produced the α-(1→4) glycoside **2** approximately 50 times more efficiently than S2 control cells, and smaller amounts of the α-(1→6) derivative **1** (Figure 6 A). No higher order glycosides were formed by BtSUC2 (Figure 6 B). BtSUC5 on the other hand produced much larger amounts of **2** (more than 50,000 times more efficient than S2 control cells) (Figure 6 A), as well as a glycoside corresponding to a further glucose addition (**5**) (Figure 6B) with a similar efficiency. BtSUC5 also produced a smaller amount of **1** (approximately 7 times more active than controls) and a peak corresponding to an unknown derivative **4 (**Figure 6 A & B respectively). BtSUC5 displayed apparent Michaelis-Menten kinetics with an estimated K_M of ~0.5 mM towards linamarin and ~0.13 M towards sucrose (Supplemental Figure S7). For BtSUC2, however, the low activity towards linamarin prevented an estimation of K_M values for this enzyme. BtSUC2 and BtSUC5 activities towards the purified phosphorylated linamarin derivatives **6** and **7** were also investigated, although the precise amounts of these compounds added were not determined due to their low abundance. Nevertheless, both enzymes seemed to glucosylate **6** forming a product with very slightly different retention time compared to **7** (Supplemental Figure S8 A), suggesting a positional isomer. While BtSUC2 could apparently not use **7** as substrate, BtSUC5 produced a compound corresponding by mass to **8,** but with a different retention time (Supplemental Figure S8 B).

BtSUC2 and BtSUC5 were also assayed with other plant-derived glucosides as potential substrates. Secondary metabolites from various two-component defensive compound classes were tested including benzoxazinoid, cyanogenic, phenolic, iridoid, flavonoid and other cyanogenic glycosides, as well as glucosinolates. BtSUC2 and 5 had detectable transglucosidase activity with all substrates (Supplemental Figure S9), with BtSUC5 having the greatest activity.

Discussion

Phloem feeding insects have long been thought to be largely unaffected by activated two-component plant defenses, as the feeding mode of these herbivores does not appear to cause enough macroscopic tissue damage to initiate hydrolytic activation of these compounds. Hence these insects may be assumed to not need mechanisms which detoxify these compounds. This report, however, offers further evidence of plant glycoside two-component defense activation during feeding by a piercing-sucking herbivore [25,27,28,30]. The production of aliphatic glucosinolate hydrolysis products, isothiocyanates [25,30] and subsequent glutathione conjugates [30], in *B. tabaci* feeding on Brassica plants offers evidence of tissue disruption during phloem feeding since aliphatic glucosinolates are not spontaneously activated as is the case with indole-derived glucosinolates [28]. Here, we report whitefly accumulation of beta-cyanoalanine, a metabolite derived from the detoxification of cyanide. Beta-cyanoalanine is also produced by plants upon detoxification of the cyanide produced by the enzyme ACC oxidase during production of the hormone ethylene [34,35]. However, while undamaged cassava and eggplant tissues contained similar amounts of beta-cyanoalanine (Figure 1 B), insects fed on these plants exhibited stark contrasts with cassava-fed whiteflies containing 50 times higher amounts of this detoxification metabolite than their eggplant-fed counterparts (Figure 1 C). Activation of cyanogenic glycosides is presumed to happen upon stylet penetration, navigation to the phloem, or in the gut of the whitefly post ingestion. A beta-cyanoalanine synthase has been identified in *Pieris rapae* [38] and although an exact homologue has not been characterized in *Bemisia tabaci*, a cystathione beta-synthase (Bta12658 or Ssa04689)

is predicted to have similar functionality [21,35].

Given the activation of cyanogenic glycosides upon whitefly feeding, we explored whether this insect possessed any pre-emptive detoxification measures. The glycosylated linamarin derivatives found in the honeydew of whiteflies feeding on cassava with the additions of α-glycosides in either an α-(1→6) or α-(1→4) fashion are mirrored in the glucosylation of the glucosinolate 4msob [30]. However, subtle differences in the ability of these insects to modify cyanogenic glucosides are clear by the appearance of many more peaks than for glucosylated glucosinolate metabolites. Multiple sugar additions occur that are hypothesized to be combinations of various linkage modes (α-(1→6) or α-(1→4)) of glucose. For example, chromatographically **1** (an α-(1→6) addition) shows an earlier elution time than the α-(1→4) derivative **2**. The most abundant peak from two glucose additions **3** with the earliest elution time was determined to be α-(1→6); α-(1→6). The additional two peaks may therefore correspond to a compound with both α-(1→6);α-(1→4) mixture **4** and a compound with two α-(1→4);α-(1→4) additions **5**. This is supported by the fact that the enzyme BtSUC5, which produces almost exclusively α-(1→4) glucosylated linamarin **2** in incubations with sucrose and linamarin, also produces the latest eluting also produces the latest **5** of the diglycoside peaks (Figure 6 B). This catalytic flexibility seems to be heavily influenced by the "R group" which is hypothesized to allow for more promiscuity in the pocket of the responsible modifying transglucosidases. When looking at the various substrates tested with enzymes BtSUC2 and 5, rutin (the largest substrate) was the poorest acceptor (Supplemental Figure S8) as seen by the retention of substrate peak intensity in all assays. A similar size-oriented influence of activity is observed in studies with a dextransucrase where the velocity of the reaction was reduced when switching the acceptor from maltose to maltotriose [39].

Phosphorylation of plant and insect derived glycosides was also detected as part of whitefly metabolism of linamarin. Phosphorylation of sugars is often a marker for catabolism or breakdown, as in the case of glycolysis [40,41]. However, phosphorylation may also be utilized in cases of transglycosylation such as in the process of glycogen synthesis [42,43] where C3 phosphorylation (a rare addition) [44-46] plays an important role in glycogen synthesis and storage [42]. Phosphorylation of plant glycosides was also reported in the gypsy moth *Lymantria dispar* and its close relative *Orgyia antigua*, where salicinoids were phosphorylated on the β-linked glucose at position C3 [47]. Here we also observed the phosphorylation of linamarin at the same position **6**.

Phosphorylation in *B. tabaci* is hypothesized to be independent from the glucosylation mechanism due to the capability of BtSUC2 and BtSUC5 to glucosylate both phosphorylated and non-phosphorylated substrates (Supplemental Figure S8). The biological role of phosphorylation is likely as a detoxification process since phosphorylation stabilizes glycosides against activation by linamarase. Detoxification has already been proposed as an explanation for the formaton of phosphorylated phenolic glycosides in *Lymantria dispar* [47]. Although phosphorylation represents an atypical mammalian phase II detoxification process, reports in insects are growing in number [48-50] and this might represent a typical conjugation in this order of animals.

BtSUC2 and BtSUC5 both transglucosylated the cyanogenic glycoside linamarin *in vitro,* showing their promiscuity to modify both glucosinolates and cyanogenic glycosides. BtSUC2 managed to produce α-(1→6) **1** and α-(1→4) **2** linked monoglycosylated products above control levels, with α-(1→6) not previously observed with glucosinolates [30]. BtSUC5 also displayed other activities, being capable of producing a not yet isolated compound **5** with a second glucose addition, possibly corresponding to the abovementioned α-(1→4);α-(1→4) diglycoside. K_M values for BtSUC5 were within physiologically expected levels for both linamarin and sucrose. The role of both of these enzymes in detoxifying cyanogenic glycosides is supported by the greater expression of their corresponding

genes when flies fed on cassava, which contains cyanogenic glycosides, than when fed on eggplant, which does not contain any cyanogenic glycosides.

Further incubations of BtSUC2 and 5 with various substrates resulted in the transglucosidation of virtually all tested metabolites (Supplemental Figure S9), aside from rutin for which activity was very low. However, when some of these substrates (e.g. benzoxazinoids) were fed directly to whiteflies, no transglucosidase products were detected in the honeydew [30]. Given that whiteflies use a variety of host plants, many of which have glucosylated two-component defenses, the formation of pre-emptive detoxification metabolites may be responsible for preferences of whiteflies for certain plants. Much more work needs to be done, however, in identifying potential defenses in whitefly host plants and determining how they are processed by the detoxification machinery. Investigations into the honeydew of other phloem feeding insects that regularly feed on plants containing such defenses would also be interesting, in order to determine if these types of detoxification reactions differ between specialist and generalist phloem feeders.

The inducibility or relative expression of these BtSUC2 and 5 was variable with respect to both the whitefly species and the plant which they fed upon. Perhaps the most striking observation was the expression levels of BtSUC5 in the cassava associated whiteflies SSA1-SG3 in comparison to MEAM1 whiteflies (Supplemental Figure S6B). BtSUC5 in SSA1-SG3 was seemingly not induced when feeding on cassava versus eggplant (Supplemental Figure S6C) and may speak to the observation of the constitutive nature of *B. tabaci* detoxification genes [51], however it may also be evidence of the lack of involvement of this specific GH13 genes for detoxification of cyanogenic glycosides.

In conclusion, the investigations into whitefly metabolism of cyanogenic glycosides illustrate the plethora of chemical transformations carried out by this phloem feeding insect. Although cyanogenic glycosides and other two component defenses were previously believed not to be activated by this feeding guild, our results suggest that whiteflies are indeed susceptible to two-component defenses and developed metabolic adaptations to pre-emptively detoxify them. Further research on phloem-feeders should discover many more metabolic adaptations to plant defenses.

Materials and Methods

Plants

Eggplant (*Solanum melongena*, cv. Black Beauty), and cassava (*Manihot esculenta*, cv. MCol22) plants were grown under standard greenhouse conditions at 26±2 ^{0}C with supplemental lighting and a photoperiod of 14:10 h (light:dark).

Insects

Bemisia tabaci (Hemiptera: Aleyrodidae) species MEAM1 were collected in southern Israel in 2003 and from Sudan in the late 1990s, and reared continuously on cotton . *Bemisia tabaci* of species SSA1-SG3 (sub-Saharan Africa species group) was collected on Bagamoyo Road, Tanzania in 2013 and reared on cassava plants.

Insect feeding on *Manihot esculenta* and other species

Groups of adult whiteflies (50-100 individuals), were collected from their host plants and switched to fresh cassava (SSA1-SG3) or eggplant (MEAM1 and SSA1-SG3). Insects were fed on plants enclosed within glass clip-cages. After 72-96 h, the honeydew/feces deposited on the glass feeding chamber was washed off with water : methanol (20:80, v:v). Samples were dried down using nitrogen gas and resuspended in water prior to analysis.

B. tabaci feeding on artificial diets with linamarin

Groups of 150 *B. tabaci* (MEAM1 species) adults were collected from host plants and switched

to artificial diet feeders (consisting of a glass tube, 3 cm height × 2 cm diameter, with a liquid diet covered with a double layer of Parafilm). Insects were allowed to feed through the Parafilm on a 10 % sucrose solution containing no additives (control) or the cyanogenic glycoside linamarin (Sigma) at a concentration of 5 mM. After 96 h, the honeydew deposited on the glass tubes was washed off with water : methanol (20:80, v:v). Honeydew was then dried down under nitrogen gas and resuspended in water prior to analysis.

Bemisia tabaci feeding on artificial diets containing isotopically-labeled sugars

Four different sucrose isotopomers were added to artificial diets: [$^{12}C_{12}$]sucrose, [$^{13}C_{12}$] sucrose, [glucose-$^{13}C_6$]sucrose, and [fructose-$^{13}C_6$]sucrose. The monosaccharides [$^{13}C_6$]fructose and [$^{13}C_6$]glucose were also fed. The artificial feeding devices consisted of a glass tube (5 cm high x 2.5 cm diameter) with the liquid diet (50 µL) held within a double layer of Parafilm. About 50 *Bemisia tabaci* MEAM1 adults were placed in each tube. Feeding assays were performed for 72 h on diets that contained 5 mM linamarin and 0.29 M of the labeled sugars. After feeding, vials were placed at -80 ºC in order to kill whiteflies, and bodies were transferred and stored for analysis. The honeydew deposited on the glass tubes was washed with water : methanol (20:80, v:v) and stored at -20 ºC until processing and analysis. A full summary of artificial diet constituents is outlined in Supplementary Table S2.

Purification and LC-MS analysis of glucosylated and phosphorylated cyanogenic glycosides, subsequent hydrolysis products and sugars

Purification of glucosylated or phosphorylated cyanogenic glycosides was performed via fractionation on a Nucleodur Sphinx RP column (250 × 4.6 mm, 5 µm, Macherey–Nagel, Düren, Germany) using an HP 1200 HPLC (Agilent Technologies, Santa Clara, CA, USA) coupled to a fraction collector (Advantec, Dublin, CA, USA). Chromatographic separation was attained using a gradient of 0.05 % aqueous formic acid (Solvent A) and acetonitrile (Solvent B) at a flow rate of 1 mL min^{-1} at 20°C as follows: 5-29 % B (12 min), 29-100 % B (0.1 min), a 2.9 min hold at 100% B, 100-5 % B (0.1 min), and a 3.9 min hold at 5 % B. Linustatin was purified from flax seeds by crushing 10 g of seeds in liquid nitrogen and extracting with 80% methanol prior to centrifugation at 10000 x g. The methanol supernatant was then dried down and concentrated before resuspension in water for purification.

Qualitative analysis of glucosylated cyanogenic glycosides in feces and honeydew extracts was also performed on an HP 1100 series HPLC. Separation was achieved on a Nucleodur Sphinx RP column (250 × 4.6 mm, 5µm, Macherey-Nagel, Germany) with a gradient of 0.2 % aqueous formic acid (solvent A) and acetonitrile (solvent B) with a flow rate of 1 mL min^{-1} at 25°C as follows: as follows 5-55 % B (25 min), 55-100 % B (0.1 min), 100 % B 0.9 min hold, 100-5 % B (0.1 min), 5 % B 3.9 min hold. The HPLC was coupled to an Esquire 6000 ESI-Ion Trap mass spectrometer (Bruker Daltonics, Bremen, Germany) operated in both positive and negative modes in the range of *m/z* 60–1500 with skimmer voltage -40 V; capillary exit voltage -146.7 V; capillary voltage 4000 V; nebulizer pressure 35 psi; drying gas 11 L min^{-1}; and gas temperature 330°C. DataAnalysis software V4 (Bruker Daltonics) was used for chromatogram analysis.

Qualitative analysis of isotopically labeled glucosylated cyanogenic glycosides in feces and honeydew extracts was performed on an HP 1100 series HPLC. Separation was achieved on a Nucleodur Sphinx RP column (250 × 4.6 mm, 5µm, Macherey-Nagel, Germany) with a gradient of 0.2 % aqueous formic acid (solvent A) and acetonitrile (solvent B) with a flow rate of 1 mL min^{-1} at 25°C as follows 5-55 % B (25 min), 55-100 % B (0.1 min), 100 % B 0.9 min hold, 100-5 % B (0.1 min), 5 % B 3.9 min hold. The HPLC was coupled to an Esquire 6000 ESI-Ion Trap mass spectrometer (Bruker Daltonics, Bremen, Germany) operated in both positive and negative modes in the range

of *m/z* 60–1000 with skimmer voltage -40 V; capillary exit voltage -113.5 V; capillary voltage 4000 V; nebulizer pressure 35 psi; drying gas 11 L min^{-1}; and gas temperature 330°C. DataAnalysis software V4 (Bruker Daltonics) was used for chromatogram analysis.

High resolution mass spectrometry of phosphorylated compounds was achieved on an Thermo Scientific UltiMate 3000 UHPLC coupled to Bruker TIMS-TOF mass spectrometer. Separation was achieved on a Nucleodur Sphinx RP column (250 × 4.6 mm, 5µm, Macherey-Nagel, Germany) with a gradient of 0.2 % aqueous formic acid (solvent A) and acetonitrile (solvent B) with a flow rate of 1 mL min^{-1} (split 1:3 source: waste) at 25°C as follows: as follows 5-55 % B (25 min), 55-100 % B (0.1 min), 100 % B 0.9 min hold, 100-5 % B (0.1 min), 5 % B 3.9 min hold. The MS was operated in negative mode scanning from *m/z* 50-1500 with the following parameters. **Source:** End plate offset: 500 V, capillary: 3500 V, Neubilizer: 3.5 bar, Dry gas: 11.0 L min^{-1}, Dry temperature 330 °C. **Tune General:** Funnel 1RF: 150 Vpp, Funnel 2 RF: 200 Vpp, isCID energy: 0.0 eV, Multipole RF: 50 Vpp, Deflection Delta: -70 V, Quadrupole energy: 4.0 eV, Low mass: 90 *m/z* Collision energy: 7.0 eV, Collision RF: 400 Vpp, Transfer time: 80.0 µs, Pre-pulse storage: 5.0 µs. Calibration took place externally immediately before the samples were run using Agilent ESI-L Low Concentration Tune Mix and an enhanced quadratic calibration curve.

Quantification of the glucosylated cyanogenic glycosides in transglucosidation and linamarase/linustatinase assays and in honeydew was accomplished via an HP 1260 series HPLC coupled to an AB Sciex API 5000 mass spectrometer (Applied Biosystems, Darmstadt, Germany). The column utilized was a Nucleodur Sphinx RP column (250 × 4.6 mm, 5µm, Macherey-Nagel, Germany) using a chromatographic gradient of 0.05 % aqueous formic acid (Solvent A) and acetonitrile (Solvent B) at a flow rate of 1 mL min^{-1} at 20°C as follows: 5-29 % B (12 min), 29-100 % B (0.1 min), a 2.9 min hold at 100% B, 100-5 % B (0.1 min), and a 3.9 min hold at 5 % B.The MS was operated in the negative mode with collision gas value 7, curtain gas pressure 35 psi, spray gas pressures 60 psi, ion spray voltage -4500 V, and turbogas temperature 600°C. Compounds were detected using multiple reaction monitoring (MRM) detection with the parameters outlined in Supplementary Table S3. Quantification was achieved using external calibration curves constructed from solutions of purified glucosylated cyanogenic glycosides of known concentrations (determined in solution via NMR as described below). Analyst 1.5 software (Applied Biosystems) was used for data acquisition and processing. All averages and standard errors were calculated from three independent biological replicates. No other statistical tests were performed.

Analysis of betacyanoalanine accumulation in the bodies of *Bemisia tabaci* SSA1-SG3 fed on cassava and eggplant was performed on an HP 1260 HPLC coupled to an AB Sciex API 5000 mass spectrometer. The column utilized was a Agilent XDB-C18 column (50 × 4.6 mm, 1.8µm, Agilent Technologies, Boeblingen, Germany) using a chromatographic gradient of 0.05 % aqueous formic acid (Solvent A) and acetonitrile (Solvent B) with a flow rate of 1.1 mL min^{-1} at 25°C as follows: 0.5 min hold at 10 % B, 10-45 % B (3.5 min), 45-100 % B (0.02 min), 0.98 min hold at 100 % B, 100-10 % B (0.02 min), 1.98 min hold at 10 % B. The mass spectrometer was operated in the negative mode with collision gas value 8, curtain gas pressure 25 psi, spray gas pressures 60 psi, ion spray voltage -4500 V, and turbogas temperature 700°C. Compounds were detected using scheduled multiple reaction monitoring (MRM) detection with the parameters outlined in Supplementary Table S3. Analyst 1.5 software was used for data acquisition and processing.

Sugar products from the transglucosidase enzyme assays were analyzed on an HP 1260 coupled to an AB Sciex API 5000 mass spectrometer with an apHeraNH2 Polymer column (150 × 4.6 mm, 5µm, Supelco Analytical, Munich, Germany) with a chromatographic gradient of water (Solvent A) and acetonitrile (Solvent B) at a flow rate of 1 mL min^{-1} at 20°C as follows: 80 % B (0.5 min), 80-

55 % B (12.5 min), 55-80 % B (1 min), hold at 80 % for 4 min. The mass spectrometer was operated in the negative mode with collision gas value 2, curtain gas pressure 35 psi, spray gas pressure 70 psi, ion spray voltage -4500 V, and turbogas temperature 700°C. Compounds were detected using scheduled multiple reaction monitoring (MRM) detection with the parameters outlined in Supplementary Table S3. Analyst 1.5 software was used for data acquisition and processing.

NMR spectroscopy

NMR spectra (^{1}H, ^{1}H-^{1}H COSY, ^{1}H-^{13}C HSQC, ^{1}H-^{13}C HMBC and ^{1}H-^{1}H SELTOCSY) were acquired on a 700 MHz Avance III HD spectrometer equipped with a 1.7 mm cryoprobe (Bruker Biospin, Rheinstetten, Germany). Further information is available in the supplementary NMR data files. Data acquisition and processing was accomplished using TopSpin ver. 3.2 (Bruker Biospin, Rheinstetten, Germany). Samples were measured in MeOH-d_3 or D_2O as indicated at 293 K. For quantification via ^{1}H NMR (10 s delay between scans), purified compounds were dried under N_2 flow and resuspended in D_2O, and a sucrose solution (3.13 mM in D_2O) was used as an external quantification standard.

Cloning and Expression in *Drosophila* S2 cells

Vectors, cells and cell media were obtained from ThermoFisher Scientific (Waltham, MA, USA). Full-length ORFs from previously characterized GH13 genes (*BtSUC2* and *BtSUC5*) were amplified from *Bemisia tabaci* MEAM1 whole-body cDNA using primer sets outlined in Supplemental Table S4, cloned into PCR4 Blunt TOPO and digested using specific restriction enzymes according to each primer as previously described [30]. The digested fragment was further ligated into the pAc5.1/V5-His A vector for expression in *Drosophila* S2 cells using Schneider's medium. Insect cells were co-transformed with pCOBLAST vector for selection via blasticidin. Transformation was achieved via $CaCl_2$ incubation following the manufacturer's standard protocols and selected with 50 µg mL^{-1} blasticidin in 6 well plates. After one week of selection, cells were maintained in T-flasks at 27°C with 10 µg mL^{-1} blasticidin. Cells were regularly maintained by splitting 1:20 when cells had achieved confluency of 90 % (approximately every 4-5 days). Cells and media were harvested for enzyme activity studies when cells reached 90 % confluency.

Enzyme Assays

Drosophila S2 cells expressing candidate GH13 enzymes as well as non-transfected control cells were centrifuged at 100 ×*g* for 5 min. The resulting supernatant was utilized as the secreted protein fraction for enzyme assays. For determination of linamarin transglucosidation activity, 5 µL of supernatant from each culturecontaining an equivalent total protein content (verified by Bradford assays) was mixed with a 5 µL 50 mM phosphate buffer at pH 7.0 containing 2 M sucrose and 5 mM linamarin, and reacted for 6 h at 25°C with no stirring. These assay conditions were also repeated for other metabolites (I3M-GSL (Phytoplan Diehm & Neuberger GmbH), pOHBz-GSL(Phytoplan), DIMBOA-Glc , amygdalin (Roth), dhurrin (Roth), salicin (Sigma), arbutin (Roth), aucubin (Roth), and rutin (Sigma)), to test the breadth of substrate activity of these enzymes. Assay conditions were chosen after screening a range of substrate concentrations and pH values that reflect previous work on glucohydrolases and phloem-feeding insects. The reaction was stopped using 15 µL of methanol and immediately stored at -20°C. Before analysis, enzyme assays were centrifuged at 5200 ×*g* for 5 min and the supernatant obtained was analyzed by LC-MS. A dilution of 1:10 in water was used for analysis of glucosylated linamarin, and a dilution of 1:1000 in water used for analyzing sugars such as glucose, fructose and trisaccharides formed from these incubations.

For K_M estimation, crude enzyme preparations of control cells were demonstrated to catalyze less than 1% of the supplied linamarin to

transglucosidated products and also catalyze low level of sucrose hydrolysis. Enzyme assays were performed in the same manner as described above, except incubation times were reduced to 30 min. Peak areas were integrated and compared based on duplicate analysis. Sucrose K_M determination was performed at a constant linamarin concentration of 2.5 mM with sucrose concentrations ranging from 0.05 – 1 M. The linamarin K_M determination utilized sucrose at a final concentration of 1 M and a range of linamarin concentrations from 25 µM to 20 mM. All assays were carried out under linear reaction conditions with respect to time and protein concentration. Substrate concentration was never reduced below 95 % of the initial level during the 30 min assay period.

For tests on the reactivity of glucosylated cyanogenic glycosides with plant myrosinase, pure standard linamarin (sigma) (10 µL of a 5 mM solution), linustatin purified from seed extracts of flax (see above section on cyanogenic glycoside purification) and honeydew from cassava-reared adult SSA1-SG3 whiteflies containing insect-derived glucosylated linamarin glycosides were mixed with a crude leaf enzyme extract from cassava with native linamarase activity or crude flax seed extracts with native linustatinase activity in 20 mM phosphate buffer solution (10 µL), pH 7.0. Approximately 10 g of cassava leaves and 10 grams of flax seeds were flash frozen in liquid nitrogen and crushed using a mortar and pestle. These crushed materials were extracted with phosphate buffer (20mM pH 7.0) and filtered using vacuum filtration. The resultant extracts were utilized as cassava crude leaf and flax crude seed extracts for enzyme assays as described above. Negative control reactions were supplemented with 5 µL of water instead of the cassava or linseed extract solutions. Reactions were incubated at room temperature with no stirring, stopped after 1 h with 20 µL acetic acid and stored frozen until LC-MS analysis. Phosphorylated derivatives were tested for stability to cassava crude enzyme under the same reaction conditions; however flax crude enzyme preparation was not utilized.

Acknowledgments/ Funding

MPI-CE, DSMZ, and HUJI greenhouse teams for plant and insect maintenance, and other members of the African Cassava Whitefly Project (cassavawhitefly.org) for helpful discussions. This work was supported financially by the Max Planck Society, the Deutsche Forschungsgemeinschaft (DFG Collaborative Research Center 1127 ChemBioSys), and the Natural Resources Institute, University of Greenwich from a grant provided by the Bill & Melinda Gates Foundation (OPP1058938).

Competing Interests

The authors declare no competing interests

References

1 Morant, A. V. *et al.* ß-glucosidases as detonators of plant chemical defense. *Phytochemistry* **69**, 1795-1813, doi:10.1016/j.phytochem.2008.03.006 (2008).

2 Pentzold, S., Zagrobelny, M., Rook, F. & Bak, S. How insects overcome two-component plant chemical defence: plant ß-glucosidases as the main target for herbivore adaptation. *Biol Rev* **89**, 531-551, doi:10.1111/brv.12066 (2014).

3 Wang, X. W., Li, P. & Liu, S. S. Whitefly interactions with plants. *Curr Opin Insect Sci* **19**, 70-75, doi:10.1016/j.cois.2017.02.001 (2017).

4 Hansen, C. H. *et al.* CYP83B1 is the oxime-metabolizing enzyme in the glucosinolate pathway in Arabidopsis. *J Biol Chem* **276**, 24790-24796, doi:DOI 10.1074/jbc.M102637200 (2001).

5 Ratzka, A., Vogel, H., Kliebenstein,

D. J., Mitchell-Olds, T. & Kroymann, J. Disarming the mustard oil bomb. *P Natl Acad Sci USA* **99**, 11223-11228, doi:10.1073/pnas.172112899 (2002).

6 Bak, S. *et al.* Cyanogenic glycosides: a case study for evolution and application of cytochromes P450. *Phytochemistry Reviews* **5**, 309-329 (2006).

7 Poulton, J. E. Cyanogenesis in Plants. *Plant Physiology* **94**, 401-405 (1990).

8 Niedzwiedz-Siegien, I. Cyanogenic glucosides in Linum usitatissimum. *Phytochemistry* **49**, 59-63, doi:Doi 10.1016/S0031-9422(97)00953-9 (1998).

9 Selmar, D. Transport of cyanogenic glucosides: linustatin uptake by Hevea cotyledons. *Planta* **191**, 191-199 (1993).

10 Olsen, K. M. & Schaal, B. A. Evidence on the origin of cassava: Phylogeography of Manihot esculenta. *P Natl Acad Sci USA* **96**, 5586-5591 (1999).

11 Jacobson, A. L., Duffy, S. & Sseruwagi, P. Whitefly-transmitted viruses threatening cassava production in Africa. *Curr Opin Virol* **33**, 167-176, doi:10.1016/j.coviro.2018.08.016 (2018).

12 Cooperation, O. f. E., Development, Food & Organization, A. Agriculture in sub-Saharan Africa: Prospects and challenges for the next decade. *OECD-FAO agricultural outlook 2016-2025* (2016).

13 Mkpong, O. E., Yan, H., Chism, G. & Sayre, R. T. Purification, Characterization, and Localization of Linamarase in Cassava. *Plant Physiology* **93**, 176-181 (1990).

14 Nambisan, B. & Sundaresan, S. Distribution of Linamarin and Its Metabolizing Enzymes in Cassava Tissues. *J Sci Food Agr* **66**, 503-507, doi:DOI 10.1002/jsfa.2740660413 (1994).

15 Cooke, R. D., Blake, G. G. & Battershill, J. M. Purification of Cassava Linamarase. *Phytochemistry* **17**, 381-383 (1978).

16 Hughes, J., Decarvalho, J. P. C. & Hughes, M. A. Purification, Characterization, and Cloning of Alpha-Hydroxynitrile Lyase from Cassava (Manihot-Esculenta Crantz). *Arch Biochem Biophys* **311**, 496-502, doi:DOI 10.1006/abbi.1994.1267 (1994).

17 Calvert, L. & Thresh, J. M. The viruses and virus diseases of cassava. *Cassava: biology, production and utilization*, 237-260 (2002).

18 Legg, J. P. *et al.* Cassava Virus Diseases: Biology, Epidemiology, and Management. *Adv Virus Res* **91**, 85-142, doi:10.1016/bs.aivir.2014.10.001 (2015).

19 Boykin, L. M. Bemisia tabaci nomenclature: lessons learned. *Pest Manag Sci* **70**, 1454-1459 (2014).

20 Oliveira, M. R. V., Henneberry, T. J. & Anderson, P. History, current status, and collaborative research projects for *Bemisia tabaci*. *Crop Prot* **20**, 709-723, doi:Doi 10.1016/S0261-2194(01)00108-9 (2001).

21 Chen, W. *et al.* The draft genome of whitefly Bemisia tabaci MEAM1, a global crop pest, provides novel insights into virus transmission, host adaptation, and insecticide resistance. *BMC biology* **14**, 1-15 (2016).

22 Mugerwa, H. *et al.* African ancestry of New World, Bemisia tabaci-whitefly species. *Scientific reports* **8**, 1-11 (2018).

23 Perez-Fons, L. *et al.* A metabolomics characterisation of natural variation in the resistance of cassava to whitefly. *BMC plant biology* **19**, 1-14 (2019).

24 Antony, B. *et al.* Bemisia tabaci (Homoptera: Aleyrodidae) and Indian cassava mosaic virus transmission. *International Journal of Tropical Insect Science* **26**, 176-182 (2006).

25 Danner, H., Desurmont, G. A., Cristescu, S.

M. & van Dam, N. M. Herbivore-induced plant volatiles accurately predict history of coexistence, diet breadth, and feeding mode of herbivores. *New Phytol* **220**, 726-738, doi:10.1111/nph.14428 (2018).

26 Kempema, L. A., Cui, X. P., Holzer, F. M. & Walling, L. L. Arabidopsis transcriptome changes in response to phloem-feeding silverleaf whitefly nymphs. Similarities and distinctions in responses to aphids. *Plant Physiology* **143**, 849-865 (2007).

27 Kim, J. H. & Jander, G. *Myzus persicae* (green peach aphid) feeding on Arabidopsis induces the formation of a deterrent indole glucosinolate. *Plant Journal* **49**, 1008-1019, doi:10.1111/j.1365-313X.2006.03019.x (2007).

28 Kim, J. H., Lee, B. W., Schroeder, F. C. & Jander, G. Identification of indole glucosinolate breakdown products with antifeedant effects on *Myzus persicae* (green peach aphid). *Plant Journal* **54**, 1015-1026, doi:10.1111/j.1365-313X.2008.03476.x (2008).

29 Markovich, O. *et al. Arabidopsis thaliana* plants with different levels of aliphatic and indolyl-glucosinolates affect host selection and performance of *Bemisia tabaci*. *J Chem Ecol* **39**, 1361-1372, doi:10.1007/s10886-013-0358-0 (2013).

30 Malka, O. *et al.* Glucosylation prevents plant defense activation in phloem-feeding insects. *Nat Chem Biol*, doi:10.1038/s41589-020-00658-6 (2020).

31 Kim, J. H., Lee, B. W., Schroeder, F. C. & Jander, G. Identification of indole glucosinolate breakdown products with antifeedant effects on Myzus persicae (green peach aphid). *The Plant Journal* **54**, 1015-1026 (2008).

32 Markovich, O. *et al.* Arabidopsis thaliana plants with different levels of aliphatic-and indolyl-glucosinolates affect host selection and performance of Bemisia tabaci. *J Chem Ecol* **39**, 1361-1372 (2013).

33 Malka, O. *et al.* Glucosinolate desulfation by the phloem-feeding insect Bemisia tabaci. *J Chem Ecol* **42**, 230-235 (2016).

34 Machingura, M., Salomon, E., Jez, J. M. & Ebbs, S. D. The beta-cyanoalanine synthase pathway: beyond cyanide detoxification. *Plant Cell Environ* **39**, 2329-2341, doi:10.1111/pce.12755 (2016).

35 Hatzfeld, Y. *et al.* beta-cyanoalanine synthase is a mitochondrial cysteine synthase-like protein in spinach and Arabidopsis. *Plant Physiology* **123**, 1163-1171, doi:DOI 10.1104/pp.123.3.1163 (2000).

36 Svensson, B. *et al.* Fascinating facets of function and structure of amylolytic enzymes of glycoside hydrolase family 13. *BIOLOGIA-BRATISLAVA-* **57**, 5-20 (2002).

37 Selmar, D. *et al.* Occurrence of the cyanogen linustatin in *Hevea brasiliensis*. *Phytochemistry* **26**, 2400-2401, doi:Doi 10.1016/S0031-9422(00)84729-9 (1987).

38 van Ohlen, M., Herfurth, A. M., Kerbstadt, H. & Wittstock, U. Cyanide detoxification in an insect herbivore: Molecular identification of beta-cyanoalanine synthases from Pieris rapae. *Insect Biochem Molec* **70**, 99-110 (2016).

39 Mayer, R. M., Matthews, M. M., Futerman, C. L., Parnaik, V. K. & Jung, S. M. Dextransucrase: acceptor substrate reactions. *Arch Biochem Biophys* **208**, 278-287 (1981).

40 Rose, I. A. & Warms, J. V. B. Mitochondrial Hexokinase - Release Rebinding and Location. *J Biol Chem* **242**, 1635-& (1967).

41 Thimm, J. & Thiem, J. Enzymes of the Carbohydrate Metabolism and Catabolism for Chemoenzymatic Syntheses of Complex Oligosaccharides. *Carbohydrate-*

Modifying Biocatalysts, 507-534 (2012).

42 Blennow, A., Engelsen, S. B., Nielsen, T. H., Baunsgaard, L. & Mikkelsen, R. Starch phosphorylation: a new front line in starch research. *Trends Plant Sci* **7**, 445-450, doi:Pii S1360-1385(02)02332-4 Doi 10.1016/S1360-1385(02)02332-4 (2002).

43 Slabnik, E. & Frydman, R. B. A Phosphorylase Involved in Starch Biosynthesis. *Biochem Bioph Res Co* **38**, 709-&, doi:Doi 10.1016/0006-291x(70)90639-X (1970).

44 Szwergold, B. S., Kappler, F. & Brown, T. R. Identification of Fructose 3-Phosphate in the Lens of Diabetic Rats. *Science* **247**, 451-454 (1990).

45 Szwergold, B. S., Kappler, F., Brown, T. R., Pfeffer, P. & Osman, S. F. Identification of D-Sorbitol 3-Phosphate in the Normal and Diabetic Mammalian Lens. *J Biol Chem* **264**, 9278-9282 (1989).

46 Smith, B. A. *et al.* Potential roles of inorganic phosphate on the progression of initially bound glucopyranose toward the nonenzymatic glycation of human hemoglobin: Mechanistic diversity and impacts on site selectivity. *Cogent Biol* **4**, doi:Artn 1425196 10.1080/23312025.2018.1425196 (2018).

47 Boeckler, G. A., Paetz, C., Feibicke, P., Gershenzon, J. & Unsicker, S. B. Metabolism of poplar salicinoids by the generalist herbivore Lymantria dispar (Lepidoptera). *Insect Biochem Molec* **78**, 39-49 (2016).

48 Ngah, W. & Smith, J. Acidic conjugate of phenols in insects: glucoside phosphate and glucoside sulphate derivatives. *Xenobiotica* **13**, 383-389 (1983).

49 Olsen, L. R. *et al.* Characterization of midazolam metabolism in locusts: the role of a CYP3A4-like enzyme in the formation of 1′-OH and 4-OH midazolam. *Xenobiotica* **46**, 99-107 (2016).

50 Scanlan, J. L., Gledhill-Smith, R. S., Battlay, P. & Robin, C. Genomic and transcriptomic analyses in Drosophila suggest that the ecdysteroid kinase-like (EcKL) gene family encodes the ‘detoxification-by-phosphorylation’enzymes of insects. *Insect Biochem Molec*, 103429 (2020).

51 Elbaz, M. *et al.* Asymmetric adaptation to indolic and aliphatic glucosinolates in the B and Q sibling species of *Bemisia tabaci* (Hemiptera: Aleyrodidae). *Mol Ecol* **21**, 4533-4546, doi:10.1111/j.1365-294X.2012.05713.x (2012).

Supplementary Information

Supplemental Figures

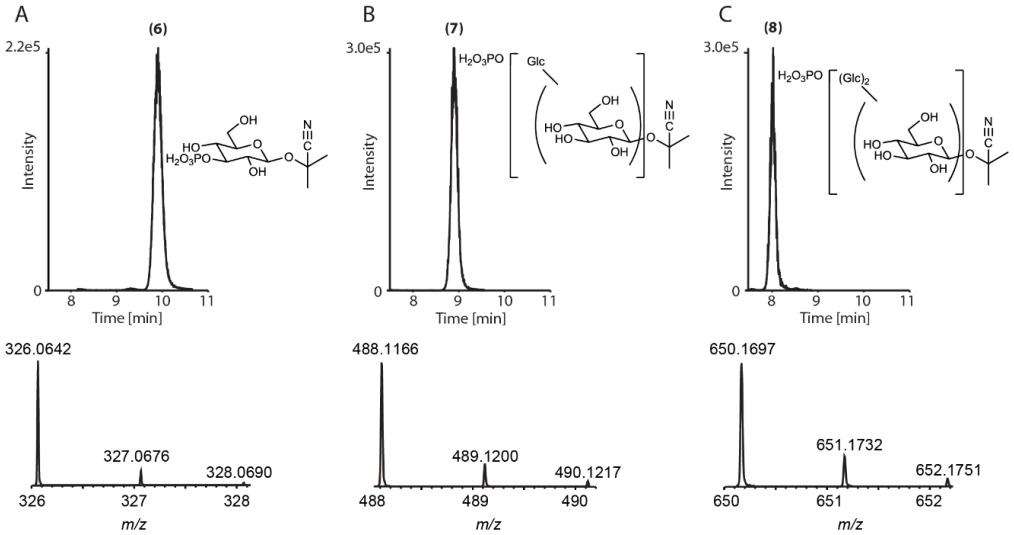


Figure S1: HRMS supports the addition of a phophate. HRMS data of a honeydew sample run on a Bruker-TOF showing mass spectra and retention time for compounds **(6, 7 and 8)** (panels A, B and C respectively).

1

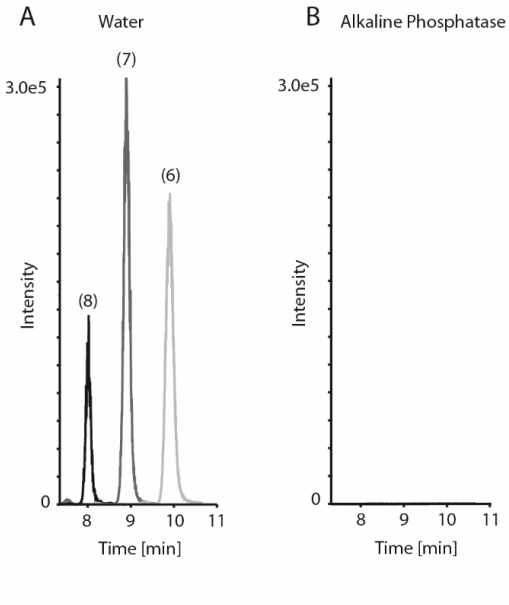


Figure S2: Alkalkine phosphatase incubation results in the disappearance of mass signals for phopshorylated compounds. Phosphoryalated compounds **(6, 7 and 8)** in the honeydew of the whitefly feedign on cassava were incubated with water (A) and alkaline phosphatase (B). The disappearance of the mass signals supports the compounds being phosphorylated products.

2

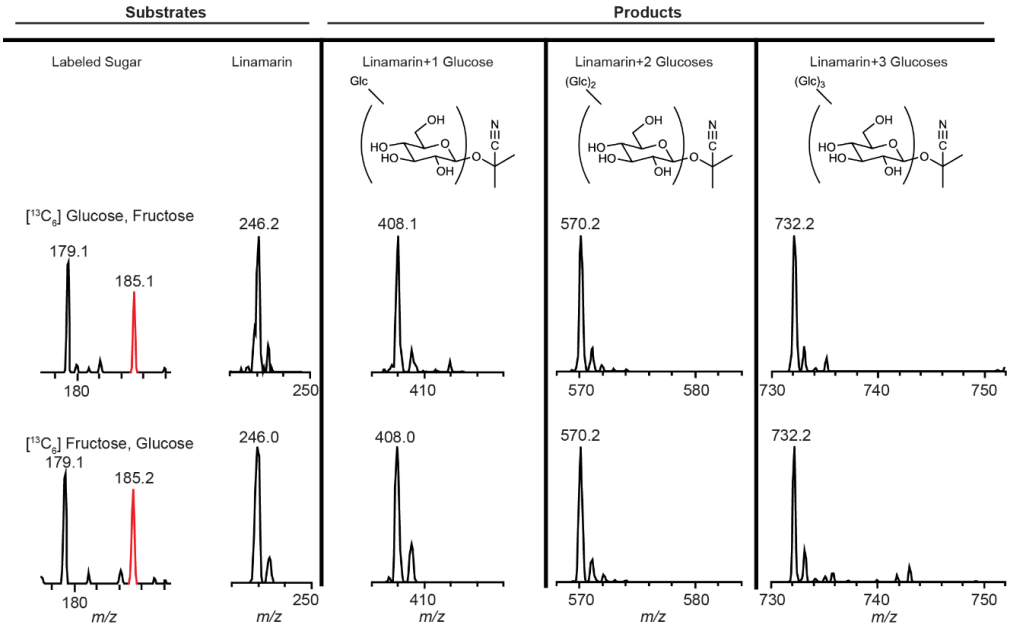


Figure S3: CN-Glc glucosylation in the whitefly *B. tabaci* is catalyzed by a transglucosidase activity. Depiction of the results from two of the five diets not shown in Figure 4, those diets with the ^{13}C-labeled monosaccharides glucose and fructose. None gave labeled glycosylated linamarin products, unlike feeding with sucrose labeled in the glucose portion. The results are consistent with a transglucosidase activity that initially hydrolyzes sucrose and links the resulting glucose moiety to linamarin.

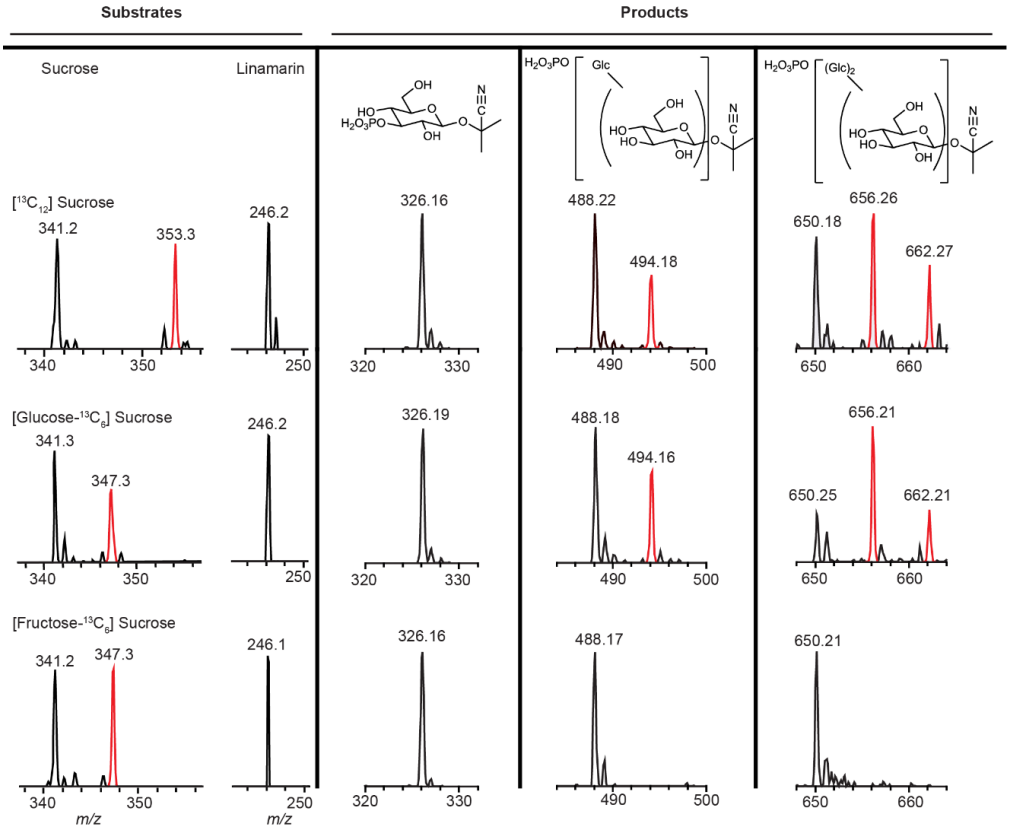


Figure S4: Glucosylation of phosphorylated linamarin derivatives is also catalyzed by a transglcusoidase. Shown are regions of mass spectra from LC-MS analyses of substrates and products of feeding experiments on three different diets. Feeding fully ^{13}C-labeled sucrose and sucrose ^{13}C-labeled in the glucose moiety gave labeling in the newly added glucose(s) of linmarin phosphate derivatives (**7 and 8**).

4

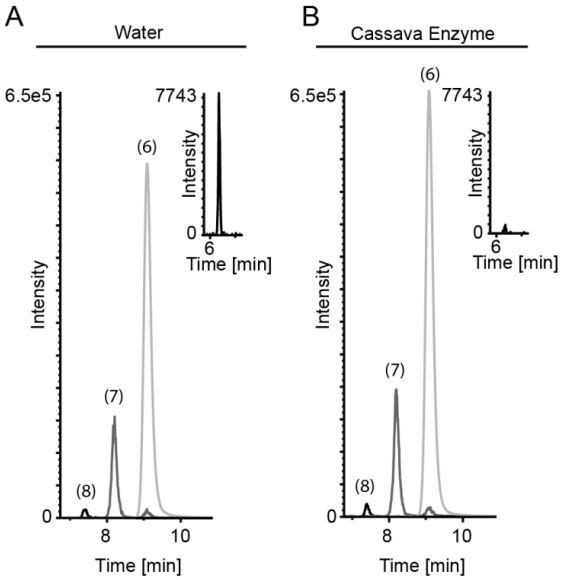
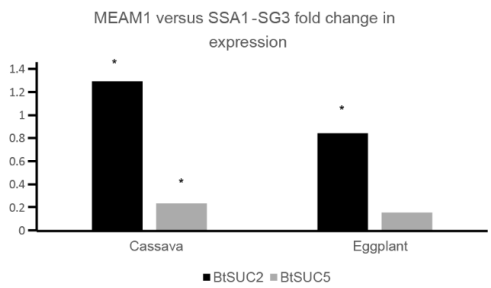


Figure S5: Phosphorylated linamarin metabolites in *B. tabaci* whitefly honeydew are resitant to linamarase activation as well. (**A**) Phosphorylated metabolites **(6, 7 and 8)** in honeydew were incubated with water (A) and crude cassava leaf enzyme containing linamarase activity (B). Compounds (**6**), (**7**) and (**8**) were resistant to activation by cassava linamarase, while the parent linamarin is readily hydrolyzed. Samples were incubated for 1 h with cassava crude enzyme extract or water in phosphate buffer at pH 7.0.

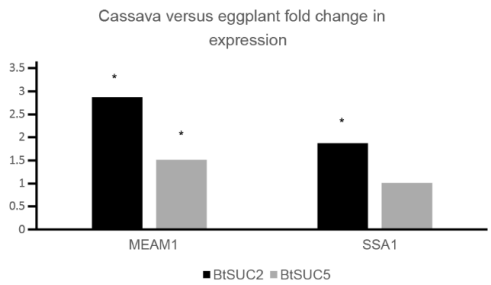
A

Protein name	% identity at protein level	% identity at nucleotide level
BtSUC2	95.69	94.70
BtSUC5	94.77	95.21

B



C



Supplemental Figure S6: BtSUC2 and 5 in SSA1-SG3 whiteflies and relative expression levels. (A) The percent identity of SSA1-SG3 homologues of BtSUC2 and 5 was determined at the DNA and protein level. Fold changes were calculated for BtSUC2 and 5 based on differences between species feeding on different plants (B) and within a species of whitefly feeding on different plants (C). Asterisks denote statistically significant differences (p value < 0.05) between sample ratios (N=3).

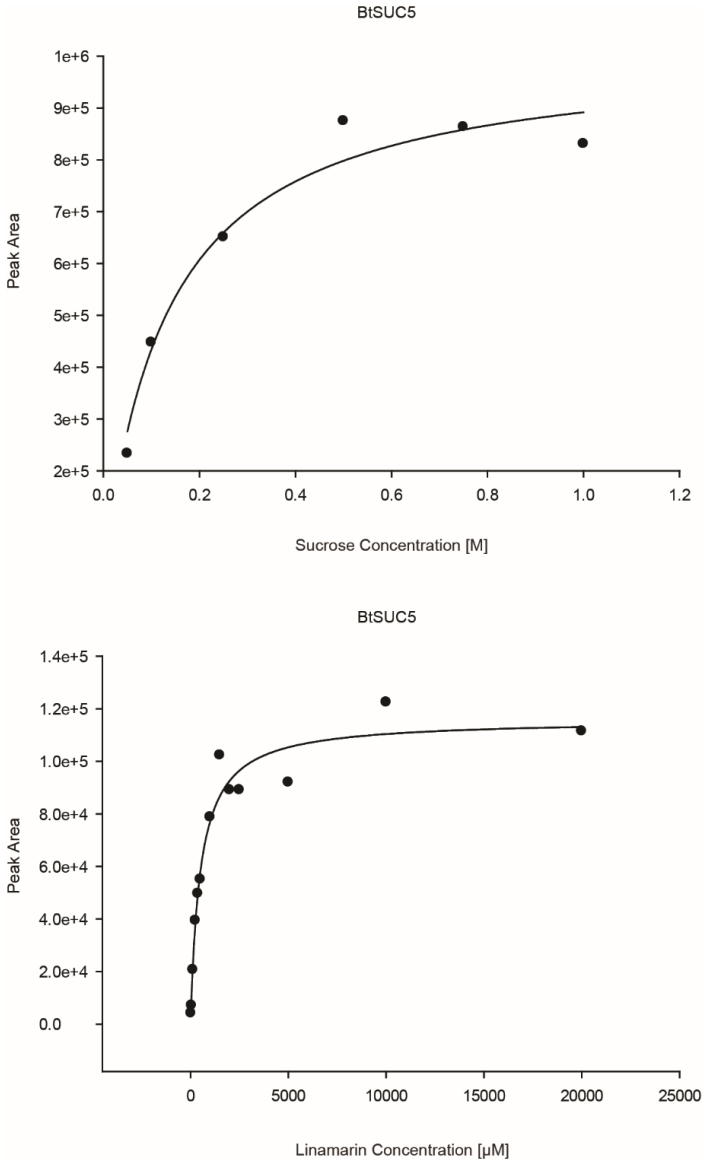


Figure S7: Kinetic characterization of BtSUC5 with linamarin and sucrose. Depicted are the relations between substrate concentration and reaction rate (product peak area) for both sucrose and linamarin.GSL. Assays were carried out as described in the methods section. When the sucrose concentration was varied from 0.05 to 1.0 M, linamarin concentration was held at 2.5 mM. When the linamarin concentration was varied from 25 µM to 20 mM, sucrose concentration was held at 1 M.

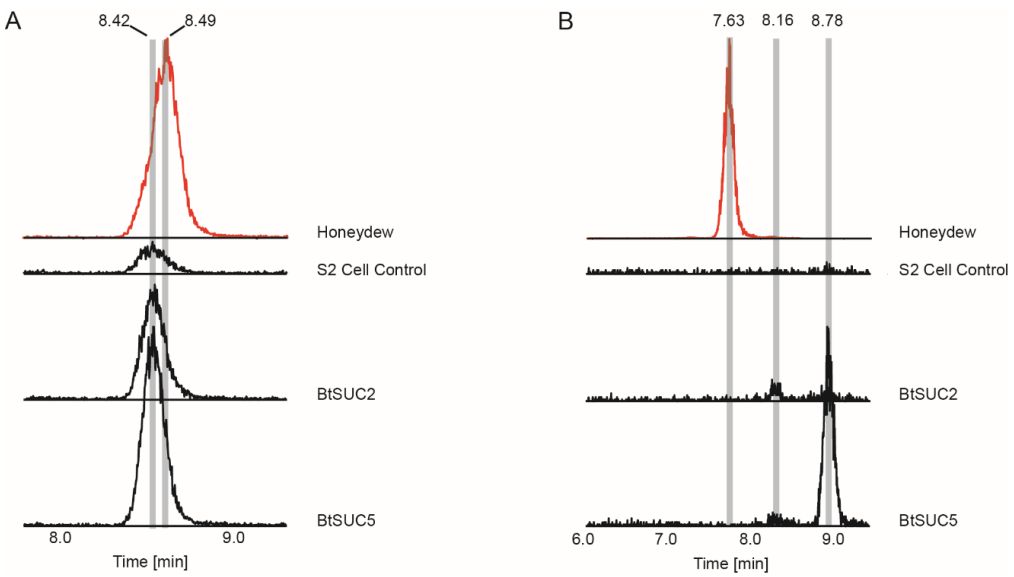


Figure S8: Chromatographic analyses of products from BtMEAM1 BtSUC2 and 5 enzymes heterologously produced in *D. melanogaster* S2 cells. Cell medium actviity was assayed with phoshorylated linamarin derivatives and sucrose. (**A**) Depicted are extracted multiple reaction monitoring (MRM) LC-MS chromatograms for **(7)** upon incubating **(6)** with sucrose and enzymes BtSUC2 and BtSUC5. The enzymes BtSUC2 and 5 showed transglucosidation activity, producing a glycosylated derivative of **(6)** however the retention differed from that of the **(7)** found within the honeydew. (**B**) Repeating the assay with purified **(7)** as a starting substrate, enzyme BtSUC2 produced low levels of a glycoside of **(7)** but again differing from honeydew peaks and BtSUC5 produced a glycoside of even greater retention difference. S2 cell control assay was of cell medium extracts of untransformed cells.

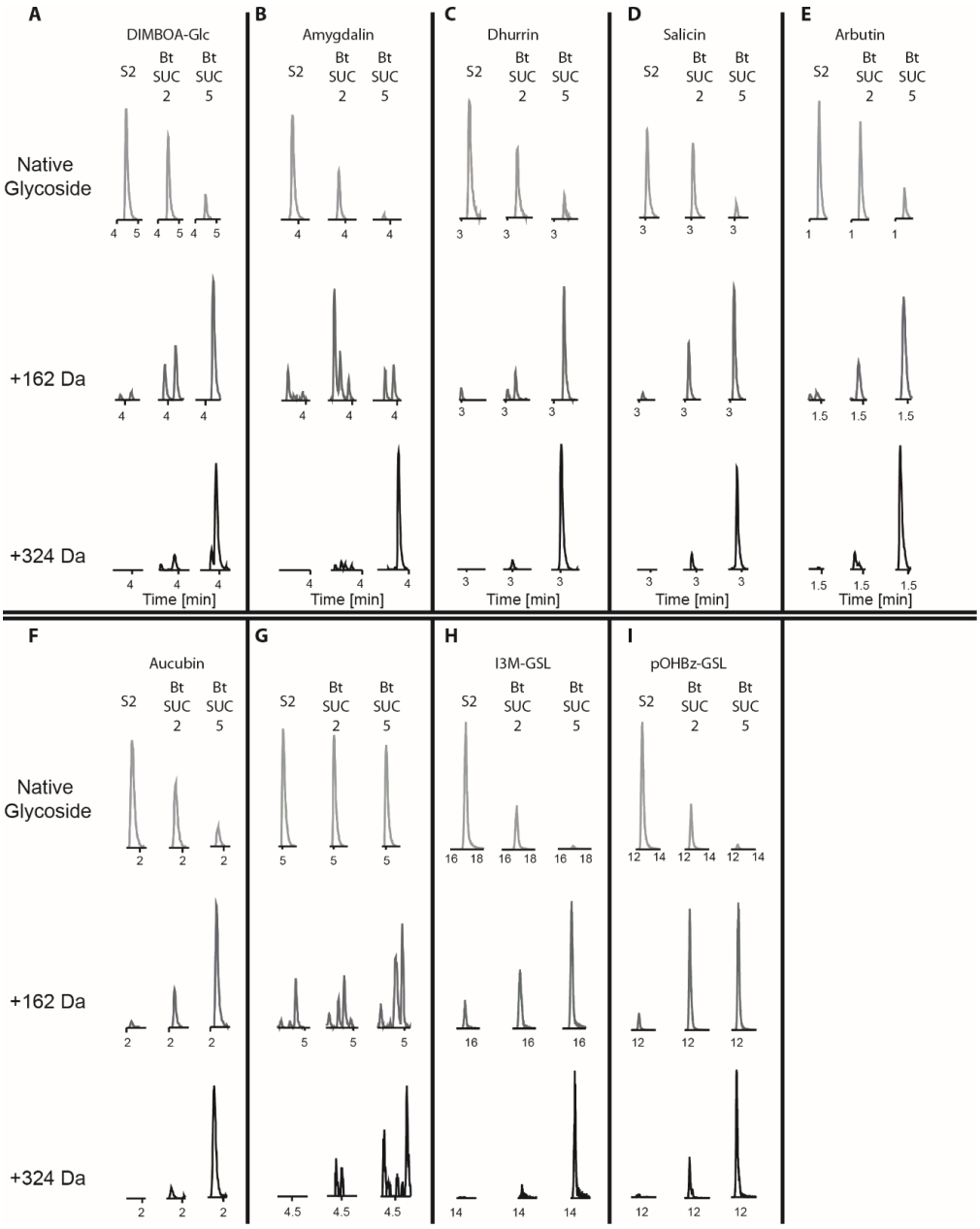


Figure S9: Substrate tests for BtMEAM1 BtSUC2 and 5 enzymes heterologously produced in *D. melanogaster* S2 cells. Cell medium actviity was assayed with various secondary

9

metabolites and sucrose. Depicted are masses for the native glycoside and the addition of one glucose (+162 Da) and 2 glucose units (+324 Da). BtSUC2 and BtSUC5 showed transglucosdiation activity with all substrates above control levels, with rutin poorest glucose acceptor based on native glycoside peak depletion. Compounds used were (**A**) DIMBOA-Glucose (benzoxazinoid), (**B**) amygdalin (cyanogenic diglycoside), (**C**) dhurrin (cyanogenic mono-glycoside), (**D**) salicin (phenolic glycoside), (**E**) arbutin (phenolic glycoside), (**F**) aucubin (irridoid glycoside), (**G**) rutin (flavanoid diglycoside), (**H**) indoyl-3-methyl glucosinolate (glucosinolate), and (**I**) para-hydroxybenzyl glucosinolate (glucosinolate). S2 cell control assay was of cell medium extracts of untransformed cells. All peaks are normalized within each set of assays to the largest peak of the respective metabolite.

Supplementary Tables

Table S1: HRMS data for phosphorylated linamarin derivatives showing measured and expected masses of the metabolites.

Measured *m/z*	Ion Formula	*m/z*	err[ppm]	mSigma	rdb	e^- confidence	N-rule
326.0642	$C_{10}H_{17}NO_9P$	326.0646	1.2	0.8	4	even	ok
488.1166	$C_{16}H_{27}NO_{14}P$	488.1175	1.7	2.7	5	even	ok
650.1697	$C_{22}H_{37}NO_{19}P$	650.1703	1.0	3.5	6	even	ok

Table S2: Constituents of artificial diets fed to *Bemisia tabaci* MEAM1.

Diet number	Isotopically-labeled compound (concentration)	Unlabeled constituents (concentration)
1	[$^{13}C_{12}$] sucrose (0.145 M)	Sucrose (0.145 M) Linamarin (5 mM)
2	[glucose-$^{13}C_6$] sucrose (0.145 M)	Sucrose (0.145 M) Linamarin (5 mM)
3	[fructose-$^{13}C_6$] sucrose (0.145 M)	Sucrose (0.145 M) Linamarin (5 mM)
4	[$^{13}C_6$] glucose (0.0725 M)	Sucrose (0.145 M) Linamarin (5 mM) Fructose (0.0725 M)
5	[$^{13}C_6$] fructose (0.0725 M)	Sucrose (0.145 M) Linamarin (5 mM) Glucose (0.0725 M)

Table S3: List of multiple reaction monitoring (MRM) parameters for individual compounds analyzed by LC-MS.

Compound	Quadrupole 1 (precursor mass Da)	Quadrupole 2 (product mass Da)	Declustering potential (V)	Entrance potential (V)	Collision energy (V)	Collision cell exit potential (V)
Linamarin-59	246	59	-50	-10	-46	-2
Linamarin + Glucose	408.129	178.9	-50	-3.5	-20	-4
Linamarin + 2 Glucoses	570	179	-50	-7.75	-30	-4
Linamarin Phosphate	326	79	-50	-7	-48	-2
Linamarin Glucose Phosphate	488	79	-50	-7	-48	-2
Linamarin + 2 Glucose Phosphate	650	79	-50	-7	-48	-2
Betacyanoalanine	113	96	-50	-10	-13	-5
Monosaccharide	178.8	89	-50	-9.5	-10	0
Disaccharide	340.9	59	-65	-10	-46	0
Trisaccharide	503.1	179	-95	-10	-28	-4
Tetrasaccharide	665.2	179	-100	-10	-48	-4

Table S4: List of primers used for cloning of GH13 genes from *B. tabaci*

Primer name	Sequence
SUC2 pACV5-HIS Full KpnI Forward S2	GCACGTGGTACCATGAGTCGGAATTTGACAATACTGC
SUC2 pACV5-HIS Full XbaI Reverse S2	CACCGGCTCTAGAATTGGATTTACCGGTTAATCTACTGC
SUC5 pACV5-HIS Full EcoRI Forward S2	GCACGTGAATTCACCATGGAGAGACTACTCTACTTTGTTGTG
SUC5 pACV5-HIS Full XbaI Reverse S2	CACCGGCTCTAGAGCTGATGAAGATTTTTAAATAAATTGATAGTAGAAG

Supplementary Note 1

Supporting information NMR

NMR spectra were recorded on a Bruker Avance III HD 700 MHz spectrometer, equipped with a cryoplatform and a 1.7 mm cryoprobe (Bruker Biospin GmbH, Rheinstetten, Germany). Spectrometer control and data processing was accomplished using Bruker TopSpin ver. 3.2. ^{13}C chemical shifts were determined indirectly by means of ^{1}H-^{13}C HSQC (heteronuclear single quantum coherence) and ^{1}H-^{13}C HMBC (heteronuclear multiple bond correlation) experiments. The glucosidic substitutions were elucidated with the help of 1D ^{1}H-^{1}H SELTOCSY (selective total correlation spectroscopy) experiments. The resulting spectra served as projection spectra for the 2D homo- and heteronuclear experiments. The positions of the chemical shifts in the molecules were furthermore determined based on $^{3}J_{HH}$ coupling constants and ^{13}C chemical shifts. Important spectral details have been assembled for deeper understanding of the structure elucidation as follows.

Figure SN1-1: Structure of **2** with chemical shifts in D_2O

Table SN1-1: Table of chemical shifts of **1, 2** and **3** in D_2O

Figure SN1-2: Structure of **2** with chemical shifts in D_2O

Figure SN1-3: Structure of **1** with chemical shifts in D_2O

Figure SN1-4: Structure of **3** with chemical shifts in D_2O

Figure SN1-5: Table of chemical shifts of **6** in D_2O with numbered structure

Figure SN1-6: Structure of **6** with chemical shifts in D_2O

Figure SN1-7: 1dNOESY (black trace) and a SELTOCSY spectrum of **6**

Figure SN1-8: DQFCOSY of **6**

Figure SN1-9: ^{1}H-^{13}C HSQC spectrum of **6**, glucosidic range

Figure SN1-10: ^{1}H-^{13}C HSQC spectrum of **6**, CH_3 range

Figure SN1-11: ^{1}H-^{13}C HSQC spectrum of linamarin in MeOH-d_3

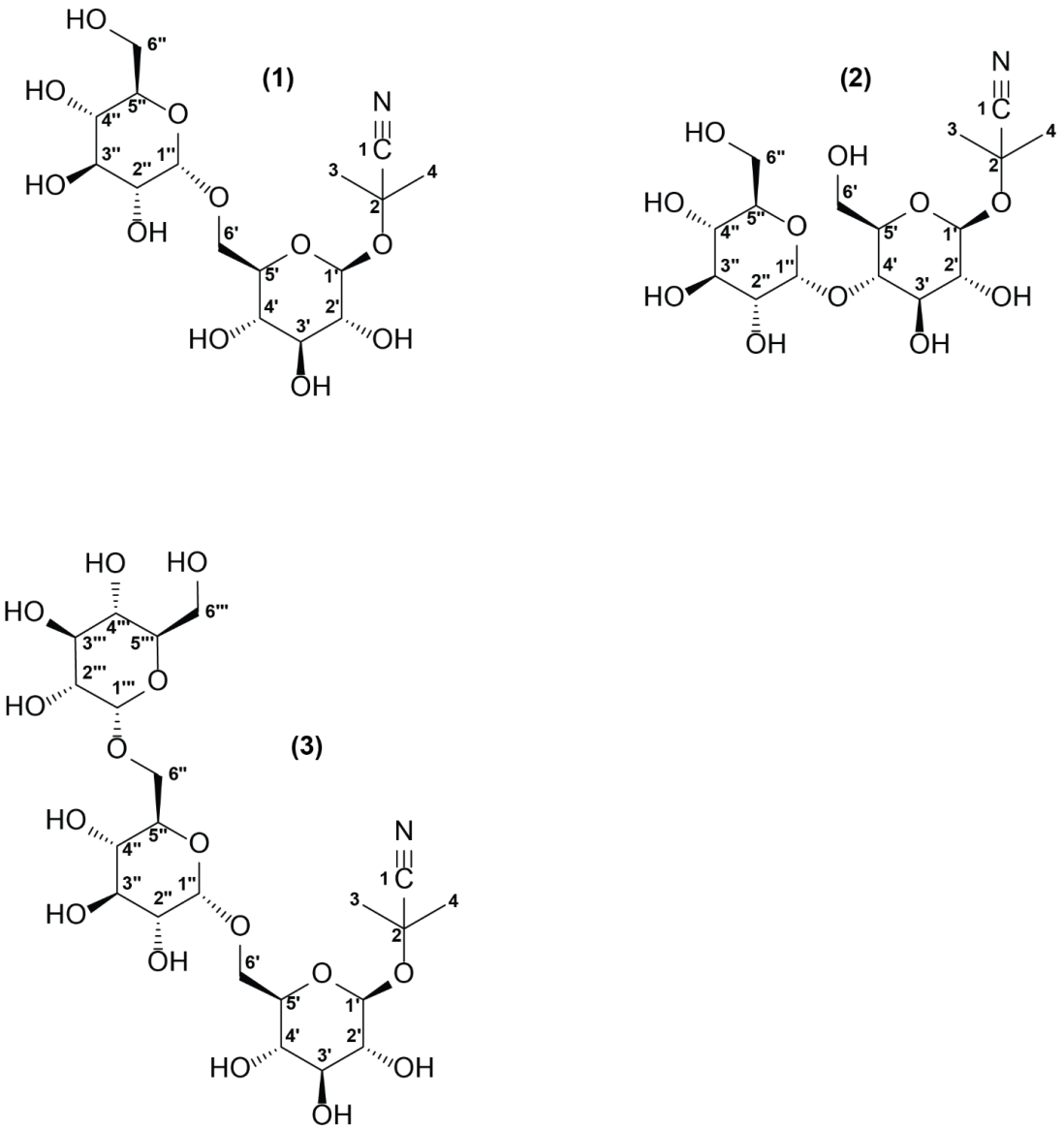


Figure SN1-1: Structure and of glucosylated linamarin derivatives (compounds **1**, **2** and **3**) with numbering.

Table SN1-1: ^{1}H and ^{13}C NMR shifts and coupling constants for glucosylated linamarin derivatives.

	Monoglucosylated α-(1→6)-linamarin **(1)**			Monoglucosylated α-(1→4)-linamarin **(2)**			Diglucosylated α-(1→6), α-(1→6)-linamarin **(3)**		
	δH	mult., *J* [Hz]	δC	δH	mult., *J* [Hz]	δC	δH	mult., *J* [Hz]	δC
Aglycone									
1	-	-	121.7	-	-	121.7	-	-	121.9
2	-	-	72.6	-	-	72.6	-	-	72.9
3	1.69	s	28.5	1.67	s	28.3	1.69	s	28.6
4	1.67	s	27.5	1.66	s	27.6	1.67	s	27.5
Original glucose of linamarin									
1'	4.65	d, 7.8	101.0	4.64	d, 7.9	101.0	4.66	d, 7.6	101.2
2'	3.21	dd, 9.5/7.8	74.6	3.24	dd, 9.5/7.9	74.3	3.22	dd, 9.5/7.6	74.6
3'	3.41	dd, 9.5/9.5	77.8	3.67	dd, 9.5/9.5	77.6	3.42	dd, 9.5/9.5	77.8
4'	3.38	dd, 9.5/9.5	71.5	3.57	dd, 9.5/9.5	80.6	3.40	dd, 9.5/9.5	71.4
5'	3.56	ddd, 9.5/5.6/2.1	76.3	3.44	m	76.6	3.58	m	76.3
6a'	3.93	dd, 11.0/5.6	67.3	3.86	m	61.9	3.96	dd, 10.7/5.6	67.4
6b'	3.70	dd, 11.0/2.1	67.3	3.82	m	61.9	3.71	dd, 10.7/1.6	67.4
First added glucose moiety									
1''	4.84	d, 3.6	99.7	5.17	d, 3.6	102.6	4.85	d, 3.6	99.8
2''	3.38	dd, 9.6/3.6	73.5	3.44	dd, 9.5/3.6	74.0	3.39	dd, 9.6/3.6	73.5
3''	3.65	dd, 9.5/9.5	75.0	3.61	dd, 9.5/9.5	74.9	3.65	m	75.3
4''	3.31	dd, 9.5/9.5	71.5	3.26	dd, 9.5/9.5	71.4	3.36	dd, 9.5/9.5	71.5
5''	3.65	m	75.0	3.68	m	74.6	3.65	m	75.3
6a''	3.79	bd, 11.0	62.4	3.83	bd, 10.9	62.4	3.90	dd, 10.7/5.2	67.4
6b''	3.68	dd, 11.0/5.2	62.4	3.65	m	62.4	3.73	bd, 10.7	67.4
Second added glucose moiety									
1'''	-	-	-	-	-	-	4.87	d, 3.6	97.6
2'''	-	-	-	-	-	-	3.41	dd, 9.5/3.6	71,4
3'''	-	-	-	-	-	-	3.65	m	73.2
4'''	-	-	-	-	-	-	3.32	dd, 9.5/9.5	69.4
5'''	-	-	-	-	-	-	3.87	m	71.8
6a'''	-	-	-	-	-	-	3.79	bd, 11.6/5.0	60.4
6b'''	-	-	-	-	-	-	3.68	bd, 11.6	60.4

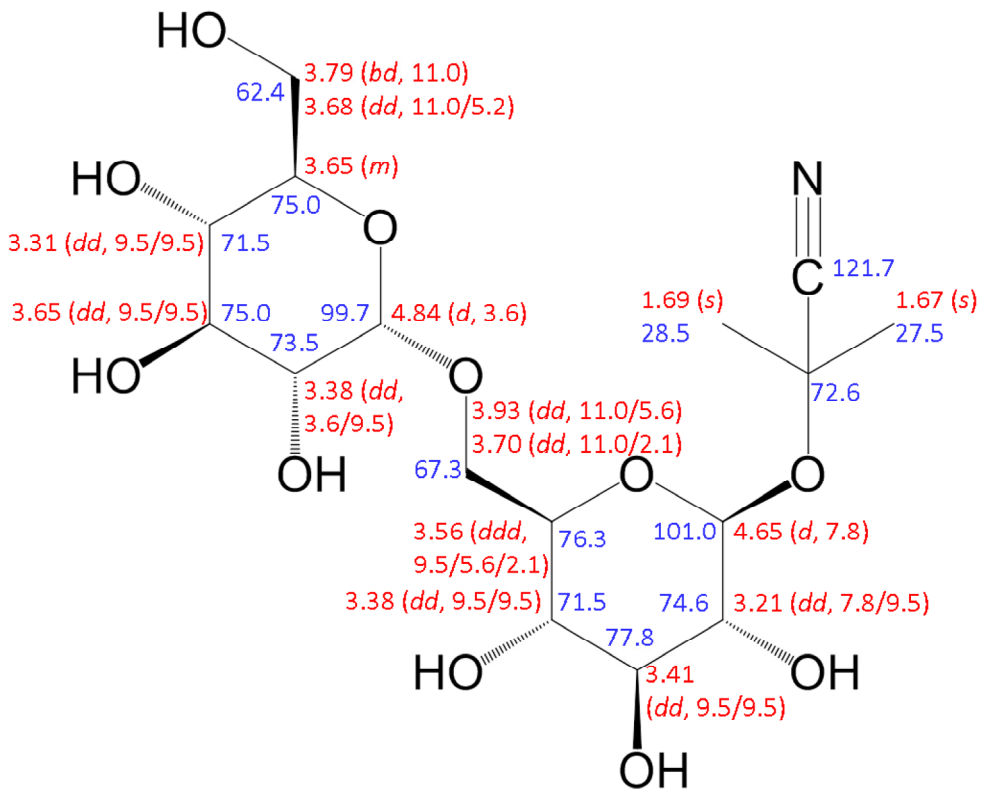


Figure SN1-2: Structure and chemical shifts of monoglucosylated α-(1→6)-linamarin (**1**). ^{1}H chemical shifts in red, ^{13}C chemical shift in blue.

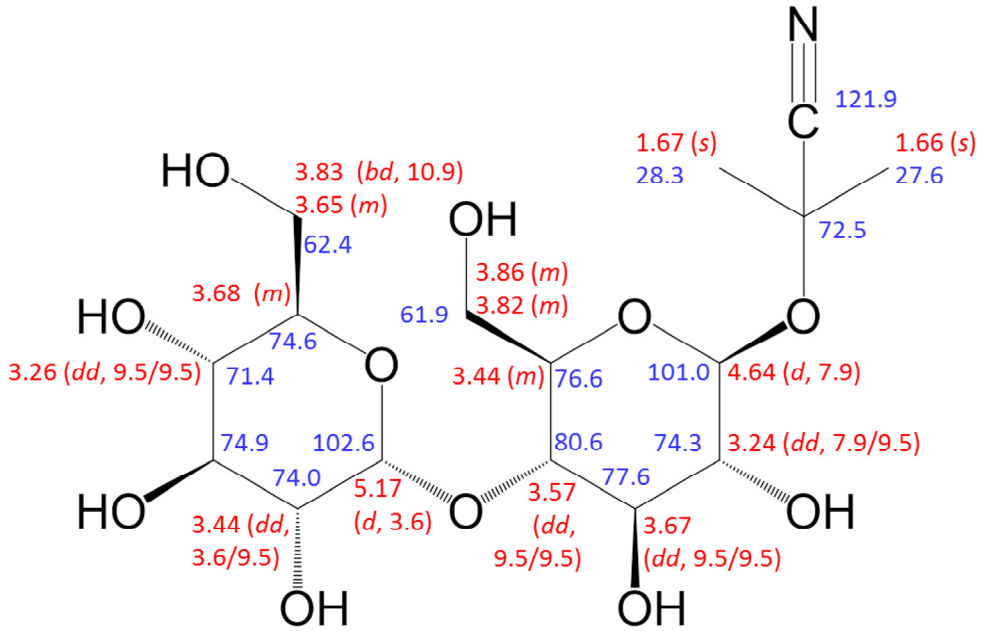


Figure SN1-3: Structure and chemical shifts of monoglucosylated α-(1→4)-linamarin (**2**). ^{1}H chemical shifts in red, ^{13}C chemical shift in blue.

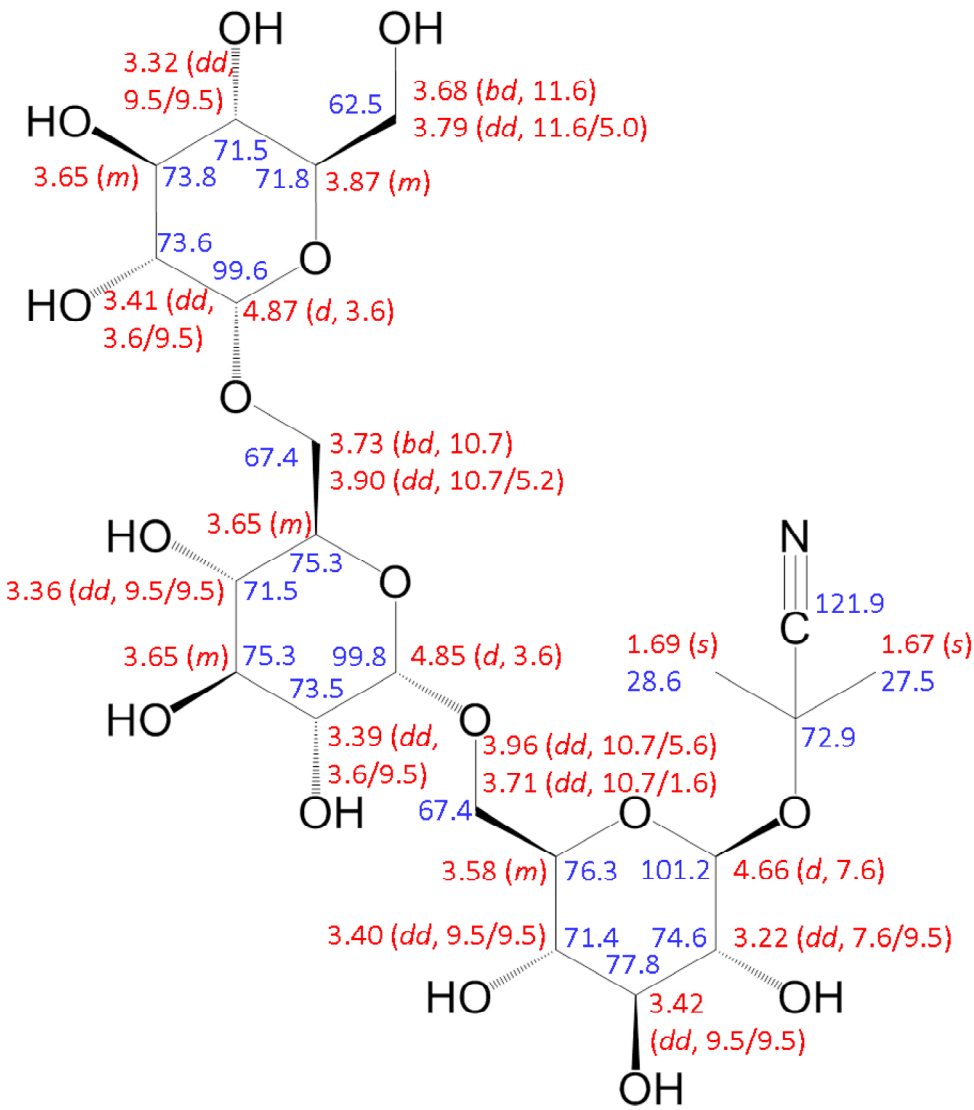


Figure SN1-4: Structure and chemical shifts of diglucosylated α-(1→6), α-(1→6)-linamarin (**3**). ^{1}H chemical shifts in red, ^{13}C chemical shift in blue.

A

Linamarin Phosphate **(6)**			
	δH	mult., *J* [Hz]	δC
Aglycone			
1	-	-	121.7
2	-	-	72.6
3	1.61	s	26.9
4	1.60	s	25.8
Original glucose of linamarin			
1'	4.82	d, 7.8	98.5
2'	3.30	dd, 9.5/7.8	72.7
3'	3.94	dd, 9.5/9.5	78.9
4'	3.43	dd, 9.5/9.5	69.5
5'	3.49	ddd, 9.5/5.6/2.1	75.7
6a'	3.80	dd, 11.0/5.6	60.6
6b'	3.63	dd, 11.0/2.1	60.6
First added glucose moiety			
1''	-	-	-
2''	-	-	-
3''	-	-	-
4''	-	-	-
5''	-	-	-
6a''	-	-	-
6b''	-	-	-

B

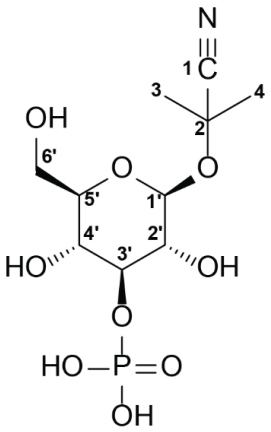


Figure SN1-5: (**A**) ^{1}H and ^{13}C NMR shifts and coupling constants for phopshorylated linamarin (**6**) and its structure (**B**) with numbering.

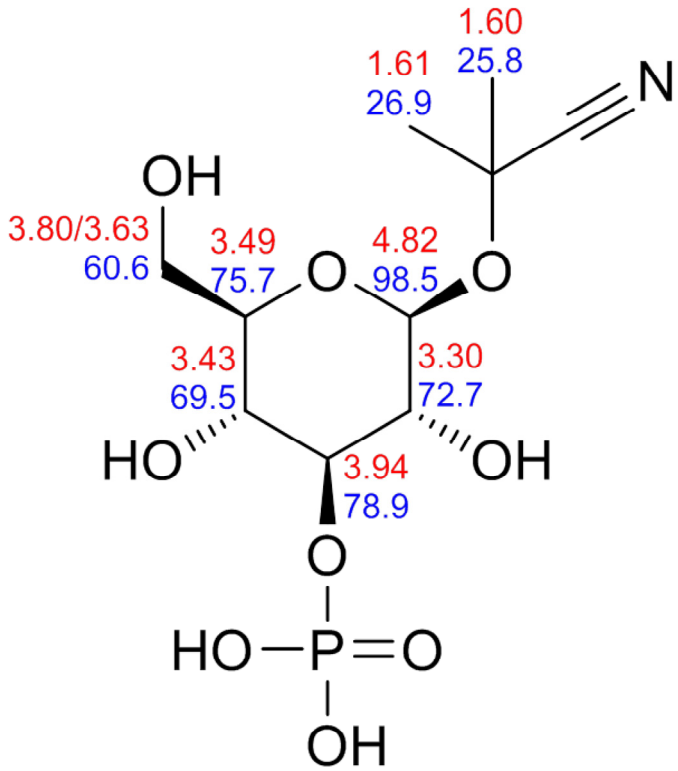


Figure SN1-6: Structure and chemical shifts of linamarin-3'-*O*-phosphate (**6**). ^{1}H chemical shifts in red, ^{13}C chemical shift in blue.

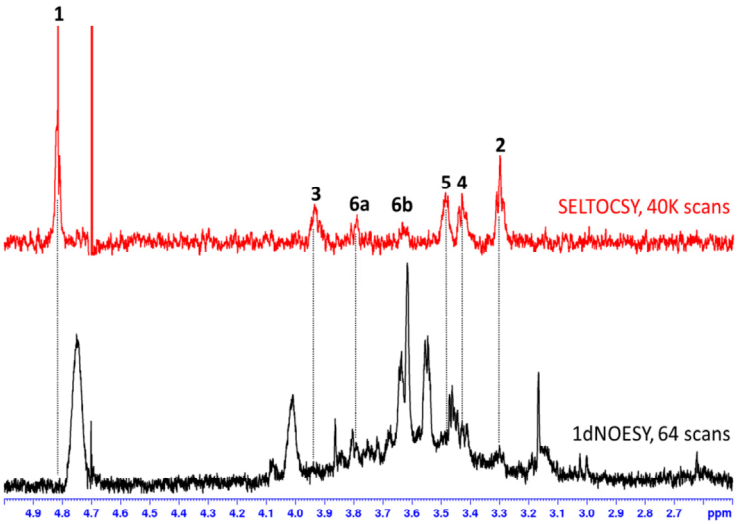


Figure SN1-7: Details of the 1dNOESY (black trace) and a SELTOCSY spectrum of linamarin-3'-*O*-phosphate (700 MHz, in D_2O). Numbers indicate the position in the glucose part of the molecule. The transmitter frequency for the SELTOCSY experiments (position one of the glucose part) was extracted from an HSQC experiment (figure SN1-9).

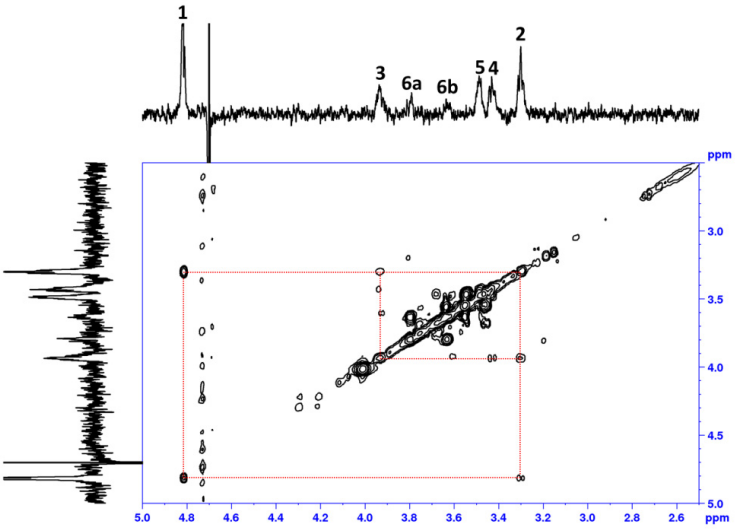


Figure SN1-6: Detail of the DQFCOSY of linamarin-3'-*O*-phosphate (700 MHz, magnitude mode, in D_2O). The SELTOCSY from Fig. SN1-7 served as F1- and F2-projection

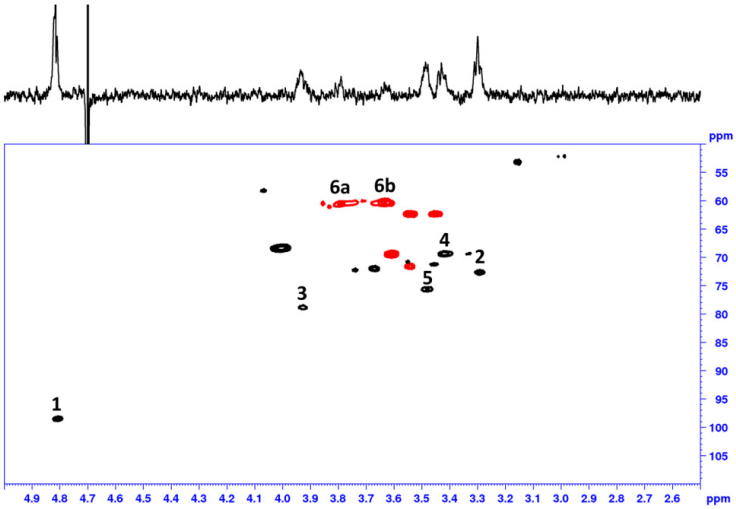


Figure SN1-9: Detail of the ^{1}H-^{13}C HSQC spectrum of linamarin-3'-*O*-phosphate. The SELTOCSY from Fig. SN1-7 served as F2-projection.

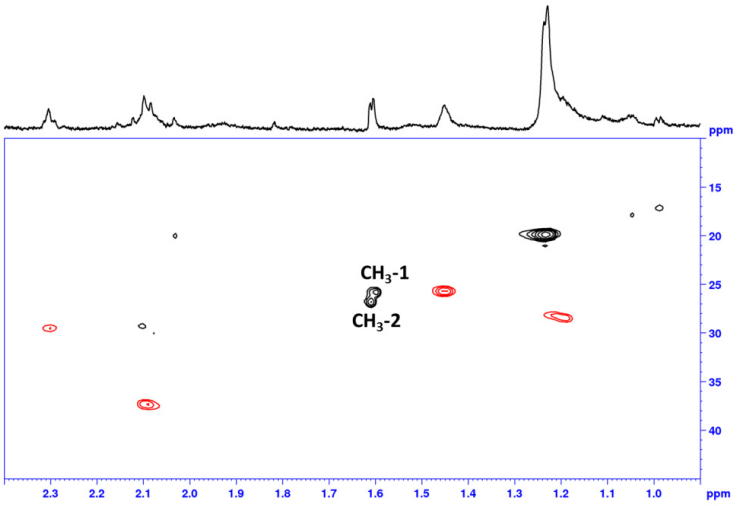


Figure SN1-10: Detail of the ^{1}H-^{13}C HSQC spectrum of linamarin-3'-*O*-phosphate. The 1dNOESY from Fig. SN1-7 served as F2-projection.

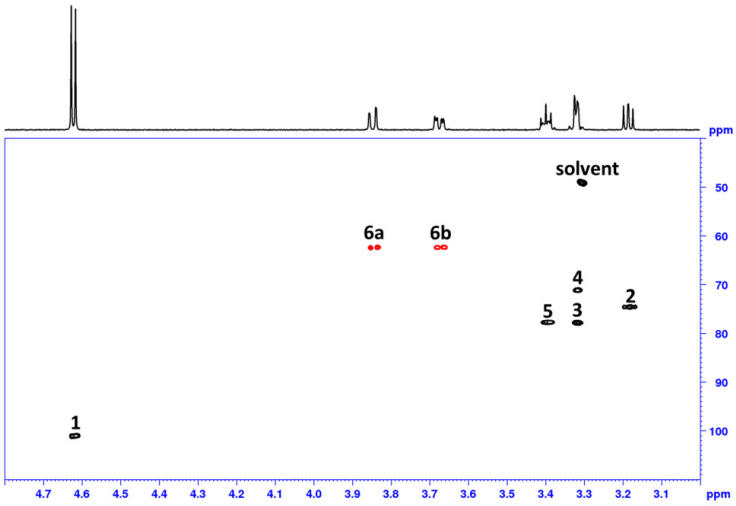


Figure SN1-11: Detail of the ^{1}H-^{13}C HSQC spectrum of linamarin (in MeOH-d_3). A SELTOCSY (transmitter on pos.1 of the glucose part) was used as F2-projection. The spectrum illustrates the chemical shifts of position 3 of the glucose part without phosphorylation.

Manuscript III

Isomerization and oligomerization of dietary disaccharides by *Bemisia tabaci* transglucosidases

Articles Preprint **(ISSN 2470-8224)**

Isomerization and oligomerization of dietary disaccharides by *Bemisia tabaci* transglucosidases

Michael L.A.E. Easson[1], Jonathan Gershenzon[1], Daniel G. Vassão[1]

[1] Max Planck Institute for Chemical Ecology, Jena 07745, Germany

Corresponding author: Daniel G. Vassão

email:

Abstract

The importance of the interaction between sugars in the phloem sap and the enzymes of phloem feeding insects that modify them has been an area of great interest in understanding the metabolic adaptations of phloem feeding insects. The whitefly Bemisia tabaci is capable of feeding on numerous plant species, and research into the sugar modifications it performs when feeding on phloem is well documented. However the responsible enzymes for these modifications/ transformations remain all but elusive. Here we document the characterization of three B.tabaci glucohydrolase (GH) enzymes which perform unique sugar isomerization reactions to form trehalulose from sucrose, as well as a plethora of other sugar transglucosidation reactions. This marks the first entry of characterized genes encoding sucrase-transglucosidases with specific roles in osmoregulation or the formation of sugar oligomers in insects.

Keywords: Sucrase-transglucosidase, Glycoside Hydrolase, Trehalulose Synthase, , Isomerization, Osmoregulation, *B. tabaci*

Introduction

Bemisia tabaci is a cryptic species complex composed of morphologically indistinguishable populations and represents an important pest for agricultural and ornamental plants (1). It is a phloem feeding insect, and therefore has specific adaptations for feeding on this specialized tissue. First and foremost is the navigation of the plant tissue by their unique stylet in order to find their sugary food source. Once within the phloem, the insect must be able to overcome the extreme osmotic pressures imposed by the high concentrations of sugars present therein, mainly in the form of the disaccharide sucrose.

One of the main ways that the insect can overcome this sugar barrier (2, 3) is via sugar modifying enzymes known as sucrase-transglucosidases (3-5) of the glucohydrolase (GH) family of enzymes (3). Phloem feeding insects produce sugar oligomers via the action of these enzymes (2, 4, 6) and by doing so reduce the solute concentration, and thus osmotic pressure of the ingested phloem sap. Mechanistically, sucrose

will enter the catalytic pocket and form a glycosyl-enzyme intermediate with the release of fructose (7), which is absorbed for energy purposes (8). From here the bound glucose has two fates which are largely dictated by the concentration of sucrose. If the concentrations are low enough (and osmoregulation is not important), water may enter the catalytic pocket and release glucose, resulting in a net total hydrolysis of sucrose (7). However, when sucrose or sugar concentrations are high, then another molecule may enter the catalytic pocket such as sucrose, resulting in glucose being transferred to this molecule. In the most relevant example being sucrose, the result is a trisaccharide, and an overall halving of sugar concentration exerted by the two previous molecules of the disaccharide. This process may repeat, thus lowering the sugar concentration of the solution until it reaches a physiologically tolerable level. Genes encoding enzymes from *Bemisia tabaci* of the GH family have been cloned and characterized previously (9, 10), however genes encoding enzymes responsible for sucrose transglucosidation processes have not been observed to date.

Transglucosidation is of general importance to the phloem feeding guild (3-6, 8, 11), however there are other processes particularly present in *Bemisia tabaci* hypothesized to be carried out by GH enzymes. One of the major chemical transformations that *B. tabaci* performs by evidence of investigations of their honeydew is the isomerization of sucrose (α-D-fructofuranosyl-(2↔1)-D-glucose) to the α-(1→1) isomer known as trehalulose (Figure 1) (12-16). The benefit for this isomerization has been hypothesized to be in order to mitigate the hydrolysis of sucrose, by converting it to a less hydrolysable form, thus lowering the potential threat of increasing overall sugar concentrations while feeding (16). Enzymes responsible for this transformation are of the GH family of enzymes (more specifically GH13)(17) and akin to enzymes responsible for transglucosidation, none have been formally characterized with this activity profile in insects. The activity of trehalulose synthases have largely been researched in bacteria (18-22), where they share homology and activity profiles with enzymes known as isomaltulose (sucrose α-(1→6) isomer) synthases (18, 21, 22). Enzymes that are characterized as trehalulose synthases have minor activities in producing isomaltulose and isomaltulose synthases produce small amounts of trehalulose.

The mechanism of isomerization follows a very similar path to transglucosidation with the formation of the glucosyl-enzyme intermediate being identical (19, 22). However, the fructose which would otherwise be released from sucrose following the formation of this glucosyl-enzyme intermediate may then be utilized in the formation of isomerized products. One of the main factors that controls the specificity of these enzymes in bacteria is the motif of amino acids ^{325}RLDRD^{329} (21, 22). Changes to this motif results in the stabilization of a glucosylpyranose intermediate state which results in the (1→1) sugar linkage observed (21). This activity in insects seems to be particularly present in species of whiteflies (12, 14, 23) and recent evidence supports this activity

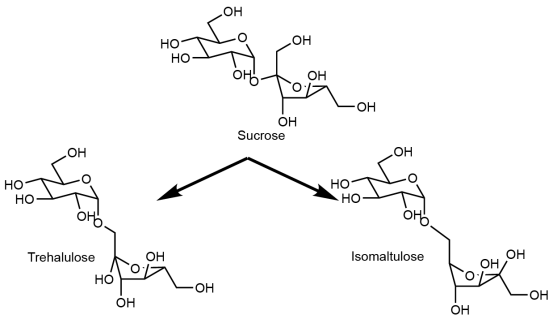


Figure 1: Sucrose is isomerized to form α-(1→1) linked trehalulose and α-(1→6) isomaltulose. Depicted is the typical activity of enzymes known as trehalulose and isomaltulose synthases. In bacteria, trehalulose synthesizing enzymes produce small amounts of isomaltulose and vice versa.

originating from an insect derived enzyme rather than a possible endosymbiotic bacteria (15).
In the realm of phloem feeders, the enzymes responsible for the modification of their sugary diet are an area of high interest for understanding the evolution of these specialized feeders. The whitefly *Bemisia tabaci* is a prolific phloem feeder (1) and represents an interesting study species; possessing a large number of GH family enzymes (3), and being well-studied in it's sugar modifying activities (12-14, 23-26). Here we explore and characterize for the first time multiple *B. tabaci* genes encoding enzymes capable of modifying sucrose to trehalulose. We demonstrate the breadth and utility of these enzymes for the formation of sugar oligomers, illustrating the function of these enzymes in osmoregulation in this phloem feeding insect. Incubation of isotopically labeled sugars with these enzymes also provides a window into the mechanistic action of these isomerases and transglucosidases *in vitro*.

Results

Cloning of a Trehalulose Synthase from *Bemisia tabaci* MEAM1 whitefly

Trehalulose (α-D-fructofuranosyl-(1→1)-D-glucose) has been identified in the honeydew of *Bemisia tabaci* since metabolic characterizations of its contents began. However, the enzymes responsible for the conversion of sucrose to trehalulose in this insect had not yet been identified. In fact, any isolated enzyme and the gene encoding it which are responsible for the conversion of sucrose to the α-(1→1) linked isomer had not been identified in insects thus far, and only genes encoding enzymes with trehalulose synthase activity have been identified in bacteria. Parallel activities to trehalulose synthase are the isomaltulose (α-D-fructofuranosyl-(1→6)-D-glucose) synthases. In bacteria, usually the activities of these enzymes are characteristically linked, with trehalulose synthases producing small amounts of isomaltulose as a side product and vice versa. In an effort to find candidate enzymes from *Bemisia tabaci* with trehalulose synthase activity, protein sequences of a previously characterized trehalulose synthase from *Pseudomonas mesoacidophila* (GenBank: ACO05018.1) and an isomaltulose synthase from *Erwinia rhapontici* (GenBank: HM461324.2) were utilized in separate BLAST searches on the compiled whitefly genome database (whiteflygenomics.org) for *Bemisia tabaci* MEAM1 (BtMEAM1). The results of the BLAST (Supplemental Figure S1) provided 8 different gene candidates. Surprisingly, *Bta03818* previously characterized (BtSUC1) (10) appeared in both the trehalulose and isomaltulose BLAST searches and previously characterized *Bta14419* (BtSUC5) from BtMEAM1 (10) was annotated as having potential isomaltulose synthase activity. All candidates were cloned into expression vectors for S2 insect cell transformation except *Bta01478* as it was not possible to amplify from whitely cDNA. Analysis by SDS-PAGE under denaturing conditions showed single bands for the expressed proteins at the expected size for each amino acid sequence demonstrating that no extensive additional post-translational modifications took place. For all genes cloned, activity of the encoded proteins seemed to exist in the media of expressing S2 insect cells, as had been previously observed (10). Phylogenetic analysis was also performed on all glycoside hydrolase (GH) genes in the BtMEAM1 genome (Figure 2A and supplemental Figure S2). A majority of the gene candidates can be seen as originating from a clade of GH genes marked with an asterisk (Figure 2A), and all candidates were of the GH13 family of GH genes, even further categorized in the GH13-17 subfamily (10).

Sucrose Hydrolysis and Isomerization Activity of Candidate Trehalulose Synthases

Of all the candidate genes expressed and assayed, only three showed any isomerization activity, or any detectable activity with sucrose (**1**) as a substrate (Figure 2 B and Supplemental Figure S3). Among these was the previously characterized

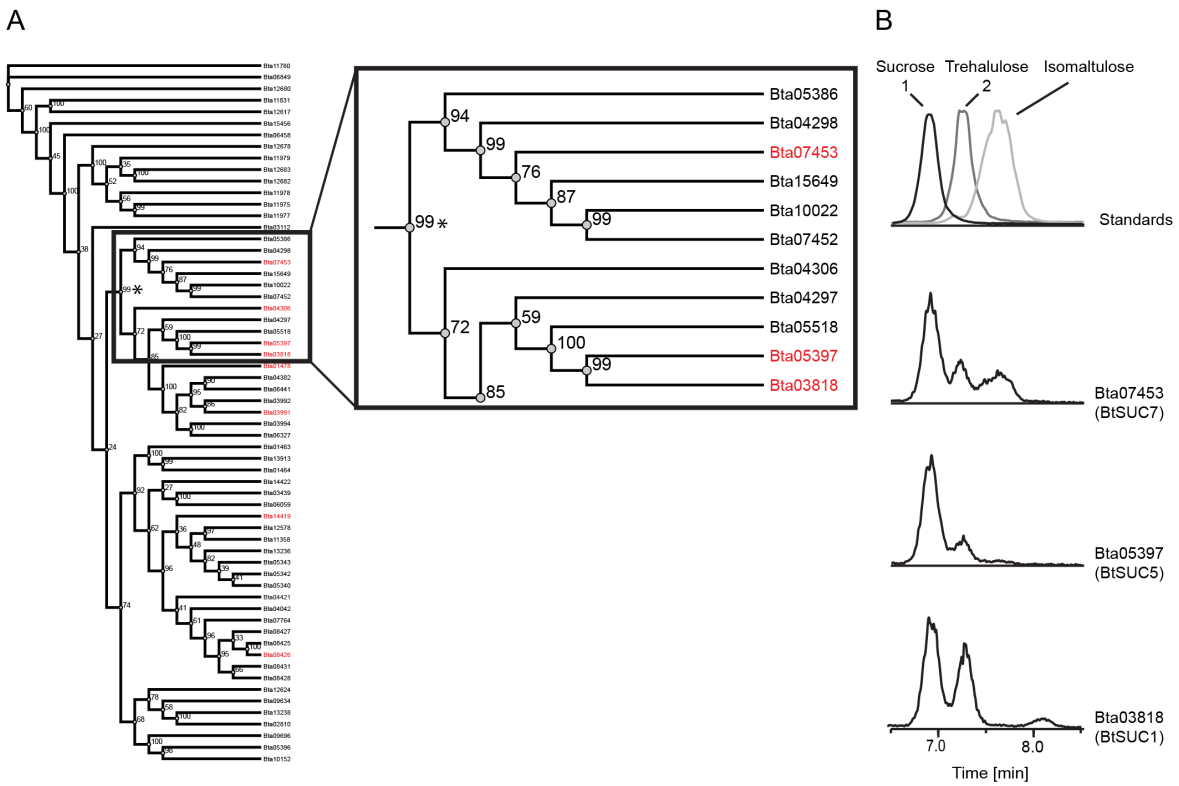


Figure 2: Glucohydrolase genes *Bta03818* (BtSUC1), *Bta05397* (BtSUC6) and *Bta07453* (BtSUC7) are caplable of sucrose isomerization. (A) A phylogeney of the glucohydrolase (GH) enzymes in Bemisia tabaci highlighting the phlogenetic relationship of BtSUC1,6 and 7 with (B) isomerization activity utilizing sucrose and producing trehalulose and other dissaccharides. Phylogeny and bootstrap was performed based on 1000 repetitions utilizing PhyML online phylogenetic tool.

BtSUC1 which has already been shown to have functionality as a hydrolase, but now also revealed is its isomerization potential, which was only detectable after sucrose hydrolysis using yeast inveratse due to the similar retention time. Two additional GH genes encoding enzymes *Bta05397* (BtSUC6) and *Bta07453* (BtSUC7) displayed seemingly lower trehalulose (**2**) synthase activity than BtSUC1 with SUC7 producing a peak with similar retention time to isomaltulose.

TIMS-ToF (Trapped Ion Mobility Spectromotry-Time of Flight) mass spectrometry is a method of mass spectral analysis which allows for an extra dimension of separation and discrimination in metabolite analysis. This method of separation subjects ions from a sample to an electric field in which they are trapped within a moving column of gas (27). The electric field is then slowly lowered and ions which are more influenced by the gas (larger ions) are eluted towards a time of flight analyzer where their mass is recorded. The information from this type of analysis allows one to collect mobilograms for each ion and assign momentum collision cross-sections (CCS) (28) values which are unique to these ions and can thus be diagnostic in combination with retention time in an HPLC column. When analyzed via TIMS-ToF, BtSUC1, 6 and 7 were shown to be able to

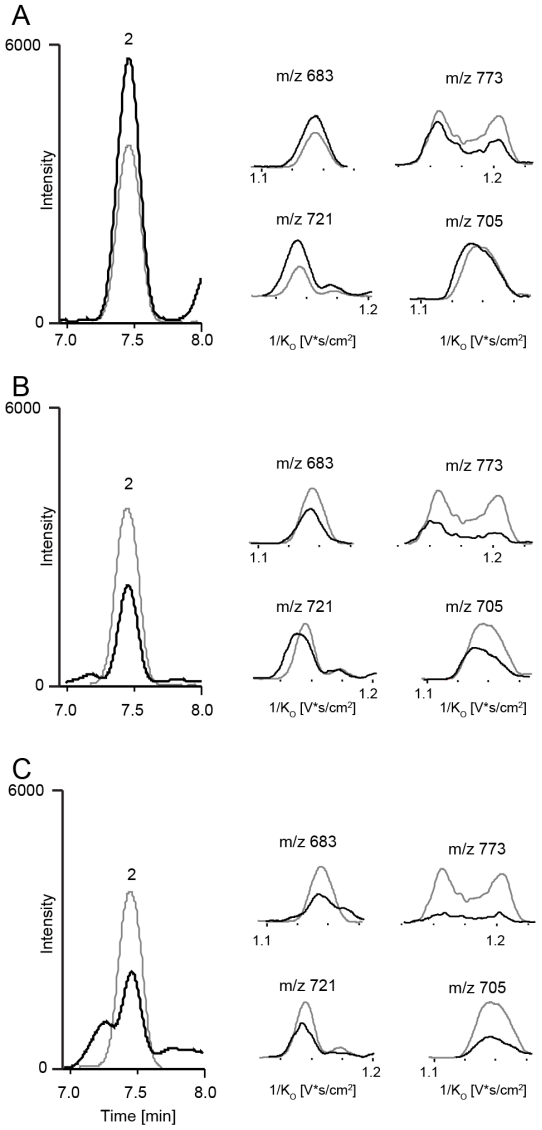


Figure 3: Trehalulose isomerization is confirmed with authentic standard and TIMS-ToF analysis. Common ions formed in the standard injection of trehalulose (grey trace) and other disaccharides were utilized in the comfirmation of trehalulose synthesis activity in BtSUC1 (A), BtSUC6 (B) and BtSUC7 (C) assays with sucrose as a substrate (black trace). Further detailed analysis of ion mobility values can be found in Supplemental Table S1.

perform the isomerization of sucrose to trehalulose by retention time and the identity to the mobility of multiple ions (Figure 3 A, B and C respectively) confirming them as an authentic insect GH enzyme with trehalulose synthase activity. BtSUC1 showed the greatest of all trehalulose synthase activity (Figure 3 A) with accumulation being the greatest in assays with this enzyme. In order to gain further mechanistic insight into the formation of trehalulose *in vitro* by these enzymes and to confirm its chemical composition, enzyme assays utilizing various labeled sugars were performed. Full hexose label incorporation was seen when the enzymes were given $^{13}C_{12}$-sucrose as a donor sugar (Supplemental Figure S4) and only one hexose label was seen to incorporate when given [glucose]$^{13}C_6$- sucrose and [fructose]$^{13}C_6$- sucrose (Supplemental Figure S4 C and D respectively). It was unexpected that label was incorporated in both $^{13}C_6$-glucose with unlabeled sucrose and $^{13}C_6$-fructose with unlabeled sucrose containing assays (Supplemental Figure S4 E and F), with more hexose label being incorporated (by isotopic abundance of the peak) in the $^{13}C_6$-fructose unlabeled sucrose assay. This provides evidence for the enzymes to utilize hydrolysis products of sucrose (free glucose and fructose) to form trehalulose *in vitro*.

The activity of BtSUC7 for the formation of the other notable disaccharide (**3**) in its assay was tested against the authentic standard of isomatlulose. However, the peak showed poor mobility and was unable to be matched to any sugar run in this study (Supplemental Table S1). Labeling studies indicates that this is compound is composed of two units of glucose, and dual hexose-label incorporation in [glucose]$^{13}C_6$-sucrose assays indicates the ability of BtSUC7 to also bind free glucose in transglucosidation reactions (Supplemental Figure S5).

Sucrose Transglucosidation Activity of Trehalulose Synthases

When incubated with sucrose, the three encoded enzymes were also assayed for

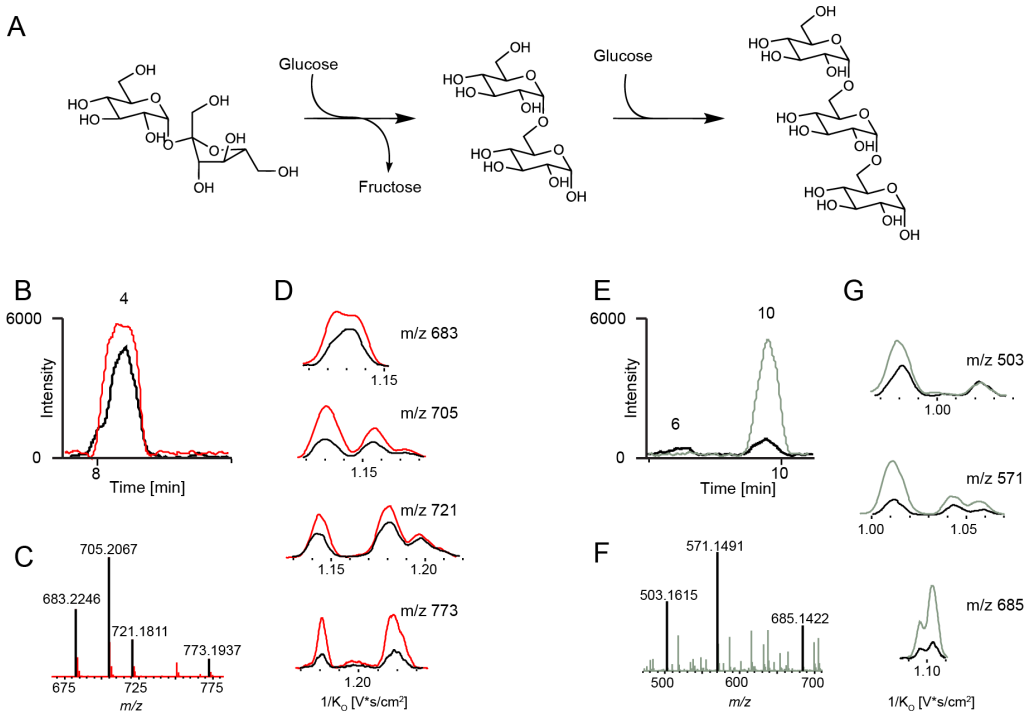


Figure 4: BtSUC1 is capable of α-(1→6) transglucosidation reactions resulting in isomaltose and isomaltotriose synthesis. (A) Depicted is the formation of isomaltose (**4**) and isomaltotriose (**10**) via subsequent transglucosidation. Through labeling studies (Supplemental Figure S6 and S8), it was shown that the source of glucose can be from either free glucose or sucrose. BtSUC1 (black trace) was confirmed by TIMS-ToF with authentic standards to be capable of the formation of the disaccharide isomaltose (panels B-D) and isomaltotriose (panels E-G) illustrating the ability of this enzyme to perform α-(1→6) transglucosidation reactions. Further detailed CCS and ion mobility values can be found in Supplemental Table S1.

potential sucrose transglucosidation activity, or the formation of larger order sugar oligomers, a process associated with osmoregulation. After concentration of protein fractions and incubation with sucrose, all enzymes were shown to be able to form a various trisaccharides from sucrose.

BtSUC1 seemed to produce an unexpected disaccharide product with retention further than 8.0 min (Figure 4A peak number **4**). When analyzed via TIMS, this metabolite seems to correspond to isomaltose (Figure 4A) and no other sugar analyzed (Supplemental Table S1). In the labeled assay studies, ^{13}C was seen to be incorporated as expected in assays containing $^{13}C_{12}$ sucrose and [glucose]$^{13}C_6$- sucrose, but not in [fructose]^{13}C6-sucrose or $^{13}C_6$-fructose and unlabeled sucrose assays (Supplemental Figure S6 B, C and D and E respectively). Interestingly, hexose label was detected in various levels of incorporation (Supplemental Figure S5 F) in assays which contained $^{13}C_6$-glucose with unlabeled sucrose, even full label incorporation, corresponding to two labeled glucose units in the final molecule showing the ability of this enzyme to bind free glucose *in vitro*.

BtSUC1 was also able to produce two peaks corresponding respectively to an unknown trisaccharide (Figure 4 B peak **6** and supplemental Table S1) and isomaltotriose (Figure 4 B peak **10**) confirmed by TIMS analysis (Figure 3 B). Through

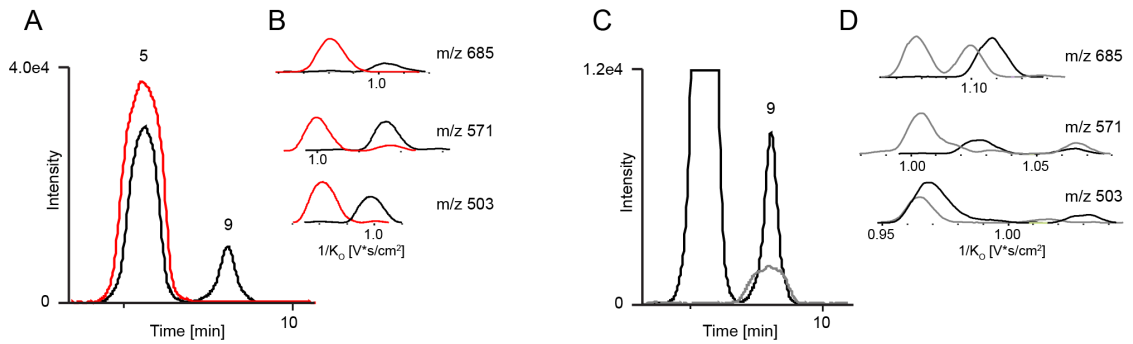


Figure 5: BtSUC6 produced trisaccharide peaks of unknown identity from incubations with sucrose. Incubations of BtSUC6 and sucrose (black trace) resulted in the formation of two prominent peaks corresponding to trisaccharides. Despite the similarity in retention time to known standards raffinose (A, red trace) with (**5**) and panose (B, grey trace) with (**9**), these sugars formed by the enzyme displayed drastically different mobilities for common trisaccharide ions (B and C). Labeling studies with these enzymes (Supplemental Figure S9) confirms (**5**) has similar composition to maltotriulose, containing two units of glucose and one of fructose and (**9**) is composed entirely of glucose units (Supplemental Figure S10). Further detailed analysis of ion mobility values can be found in Supplemental Table S1.

analysis of the labeling assays, it was seen that labeled hexose units were incorporated into the unknown peak in $^{13}C_{12}$ sucrose resulting in full label incorporation, [glucose]$^{13}C_6$- sucrose resulting in two labeled hexoses added and in [fructose]$^{13}C_6$-sucrose with only one labeled hexose (Supplemental Figure S7 A,B and C). Label was also incorporated in assays containing $^{13}C_6$-glucose with unlabeled sucrose, displaying isotope peaks corresponding to no label incorporation, one and two hexose labels. Again, only one hexose label was incorporated in $^{13}C_6$-fructose with unlabeled sucrose composition assays. Isomaltotriose was labeled in assays containing $^{13}C_{12}$ sucrose, [glucose]$^{13}C_6$- sucrose and $^{13}C_6$-glucose with unlabeled sucrose (Supplemental figure S8). The labels were absent in [fructose]$^{13}C_6$- sucrose and $^{13}C_6$-fructose with unlabeled sucrose assays (Supplemental figure S8) as expected for a molecule containing only glucose units. This further illustrates the ability of this enzyme to transglucosidate in an α-(1→6) fashion with additional activity for the binding of free glucose. No further higher order glycosides were detected in these assays with BtSUC1.

With the assumption of similar ionization potential for all compounds within a single class (di, tri and tetrasaccharides) BtSUC6 showed the greatest potential for transglucosidation activity *in vitro*, producing two trisaccharide peaks of unknown identity (Figure 5 A and Supplemental Table S1). Both of these peaks through labeling studies showed label incorporation at varying levels, giving insight into their chemical composition. For the first eluting peak (peak **5**) full label incorporation was seen in assays with $^{13}C_{12}$-sucrose, and importantly two and one hexose label was incorporated for [glucose]$^{13}C_6$- sucrose and [fructose]$^{13}C_6$- sucrose respectively (Supplemental Figure S9). This pattern of addition shows that the molecule consists of two units of glucose and one of fructose. For the second eluting peak (Figure 5 B peak **9**) full label incorporation in $^{13}C_{12}$ sucrose, and [glucose]$^{13}C_6$- sucrose as well as various levels of label incorporation were seen for $^{13}C_6$-glucose with unlabeled sucrose indicates the molecule consists of entirely glucose units (Supplemental Figure S10). Larger oligosacchairdes were also seen for BtSUC6 (Supplemental Figure S11 A) corresponding to 4 hexose units polymerized (**11** and **13**), consisting of an earlier eluting peak

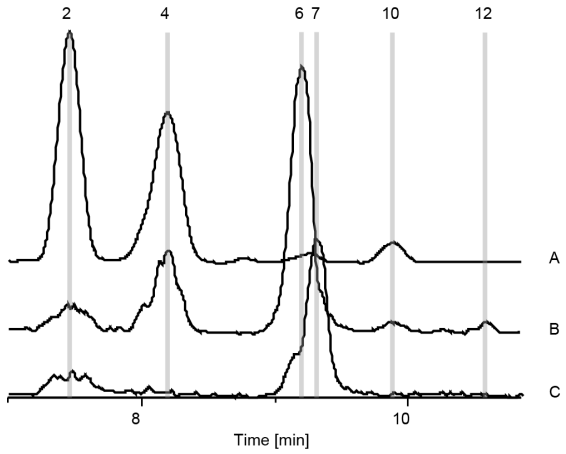


Figure 6: Incubations of BtSUC1 result in different products depending on the starting sugar substrate. BtSUC1 was incubated in parallel assays with sucrose (A), trehalulose (B) and isomaltulose (C) in order to visualize substrate promiscuity. Between sucrose and trehalulose, identical products were formed with greater amounts of higher order products (**6**) and (**12**) being formed in trehalulose assays. In incubations with isomaltulose, a trisaccharide (**7**) was produced by BtSUC1 with similar retention to (**6**) but different TIMS for ions used diagnosis (Supplemental Table S1). Interestingly, isomaltose (**4**) and isomaltotriose (**10**) were not detected in isomaltulose incubations.

showing patterns of incorporation for three units of glucose and one unit of fructose (Supplemental Figure S11 B-G) and a later eluting peak who's composition was unable to be determined due to abundance in labeling studies.

BtSUC7 produced yet another peak for trisaccharides unable to be identified by any standards utilized in this study (Supplemental Figure S12 A peak **8**). Labeling studies for this peak provide evidence for the presence of three glucose units as $^{13}C_{12}$ sucrose and [glucose]$^{13}C_6$- sucrose result in full hexose labeling but [fructose]$^{13}C_6$-sucrose results in no incorporation (Supplemental Figure S12 B).

Trehalulose Incubation with Trehalulose Synthases

Trehalulose incubation was also an important process to understand with regards to enzyme activities, as this is metabolite represents the most abundant compound in the honeydew of whiteflies feeding on virtually every plant. Interestingly, the first activity that was apparent was the hydrolysis of trehalulose to glucose and fructose by both BtSUC1 and 6, but absent in SUC7 (Supplemental Figure S13). Following a similar pattern, BtSUC1 also showed transglucosidation activity with trehalulose and formed the same peaks that were formed in the sucrose incubation assays (Figure 6A), with variable efficiency and an additional peak corresponding to 4 hexose additions (Figure 6 B peak **12**). The inability of BtSUC1 to form detectable quantities of tetrasaccharide product utilizing sucrose perhaps marks trehalulose as the better substrate for transglucosidation and osmoregulation. BtSUC6 was also able to produce peaks which corresponded identically to those formed in sucrose assays (Supplemental Figure S14), but interestingly in lower levels than seen previously. BtSUC7 was unable to utilize trehalulose as a substrate to form any other products.

Isomaltulose Incubation with of Trehalulose Synthases

In a parallel set of experiments, incubations with isomaltulose provided insight into the perhaps inevitable promiscuity of these enzymes, but also observed specificity. BtSUC1 and 6 were capable of hydrolyzing this metabolite, despite the fact that they seemingly do not form it *in vitro* (Supplemental Figure S13 C). The enzyme that forms an unknown disaccharide with similar retention to isomaltulose (BtSUC7) does however not hydrolyze this substrate, mimicking the observed inactivity towards trehalulose (Supplemental Figure S13).

Interestingly in these isomaltulose

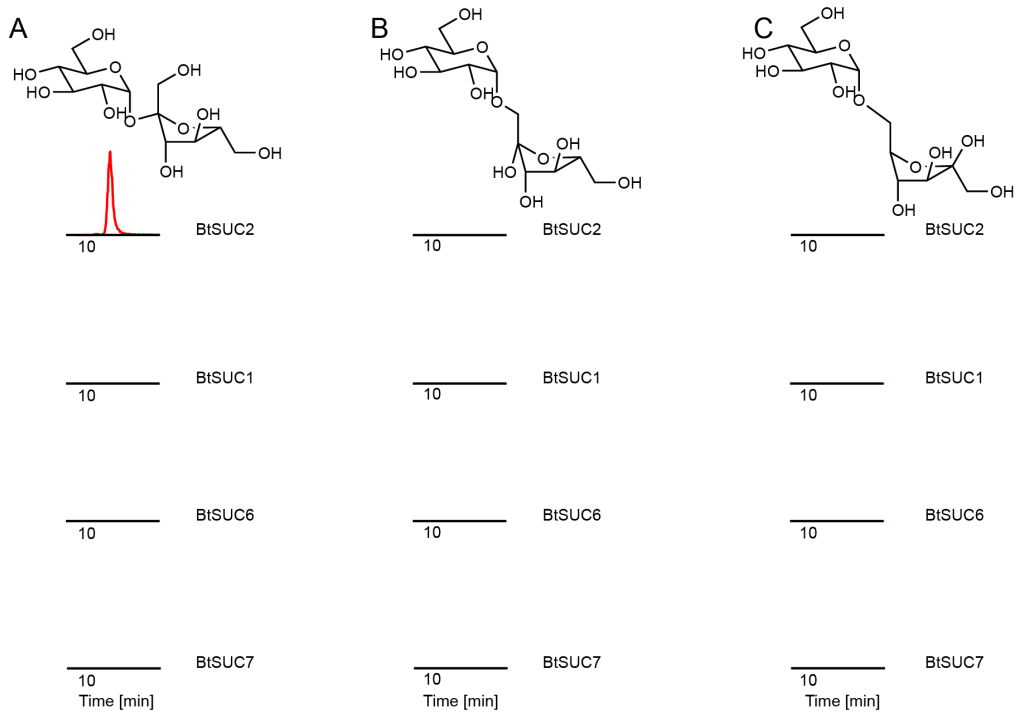


Figure 7: BtSUC1, 6 and 7 are unable to glycosylate secondary metabolite 4msob using substrate sugars sucrose, trehalulose and isomaltulose. Incubations of BtSUC2 (previously shown to have transglucosidase activity with 4msob), 1, 6 and 7 with sucrose (A), trehalulose (B) and isomaltulose (C) containing 2.5mM 4msob-GSL resulted in the formation of glycosylated 4msob-GSL products only in BtSUC2 incubations with sucrose. Depicted are 4msob-GSL (black trace) and 4msob+1 glucose (red).

incubations, it was surprising that BtSUC1 was unable to form isomaltose or subsequent isomaltotriose (Figure 6 C), despite the observed hydrolysis of isomaltulose and the observation that these metabolites can be formed via transglucosidation reaction utilizing free glucose (Supplemental Figure S6 and S8). The absence of these peaks from incubations with isomaltulose is unexpected and perhaps gives insight into new nuances regarding the formation of glucosyl-enzyme intermediates with various substrates.

BtSUC1 displayed further differential transglucosidase activity with this substrate; forming a peak with slightly different retention to the dominant trisaccharide in trehalulose incubations and a distinct TIMS separation (Figure 6 C peak **7**). BtSUC6 however formed identical products to the trehalulose and sucrose incubations (Supplemental Figure S14), however to a lesser extent than sucrose incubation once again. BtSUC7 was unable to utilize this substrate in any capacity.

Transglucosidation of Secondary Metabolite 4msob

Since the discovery of modified secondary metabolites, both 4msob-glucosinolate (GSL) and the cyanogenic glycoside linamarin (10) (and in preparation), it is still not entirely determined whether the mechanism or enzymes involved in

osmoregulation (isomerization and oligosaccharide formation) is the same utilized in the modification of secondary metabolites. With the current data supporting enzymes isomerizing sucrose to form trehalulose and this metabolite being utilized in trisaccharide and larger saccharide formation, sucrose and trehalulose were utilized in enzyme assays containing 4msob-GSL and expressed trehalulose synthases BtSUC1, 6 and 7 as well as BtSUC2 which showed transglucosidation of 4msob-GSL previously (10). Isomaltulose was also used in a similar assay in order to test if there is any promiscuity in its utilization by these enzymes. Only transglucosidation of 4msob-GSL was accomplished in BtSUC2 assays utilizing sucrose and none of the enzyme-substrate combinations were capable of producing the transglucosidation product of 4msob-GSL (Figure 7A, B and C respectively). This observation provides further evidence for a division between secondary metabolite modification and the reactions important for osmoregulation.

K_M estimation for sucrose utilization

Due to the multi-product nature of the transformations highlighted by these enzymes, K_M determination was accomplished via the reduction of sucrose peak areas in assays (Supplemental Figure S15). Here it was found that BtSUC1 has a K_M estimated for sucrose utilization of ~0.2 M, BtSUC6 being somewhat higher at ~0.5 M and finally BtSUC7 the lowest at 0.04 M. With BtSUC7 being the best of all enzymes in terms of hydrolysis, it seems to make sense that it too has the lowest K_M as it has been previously hypothesized that greater hydrolysis is seen with lower sucrose concentrations in aphids (29), while BtSUC1 and 6 have higher K_Ms, more reflective of the processes and concentrations seen with transglucosidating activites (6).

Discussion

Phloem feeders are a unique feeding guild among herbivores, choosing to feed on the sugar-rich phloem through navigation of their specialized stylets. As a consequence, phloem feeders such as *Bemisia tabaci* must be able to utilize their food source without ill effects of feeding on such a sugar dense mixture. *Bemisia tabaci* (and other phloem feeders) accomplish this task through the tight regulation of water transporters (2) and the transformation of their food into a less threatening composition (osmotically) via sucrase-transglucosidases. These enzymes, of the GH superfamily of enzymes (17) seem to have undergone a huge expansion in terms of the number of individual genes present in the genome compared to non-phloem feeders(10), with *B. tabaci* having the most of phloem feeders surveyed to date (3, 10). This research aimed to identify candidate enzymes responsible for the impressive polymerizations and isomerizations in *Bemisia tabaci*.

Trehalulose is one of the major components of whitefly honeydew (12) and seems to be particularly present in whitefly metabolism (12, 14, 23), where aphids and other phloem feeders do not seem to produce this sucrose isomer at all (12). Trehalulose has been previously thought to serve as a quick sugar transformation of sucrose to a less or non-hydrolyzable isomer (16). This strategy would effectively allow a majority of the incoming disaccharide (sucrose) (30, 31), in phloem to become more resistant to hydrolyzing enzymes that may increase or exacerbate the problem of osmoregulation. Interestingly, BtSUC1 and 6 were shown to have significant hydrolysis activity towards trehalulose (Supplemental Figure S13) while also being able to form it *in vitro*, but BtSUC7 who also formed trehalulose to a lesser extent *in vitro* (Figure 2 C) was unable to perform any chemical transformation of this substrate, including hydrolysis. The hydrolysis of trehalulose was unexpected and perhaps speaks against the aforementioned hypothesis that the benefit of this compound is lowering potential hydrolysis, or at the very least it may still reduce this potential in other enzymes. However, the hydrolytic action of the observed may be an *in vitro* construct as these enzymes are also known to form what appears to be

higher order protein complexes in the whitefly (15), which were not seen here, and may have drastic effects on catalysis such as transglucosidation and hydrolysis.

Activities of further interest are provided by labeling studies with trehalulose synthases, demonstrating that these enzymes were capable of incorporating free glucose and fructose in the formation of trehalulose from sucrose with free fructose being incorporated to a greater extent by isotopic abundance (Supplemental Figure S4). This catalytic flexibility may indicate that the enzymes which form trehalulose are also capable of synthesizing it from sucrose hydrolysis products in the whitefly gut. This flexibility would also help in osmoregulatory functions within the insect in order to already produce disaccahrides from glucose and fructose alone, thus mitigating effects of hydrolysis.

Originally, the genes were chosen by amino acid sequence shared homology to bacterial trehalulose and isomaltulose synthases with the hypothesis that catalytic motifs would be similar, however the sequences are fairly distinct (Supplemental Figure 15). The motif which is of noted importance for direction towards trehalulose and isomaltulose, being the ^{325}RLDRD^{329} (21, 22) was present in the isomaltulose synthase from *Erwinia*, and a sequence of similarity "RYDRA" present in trehalulose synthase from Pseudomonas did not align with our candidate enzymes (Supplemental Figure S15). Upstream of this sequence in the alignment however, exists a motif of RYXR present in BtSUC1 and 6 illustrating their similarity to the trehalulose synthase with tyrosine being possibly playing an important role for direction towards trehalulose. Interestingly, none of the enzyme here seemed to produce isomaltulose, which is an expected side-product always present with bacterial enzyme activities (18-22), but is consistent with reports of whitefly gut activities (15). Although the activity has been described (15), any gene encoding an enzyme that produced trehalulose from sucrose in insects has not been cloned and characterized, marking this as the first full description of this activity in insects to date.

Transglucosidation is of paramount importance to whitefly and phloem feeder metabolism, being one of the two pillars for osmoregulation (2). All of the enzymes tested here were multifunctional in the sense that they were able to produce trehalulose from sucrose, but also transglucosidation products. The most effective of all three enzymes by far in terms of transglucosidation efficiency using sucrose was BtSUC6 (Figure 4 A and B). The activity for transglucosidation was incredibly variable between each of the enzymes tested, producing unique peaks corresponding by mass to trisaccaharides. BtSUC1 was interestingly able to produce two compounds that were identifiable as isomaltose and isomaltotriose (Figure 3 A and B respectively), an activity that was not observed previously (10), most likely due to the concentration of enzymes extracts (approximately 10 times greater here). Both of these metabolites are composed entirely of glucose units with α-(1→6) linkages and thus are prime examples of transglucosidation activity. Similar to the observed formation of trehalulose in labeling studies, this enzyme was able to bind free units of glucose as donors in the polymerization process (Supplemental Figure S6 and S7). This process was approximately 33% as efficient in the formation of trehalulose than using sucrose as the glucose donor based on isotopic abundance of the labeled products (Supplemental Figure S4), but was utilized more so in the formation of isomaltose (Supplemental Figure S6). This activity is unprecedented as it is not expected from this family of GH enzymes, however similar activity is observed for the polymerization of glucose units with glucoamylases of *Aspergillus niger* (32). It can perhaps be expected for enzymes with significant hydrolyase activity as the equilibrium exists for the reaction with water and naturally vice versa where glucose may enter the catalytic pocket (32). The formation of isomaltose and isomaltotriose is also observed when BtSUC1 was incubated with trehalulose. An explanation for why this activity is inherently absent from incubations with isomaltulose is not clear, as it is

effeciently hydrolyzed by BtSUC1 to produce free monosaccharides to a degree similar to sucrose (Supplemental Figure S13).

BtSUC6 produced trisaccharides of variable composition, being of similar composition to maltotriulose (Figure 5 A) consisting of two units of glucose and one unit of fructose and those consisting of only glucose units (Figure 5 B). It is expected that the first of the eluting peaks is a transgluosidation product of sucrose directly, which places the activity well in line with osmoregulatory function in lowering sugar concentration in the insect gut. As seen with BtSUC1; BtSUC6 is capable of utilizing free glucose in the formation of glucosyl-enzyme intermediates for transglucodiation reactions (Supplemental Figure S7 and S8). The peak formed for the glucose only polymer (peak #**9**) is different than isomaltotriose, and at this point unidentified. There are reports however, for unusual trisaccharides in whitefly honeydew that have been identified as bemisiose (A glucose trisaccahride) (13) and this peak may correspond to this metabolite. Additionally, BtSUC6 was able to form detectable transglucosidation products corresponding to tetrasaccharides (Supplemental Figure S11 A). Once more it was seen that for peak # **11** (Supplemental Figure S11 A), sucrose may be the glucose unit acceptor as indicated by labeling studies (Supplemental figure S11 B), notably the incorproration of three labeled hexose units in [glucose]$^{13}C_6$- sucrose assays. The composition of the later eluting peak was not able to be determined through labeling studies due to abundance in these assays, however it is expected to be yet another glucose-only polymer due to the duality of transglucosidase products seen thus far by these enzymes. BtSUC6 was also able to form the same products in incubations with trehalulose and isomaltulose (Supplemental Figure S14) as seen in the enzyme assays with sucrose, however to a lesser extent, perhaps showing a preference for the utilization sucrose rather than trehalulose in transglucosidation activities, seemingly the opposite of BtSUC1.

BtSUC7 has the highest activity as a hydrolase for sucrose of the three GH enzymes tested here, however shows no activity as a hydrolase with trehalulose or isomaltulose (Supplemental Figure S13). The formation of trisaccharide by this enzyme in sucrose assays was seen as being the result of glucose unit polymerization by label incorporation (Supplmental Figure S12), and the compound remains unidentified at this point. The inactivity of this enzyme towards trehalulose and isomaltulose in any capacity (both hydrolysis and transglucosidation) is unexpected, however it perhaps gives credence to the theory of trehalulose isomerization lowering the overall hydrolysis in this insect's gut with other GH-13 enzymes aside from BtSUC1 and 6.

Finally, the incubation of BtSUC1, 6 and 7 with sucrose, trehalulose, and isomaltulose in assays containing the secondary metabolite 4msob-GSL provides further evidence for a division between osmoregulatory functions within this insect and enzymes dedicated to secondary metabolite modification (10). This division in many areas seems to make sense as the K_M for sugar oligomer formation is many factor higher in terms of concentration as µM concentrations of sugar (corresponding to secondary metabolite concentration) in the insect gut are of no concern from an osmoregulatory perspective. Having dedicated enzymes for both processes is a good strategy, which mechanistically allows the two separate paths to operate on drastically different concentration ranges such as molar, for sugar modification and µmolar for secondary metabolite detoxification.

In conclusion, the analysis of candidate trehalulose synthases from the whitefly *Bemisia tabaci* has resulted in the characterization of three genes encoding enzymes with trehalulose synthase activity; catalysis previously only characterized in bacteria. These enzymes were discovered to be multi-product forming, and capable of a vast array of sugar modifications, including hydrolysis of sucrose and its isomers, isomerization and transglucosidation, placing these enzymes as potentially important in the process for osmoregulation within the insect.

This research highlights the utility of glycoside hydrolase enzymes in the metabolism of *B. tabaci* where further analysis into whitefly honeydew will undoubtedly reveal what enzymes are responsible for certain activities in the profile of this phloem feeder.

Materials and Methods

Phylogenetic analysis

Amylosucrase protein sequences from Bemisia tabaci (MEAM1) were downloaded from open database (www.whiteflygenomics.org). Putative GH protein sequences were then aligned using MUSCLE with default parameters. Maximum likelihood tree inference was conducted with PhyML with 1000 bootstrap as node support values. Tree representation was performed in Figtree version 1.4.3. Other alignment tasks were performed using BioEdit using default MUSCLE parameters.

Cloning and Expression in *Drosophila* S2 cells

Vectors, cells and cell media were obtained from ThermoFisher Scientific (Waltham, MA, USA). Full-length ORFs from candidate GH13 genes (*SUC1, SUC6,* and *7*) were amplified using using primer sets outlined in Supplemental Table S2 from *Bemisia tabaci* MEAM1 whole-body cDNA synthesized using Superscript IV (Invitrogen) on RNA extracted with TriZol; cloned into PCR4 Blunt TOPO and released using corresponding restriction enzymes according to each primer. The digested fragment was further ligated into the pAc5.1/V5-His A vector for expression in *Drosophila* S2 cells using Schneider's medium. Insect cells were co-transformed with pCOBLAST vector for selection via blasticidin. Transformation was achieved via $CaCl_2$ incubation following the manufacturer's standard protocols and selected with 50 µg mL^{-1} blasticidin in 6 well plates. After one week of selection, cells were maintained in T-flasks at 27°C with 10 µg mL^{-1} blasticidin. Cells were regularly maintained by splitting 1:20 when cells had achieved confluency of 90 % (approximately every 4-5 days). Cells and media were harvested for enzyme activity studies when cells reached 90 % confluency.

Enzyme Assays

Drosophila S2 cells expressing candidate GH13 enzymes as well as non-transfected control cells were centrifuged at 100 ×*g* for 5 min. The resulting supernatant was concentrated using Amicon Ultra 0.5 mL Ultracel® -30 K centrifugal filters a factor of 10x and then utilized as the secreted protein fraction for enzyme assays. For determination of sugar isomerization and transglucosidation activity, 5 µL of each supernatant containing an equivalent total protein content (verified by Bradford assays) was mixed with a 5 µL 50 mM phosphate buffer at pH 7.0 containing 2 M sucrose concentration, and reacted for 6 h at 25°C with no stirring. These assay conditions were also repeated for other sugars, trehalulose (BOC Sciences) and isomaltulose (Roth), as well as various labeled sucrose isotopomers (Supplementary table S3). Assay conditions were chosen after screening a range of substrate concentrations and pH values that reflect previous work on glucohydrolases and phloem-feeding insects. The reaction was then stopped using 15 µL of methanol and immediately stored at -20°C. Before analysis, enzyme assays were centrifuged at 5200 ×*g* for 5 min and the supernatant obtained was analyzed by LC-MS API5000. A dilution of 1:1000 in water was used for analysis via LC-MS. When testing for transglucosidation of the secondary metabolite 4msob-GSL, assay conditions were repeated identically as in (10), with sucrose, as well as assays with trehalulose or isomatlulose and 4msob-GSL. For analysis on the Bruker TIMS-TOF, assay conditions were repeated identically, however reactions were not stopped with methanol, and instead 3 units of yeast invertase (Sigma Aldrich) in 50mM phosphate buffer pH 7.0 was added to the reaction for 30 minutes to invert sucrose to glucose and fructose

in order to better visualize isomerized trehalulose and other metabolites with similar retention time. This sample was then filtered through a 30 kDa size exclusion filter (Amicon Ultracel®) to remove protein before normal 1:1000 dilution and injection.

LC-MS analysis of sugar hydrolysis, isomerization and transglucosdiation activties

Analysis of sugar hydrolysis products from the enzyme assays was perfromed on an HP 1260 coupled to an AB Sciex API 5000 mass spectrometer with an apHeraNH2 Polymer column (150 × 4.6 mm, 5µm, Supelco Analytical, Munich, Germany) with a chromatographic gradient of water (Solvent A) and acetonitrile (Solvent B) at a flow rate of 1 mL min^{-1} at 20°C as follows: 80 % B (0.5 min), 80-55 % B (12.5 min), 55-80 % B (1 min), hold at 80 % for 4 min. The mass spectrometer was operated in the negative mode with collision gas value 2, curtain gas pressure 35 psi, spray gas pressure 70 psi, ion spray voltage -4500 V, and turbogas temperature 700°C. Compounds were detected using scheduled multiple reaction monitoring (MRM) detection with the parameters outlined in Supplementary Table S4. Analyst 1.5 software was used for data acquisition and processing.

Qualitative analysis of sugar hydrolysis products from the enzyme assays was perfromed on an Thermo Scientific UltiMate 3000 UHPLC coupled to Bruker TIMS-TOF mass spectrometer with an apHeraNH2 Polymer column (150 × 4.6 mm, 5µm, Supelco Analytical, Munich, Germany) with a chromatographic gradient of water (Solvent A) and acetonitrile (Solvent B) at a flow rate of 1 mL min^{-1} (split 1:3 source: waste) at 20°C as follows: 80 % B (0.5 min), 80-55 % B (12.5 min), 55-80 % B (1 min), hold at 80 % for 4 min. The MS (Bruker Daltonics, Bremen, Germany) was operated in negative mode with TIMS enabled with the following sectional settings. **TIMS Mode** mass scan range: *m/z* 50-1500, TIMS Detect range: 0.60-1.40 V*s/cm^2, ramp time: 228.6 ms, spectral rate: 4.28 Hz, duty cycle: lock to 100%, rolling average: 3x. **Source** End plate offset: 500 V, capillary: 4500 V, Neubilizer: 1.8 bar, Dry gas: 8.0 l/min, Dry temperature 200 °C. **Tune General** Funnel 1RF: 150 Vpp, Funnel 2 RF: 300 Vpp, isCID energy: 0.0 eV, Multipole RF: 200 Vpp, Deflection Delta: -70 V, Quadrupole energy: 5.0 eV, Low mass: 100 *m/z* Collision energy: 10 eV, Collision RF: 400 Vpp, Transfer time: 62.5 µs, Pre-pulse storage: 5.0 µs. **Tune TIMS** Δ1: 20.0 V, Δ2: 120.0 V, Δ3: -70.0 V, Δ4: -60.0 V, Δ5: 0.0 V, Δ6: -70.0 V.

Ion mobility and m/z measurements were internally calibrated using Agilent ESI-L Low Concentration Tune Mix using a 20 µL injection loop with a calibration segment at the beginning of each run. Three calibration points *m/z* 302, 602 and 1034 were used for a linear TIMS calibration.

Quantification of the 4msob-GSL transglucosidation assays was accomplished via an HP 1260 series HPLC coupled to an AB Sciex API 5000 mass spectrometer (Applied Biosystems, Darmstadt, Germany). The column utilized was a Nucleodur Sphinx RP column (250 × 4.6 mm, 5µm, Macherey-Nagel, Germany) using a chromatographic gradient of 0.05 % aqueous formic acid (Solvent A) and acetonitrile (Solvent B) with a flow rate of 1mL min^{-1} at 25°C as follows: 1.5 % B (2 min), 1.5-10 % B (2.5 min), 10-40 % B (7.5 min), 40-70 % B (5 min), 70-100 % B (0.1 min), hold at 100 % B (2.4 min), 100-1.5 % B (0.1 min), and hold at 1.5 % B (3.9 min). The MS was operated in the negative mode with collision gas value 2, curtain gas pressure 35 psi, spray gas pressures 70 psi, ion spray voltage -4500 V, and turbogas temperature 700°C. Compounds were detected using multiple reaction monitoring (MRM) detection with the parameters outlined in Supplementary Table S4. Analyst 1.5 software (Applied Biosystems) was used for data acquisition and processing.

References

1. Oliveira MRV, Henneberry TJ, & Anderson P (2001) History, current status, and collaborative research projects for *Bemisia tabaci*. *Crop Prot* 20(9):709-723.

2. Douglas AE (2006) Phloem-sap feeding by animals: problems and solutions. *J Exp Bot* 57(4):747-754.

3. Jing X, *et al.* (2016) Evolutionary conservation of candidate osmoregulation genes in plant phloem sap-feeding insects. *Insect Mol Biol* 25(3):251-258.

4. Fisher DB, Wright JP, & Mittler TE (1984) Osmoregulation by the aphid *Myzus persicae*: a physiological role for honeydew oligosaccharides. *J Insect Physiol* 30(5):387-&.

5. Walters FS & Mullin CA (1988) Sucrose-dependent increase in oligosaccharide production and associated glycosidase activities in the potato aphid *Macrosiphum-Euphorbiae* (Thomas). *Arch Insect Biochem* 9(1):35-46.

6. Cristofoletti PT, Ribeiro AF, Deraison C, Rahbe Y, & Terra WR (2003) Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid *Acyrthosiphon pisum*. *J Insect Physiol* 49(1):11-24.

7. Sinnott ML (1990) Catalytic mechanism of enzymic glycosyl transfer. *Chemical Reviews* 90(7):1171-1202.

8. Price D, *et al.* (2007) Molecular characterisation of a candidate gut sucrase in the pea aphid, Acyrthosiphon pisum. *Insect Biochem Molec* 37(4):307-317.

9. Jing X, *et al.* (2016) Evolutionary conservation of candidate osmoregulation genes in plant phloem sap-feeding insects. *Insect Mol Biol* 25(3):251-258.

10. Malka O, *et al.* (2020) Glucosylation prevents plant defense activation in phloem-feeding insects. *Nat Chem Biol*.

11. Karley A, Ashford D, Minto L, Pritchard J, & Douglas A (2005) The significance of gut sucrase activity for osmoregulation in the pea aphid, Acyrthosiphon pisum. *J Insect Physiol* 51(12):1313-1319.

12. Byrne DN, Hendrix DL, & Williams LH (2003) Presence of trehalulose and other oligosaccharides in hemipteran honeydew, particularly *Aleyrodidae*. *Physiol Entomol* 28(2):144-149.

13. Hendrix DL & Wei Y-a (1994) Bemisiose: an unusual trisaccharide in Bemisia honeydew.

14. Hendrix DL, Wei YA, & Leggett JE (1992) Homopteran honeydew sugar composition is determined by both the insect and plant species. *Comp Biochem Phys B* 101(1-2):23-27.

15. Salvucci ME (2003) Distinct sucrose isomerases catalyze trehalulose synthesis in whiteflies, Bemisia argentifolii, and Erwinia rhapontici. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* 135(2):385-395.

16. Salvucci ME, Wolfe GR, & Hendrix DL (1997) Effect of sucrose concentration on carbohydrate metabolism in Bemisia argentifolii: biochemical mechanism and physiological role for trehalulose synthesis in the silverleaf whitefly. *J Insect Physiol* 43(5):457-464.

17. Svensson B, *et al.* (2002) Fascinating facets of function and structure of amylolytic enzymes of glycoside hydrolase family 13. *BIOLOGIA-BRATISLAVA-* 57(SUP/2):5-20.

18. Ravaud S, Watzlawick H, Haser R, Mattes R, & Aghajari N (2005) Expression, purification, crystallization and preliminary X-ray crystallographic studies of the trehalulose synthase MutB from Pseudomonas mesoacidophila MX-45. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* 61(1):100-103.

19. Watzlawick H & Mattes R (2009) Gene cloning, protein characterization,

and alteration of product selectivity for the trehalulose hydrolase and trehalulose synthase from "Pseudomonas mesoacidophila" MX-45. *Applied and environmental microbiology* 75(22):7026-7036.

20. Xu Z, *et al.* (2013) The Structural Basis of Erwinia rhapontici Isomaltulose Synthase. *Plos One* 8(9).

21. Zhang D, Li N, Swaminathan K, & Zhang L-H (2003) A motif rich in charged residues determines product specificity in isomaltulose synthase. *Febs Lett* 534(1-3):151-155.

22. Zhang DH, Li N, Lok SM, Zhang LH, & Swaminathan K (2003) Isomaltulose synthase (PalI) of Klebsiella sp LX3 - Crystal structure and implication of mechanism. *J Biol Chem* 278(37):35428-35434.

23. Byrne DN & Miller WB (1990) Carbohydrate and amino acid composition of phloem sap and honeydew produced by *Bemisia tabaci*. *J Insect Physiol* 36(6):433-439.

24. Hendrix DL & Salvucci ME (2001) Isobemisiose: an unusual trisaccharide abundant in the silverleaf whitefly, Bemisia argentifolii. *J Insect Physiol* 47(4-5):423-432.

25. Wei YA, Hendrix DL, & Nieman R (1996) Isolation of a novel tetrasaccharide, bemisiotetrose, and glycine betaine from silverleaf whitefly honeydew. *J Agr Food Chem* 44(10):3214-3218.

26. Wei YA, Hendrix DL, & Nieman R (1997) Diglucomelezitose, a novel pentasaccharide in silverleaf whitefly honeydew. *J Agr Food Chem* 45(9):3481-3486.

27. Michelmann K, Silveira JA, Ridgeway ME, & Park MA (2015) Fundamentals of trapped ion mobility spectrometry. *J Am Soc Mass Spectrom* 26(1):14-24.

28. Gabelica V, *et al.* (2019) Recommendations for reporting ion mobility Mass Spectrometry measurements. *Mass Spectrom Rev* 38(3):291-320.

29. Cristofoletti PT, Ribeiro AF, Deraison C, Rahbé Y, & Terra WR (2003) Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid Acyrthosiphon pisum. *J Insect Physiol* 49(1):11-24.

30. Rennie EA & Turgeon R (2009) A comprehensive picture of phloem loading strategies. *P Natl Acad Sci USA* 106(33):14162-14167.

31. Riens B, Lohaus G, Heineke D, & Heldt HW (1991) Amino acid and sucrose content determined in the cytosolic, chloroplastic, and vacuolar compartments and in the phloem sap of spinach leaves. *Plant Physiology* 97(1):227-233.

32. Nikolov ZL, Meagher MM, & Reilly PJ (1989) Kinetics, equilibria, and modeling of the formation of oligosaccharides from D-glucose with Aspergillus niger glucoamylases I and II. *Biotechnology and bioengineering* 34(5):694-704.

Supplementary Information

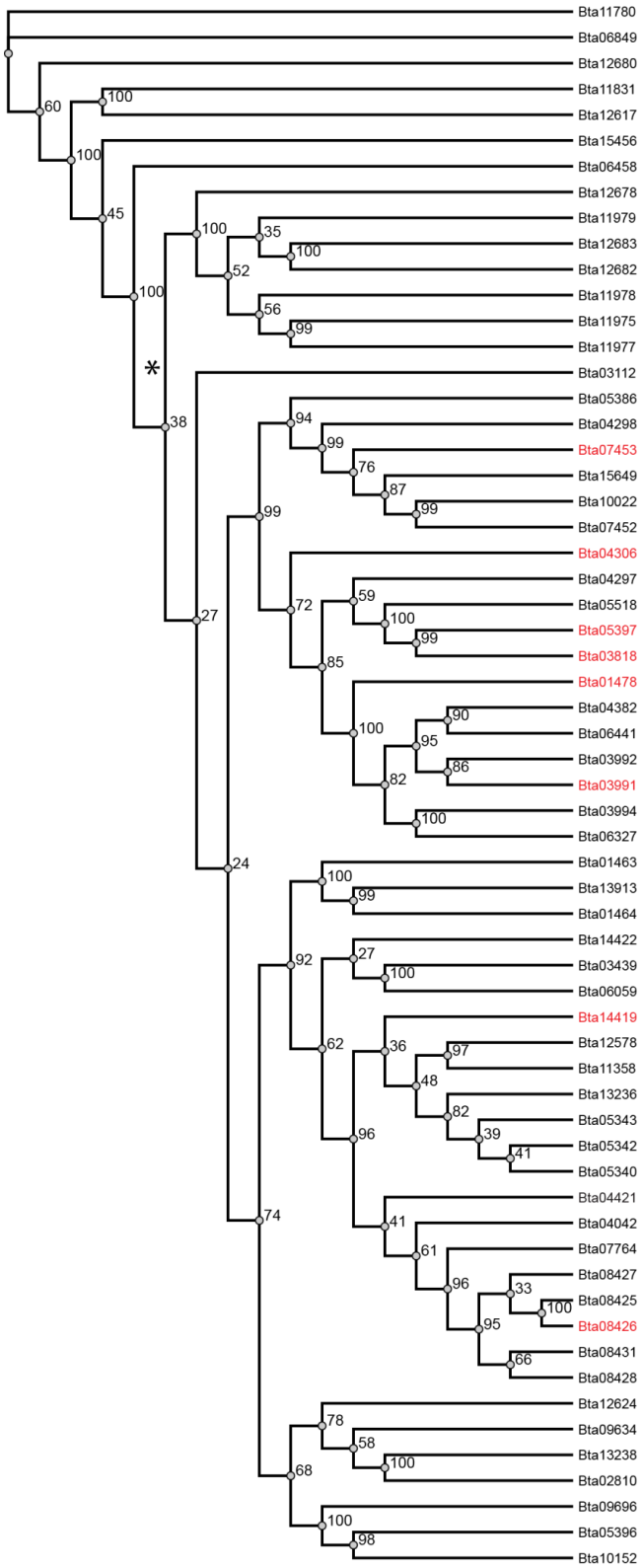
Whitefly Genome BLAST results for Trehalulose Synthase

Bta04306	307	2e-95
Bta03818	296	1e-91
Bta01478	298	2e-91
Bta07453	295	6e-91
Bta03991	292	2e-90

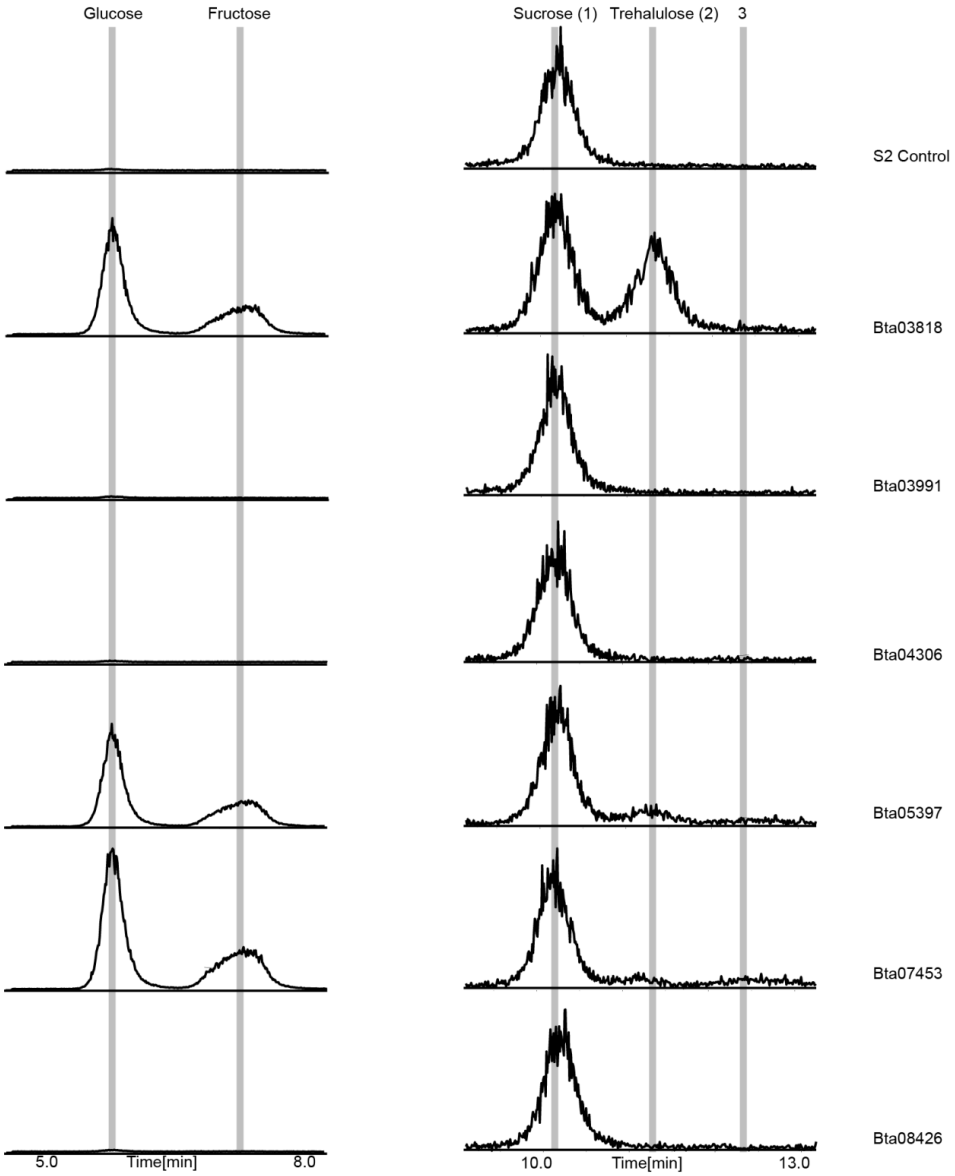
Whitefly Genome BLAST results for Isomaltulose Synthase

Bta01478	300	3e-92
Bta14419	292	1e-89
Bta03818	289	8e-89
Bta08426	288	1e-88
Bta05397	289	2e-88

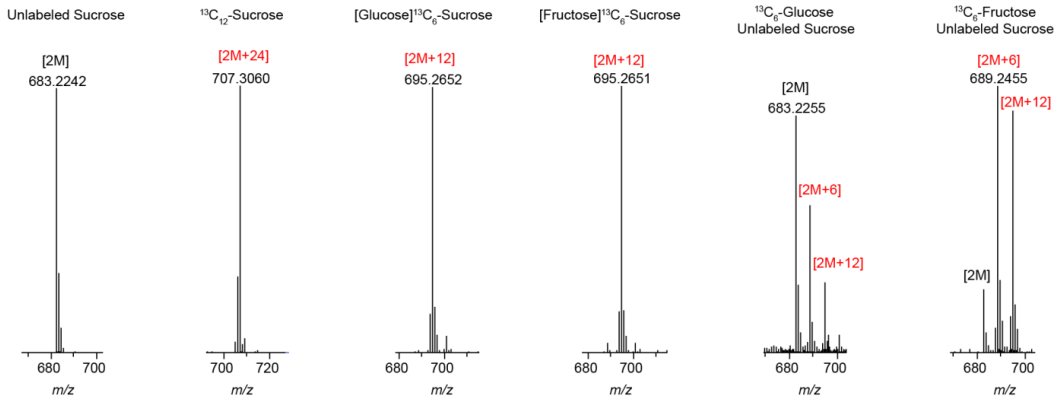
Supplemental Figure S1: BLAST search of B.tabaci genom using bacterial trehalulose and isomaltulose synthases. Protein sequences for a trehalulose synthase from *Pseudomonas mesoacidophila* and an isomlatulose synthase from *Erwinia rhepontica* were used in a blastp with default parameters on www.whiteflygenomics.org resulting in the above gene candidates.



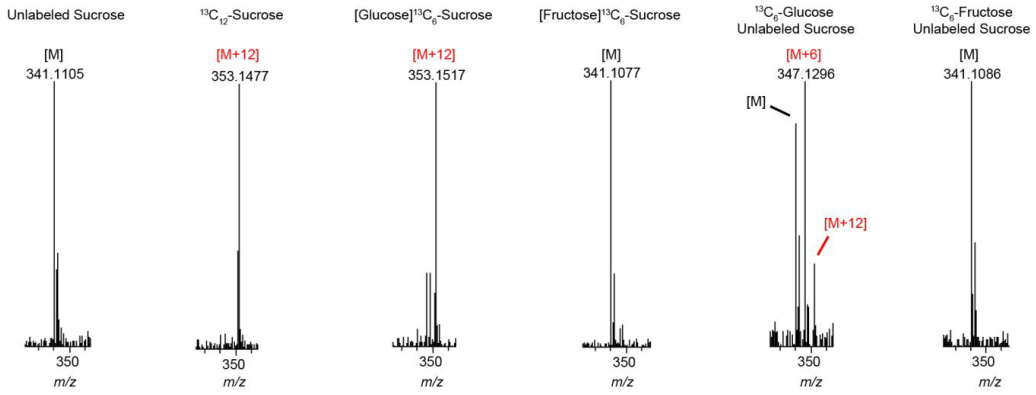
Supplemental Figure S2: Expanded phylogenetic analysis of GH genes in *B. tabaci*. Depicted is the full phylogenetic analysis from Figure 2 displaying all annotated genes in *B. tabaci*. Cloned candidates are highlighted in red.



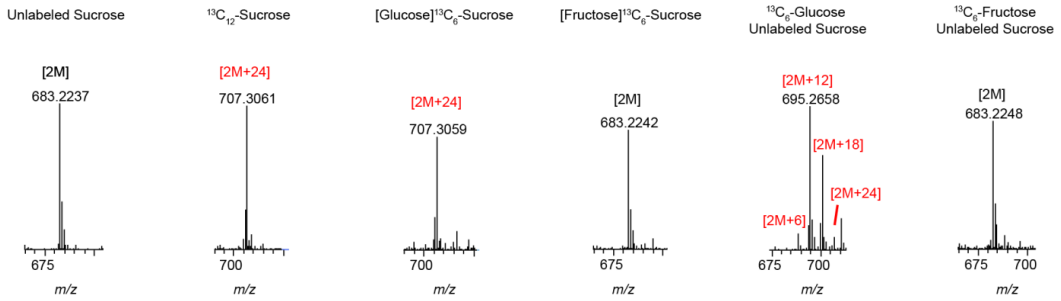
Supplemental Figure S3: Sucrose modifying activity of all cloned gene candidates. All cloned candidates were tested for activity in the modification of sucrose, being both hydrolysis to glucose and fructose as well as isomerization to trehalulose. The only candidates which showed any activity were Bta03818 (BtSUC1), Bta05397 (BtSUC6) and Bta07453 (BtSUC7).



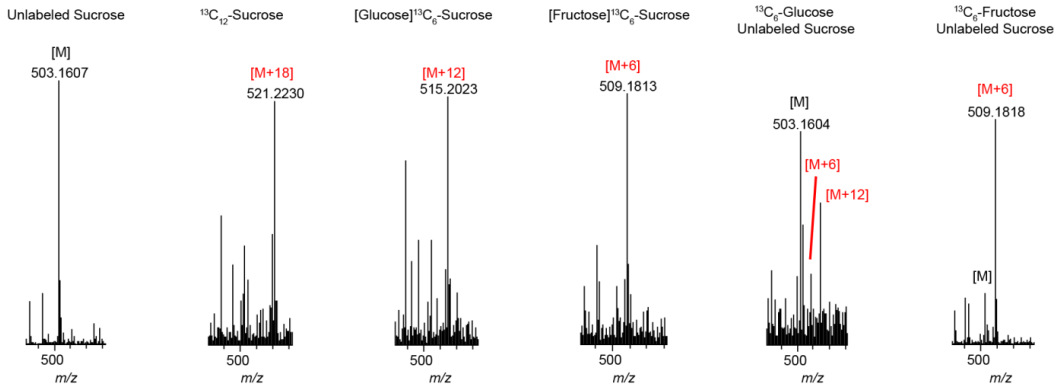
Supplemental Figure S4: MS patters for trehalulose (peak 2) BtSUC1 incubations with various sugar isotopomers. Depicted is the MS spectra for the dimer of trehalulose formed in incubations with sucrose and candidate enzymes BtSUC1,6 and 7. Fully label incorporation is seen with fully labelled sucrose, and half incorporation with glucose-labeled sucrose and fructose-labeled sucrose. Various label incorporation was seen with free labeled glucose and fructose, specifically showing the binding of free glucose to transglucosidate. Higher isopopic ratios were seen in free labeled fructose assays giving insight into the migration of free fructose into the catalytic pocket.



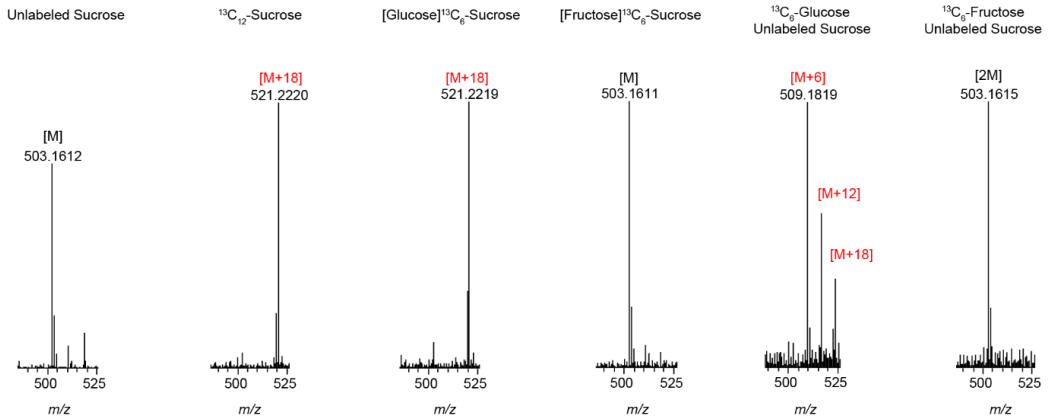
Supplemental Figure S5: MS patterns for the unknown disaccharide (peak 3) formed by BtSUC7 in incubations with sucrose and various sugar isotpomers. The unknown disaccharide formed by SUC7 is shown to be composed of only glucose units by label incorporation as seen by the absence of label C13 incorporation in fructose labeled sugar incubations.



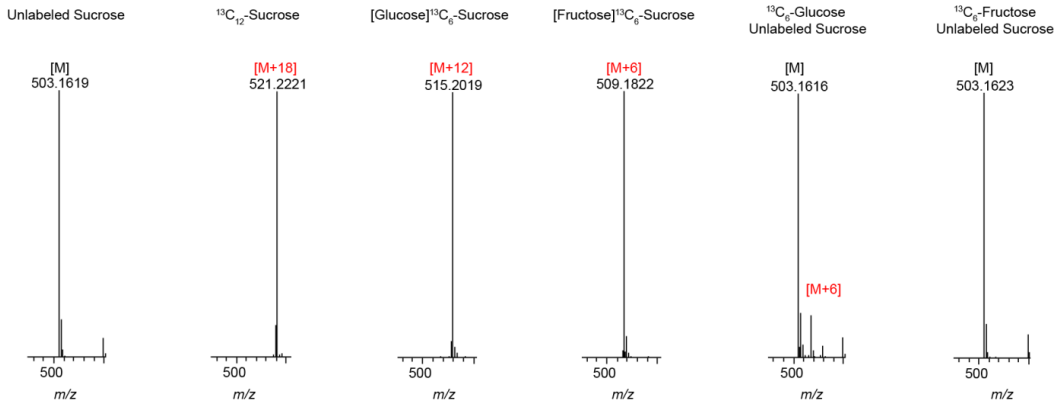
Supplemental Figure S6: MS patterns for isomaltose formed in BtSUC1 incubations with sucrose and various sugar isotopomers. Depicted is the MS pattern for the dimer of isomaltose showing no label incorporation in any assays with fructose labeled saccharides. Various levels of incorporation in labeled free glucose incubations shows the ability of this enzyme to bind free glucose in the catalytic pocket.



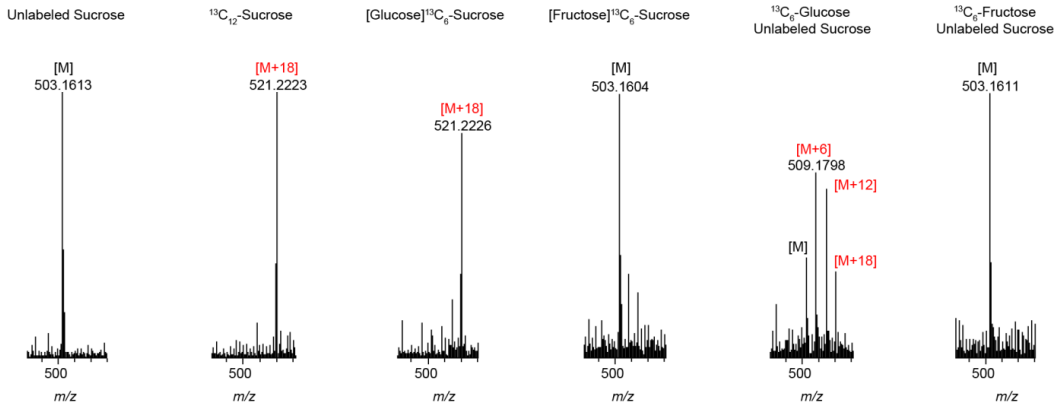
Supplemental Figure S7: MS patterns for an unknown trisaccharide (peak 6) formed by BtSUC1 in incubations with sucrose and various sugar isotopomers. Depicted is the monomer of peak 6 formed by BtSUC1 in sucrose incubations showing mechanistic label incorporations of +18 Da for fully labeled, +12 for glucose labeled sucrose and +6 for fructose labeled sucrose, showing the composition of the molecule being two glucose units and one fructose unit. This molecule can also be formed from the transglucosidation of free fructose.



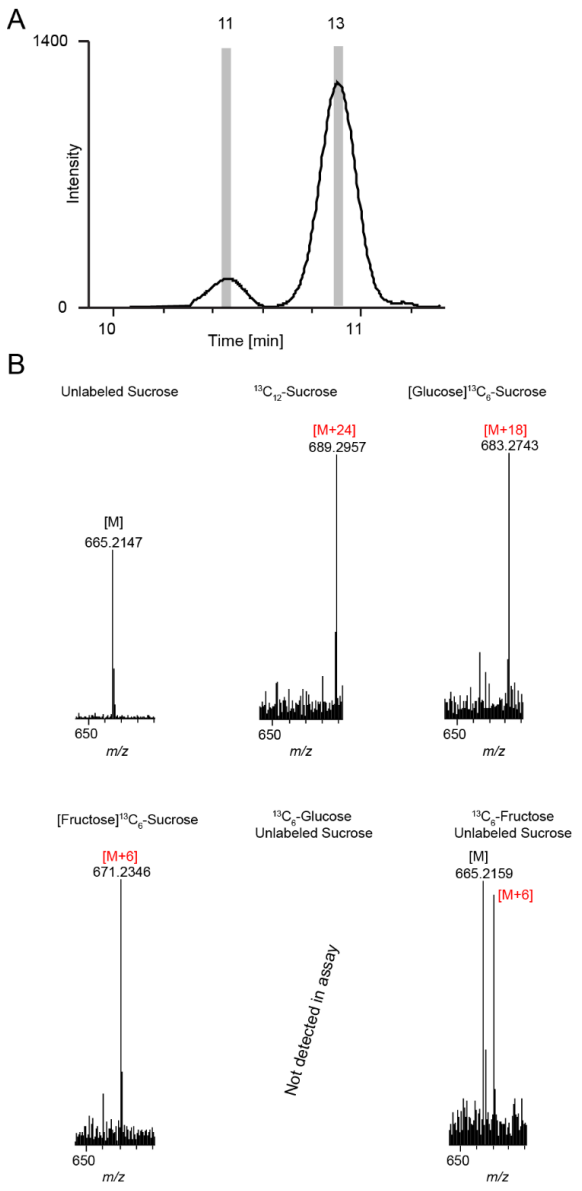
Supplemental Figure S8: MS patterns for isomaltotriose (peak 10) formed by BtSUC1 in incubations with sucrose and various sugar isotompomers. Depicted is the MS pattern for the monomer of isomaltotiose showing label incorporations patterns consistent with the molecule consisting of only glucose units. Once again BtSUC1 shows the ability to bind free glucose in this reaction.



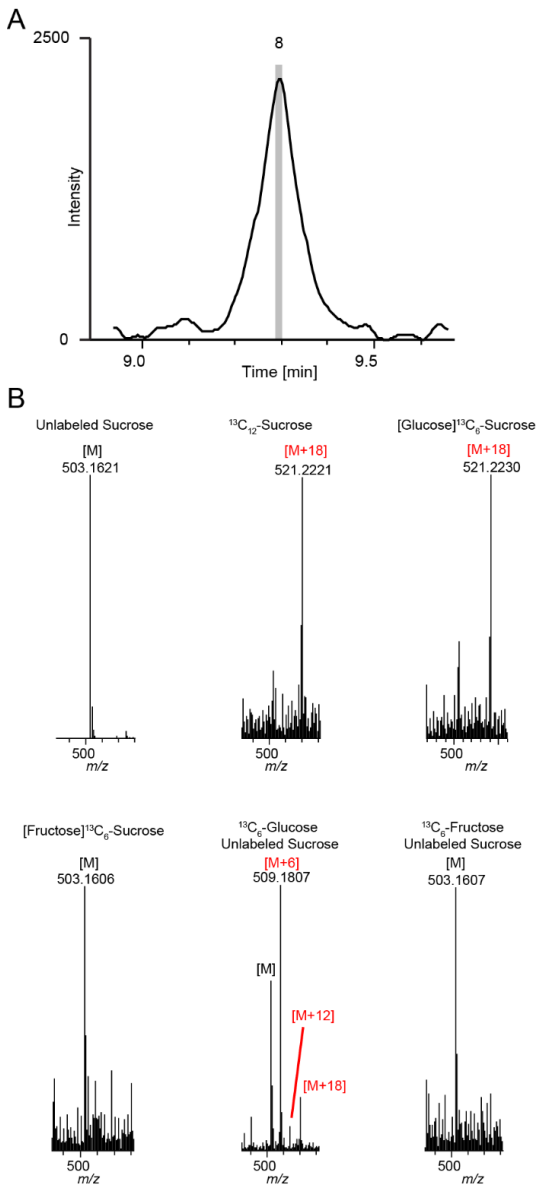
Supplemental Figure S9: MS patterns for an unknown trisaccharide (peak 5) formed by BtSUC6 in incubations with sucrose and various sugar isotopomers. Depicted is the MS pattern for the monomer of peak 5 formed by BtSUC6 showing a label incorporation pattern consistent with a molecule composed of two molecules of glucose and one fructose unit. By isotopic abundance, the binding efficiency of free glucose is notably lower for BtSUC6 than for BtSUC1.



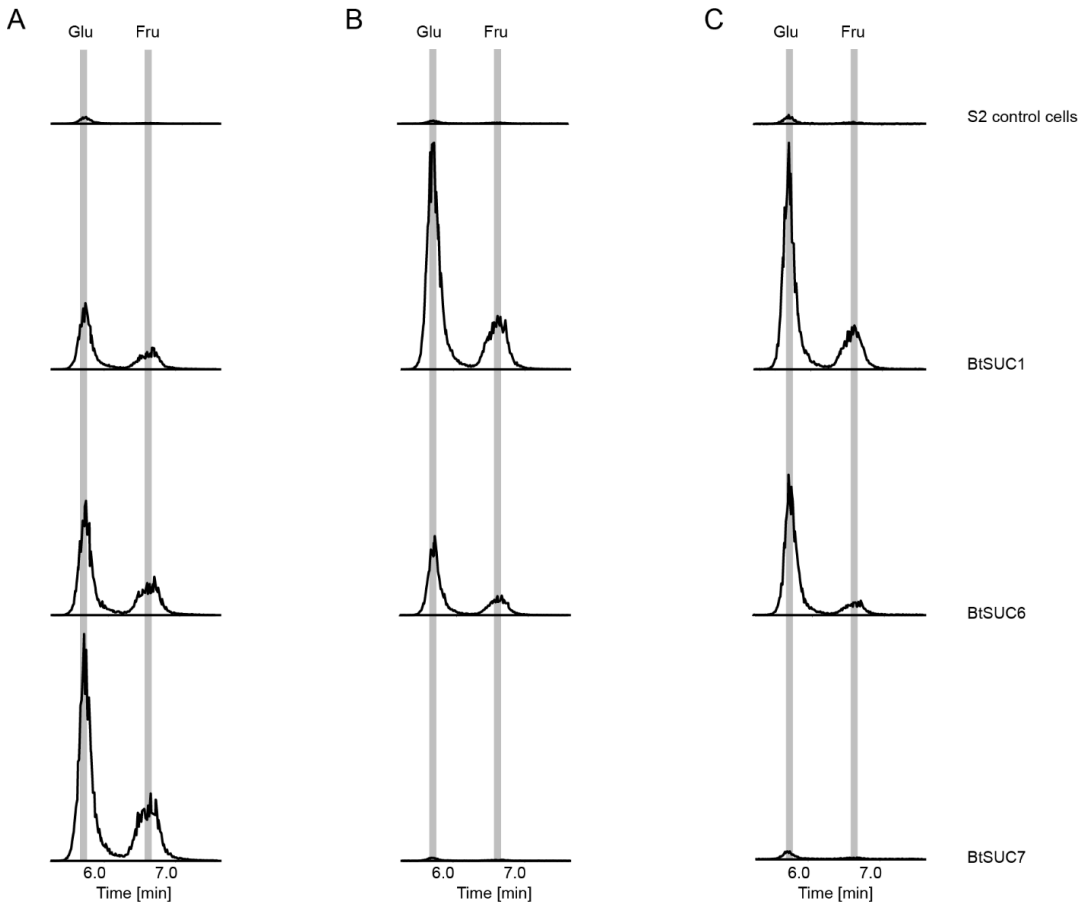
Supplemental Figure S10: MS patterns for an unknown trisaccharide (peak 9) formed by BtSUC6 in incubations with sucrose and various sugar isotopomers. Depicted is the MS pattern for the monomer of peak 9 formed by BtSUC6 showing a label incorporation pattern consistent with a molecule composed entirely glucose units.



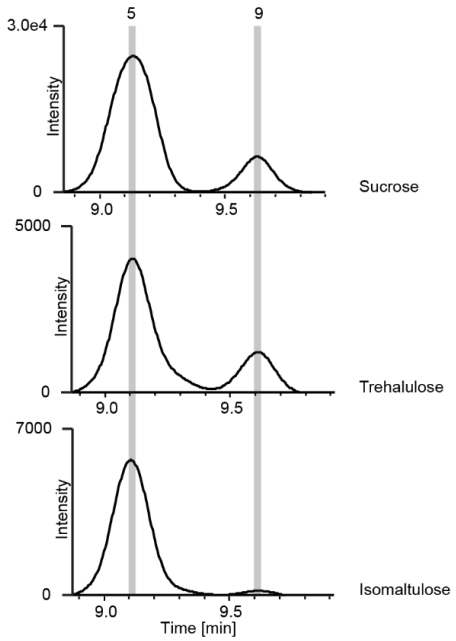
Supplemental Figure S11: Chromatograms of two tetrasacchrides (peaks 11 and 13) formed by BtSUC6 with MS patterns for peak 11following sucrose and sugar isotopomer incubations. (A) BtSUC6 was an efficient transglucosdase with sucrose producing peaks for unknown tetrasaccharides (peak 11 and 13). (B) MS patterns were only available for peak 11 due to low abundance peak 13 in labeled assays, showing a molecule composition of three units of glucose and 1 unit of fructose. Since it was not detected in free glucose assays, it is perhaps an inhibitor for the formation of this compound.



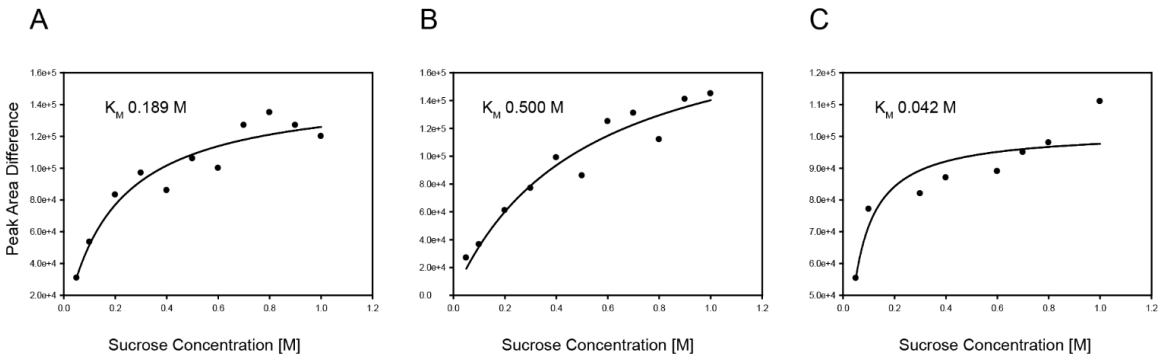
Supplemental Figure S12: Chromatograms and MS patterns for an unknown trisacchairide (peak 8) produced by BtSUC7 in incubations with various sucrose and sugar isotopomers. (A) Chromatogram of peak 8 produced by BtSUC7 in incubations with sucrose. (B) This compound has an MS pattern consistent with a compound composed of entirely glucose units.



Supplemental Figure S13: Hydrolysis activity of candidate Gh13 enzymes with various disaccharides. (A) Incubations with sucrose resulted in hydrolysis to glucose and fructose in all GH enzyme incubations; however BtSUC7 was unable to hydrolyze trehalulose (B) and isomlatulose (C).



Supplemental Figure S14: Incubations of BtSUC6 with sucrose, isomaltulose, and trehalulose resulted in the same transglucosidation products. Depicted are the chromatograms of incubations of BtSUC6 with sucrose, trehalulose and isomaltulose, showing the production of the same products, which is opposed to BtSUC1 which formed different trisaccharide products.



Supplemental Figure S15: K_M determination by sucrose utilization. Kinetic parameters were estimated for each enzyme based on the utilization of sucrose or difference between sucrose peak areas in control cells and cells expressing trehalulose synthases BtSUC1 **(A)**, BtSUC6 **(B)** and BtSUC7 **(C)**.

```
Trehalulose_synthase_ACO05018.1      ---------------------MLMKRLF-----AASLMLAFSSVSS---VRAEEAVKPGA      31
Isomaltulose_Synthase_HM461324.2     ----------MDSQGLKTAVAIFLATTF-----SATSYQACSAGPDTAPSLTVQQSNALP      45
BtSUC7                               ----------------MIILNNFIKELYFKNMYKFALYLAI-FGYATNLSEALALKNQAL      43
BtSUC1                               ------------------------------MKIAVLSFLV-FTTAWCREEMEEQPQRDL      28
BtSUC6                               MHQPEIWLNQSDVTDLKVLLTGA-TKLL---MKLVAVFCLV-VASAT-CEVMKAYPQRDL      54
                                                                                                :

Trehalulose_synthase_ACO05018.1      PWWKSAVFYQVYPRSFKDTNGDGIGDFKGLTEKLDYLKGLGIDAIWINPHYASPNTDNGY      91
Isomaltulose_Synthase_HM461324.2     TWWKQAVFYQVYPRSFKDTNGDGIGDLNGIIENLDYLKKLGIDAIWINPHYDSPNTDNGY     105
BtSUC7                               DWWQRGVFYQVYPRSFRDSNADGIGDLRGIAEKVWYLKDLGVTAVWLNPIFESPNLDLGY     103
BtSUC1                               AWWEKGVIYQIYPRSFKDSDGDGIGDLKGIAEKIDYLSKLGVGAVWISPIFRSPMADFGY      88
BtSUC6                               AWWERGVIYQIYPRSFKDSDGDGVGDLKGIAEKVDYLSNLGIKAVWLSPIFRSPMVDFGY     114
                                      **: .*:**:*****:*::.**:**:.*: *:: **. **: *:*:.* : **  * **

Trehalulose_synthase_ACO05018.1      DISDYREVMKEYGTMEDFDRLMAELKKRGMRLMVDVVINHSSDQHEWFKSSRASKDNPYR     151
Isomaltulose_Synthase_HM461324.2     DIRDYRKIMKEYGTMEDFDRLISEMKKRNMRLMIDIVINHTSDQHAWFVQSKSGKNNPYR     165
BtSUC7                               DISNYTKIDPVYGTLEDLEYLKNALHKAGLRLLLDFVPNHTSDQHPWFQKSIK-KIEPYT     162
BtSUC1                               DISDFRAIEPMFGTMKDFERLKRLFHKNGLKMILDFVPNHTSDEHDWFKKSVA-RVDPYT     147
BtSUC6                               DISDFRAIEPMFGTMEDFEVLKELFQKNGLKIILDFVPNHTSDQHDWFVRSVA-REDPYT     173
                                     ** ::  :   :**:.:** :*   ::* .:::::*.* **:**:* **  *   : :**

Trehalulose_synthase_ACO05018.1      DYYFWRDGKD-----GHEPNNYPSFFGGSAWEKDPVTGQYYLHYFGRQQPDLNWDTPKLR     206
Isomaltulose_Synthase_HM461324.2     DYYFWRDGKD-----GHAPNNYPSFFGGSAWEKDDKSGQYYLHYFAKQQPDLNWDNPKVR     220
BtSUC7                               NYYVWKDAKFNAKGERMPPTNWISEFGGYTWSWNEERQQYYLHLFHYKQPDLNYNSKELV     222
BtSUC1                               NYYNWVDGKASENGTIQPPNNWLSYFGGSAWTWNEKRGQYYLHQFHPKQPDLNYRNPLVV     207
BtSUC6                               NYYVWVDAKY-VNGTRSPPNNWLSDFGGSAWTWNEKRKQYFLHQFHHKQPDLNYRNPLVV     232
                                     :** * *.*         *.*: * *** :*  :    **:** *  :*****: .  :

Trehalulose_synthase_ACO05018.1      EELYAMLRFWLDKGVSGMRFDTVATYSKTPGF--------PDLTPEQMKNFAEAYTQGPN     258
Isomaltulose_Synthase_HM461324.2     QDLYDMLRFWLDKGVSGLRFDTVATYSKIPNF--------PDLSQQQLKNFAEEYTKGPK     272
BtSUC7                               KEMTDVIRFWLDRGIDGLRMDAVTFLYEDPQYRDQELL-SPDLDPWDFKSYNRCHTMDQP     281
BtSUC1                               QEMKDVLTYWMDKGVDGFRMDAVMTIMEDIKLRDEPLSGKTDVLPTDEEYLNHIYTRDQP     267
BtSUC6                               QEMKDVLTYWMDKGVDGFRMDAVMTIIEDDQLRDEPPSNKPDVLPTDVEFLTHIHTRDQP     292
                                     :::  :: :*:*:*:.*:*:*:*    :             *:   : :   . :* .

Trehalulose_synthase_ACO05018.1      LHRYLQEMHEKVFDHYDAVTAGEIFGA-----PLNQVPLFI-DSRRKELDMAFTFDLIRY     312
Isomaltulose_Synthase_HM461324.2     IHDYVNEMNREVLSHYDIATAGEIFGV-----PLDKSIKFF-DRRRNELNIAFTFDLIRL     326
BtSUC7                               ATYRLITKFREIFDAYSKKDGETKVMITEAYTSLDKTMEYYKFGDKPGAHMPFNFLFITS     341
BtSUC1                               GTYEIIKQFRKHLDDYSRKTHTVKFMATEAYSNLTSTMKYYGTKDNPGAHFTFNFETIEA     327
BtSUC6                               GTYDIIKQFRKLLDDYSAKTGTVKFMATEAYSNLTNTMKYYGDTYYPGAHFTFNFGPITG     352
                                        :   .: :. *.       .        * .   :         .: *.*  *

Trehalulose_synthase_ACO05018.1      DRAL-DRWHTIPRTLADFRQTIDKVDAIAGEYGWNTFFLGNHDNPRAVSHFGDDRPQWRE     371
Isomaltulose_Synthase_HM461324.2     DRDADERWRRKDWTLSQFRKIVDKVDQTAGEYGWNAFFLDNHDNPRAVSHFGDDRPQWRE     386
BtSUC7                               LDG--------RSPAKEYQRIIESWMSGMPEGQWPNWVIGNHDNHRLASRYGTDLV----     389
BtSUC1                               LTP--------QSNAPDFKEVIEDWYAALPASKWSNWVLGNHDKRRVGSRYGIDML----     375
BtSUC6                               LSA--------QSNAKKFKEVIENWYRYLPEGKWSNWVLGNHDQQRVASRYGIDML----     400
                                                     .::. ::.         *  :.:.***: *  *::* *

Trehalulose_synthase_ACO05018.1      ASAKALATVTLTQRGTPFIFQGDELGMTNYPFKTLQDFDDIEVKGFFQDYVETGKATAEE     431
Isomaltulose_Synthase_HM461324.2     HAAKALATLTLTQRATPFIYQGSELGMTNYPFKKIDDFDDVEVKGFWQDYVETGKVKAEE     446
BtSUC7                               ---DGLNMIGLLLPGTGNTYYGDEIGMEDTNVRFDEGVDPQ-------ACEFG------     432
BtSUC1                               ---DGLHMLQMCLHGTSVTYAGDEIGMVDTFIRWDQTKDPP--------ALNVG------     418
BtSUC6                               ---DGLHMLQMCLHGTSVSYSGDEIGMTGPFIRWDQTKDPP--------ALNVG------     443
                                        ..*  : :   .*   : *.*:** .  .:  :  *           : *

Trehalulose_synthase_ACO05018.1      LLTNVALTSRDNARTPFQWDDSANAGFTT-GKPWLKVNPNYTEINAAREI-GDPKSVYSF     489
Isomaltulose_Synthase_HM461324.2     FLQNVRQTSRDNSRTPFQWDASKNAGFTS-GTPWLKINPNYKEINSADQI-NNPNSVFNY     504
BtSUC7                               -KEKYRLVTRDPERTPFQWDTTLNSGFSASLKTWLPVNSNFWRLNLKAQVQSEGNSHYKV     491
BtSUC1                               -PERYQRFTRDPARTPFQWNASTSAGFSTNPKTWLPVNPDYWSHNLVTEK-KKASSHFKN     476
BtSUC6                               -PERYLRFTRDPARTPFQWASCTSAGFSTNPKTWLPVNPNYWSHNLVTEK-KKNRSHLKN     501
                                        .    :**  ******    .:**::  . ** :* ::    *   :   .   *  .

Trehalulose_synthase_ACO05018.1      YRNLISIRHETPALSTGSYRDIDPSNADVYAYTRS-QDGETYLVVVNFKAEPRSFT----     544
Isomaltulose_Synthase_HM461324.2     YRKLINIRHDIPALTYGSYIDLDPDNNSVYAYTRT-LGAEKYLVVINFKEEVMHYT----     559
BtSUC7                               YKRLIDVRK-TDTMLYGALETHVLSK-WVFSFARRQNGSDTYVVVVNLGSETAPVDLSAF     549
BtSUC1                               YRRLLTLKK-SPVIQFGSVNVYTLSD-WVLVITRTLKDHPTYIVILNLGTELEDTKGLRK     534
BtSUC6                               YQKLLALKE-SPVIQFGNLNVYTLSD-WVLVVTRALEGHPTYVVVLNIGSEIEYTTKLSA     559
                                     *:.*: ::.   .:  *       ..  *   :*   .  .*:*.:*:  *

Trehalulose_synthase_ACO05018.1      ---LPDGMHIAETLIESSSPAAPAAGAASLELQPWQSGIYKVK----------------     584
Isomaltulose_Synthase_HM461324.2     ---LPGDLSINKVITENNSHTIVNKNDRQLRLEPWQSGIYKLN----P------------     600
BtSUC7                               MNDIPDTLTVHTSSINSQHQPGDKVAINEFMMRPKSSLLLTTASEVPPPSYKTITSSSSK     609
BtSUC1                               IANLPDQIRLHTCSINCGYSPGAQLRTDEIQLRPKAGFVCIARGGLKTS-----------     583
BtSUC6                               VANLPEQLKLHICSLNCGYTPGVQLHTEKIQLRPKAGMVLSTQKGVQASREADES--SSE     617
                                        :*  : :    :            .: :.*  . :

Trehalulose_synthase_ACO05018.1      ------------------       584
Isomaltulose_Synthase_HM461324.2     ------------------       600
BtSUC7                               LNVPIGTLILFLFCKYLV       627
BtSUC1                               ------------------       583
BtSUC6                               LQME--------------       621
```

Supplemental Figure S15: Multiple sequence alignment of *B. tabaci* and bacterial proteins. Default MUSCLE parameters in BioEdit produced the above alignement showing divergence in sequence from *B. tabaci* and query bacterial enzymes. The "RLDRD" motif is highlighted (black box) in the isomaltulose synthase, however no identical sequence exists in any *B. tabaci* enzymes. A sequence of similar identity exists in the trehalulose synthase (red box) which is closer to an upstream sequence in BtSUC1 and 6 and 7.

Supplemental Table S1: List of analyzed sugar standards as well as peaks of unknown identity formed in enzyme assays showing the mobility of diagnostic ions for di-, tri-, and tetrasaccharides

Metabolite	Ret. Time	Diagnostic Ion 1 Mobility [1/K_O] [V*s/cm^2] (CCS)(Å^2)	Diagnostic Ion 2 Mobility [1/K_O] [V*s/cm^2] (CCS)(Å^2)	Diagnostic Ion 3 Mobility [1/K_O] [V*s/cm^2] (CCS)(Å^2)	Diagnostic Ion 4 Mobility [1/K_O] [V*s/cm^2] (CCS)(Å^2)
Disaccharides		**m/z 683**	**m/z 705**	**m/z 721**	**m/z 773**
Trehalulose (2)	7.5 min	1.136 (232.1)	1.137 (232.1)	1.157 (236.1) & 1.179 (240.6)	1.166 (237.8) & 1.204 (245.4)
Sucrose (1)	7.2 min	1.115 (227.8)	1.136 (231.9) & 1.163 (237.5)	1.125 (229.6)	1.139 (232.4) & 1.179 (240.4)
Isomaltulose	7.9 min	1.162 (237.4)	1.114 (227.6) & 1.154 (235.7) & 1.223 (249.7)	1.124 (229.5)	1.186 (241.1) & 1.211 (246.9)
Isomaltose (4)	8.2 min	1.131 (321.0)	1.131 (231.0) & 1.157 (236.2)	1.144 (233.6) & 1.181 (241.1) & 1.197 (244.3)	1.163 (237.0) & 1.194 (243.3) & 1.237 (252.1)
Lecurose	7.4 min	1.144 (233.7)	1.160 (236.8) & 1.201 (245.2)	1.089 (222.2) & 1.155 (235.8) & 1.187 (242.3) & 1.265 (258.1)	1.162 (236.8) & 1.194 (243.3)
Maltulose	7.5 min	1.158 (236.6)	1.156 (236.1) & 1.196 (244.3)	1.155 (235.8)	1.195 (243.6)
Turanose	7.3 min	1.122 (229.1)	1.148 (234.5)	1.164 (237.5)	1.144 (233.2) & 1.170 (238.6) & 1.186 (241.7) & 1.213 (247.3)
Trehalose	7.9 min	1.126 (230.1)	1.159 (236.7)	1.142 (233.0)	1.154 (235.3)
Cellobiose	7.8 min	1.126 (230.0)	1.097 (224.0) & 1.132 (231.0) & 1.185 (242.0)	1.111 (226.9)	1.143 (233.1) & 1.164 (237.4)
Melibiose	8.4 min	1.146 (234.1)	1.134 (231.6) & 1.164 (237.6) & 1.186 (242.2) & 1.211 (247.3)	1.148 (234.3) & 1.177 (240.3) & 1.195 (243.9) & 1.225 (250.1)	1.190 (242.5) & 1.235 (251.7) & 1.269 (258.6)
Gentiobiose	8.4 min	1.163 (237.7)	1.214 (247.9) & 1.242 (253.5)	1.197 (244.3) & 1.245 (254.0)	1.230 (250.8) & 1.295 (263.9)
Peak 3	7.8 min	1.128 (230.4) & 1.148 (234.5) & 1.162 (237.3) & 1.177 (240.5)	1.154 (235.5)	1.142 (233.1) & 1.174 (239.5)	N/A

Trisaccharides		**m/z 503**	**m/z 571**	**m/z 685**	**N/A**
Raffinose	9.1 min	0.963 (198.2)	1.001 (205.2) & 1.036 (212.4)	1.081 (220.9)	N/A
Kestose	8.8 min	0.984 (202.4)	1.017 (208.5) & 1.043 (213.9)	1.090 (222.8)	N/A
Isomaltotriose (10)	9.9 min	0.980 (201.5) & 1.024 (210.6)	1.012 (207.5) & 1.043 (213.9) & 1.057 (216.8)	1.094 (223.4) & 1.107 (226.1)	N/A
Panose	9.6 min	0.965 (198.6) & 1.017 (209.2)	1.005 (206.1) & 1.032 (211.7) & 1.066 (218.7)	1.078 (220.3) & 1.100 (224.8) & 1.129 (230.6)	N/A
Peak 5	9.1 min	0.998 (205.4)	1.034 (212.0)	1.103 (225.4)	N/A
Peak 6	9.3 min	0.972 (200.0) & 1.014 (208.7)	1.012 (207.5) & 1.053 (216.0)	1.064 (217.4) & 1.092 (223.1)	N/A
Peak 7	9.3 min	0.978 (201.2) & 1.041 (214.2) & 1.076 (221.3)& 1.126 (231.7)	1.013 (207.7) & 1.052 (215.8)	1.097 (224.0) & 1.139 (232.7)	N/A
Peak 8	9.3 min	0.968 (199.2) & 1.033 (212.5)	0.977 (204.5) & 1.013 (207.8) & 1.045 (214.4) & 1.087 (223.0)	1.060 (216.6) & 1.128 (230.4)	N/A
Peak 9	9.6 min	0.969 (199.3) & 1.033 (212.4)	1.028 (210.9) & 1.066 (218.6)	1.109 (226.5)	N/A
Tetrasaccharides		**m/z 665**	**m/z 711**	**m/z 733**	**m/z 779**
Peak 11	10.5 min	1.084 (221.7) & 1.135 (232.1) & 1.160 (237.2)	1.104 (225.4) & 1.122 (229.0) & 1.140 (232.8) & 1.161 (237.0)	1.170 (238.7)	1.153 (235.0) & 1.167 (237.9)
Peak 12	10.6 min	1.090 (222.8) & 1.123 (229.7) & 1.143 (233.6)	1.114 (227.5) & 1.132 (231.2)	1.154 (235.5) & 1.176 (240.0)	1.185 (241.5)
Peak 13	10.9 min	1.114 (227.7) & 1.170 (239.2)	1.152 (235.3) & 1.169 (238.7)	1.165 (237.8)	1.197 (244.0)

Supplementary Table S2: Primers used for cloning of GH13 genes from *B. tabaci*

Primer name	Sequence
Bta03818 Genome F Kozak EcoRI	GGCCGGAATTCACCATGAAAATAGCAGTGCTTTCATTTCTC
Bta03818 Genome R XbaI	GTACGTCTAGAAGAGGTTTTTAGGCCACCCC
Bta03991 Genome F Kozak NotI	GCTCAGCGGCCGCACCATGGCAAGCATAAGATATCCCATAAT
Bta03991 Genome R XhoI	TAGCACTCGAGCGAGGGAATCTTTTGGTTCGTC
Bta04306 Genome F Kozak EcoRI	GCGGTGAATTCACCATGAAGCTTATTTTTGTGGCAGCC
Bta04306 Genome R XhoI	ATAATCTCGAGTAAGGGCCAGCGCGCCTG
Bta05397 Genome F Kozak XhoI	TGCATCTCGAGACCATGCATCAGCCAGAAATATGGTTG
Bta05397 Genome R XbaI	GCTGGTCTAGATTCCATTTGAAGTTCTGAACTTGACTC
Bta07453 Genome F Kozak KpnI	GGCCGCGGGTACCACCATGATTATATTAAATAATTTCATA AAGGAACTTTATTTCAAAAAC
Bta07453 Genome R XhoI	GCGGCGCTCGAGTACTAAGTATTTACAAAAGAGGAATAA AATGAGAGT
Bta08426 Genome F Kozak EcoRI	GCGGCGAATTCACCATGACACAAGTTTTTGGTTTTTTATT ATTCTC
Bta08426 Genome R ApaI	GTACTGGGCCCCCCGATATTGACTATGGAATTGGC

Supplemental Table S3: Composition of assays used for enzyme mechanistic analysis

Assay number	Isotopically-labeled compound (concentration)	Unlabeled constituents (concentration)
1	None	Sucrose (1 M)
2	[$^{13}C_{12}$] sucrose (1 M)	None
3	[glucose]$^{13}C_6$-sucrose (1 M)	None
4	[fructose]$^{13}C_6$-sucrose (1 M)	None
5	[$^{13}C_6$]-glucose (0.5 M)	Sucrose (0.5 M)
6	[$^{13}C_6$]-fructose (0.5 M)	Sucrose (0.5 M)

Supplemental Table S4: List of multiple reaction monitoring (MRM) parameters for individual compounds analyzed by LC-MS.

Compound	Quadrupole 1 (precursor mass Da)	Quadrupole 2 (product mass Da)	Declustering potential (V)	Entrance potential (V)	Collision energy (V)	Collision cell exit potential (V)
Monosaccharide	178.8	89	-50	-9.5	-10	0
Disaccharide	340.9	59	-65	-10	-46	0
Trisaccahride	503.1	179	-95	-10	-28	-4
Tetrasaccharide	665.2	179	-100	-10	-48	-4

Discussion

General

The manuscripts pertaining to this thesis sought to investigate several major metabolic pathways of *B. tabaci* involved in the digestion of compounds from the phloem of their host plants and characterized the genes and enzymes involved. Due to the wide host range of *B. tabaci*, it serves as an excellent model for the exploration of metabolism important for primary functions as well as processing host plant-specific compounds. Highlighted throughout this thesis is the metabolism of plant defenses known as two-component defenses, which were previously thought not to be effective against phloem feeding insects. Not only do we provide evidence for the activation of these defenses by whiteflies (**Manuscripts I & II**), but we also describe various universal and explicit forms of detoxification in *B. tabaci* for glucosinolates (**Manuscript I**) and cyanogenic glycosides (**Manuscript II**), both of which are well-known examples of two-component defenses. The catalysts are transglucosidases that are members of Glucohydrolase Family 13 (GH13). We also characterized the previously elusive whitefly transglucosidase enzymes involved in the iso-energetic conversion of dietary sugars to larger molecules for osmoregulatory purposes (**Manuscript III**). Thus, transglucosidase-mediated metabolism in *B. tabaci* serves multiple functions.

I. Two-component defense activation in the phloem feeder *B. tabaci*

Due to the inconspicuous mode of feeding that phloem feeders employ, two-component defenses are thought to be ineffective (Walling, 2008), especially where tissue disruption is a prerequisite for defense activation. Therefore, in **Manuscript I & II**, we were surprised to report evidence to the contrary, showing evidence for the hydrolysis and activation of two-component defenses, including glucosinolates and cyanogenic glycosides respectively. In **Manuscript I**, we observed that honeydew contained minute amounts (only visible following concentration of *B. tabaci* samples) of known hydrolysis products of 4msob-GSL, including toxic isothiocyanate (ITC) and nitrile (CN) derivatives. Since hydrolysis of these metabolites may have taken place after excretion from the whitefly due to degradation by high temperature or other abiotic factors, it was therefore very exciting to also detect known mercapturic acid pathway detoxification metabolites of 4msob-GSL

in the honeydew of *B. tabaci*, which must have arisen from enzymatic processes. Identification of these isothiocyanate conjugates indicated that hydrolysis within the insect was followed by conjugation to the tripeptide GSH, which preceded excretion. Similar mercapturic acid pathway metabolites were also seen in the honeydew of aphids feeding on plants containing indolic glucosinolates. However, indolic glucosinolates are much less stable than aliphatic glucosinolates such as 4msob-GSL, being subject to spontaneous degradation (Kim et al., 2008). Taken together, the observation of glucosinolate hydrolysis products and their detoxified derivatives in the honeydew of *B. tabaci* implies that a mixing of glucosinolate and myrosinase took place during insect feeding.

Similarly, in the case of cyanogenic glycosides from cassava in **Manuscript II**, we observed a significant accumulation of the cyanide detoxification metabolite beta-cyanoalanine in the bodies of whiteflies feeding on cassava in comparison to those that fed on eggplant. In contrast, beta-cyanoalanine accumulation in plants was similar in the cassava and eggplant leaves analyzed. Here again implying the hydrolysis of a two-component defense, this time cyanogenic glycosides, occurred during phloem feeding. Similarly, the expression of the cyanide detoxification genes rhodanese and beta-cyanoalanine synthase increased when whiteflies fed on cyanogenic cassava plants versus non-cyanogenic sweet potato plants (Antony et al., 2006).

During phloem feeding, hydrolysis of two-component defenses can occur at multiple stages, including initial probing, navigation of the stylet through the tissue, pre-ingestion due to tissue disruption, and post-ingestion. During probing, when a phloem feeder monitors host plant suitability for consumption, the sampling of multiple cells is undertaken (Esch & Tjallingii, 1990). The glycosides and activating enzymes may then mix in the stylet and gut of the insect depending on the sequence of cells sampled. This is especially true for glucosinolate-containing plants where the two components may be in adjacent “S” and “M” cells (Nintemann et al., 2018). In addition, navigation of the stylet towards the phloem often involves sampling of individual cells along the apoplastic path, perhaps all of them (Tjallingii & Esch, 1993). Therefore, it is possible that cell content may leak into the path of the stylet and ultimately result in the assembly of the two-components in the plant defense system. Finally, hydrolysis may be the result of non-specific glucosidases in the insect gut that are normally involved in other processes such as starch or saccharide hydrolysis. The promiscuity of α and β-glucosidases from insects has been

noted (Ferreira, Torres, & Terra, 1998), and therefore may be responsible for the hydrolysis of plant defenses glucosides and their resulting activation.

II. Metabolism and detoxification of glucosinolates in *B. tabaci*

In **Manuscript I**, we report the identification of trans-glucosidases as a means of detoxification, representing an entirely new class of enzymes in phase II conjugation detoxification. The utilization of sugars as a conjugation mechanism is well characterized and known to be a major feature of phase II detoxification. However, the glucosylating enzymes previously described in this context are the UDPGTs, which are dependent on an activated form of glucose or in the case of mammals, glucuronidate for activity. The process of glucose activation is energy intensive, requiring the input of chemical energy in the form of adenosine triphosphate (ATP) for the production of the sugar cofactor. On the other hand, transglucosidation is an iso-energetic process requiring no additional input of energy, but instead using the chemical energy of donor sugar binding for glucose transfer (Ünligil & Rini, 2000). In **Manuscript I**, we explore transglucosidation as a novel form of pre-emptive glucosinolate detoxification as well as the well-known avenues for detoxification of glucosinolate hydrolysis products.

II.I Pre-emptive detoxification of glucosinolates

The novel transglucosidation of 4msob-GSL in **Manuscript I** can be considered a pre-emptive detoxification, by which the compound is modified so that it is no longer a substrate for the activating enzyme (in this case myrosinase). So too do we observe that the insect derived glycosides of 4msob-GSL are stable to enzymatic activation by myrosinase, whereas native unmodified 4msob-GSL is readily hydrolyzed. It is important to note that the linkage mode of the newly added sugar is α (both 1→6 and 1→4 linkage) instead of the native β-linked glucose in plant glycosides, which may play an important role in making these compounds inert to myrosinase. However, the simple fact that two sugars are present instead of one may also block the activating thioglucosidase activity. Interestingly, though the single transfer of one glucose moiety to 4msob-GSL is sufficient for stability towards myrosinase, it is observed that 4msob-GSL is polyglucosylated by *B. tabaci*. The enzymes

that perform this modification were shown through labeling studies to be sucrase-transglucosidases, where the transferred glucose residue originates from sucrose itself. Tentative sucrase-transglucosidases of the GH13 family of glycoside hydrolases have been shown to have higher expression in *B. tabaci* guts in comparison to the whole body (X. W. Wang et al., 2012), where they are hypothesized to have utility in the osmoregulation of sugars. Indeed, this pattern of poly-glucosylation has been previously reported in *B. tabaci* with regards to osmoregulation and the modification of saccharides ingested in the phloem, but it was never hypothesized that these enzymes may also utilize other molecules as acceptor substrates for detoxification (further discussed in section V of discussion). Cloning of the candidate GH13 enzymes BtSUC2 and 5 resulted in the transglucosidation of 4msob-GSL using sucrose as a sugar donor, with BtSUC5 being the more efficient of the two enzymes. Both of these enzymes were only capable of transglucosidating 4msob-GSL in an α-(1→4) fashion (although the α-(1→6) is most abundant in the honeydew) and did not produce higher order additions to the glucosinolate. Furthermore, these recombinant proteins functioned as poor sucrose hydrolases and were also not capable of producing detectable transglucosidation products of sucrose alone, thus suggesting that they do not participate in osmoregulation.

Interestingly, the transglucosidation of glucosinolates was not the first pre-emptive detoxification of glucosinolates reported in *B. tabaci*. It was found that whiteflies feeding on glucosinolate containing plants also excrete desulphoglucosinolates (Malka et al., 2016), a well-known detoxification product (Ratzka et al., 2002). These two pathways of detoxification seem to be dedicated to different classes of metabolites. In the honeydew, the ratio of intact 4msob-GSL : desulpho 4msob-GSL : glycosylated 4msob-GSL is approximately 1: 2 : 14, marking glucosylation as the dominant metabolic process. However, for certain glucosinolates such as indole-derived compounds; the majority of metabolism in *B. tabaci* is by means of desulphation, and only minute levels of glycosides for these metabolites were detected. It is perhaps of further importance to recall that indolic glucosinolates were previously shown to be a deterrent to aphid feeding (Kim et al., 2008), and are even shown to accumulate following aphid infestation of *Arabidopsis thaliana* (Kim & Jander, 2007). Thus, desulphation of indolic glucosinolates in whiteflies may be a more effective detoxification mechanism than poly-glucosylation.

II.II Post-hydrolysis detoxification of glucosinolates

In **Manuscript I** we additionally observed that despite these pre-emptive forms of detoxification, some glucosinolate hydrolysis takes place during phloem feeding, and the hydrolysis products are detoxified via glutathione conjugation and further metabolized via the mercapturic acid pathway. These detoxification products, which include the glutathione conjugate of 4msob-ITC and the N-acetyl cysteine product of the conjugate, were only able to be detected after concentration of honeydew samples, perhaps showing the efficiency of the pre-emptive detoxification processes. The detection of this mode of detoxification was not unexpected as indolic glucosinolate hydrolysis metabolites are detoxified in an identical manner in aphids (Kim et al., 2008).

III. Metabolism and detoxification of cyanogenic glycosides and other two-component defenses in *B. tabaci*

Following the discovery of a mechanism for pre-emptive glucosinolate detoxification in **Manuscript I** and the detection of similar peaks in the analysis of honeydew from whiteflies feeding on the cyanogenic plant cassava, the metabolism of cyanogenic glycosides was further investigated. In **Manuscript II** we were able to confirm that indeed, the related cyanogenic glycoside two-component defense linamarin is pre-emptively detoxified in a similar manner, with an additional and unexpected pathway for detoxification via the addition of phosphate. Furthermore, we also discovered evidence for hydrolysis of cyanogenic glycosides via the accumulation of cyanide detoxification products. Finally, in **Manuscript II** we also touch on the substrate breadth of characterized enzymes for glucosylation, and their role in other plant-*Bemisia tabaci* interactions.

III.I Pre-emptive detoxification of cyanogenic glycosides

III.I.I Transglucosidation

Following the observations that the honeydew of whiteflies feeding on cassava contains similar glycosides to those feeding on glucosinolate-containing plants, the investigation of cyanogenic glycoside metabolism in *B.tabaci* was pursued to see if the

patterns of glucose addition were similar to those already observed with glucosinolates. Indeed, the glucosylation of cyanogenic glycosides occurs in *B. tabaci*, with the same enzymatic mechanism as for glucosinolates as determined by labeling studies. These glycosides are also inert to the activating enzyme linamarase, and thus represent the product of a pre-emptive detoxification reaction. Interestingly, the glucosylated cyanogenic glycosides are also stable in the presence of a disaccharidase known as linustatinase. This disaccharidase is known to be capable of hydrolyzing the diglucoside of linamarin, commonly referred to as linustatin, which is a diglucoside with a β-(1→6) linkage (D. Selmar, R. Lieberei, & B. Biehl, 1988). Importantly, one of the insect diglucosides also has a 1→6 linkage to linamarin, although this time being α, allowing us to say with further confidence that one of the major factors for stability of these glycosides to breakdown by cyanogenic glycoside hydrolases is the incorporation of an α–linkage with the regiochemical differences being of less importance.

In the metabolism of the cyanogenic glycoside linamarin, *B. tabaci* is capable of adding both an α-(1→6) and α-(1→4) glucose to the originally β-linked glucose of the molecule, in a similar fashion as to 4msob-GLS. However, in the case of the trisaccharide formed after two glucose additions by transglucosidases, there appear to be multiple peaks in the LC-MS chromatogram in comparison to the one peak observed after 4msob-GSL glucosylation. The earliest eluting and most abundant trisaccharide from cassava honeydew was purified in large enough quantities for NMR analysis and found to correspond to an α-(1→6), α-(1→6) modification, similar to that found for the 4msob-GLS trisaccharide. However, two additional peaks of unknown linkage mode are also present. A multiplicity of chromatographic peaks also occurs for the products of further glucose addition to linamarin, and these are hypothesized to result from regiochemical variation in the addition of glucose by the transglucosidases. That these isomers have different retention times was already apparent from our work on the dissacharide derivative of linamarin and 4msob-GSL. Here the product of BtSUC5 that arises exclusively from α-(1→4)-glucosylation of linamarin elutes later than the α-(1→6) linked compound. The trisaccharide produced by BtSUC5 was also the latest eluting of the triglucoside peaks in the honeydew. Assuming that the specificity for α-(1→4) glucose addition is a fixed property of this enzyme, then the last of the eluting triglucoside peaks for linamarin would also be α-(1→4), α-(1→4)-substituted with the middle eluting peak possibly being a mixture of α-(1→6)- and α-

(1→4)-additions. The exact reason for the presence of multiple peaks for triglucosides of linamarin and not for the triglycosides of 4msob-GSL metabolism is unknown. It may depend on the non-sugar portion of these metabolites, with linamarin being a notably smaller molecule than 4msob-GSL.

In **Manuscript II** we utilized the same two enzymes that displayed transglucosidase activity with 4msob-GSL in enzyme assays with sucrose and linamarin. As mentioned previously, BtSUC5 was capable of producing the α-(1→4) glucoside of this plant-glycoside and was additionally able to produce a triglucoside peak, not previously observed in 4msob-GSL incubations. BtSUC2 on the other hand did not utilize linamarin as a substrate for glucosylation to the same extent as 4msob-GSL, with activity being barely above control levels.

III.I.II Phosphorylation

In **Manuscript II**, we describe the detection of an additional form of pre-emptive detoxification for linamarin in *B. tabaci*: the addition of a phosphate moiety in a phase II detoxification reaction. Phosphorylation of plant defense compounds has recently been reported in the chewing insect *Lymantria dispar* and its close relative *Orgyia antiqua* (Boeckler et al., 2016), where salicinoids including salicin and the breakdown product catechol glucoside were *O*-phosphorylated at the 3 position of the original β-linked glucose. So too in *B. tabaci* do we observe the same positional phosphorylation of linamarin. Addition to sugars at the C-3 position is rare in mammals (Szwergold, Kappler, & Brown, 1990; Szwergold, Kappler, Brown, Pfeffer, & Osman, 1989), and the use of phosphorylation is not common in mammalian detoxification processes in general (Scanlan et al., 2020). Instead mammals modify xenobiotic metabolites via the transfer of sulphate which enhances their secretion in typical phase II detoxification fashion (Karl W Bock et al., 2012). In insects however, phosphorylation has been reported since the early 1960s and is especially described in the detoxification of insecticides (Olsen et al., 2014; Olsen et al., 2016).

The conjugation of phosphate to linamarin in **Manuscript II** causes this metabolite to become inert to the activating enzyme linamarase, and thus it is interesting that there is sometimes an overlap between transglucosidation and phosphorylation of the same

substrates. This may further enhance their excretion, and a similar reaction is observed for the detoxification of the insecticide midazolam by *S. gregaria*. Here, *S. gregaria* first hydroxylates the insecticide by phase I detoxification P450s, followed by the variable addition of a phosphate alone, or a glucose catalyzed by a UDPGT, which can be further phosphorylated (Olsen et al., 2016) parallel to the glucosylation-phosphorylation sequence observed in *B. tabaci*. The transfer of multiple phosphate residues to a single metabolite was not observed despite the possibility for multiple phosphorylations following multiple transglucosidation reactions. Thus, it may be inferred that phosphorylation is restricted to the terminal glucose residue. Confirmation requires the purification of these metabolites in a large enough quantity for NMR analysis, which was not possible for the work described in **Manuscript II**.

III.II Potential pre-emptive detoxification of other two-component defense classes

The heterologously-expressed transglucosidases were tested with sucrose and various other plant defense glycosides, including benzoxazinoid, phenolic, iridoid, flavonoid and additional cyanogenic glycosides, as well as with additional glucosinolates. Interestingly, virtually every class of these metabolites apart from the flavonoid glycoside rutin was efficiently glucosylated by these enzymes, with BtSUC5 showing the greatest activity as a transglucosidase. The observed activities therefore suggest that *B. tabaci* possesses the ability to avoid the toxic effects of a variety of plant defense compounds. It may therefore be expected that these modified metabolites are present in the honeydew of *B. tabaci* or perhaps other more specialized phloem feeders which regularly feed on plants containing these defense metabolites.

III.III Post-hydrolysis detoxification of cyanogenic glycosides

One of the first lines of evidence of cyanogenic glycoside activation by *B. tabaci* was the report of an increase of hydrogen cyanide detoxification enzyme activity in whiteflies that had fed on cassava versus the non-cyanogenic plant sweet potato (Antony et al., 2006). In a parallel experiment in **Manuscript II**, we observed the increased

accumulation of beta-cyanoalanine (a cysteine conjugate of cyanide), catalyzed by beta-cyanoalanine synthase, in the bodies of whiteflies feeding on cassava versus eggplant. Although beta-cyanoalanine synthase is widespread in higher plants due to the universal formation of hydrogen cyanide as a by-product in ethylene biosynthesis (Peiser et al., 1984), this activity has been reported in only one other insect, *Pieris rapae* (van Ohlen et al., 2016). With the evidence for this activity in *B. tabaci* and the existence of likely homologs annotated as a cystathione-β-synthases in various whitefly biotypes (Bta12658 in *B. tabaci* MEAM1 or Ssa04689 in *B. tabaci* SSA1), this enzyme may represent a general whitefly adaptation for feeding on cyanogenic plants.

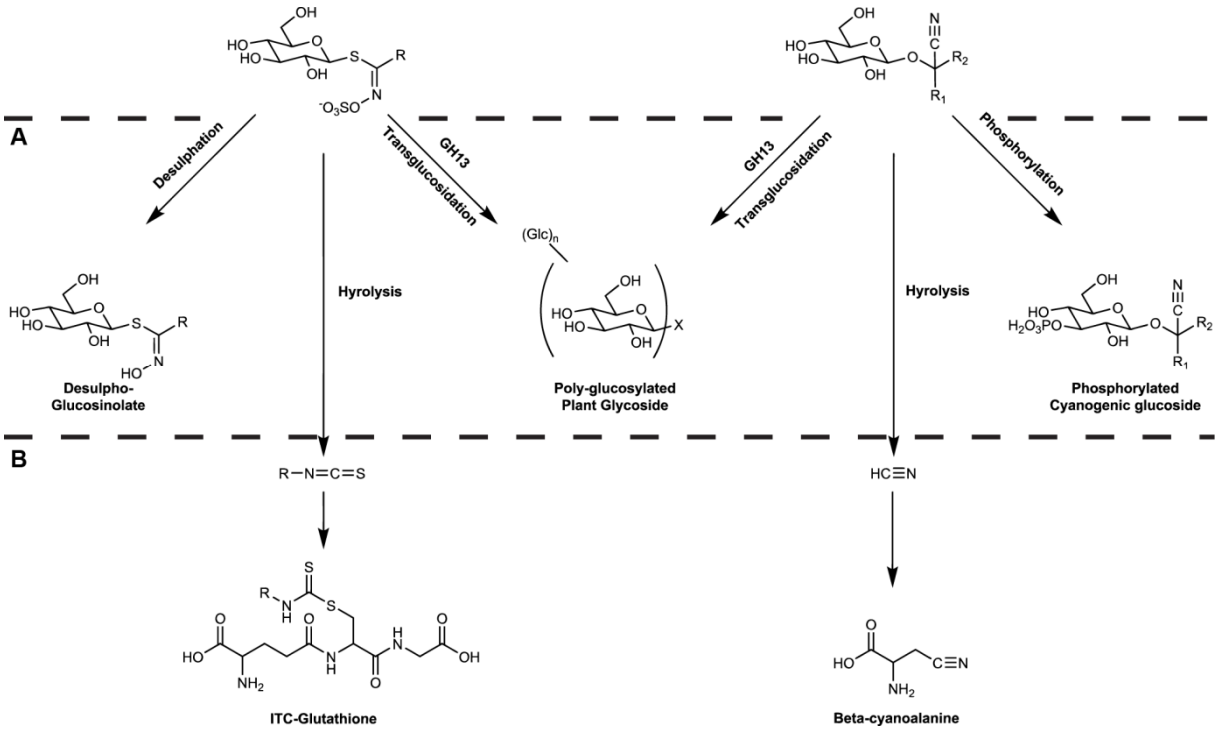


Figure 6: Overview of general and specific detoxification processes in *B. tabaci* for glucosinolates and cyanogenic glycosides. (A) Pre-emptive detoxification of plant two-component defenses may involve reactions specific for glucosinolates, such as desulphation, or specific for cyanogenic glycosides, such as phosphorylation. The utilization of GH13 transglucosidases for poly-glucosylation, however, is common to both glucosinolates and cyanogenic glycosides. (B) Nevertheless, these insects also possess post-hydrolysis detoxification strategies, including the mercapturic acid pathway for isothiocyanates derived from glucosinolate hydrolysis and the synthesis of the nitrile-

containing amino acid, beta-cyanoalanine, for detoxification of the HCN derived from cyanogenic glycosides. Abbreviations: GH (**G**lycoside **H**ydrolase), ITC (**I**so**T**hio**C**yanate).

IV. Saccharide metabolism in *B. tabaci*

Besides plant defense compounds, when phloem feeders begin ingesting the phloem contents at a high rate, sugars may become deadly as well due to the risk of dehydration. Phloem feeders must be able to control water loss associated with feeding on such a concentrated sugar source (Douglas, 2006). One of the best-known methods for osmoregulation of sugars by these insects is the use of sucrase-transglucosidases that employ glucose containing disaccharides for the formation of sugar oligomers, ultimately reducing the overall concentration of osmolytes ingested (Cristofoletti et al., 2003; Douglas, 2006). Sugar metabolism in *B. tabaci* is also distinguished by the activity of sucrose isomerizing enzymes known as trehalulose synthases. In **Manuscript III** we explored the possible genes encoding enzymes involved in osmoregulation of the same GH13 family which have remained elusive since the first enzymatic description of transglucosidase activity in aphid gut extracts (Cristofoletti et al., 2003).

IV.I Sucrose isomerases

Analysis of saccharide composition in general whitefly and specific *B. tabaci* honeydew has been extensively investigated since the early 1990s with further description obtained for individual unique saccharides formed by these insects bearing the associated insect name such as bemisiose and isobemisiose (Hendrix & Salvucci, 2001; Hendrix & Wei, 1994). One of the unique metabolisms observed in whiteflies as well as *B. tabaci* is the production of the sucrose isomer trehalulose (Byrne et al., 2003; Byrne & Miller, 1990), the former sugar being an α-(1↔2) glucose, fructose disaccharide, and the latter being an α-(1→1) glucose, fructose disaccharide. The value of this isomerization to whiteflies is not clear, but since an α-(1→1) glycosidic arrangement of these monosaccharides seems to be more stable than an α-(1↔2) arrangement towards hydrolytic enzymes (Salvucci et al., 1997), trehalulose is less likely to be hydrolyzed, which would allow for more osmotic stability in the insect gut.

Enzymes that display sucrose to trehalulose isomerization are known from bacteria (Rhimi et al., 2008), where they also have a side activity involving the formation of the α-(1→6) isomer of sucrose known as isomaltulose. Despite the obvious activity in whiteflies and in part due to the fact that no gene encoding an enzyme with trehalulose synthase activity had been characterized in insects until now, past investigations studied the possibility that the activity was derived from microbial symbionts (Salvucci, 2003), but found that it was indeed an insect activity. Three candidate enzymes (BtSUC1, 6 and 7) were generated via a blast search of the *B. tabaci* MEAM1 genome (whiteflygenomics.org) using both a trehalulose synthase from *Pseudomonas mesoacidophila* and an isomaltulose synthase from *Erwinina rhepontica* as queries. Interestingly, one of the candidate enzymes (BtSUC1) was previously characterized in **Manuscript I** as being active in sucrose hydrolysis. All three candidates were capable of trehalulose synthesis, which was confirmed via mass spectrometric analysis on an ion mobility-qToF instrument with authentic standards, marking the first entry of genes encoding enzymes capable of this activity in insects. It was shown that these enzymes can produce trehalulose from the free monosaccharides glucose and fructose in labeling studies, which may be useful in osmoregulation if monosaccharide levels are high, perhaps due to hydrolytic activity. Unexpectedly, BtSUC1 and 6, the enzymes with the greatest activity, were also capable of hydrolyzing trehalulose, seemingly in direct opposition to the hypothesis of trehalulose being a less hydrolysable disaccharide. However, it may be that this activity in binding trehalulose is important for transglucosidase activity (further discussed below in IV.II), or at least serves to reduce hydrolysis in other hydrolases present in the whitefly gut. It is also important to note that the enzymes characterized in **Manuscript III** had no detectable isomaltulose synthesis activity. Thus, it would be interesting in future work to characterize the mechanistic differences between insect GH13 enzymes catalyzing this conversion in comparison to bacterial enzymes of the same type. Already sequence motifs have been identified that may be responsible for trehalulose vs. isomaltulose synthesis. Mutational studies in this regard may provide insight into the apparent specificity of *B. tabaci* trehalulose synthases and the possible importance of larger dimers of these enzymes (Salvucci, 2003).

IV.II Sugar polymerization

Perhaps one the most important activities for phloem feeders is the ability to form sugar oligomers, effectively reducing the threat of dehydration by polymerizing ingested saccharides. Alike the enzymes which are responsible for trehalulose synthesis, characterization of the genes encoding enzymes with activities for sugar transglucosidation has been equally evasive in insects (Jing et al., 2016; Price et al., 2007). Interestingly, the same enzymes that showed trehalulose synthesis activity in **Manuscript III** were also able to synthesize sugar oligomers from sucrose and other disaccharides such as trehalulose. These enzymes carried out various transglucosidation reactions such as the synthesis of isomaltose in the case of BtSUC1, which has not been reported in whitefly honeydew, as well as numerous unknown trisaccharides, some of which may correspond to unique *B. tabaci* sugars such as bemisiose and isobemisiose. In labeling experiments, BtSUC1, 6 and 7 were capable of incorporating free labeled-glucose in these transglucosidation reactions, which has not been previously reported for sucrase-transglucosidases. However, some evidence of reversible glucose binding to GH enzymes has been noted (Nikolov, Meagher, & Reilly, 1989), which may affect the osmolarity of free monosaccharides. Additionally in incubations with sucrose, these enzymes seemed to produce two saccharides for each transglucosidation reaction, corresponding to one product composed entirely of glucose units and another with a fructose and additional glucose units.

Since trehalulose represents the majority of all excreted carbohydrate in the honeydew, enzyme assays with this disaccharide as substrate are of particular interest in order to determine if trehalulose is used in such processes as transglucosidation. BtSUC7 did not utilize trehalulose as a substrate, similar to its specificity for hydrolysis, while BtSUC1 and 6 formed the same products as on incubations with sucrose, suggesting that trehalulose could be the real substrate for these transglucosidation reactions instead of sucrose. Incubations with the related α-(1→6) isomer isomaltulose produced identical activity in BtSUC6 as from sucrose and trehalulose incubations, but different activity in BtSUC1, where a new trisaccharide was formed instead of the di-glucose metabolite isomaltose (α-D-glucopyranosyl-(1→6)-D-glucose). Since this disaccharide is clearly formed via transglucosidation (also confirmed by labeling studies), its absence in isomaltulose assays with BtSUC1 where isomaltulose is also hydrolyzed is an unexpected

observation that may offer us insight into the possibility that different catalytic residues are used in these enzymes for the formation of various transglucosidation products.

V. The division between detoxification and osmoregulation through GH13 enzyme recruitment

The mechanistic similarity between GH13 transglucosidases that can pre-emptively detoxify plant secondary metabolite glycosides (**Manuscripts I & II**) and those which isomerize and polymerize sugars (**Manuscript III**) is readily apparent. However, in *B. tabaci*, these processes seem to be parallel rather than intersecting such that enzymes which form saccharide polymers cannot modify secondary metabolites (**Manuscript III**), and vice versa (**Manuscript I & II**). The specialization of enzymatic activity from a more promiscuous ancestor is a phenomenon commonly observed in metabolic evolution, and may also be true for the evolution of GH13 enzymes in *B. tabaci*. Considering the rather large difference in metabolite concentrations when comparing sugars with secondary metabolites in the phloem, it is logical to have enzymes that divide the labour and specialize in the modification of one group or another rather than having a single enzyme that can "do it all" (Jensen, 1976). In accordance with this line of thought, there has been a rather large expansion in the number of GH13 enzymes in *B. tabaci* (**Manuscript I**) in comparison to non-phloem feeders as well as other phloem feeders. One reason for this expansion could be to accommodate the different metabolites of all the many plants that *B. tabaci* feeds on. Therefore, the many GH13 enzymes may well explain their success as phloem feeders.

VI. Potential future control measures for *B. tabaci*

The characterization of specific enzymes in *B. tabaci* involved in detoxification and osmoregulatory functions provides new targets for future control measures against this insect such as plant delivered RNA interference (RNAi). The promise of RNAi with regards to controlling phloem feeding pests has already been shown and phloem specific promotion of double stranded RNA elements (Tzin et al., 2015) and gene stacking (Eakteiman et al., 2018) have notable effects on whitefly fecundity. The silencing of detoxification enzymes could negatively impact whitefly feeding on individual host plants, while the general

silencing of osmoregulatory enzymes could serve as potent control measures against whiteflies on all hosts.

VII. Conclusion

In this dissertation, I sought to investigate the metabolic fates of plant defense compounds as well as common sugars in the polyphagous phloem feeder *B. tabaci.* Given the vast number of plants that whiteflies can utilize, it would be interesting to learn if detoxification of host defenses allows *B. tabaci* to feed with such impunity. The answer would be relevant not only for research on whiteflies but also for understanding how herbivore metabolism plays a role in in the host breadth of all herbivores. I have illustrated some general and specific detoxification strategies employed by *B. tabaci* for the detoxification of activated plant defenses and characterized some important osmoregulatory enzymes from which detoxification activities seem to have emerged.

In the first two manuscripts, I highlighted the interactions of two-component defenses with the whitefly *B. tabaci*. A major finding was that whitefly feeding activated these defenses, previously thought not to respond to the damage of phloem feeders. Likewise, we also observed the well-known detoxification via the mercapturic acid pathway for glucosinolate hydrolysis products formed after whitefly feeding on Brassica plants and the detoxification via beta-cyanoalanine synthase for hydrogen cyanide formed after whitefly feeding on cassava plants containing cyanogenic glycosides. The formation of toxic hydrolysis products of these plant defense compounds upon whitefly feeding hints about the advantages whiteflies might obtain if they were able to detoxify these plant defenses in a pre-emptive manner.

Detailed analysis of whitefly honeydew allowed my colleagues and I to investigate the metabolism of ingested phloem contents. We identified the presence of a unique glucose conjugation mechanism utilized in the detoxification of both glucosinolates and cyanogenic glycosides, with the latter compounds also being pre-emptively phosphorylated as a means of detoxification. The glucosylation reaction is a phase II conjugation mechanism not previously described in literature and is seemingly restricted to phloem-feeding herbivores, being intimately related to saccharide modification and osmoregulation. We further described enzymes involved in this modification from Family 13 of the Glucohydrolase

(GH) class, which seem to be specific for the metabolism of secondary metabolites instead of having osmoregulatory functions. In the final chapter, we characterized novel sugar transglucosidating enzymes from *B. tabaci* that readily modify saccharides for osmoregulatory purposes, but show no activity towards plant defense glycosides. All of these activities are attributable to enzymes of the GH13 family, which have undergone a large expansion within *B. tabaci* and might explain the success and radiation of hemipteran pests. Taken together, the work of this thesis highlights the metabolism of host dietary constituents in *B. tabaci* as a whole, and illustrates the importance of GH13 enzymes in the detoxification of plant defense glycosides, and the metabolism of sugars associated with osmoregulation in *B. tabaci*.

General Summary

The evolutionary arms race between plants and their herbivores has been an area of great interest to ecologists since even before this concept was clearly defined. Plants typically exert selection pressures on insects through the production of toxic chemical defenses and therefore insects have responded through specific detoxification strategies. As a result of many cycles of these co-evolutionary processes, the number of chemical defenses in plants has increased enormously. Yet, generalist insects such as the whitefly *B. tabaci* are still capable of feeding on numerous plants that differ in their chemical landscapes. The mechanisms which allow this phloem feeder to utilize so many host plants with absolute impunity may be found within the metabolism of this insect.

In this dissertation, I investigated how a generalist phloem-feeding insect, the tobacco whitefly *Bemisia tabaci*, metabolizes plant defense compounds and sugars via transglucosidation reactions. Chemical analysis of honeydew coupled to gene identification and biochemical characterization of heterologously expressed sugar modifying GH13 enzymes were employed to understand the mechanism, function and evolution of whitefly metabolism.

One of the best examples of chemical defense evolution is the production of activated chemical defenses, usually in the form of glycosides and their activating hydrolases. These plant defenses are well studied with regards to chewing herbivores due to the amount of tissue damage produced when feeding, but are thought to not be activated by phloem feeding insects due to the minimal damage made upon piercing the phloem. Within this thesis we investigated the metabolism of two major classes of activated defenses, the glucosinolates (**Manuscript I**) and the cyanogenic glycosides (**Manuscript II**).

Chemical analysis of honeydew from whitefly feeding on *Arabidopsis thaliana* (**Manuscript I**), revealed the presence of glucosinolate hydrolysis products, as well as mercapturic acid pathway conjugates of these activated toxins, showing the ingestion and activation of glucosinolates when feeding. Further investigations into the honeydew also revealed the presence of unique metabolic derivatives of the major glucosinolate present in the plant, 4msob-GSL (as well as others). After purification and extensive spectral measurements, it was found that these derivatives were products of multiple glucose

additions in either α-(1→6) or α-(1→4) orientations to the originally β-linked glucose of the plant toxin, with derivatives up to four glucose additions detected by LC-MS. These insect derived glycosides were not hydrolyzed by the activating enzyme myrosinase, and therefore constitute a completely novel phase II detoxification mechanism in *B. tabaci*. Furthermore, glycosides were detected in the honeydew of other phloem feeding insects feeding on brassica plants, but not in other herbivorous arthropods, showing the restriction of this metabolism to the guild of phloem-feeders.

Classically, glucose conjugation in detoxification reactions is catalyzed by UDP glucosyltransferases (UDPGTs). However, these enzymes usually only transfer a single glucose moiety to a molecule and it is rarely observed that sugar residues are further glucosylated. Therefore, isotopomers of sucrose were fed along with 4msob-GSL in artificial diets in order to determine if other enzymes such as sucrase-transglucosidases could be involved. Indeed the *in vivo* incorporation of label in the insect glycosides provided direct mechanistic evidence for the activity of a glucose transferring sucrase-transglucosidase, as the responsible enzyme(s). Typically, these enzymes of the GH13 family of glycoside hydrolases are associated with osmoregulatory processes within the insect where incoming dehydrating sucrose concentrations must be reduced in osmolarity by the action of these sugar polymerizing proteins. The number of GH13 enzymes in the *B. tabaci* genome was shown to be far greater than that of other insects (~ 60 coding genes), even in comparison to other phloem feeders, such as aphids (~ 30 coding genes). In an effort to characterize enzymes that can perform transglucosidation of the secondary metabolite 4msob-GSL, we cloned and characterized five genes coding enzymes which by sequence homology to other insect GH13 enzymes and their expression profiles might be responsible for this transformation *in vivo*. Two of the enzymes (named BtSUC2 and BtSUC5) which were cloned and expressed in insect cells were able to catalyze the transfer of glucose to 4msob-GSL using only sucrose as a donor sugar while the three other enzymes only performed sucrose hydrolysis.

Following this discovery of not only the hydrolysis of 4msob-GSL due to *B. tabaci* feeding activities, but also a novel form of detoxification in this phloem feeding insect, we also sought to understand the ubiquity of this detoxification strategy for other activated defenses. In **Manuscript II** the cyanogenic glycoside linamarin and the important African staple crop cassava came into focus as a new system to study a cryptic species of *B. tabaci* (SSA1) feeding on this toxic plant. In a similar fashion to the patterns observed in

Manuscript I, we found that whiteflies which feed on cassava (a cyanogenic plant) in comparison to eggplant (a non-cyanogenic plant) had extremely elevated levels of a cyanide detoxification metabolite, being beta-cyanoalanine accumulating in their bodies, demonstrating a similar although indirect measurement of cyanogenesis during feeding by this whitefly.

Further investigations into the honeydew of *B. tabaci* SSA1 feeding on cassava revealed the same poly-glucose metabolism taking place for the cyanogenic glycoside linamarin as seen previously with regards to 4msob-GSL, with varying regiochemical additions for higher order glycosides. Unsurprisingly, these insect-produced glycosides were also resistant to hydrolysis by the activating enzyme linamarase present in cassava leaf enzyme extracts, demonstrating that poly-glucosylation of this cyanogenic glycoside was a pre-emptive detoxification mechanism. Importantly, as some cyanogenic glycosides also exist in disaccharide forms in some plants such as the β-(1→6)-linked diglucoside of linamarin, linustatin, we tested the stability of the insect glycosides to these disaccharidases which are capable of hydrolyzing linustatin. The observed stability of the insect derived α-glycosides in comparison to the β-linked plant metabolites shows the detoxification function of transglucosidation and may also demonstrate the importance of the anomeric stereochemistry for stability to plant activating enzymes. The whitefly enzymes previously characterized for transglucosidation were tested for activity with linamarin and glucose-transfer from sucrose to linamarin was observed with low efficiency for BtSUC2. The enzyme BtSUC5 on the other hand demonstrated more rapid glucosylation of linamarin and the ability to even produce higher order glycosides, with a metabolite corresponding to two glucose units detected. These enzymes were also tested with various plant glycosides and it was found that BtSUC2 and BtSUC5 were able to transglucosidate all metabolites tested, thus hinting that this metabolism may be utilized on many other defense compounds.

Manuscript II also describes the unexpected detection of other derivatives of cyanogenic glycosides, not previously seen in the metabolism of 4msob-GSL, which upon purification were determined to be the result of a transfer of a phosphate residue. This addition was regiochemically determined to be at the 3-*O* position on the β-sugar of linamarin, and was also seen to decorate the insect-derived glycosides. The phosphorylated derivatives were also observed to be stable to activating enzymes in cassava and therefore also constitute a pre-emptive detoxification strategy. While phosphorylation as a detoxification strategy is uncommon in mammals, it has been reported in multiple instances

in insects for the modification of insecticides and plant toxins. The presence of multiple pre-emptive detoxification strategies for cyanogenic glycosides was surprising, and illustrates the variability of metabolism in *B. tabaci* in terms of both general detoxification in the form of transglucosidation and more specific detoxification such as phosphorylation.

Finally, **Manuscript III** we endeavoured to expand our knowledge on *B. tabaci* metabolism of sugars that are frequently associated with osmoregulation. We investigated the enzymes responsible for the isomerization of sucrose to trehalulose which is particularly present in whitefly metabolism. Utilizing bacterial enzymes which show trehalulose synthase activity as queries to search the whitefly genome, we were able to identify three genes shown to encode enzymes that catalyze the isomerization of sucrose to trehalulose. One enzyme originally characterized as a hydrolase in **Manuscript I** (BtSUC1) was now determined to also perform isomerization reactions with sucrose. TIMS-ToF analysis allowed the direct conformation of the sugar transformations performed by these enzymes following invertase treatment of sucrose incubations, marking the first identification of trehalulose synthase enzymes in insects.

Isomerization was interestingly not the only reaction that these enzymes were capable of performing. The three active trehalulose synthases also catalyzed a plethora of transglucosidation reactions with sucrose and other saccharides which has also not been demonstrated before in previous literature. Interestingly, these enzymes were capable of the transglucosidation of saccharides such as sucrose but were unable to utilize the secondary metabolite 4msob-GSL as an acceptor substrate, further demonstrating a division between osmoregulation and detoxification enzymes. TIMS-ToF analysis allowed us to identify only a fraction of the metabolic products that these enzymes produced and further investigations utilizing more sugar standards are necessary to identify the remainder. The characterization of these enzyme activities gives us greater understanding of sugar transformations important for *B. tabaci* host utilization, which may include the formation of the unique *Bemisia* sugar products such as bemisiose and isobemisiose in future studies.

Overall, the chemical analysis of whitefly honeydew provided deep insights into the complex metabolism of *B. tabaci* with respect to the detoxification of plant metabolites. Coupled with biochemical characterization of the responsible enzymes, the work demonstrates how *B. tabaci* whiteflies are adapted to utilize their host plants.

Zusammenfassung

Das evolutionäre Wettrüsten zwischen Pflanzen und Herbivoren war für Ökologen schon immer von großem Interesse. Es ist charakteristisch für Pflanzen durch toxisch-chemische Abwehrmechanismen einen Selektionsdruck auf Insekten auszuüben. Letztere hingegen haben, um dem entgegenzuwirken, spezifische Entgiftungsstrategien ausgebildet. Infolge dieses Wettbewerbs ist die Anzahl der chemischen Abwehrmechanismen in der Pflanzenwelt enorm. Dennoch gelingt es generalistischen Herbivoren wie der Weißfliege *B. tabaci*, sich von zahlreichen Pflanzen zu ernähren und damit auch unterschiedlichen chemischen Verteidigungen zu widerstehen. Hinweise auf die Mechanismen, die es diesem sogenannten Phloem-Fresser ermöglichen, so viele Wirtspflanzen zu befallen, können im Metabolismus des Insekts gefunden werden.

Mit Hilfe von sowohl analytische Untersuchungen von Honigtau, als auch durch biochemische Charakterisierung von zuckermodifizierenden GH13-Enzymen, trägt diese Dissertation dazu bei unser Wissen über generelle und spezifische Metabolisierung von pflanzlichen Verteidigungsmetaboliten und Zuckern, durch den generalistischen Phloem-Fresser *B. tabaci,* zu erweitern.

Eines der besten Beispiele für die Evolution der chemischen Abwehr ist die Erzeugung aktivierter chemischer Abwehrmechanismen, üblicherweise in Form von Glykosiden und ihren aktivierenden Hydrolasen. Klassischerweise werden diese pflanzlichen Abwehrmechanismen bei kauenden Pflanzenfressern auf Grund der beim Fressen verursachten Menge an Gewebeschäden untersucht. Bei Phloem-fressenden Insekten werden diese hingegen als inert/unbeteiligt angenommen. Aufgrund aktuellerer Forschungsergebnisse ist allerdings seit einigen Jahren ein Paradigmenwechsel auf diesem Gebiet zu beobachten. Im Rahmen dieser Dissertation untersuchten wir den Metabolismus von zwei Hauptklassen aktivierter Abwehrmechanismen, den Glucosinolaten (**Manuskript I**) und den cyanogenen Glykosiden (**Manuskript II**).

Bei der chemischen Analyse des Honigtaus der Weißfliege, welche sich zeitgleich von *Arabidopsis thaliana* ernährte (**Manuskript I**), beobachteten wir die Produktion von Glucosinolat-Hydrolyseprodukten, sowie von aus dem Mercaptursäure-Weg stammenden Konjugaten, aus den aktivierten Toxinen, was die Aufnahme und Aktivierung von Glucosinolaten nach dem Fraß des Insekts beweist. Weitere Untersuchen des Honigtaus zeigten zudem das Vorhandensein einzigartiger 4msob-GSL Derivate (neben anderen),

dem Hauptglucosinolat in der Nahrung des Insekts. Nach der Aufreinigung und Charakterisierung wurde festgestellt, dass sich diese Derivate durch mehrfache Addition von Glucose mit entweder einer α-(1→6)- oder α-(1→4)-Orientierung zur ursprünglich β-verknüpften Glucose des Pflanzentoxins ableiten ließen. Mittels LC-MS konnten hierfür bis zu vier Glucose Additionen nachvollzogen werden. Diese von Insekten abgeleiteten Glykoside waren gegenüber dem aktivierenden Enzym Myrosinase inaktiv und stellen daher Produkte eines völlig neuen Phase-II-Entgiftungsmechanismus in *B. tabaci* dar. Darüber hinaus wurden im Honigtau anderer Phloem-fressender Insekten, welche sich von Brassica Pflanzen ernähren, Glykoside nachgewiesen, die in anderen pflanzenfressenden Arthropoden nicht detektiert werden konnten.

Klassischerweise wird die Konjugation von Glucose bei der Entgiftung durch UDPGTs katalysiert. Da diese Enzyme für gewöhnlich nur ein einzelnes Glucose-Molekül übertragen, ist die weitere Glucose Konjugation der Zucker-Einheit eine seltene Beobachtung. Um zu bestimmen, ob andere Enzyme wie Sucrase-Transglucosidasen beteiligt sein könnten wurden Isotopomere von Saccharose und 4msob-GSL in künstlichen Diäten verwendet. In der Tat lieferten die in vivo Markierungen, eingebaut in die Insektenglykoside, einen direkten mechanistischen Beweis für die Aktivität einer Glucose-übertragenden Sucrase-Transglukosidase als verantwortliches Enzym. Typischerweise werde die Enzyme der GH13-Familie von Glykosid-Hydrolasen mit osmo-regulatorischen Prozessen innerhalb des Insekts assoziiert, wo große Mengen dehydratisierender Saccharose aufkommen und durch diese zuckerpolymerisierenden Proteine abgebaut werden. Beim Vergleich mit der GH13-Familie von *B. tabaci* wurde festgestellt, dass die Anzahl der Enzyme im Genom der Weißfliege weitaus größer war als bei anderen Insekten (~ 60 kodierende Gene). Deutlich wird dies besonders im Vergleich zu anderen Vertretern aus der Gilde der Phloem-Fresser, wie bspw. den Blattläusen (~ 30 kodierende Gene).

Um diese extrem große Enzymfamilie innerhalb von *B. tabaci* und um Enzyme charakterisieren zu können, welche in der Lage sind diese Transglucosidierungsreaktionen von Sekundärmetaboliten wie dem 4msob-GSL durchzuführen, klonierten und charakterisierten wir fünf Enzym kodierende Gene, die auf Grund ihrer Sequenzhomologie zu anderen Insekten-GH13-Enzymen sowie ihrer Expressionsprofile für diese *in vivo* Transformation verantwortlich sein könnten. Während zwei der Enzyme (BtSUC2 und BtSUC5), die in den Insektenzellen kloniert und exprimiert wurden, in der Lage waren, indem sie nur Saccharose als Zucker-Donor verwendeten, den Transfer von Glukose auf

4msob-GSL zu katalysieren, führten die anderen drei Enzyme nur eine Hydrolyse der Saccharose durch.

Nach dieser einzigartigen Entdeckung von sowohl der Hydrolyse von 4msob-GSL nach Fraß Aktivität von *B. tabaci*, sowie ebenfalls einer neuartigen Form der Entgiftung der Pflanzenverteidigung durch dieses Phloem-fressende Insekt, wollten wir die Allgegenwärtigkeit dieser Entgiftungsstrategie für andere verwandte aktivierte Abwehrmechanismen verstehen, welche bereits charakteristische Anzeichen einer Modifikation gezeigt hatten (**Manuskript I**). In **Manuskript II** fokussierten wir uns deshalb auf Maniok, einem wichtigen Grundnahrungsmittel in Afrika, sowie dem darin enthaltenen cyanogenem Glykosid Linamarin, als neues System zur Untersuchung kryptischer Arten von *B. tabaci* (SSA1), die sich von dieser toxischen Pflanze ernähren. Wie bereits in **Manuscript I** auf ähnliche Weise beobachtet, stellten wir fest, dass Weißfliegen, die sich von Maniok (einer cyanogenen Pflanze) ernährten, im Vergleich zu Auberginen (einer nicht cyanogenen Pflanze), extrem hohe Mengen an Beta-cyanoalanin, eines Entgiftungsmetaboliten der oben, welches sich in den Körpern der Tiere ansammelte und zugleich ein – wenn auch indirektes – Maß der während der Nahrungsaufnahme der Weißfliege stattfindenden Cyanogenese darstellt.

Weitere Untersuchungen des Honigtaus von auf Maniok fressenden *B. tabaci* SSA1 ergaben den gleichen Poly-Glukosestoffwechsel, mit ähnlich variierender regio-chemischer Dekoration der Zucker-Funktionen, für das cyanogene Glykosid Linamarin, wie er zuvor bereits bei 4msob-GSL beobachtet wurde. Es war daher auch nicht überraschend, dass diese vom Insekt produzierten Glykoside inert gegenüber dem das Linamarin aktivierenden Enzyms, gewonnen aus Maniokblättern stammenden Enzymextrakten, waren. Dies demonstrierte die präventive Detoxifizierung dieses cyanogenen Glykosids durch denselben Mechanismus, was durch eine parallel durchgeführte Isotopenmarkierungsstudie, unter Verwendung von Saccharoseisotopomeren, weiter bestätigt werden konnte wurde.

Einige cyanogene Glykoside, wie bspw. das β-(1→6)-verknüpfte Diglykosid des Linamarins – Linustatin – liegen auch in Form eines Disaccharides vor. Es ist uns ebenfalls gelungen, die Stabilität dieser Glykoside gegenüber von Disaccharidasen zu testen, welche das oben erwähnte pflanzliche Diglykosid hydrolysieren würden. Die beobachtete Stabilität der vom Insekten abgeleiteten α-Glykoside im Vergleich zu den β-verknüpften

Pflanzenmetaboliten zeigt möglicherweise die Bedeutung der Stereochemie am anomeren Kohlenstoff für deren Stabilität gegenüber der aktivierenden Wirkung der pflanzlichen Enzyme. In ähnlicher Weise wurden Enzyme auf Aktivität mit Linamarin getestet, welche zuvor als charakteristisch für die Transglukosidierung beschrieben wurden. Hierbei konnte, wenn auch nur mit geringerer Effizienz, für BtSUC2 ein Glucose-Transfer von Saccharose zum Linamarin beobachtet werden. Andererseits zeigte das Enzym BtSUC5 die Fähigkeit, sogar Linamarin-Glykoside höherer Ordnung zu erzeugen, wobei sogar ein Metabolit mit zwei zusätzlichen Glucoseeinheiten nachgewiesen werden konnte. Darüber hinaus wurden die Aktivitäten dieser Enzyme mit verschiedenen Pflanzenglykosiden mit dem Ergebnis getestet, dass BtSUC2 und BtSUC5 alle im Versuch verwendeten Metaboliten transglukosidiert werden konnten. Dies impliziert, dass eine Übertragbarkeit dieses Metabolismus auf andere Abwehrstoffe möglich sein könnte, sofern diese Enzyme für die Transformation *in vivo* verantwortlich sind.

Unerwarteter Weise konnten wir in **Manuskript II** auch andere Derivate cyanogener Glykoside nachweisen, welche zuvor nicht im Metabolismus von 4msob-GSL beobachtet und nach der Aufreinigung als Produkt aus der Übertragung eines Phosphatrests bestimmt wurden. Durch regio-chemische Untersuchungen zeigte sich schließlich, dass die Addition dieser Phosphatgruppe an Position 3-O der β-Glucose-Einheit des Linamarins stattfindet und es wurde auch gesehen, dass sie die von Insekten abgeleiteten Glykoside weiter dekorieren. Die phosphorylierten Derivate erwiesen sich als stabil, gegenüber der aktivierenden Enzyme aus Maniok und stellen daher auch eine präventive Entgiftungsstrategie dar. Phosphorylierung als Entgiftungsstrategie bei Säugetieren ist ungewöhnlich. Jedoch wurde diese Art der Detoxifizierung in mehreren Fällen bei Insektiziden und auch Pflanzentoxinen beobachtet, weshalb Phosphorylierung als bei Insekten üblicher Entgiftungsmechanismus angesehen wird. Das Vorhandensein mehrerer präventiver Entgiftungsmechanismen für cyanogene Glykoside war überraschend und veranschaulicht die metabolische Variabilität in *B. tabaci* sowohl hinsichtlich der allgemeinen Detoxifizierung in Form einer Transglukosidierung, als auch einer spezifischeren Entgiftung wie der Phosphorylierung.

In **Manuscript III** bemühten wir uns schließlich, unser Wissen über, besonders mit Osmoregulation assozierte, Metabolisierung von Zuckern durch *B. tabaci* zu erweitern. Dabei haben wir uns zu Beginn der Untersuchung auf Enzyme fokussiert, die für die einzigartige, ausschließlich in Weißfliegen beobachtete, Isomerisierung von Saccharose zu

Trehalulose verantwortlich sind. Unter Verwendung bakterieller Enzyme, welche Trehalulose-Synthaseaktivität innerhalb des Weißfliegengenoms abfragen und damit sichtbar machen, konnten wir drei Gene charakterisieren, die Enzyme, verantwortlich für die Isomerisierung von Saccharose zu Trehalulose, codieren. Für ein Enzym, welches ursprünglich in **Manuskript I** (BtSUC1) als Hydrolase charakterisiert wurde, konnte nun auch die Katalyse von Isomerisierungs-Reaktionen von Saccharose aufgezeigt werden. Die TIMS-ToF-Analyse ermöglichte die direkte Konformation der von diesen Enzymen durchgeführten Zuckertransformationen nach Invertase-Behandlung von Saccharose-Inkubationen, was die erste Identifizierung von Trehalulosesynthase-Enzymen bei Insekten markierte.

Die Isomerisierung erwies sich nicht als einzige Reaktion, die diese Enzyme katalysieren. Die drei Kandidaten konnten neben Hydrolyse, eine Vielzahl von Transglukosidierungs-Reaktionen mit Saccharose und anderen Sacchariden katalysieren, was in der bisherigen Literatur noch nicht beschrieben wurde. Interessanterweise waren diese Enzyme in der Lage Saccharide wie Saccharose zu transglukosidieren. Sie konnten jedoch den Sekundärmetaboliten 4msob-GSL nicht als Substrat verwenden, was darüber hinaus die Möglichkeit einer Separation zwischen Osmoregulations- und Entgiftungsenzymen andeutet. Mit Hilfe von TIMS-ToF-Analyse waren wir in der Lage einen Teil der Stoffwechselprodukte dieser Enzyme zu identifizieren. Weiterführende Untersuchungen machen jedoch die Verwendung zusätzlicher Zuckerstandards unerlässlich. Die Charakterisierung dieser Enzymaktivitäten gibt uns ein besseres Verständnis der Zuckertransformationen, die für die Verwendung des *B. tabaci*-Wirts wichtig sind, einschließlich der Bildung der einzigartigen *Bemisia*-Zuckerprodukte wie Bemisiose und Isobemisiose in zukünftigen Studien.

Insgesamt lieferte die chemische Analyse des Honigtaus der Weißfliege tiefe Einblicke in den variablen Metabolismus, hinblicklich der Detoxifizierung von Pflanzenmetaboliten, von *B. tabaci* und sogar den anderer Phloem-fressender Insekten. In Verbindung mit der biochemischen Enzymcharakterisierung von Proteinen, die für einen einzigartigen und dominanten Metabolismus verantwortlich sind, heben diese Ergebnisse die Nützlichkeit von zuckermodifizierenden Enzymen für *B. tabaci* zur Etablierung auf seinen vielfältigen Wirtspflanzen besonders hervor.

Literature Cited

Abd-Rabou, S., & Simmons, A. M. (2010). Survey of reproductive host plants of Bemisia tabaci (Hemiptera: Aleyrodidae) in Egypt, including new host records. *Entomological news, 121*(5), 456-465.

Acin-Perez, R., Gatti, D. L., Bai, Y., & Manfredi, G. (2011). Protein phosphorylation and prevention of cytochrome oxidase inhibition by ATP: coupled mechanisms of energy metabolism regulation. *Cell metabolism, 13*(6), 712-719.

Agerbirk, N., & Olsen, C. E. (2012). Glucosinolate structures in evolution. *Phytochemistry, 77*, 16-45.

Ali, J. G., & Agrawal, A. A. (2012). Specialist versus generalist insect herbivores and plant defense. *Trends in Plant Science, 17*(5), 293-302.

Andersen, M. D., Busk, P. K., Svendsen, I., & Møller, B. L. (2000). Cytochromes P-450 from Cassava (Manihot esculentaCrantz) Catalyzing the First Steps in the Biosynthesis of the Cyanogenic Glucosides Linamarin and Lotaustralin CLONING, FUNCTIONAL EXPRESSION IN PICHIA PASTORIS, AND SUBSTRATE SPECIFICITY OF THE ISOLATED RECOMBINANT ENZYMES. *Journal of Biological Chemistry, 275*(3), 1966-1975.

André, I., Potocki-Véronese, G., Morel, S., Monsan, P., & Remaud-Siméon, M. (2010). Sucrose-utilizing transglucosidases for biocatalysis *Carbohydrates in Sustainable Development I* (pp. 25-48): Springer.

Antony, B., Lisha, V., Palaniswami, M., Sugunan, V., Makeshkumar, T., & Henneberry, T. (2006). Bemisia tabaci (Homoptera: Aleyrodidae) and Indian cassava mosaic virus transmission. *International Journal of Tropical Insect Science, 26*(3), 176-182.

Ashford, D., Smith, W., & Douglas, A. (2000). Living on a high sugar diet: the fate of sucrose ingested by a phloem-feeding insect, the pea aphid Acyrthosiphon pisum. *Journal of Insect Physiology, 46*(3), 335-341.

Bak, S., & Feyereisen, R. (2001). The involvement of two P450 enzymes, CYP83B1 and CYP83A1, in auxin homeostasis and glucosinolate biosynthesis. *Plant Physiology, 127*(1), 108-118.

Bak, S., Nielsen, H. L., & Halkier, B. A. (1998). The presence of CYP79 homologues in glucosinolate-producing plants shows evolutionary conservation of the enzymes in the conversion of amino acid to aldoxime in the biosynthesis of cyanogenic glucosides and glucosinolates. *Plant molecular biology, 38*(5), 725-734.

Bak, S., Paquette, S. M., Morant, M., Morant, A. V., Saito, S., Bjarnholt, N., . . . Simonsen, H. T. (2006). Cyanogenic glycosides: a case study for evolution and application of cytochromes P450. *Phytochemistry Reviews, 5*(2-3), 309-329.

Bak, S., Tax, F. E., Feldmann, K. A., Galbraith, D. W., & Feyereisen, R. (2001). CYP83B1, a cytochrome P450 at the metabolic branch point in auxin and indole glucosinolate biosynthesis in Arabidopsis. *The Plant Cell, 13*(1), 101-111.

Basu, A. (2019). *Bemisia Tabaci (Gennadius): Crop Pest And The Principal Whitefly Vector Of Plant Viruses*: CRC Press.

Bazzaz, F. A., & Grace, J. (1997). *Plant resource allocation*: Elsevier.

Benn, M. (1977). Glucosinolates *Low Molecular Weight Sulphur Containing Natural Products* (pp. 197-210): Elsevier.

Bergé, J., Feyereisen, R., & Amichot, M. (1998). Cytochrome P450 monooxygenases and insecticide resistance in insects. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 353*(1376), 1701-1705.

Bhagwat, S. V., Mullick, J., Avadhani, N., & Raza, H. (1998). Differential response of cytosolic, microsomal, and mitochondrial glutathione S-transferases to xenobiotic inducers. *International journal of oncology, 13*(2), 281-289.

Björkman, R., & Janson, J.-C. (1972). Studies on myrosinases: I. Purification and characterization of a myrosinase from white mustard seed (Sinapis alba, L.). *Biochimica et Biophysica Acta (BBA)-Enzymology, 276*(2), 508-518.

Blažević, I., Montaut, S., Burčul, F., Olsen, C. E., Burow, M., Rollin, P., & Agerbirk, N. (2020). Glucosinolate structural diversity, identification, chemical synthesis and metabolism in plants. *Phytochemistry, 169*, 112100.

Bock, K. W. (2016). The UDP-glycosyltransferase (UGT) superfamily expressed in humans, insects and plants: Animal□ plant arms-race and co-evolution. *Biochemical pharmacology, 99*, 11-17.

Bock, K. W., Burchell, B., Chiba, M., Clarke, D., Creveling, C., Fenselau, C., . . . Lu, A. (2012). *Conjugation—Deconjugation Reactions in Drug Metabolism and Toxicity* (Vol. 112): Springer Science & Business Media.

Boeckler, G. A., Paetz, C., Feibicke, P., Gershenzon, J., & Unsicker, S. B. (2016). Metabolism of poplar salicinoids by the generalist herbivore Lymantria dispar (Lepidoptera). *Insect Biochemistry and Molecular Biology, 78*, 39-49.

Bones, A., & Iversen, T.-H. (1985). Myrosin cells and myrosinase. *Israel Journal of Plant Sciences, 34*(2-4), 351-376.

Borek, V., Elberson, L. R., McCaffrey, J. P., & Morra, M. J. (1998). Toxicity of isothiocyanates produced by glucosinolates in Brassicaceae species to black vine weevil eggs. *Journal of Agricultural and Food Chemistry, 46*(12), 5318-5323.

Bourgaud, F., Gravot, A., Milesi, S., & Gontier, E. (2001). Production of plant secondary metabolites: a historical perspective. *Plant science, 161*(5), 839-851.

Boykin, L. M., Shatters Jr, R. G., Rosell, R. C., McKenzie, C. L., Bagnall, R. A., De Barro, P., & Frohlich, D. R. (2007). Global relationships of Bemisia tabaci (Hemiptera: Aleyrodidae) revealed using Bayesian analysis of mitochondrial COI DNA sequences. *Molecular phylogenetics and evolution, 44*(3), 1306-1319.

Bradshaw, C. J., Leroy, B., Bellard, C., Roiz, D., Albert, C., Fournier, A., . . . Courchamp, F. (2016). Massive yet grossly underestimated global costs of invasive insects. *Nature communications, 7*(1), 1-8.

Brown, J. (1994). Current status of Bemisia tabaci as a plant pest and virus vector in agroecosystems worldwide. *FAO Plant Protection Bulletin, 42*(1/2), 3-32.

Brown, P. D., Tokuhisa, J. G., Reichelt, M., & Gershenzon, J. (2003). Variation of glucosinolate accumulation among different organs and developmental stages of Arabidopsis thaliana. *Phytochemistry, 62*(3), 471-481.

Busk, P. K., & Møller, B. L. (2002). Dhurrin synthesis in sorghum is regulated at the transcriptional level and induced by nitrogen fertilization in older plants. *Plant Physiology, 129*(3), 1222-1231.

Bussy, A. (1840). Sur la formation de l'huile essentielle de moutarde. *J Pharm, 27*, 464-471.

Byrne, D. N., & Bellows Jr, T. S. (1991). Whitefly biology. *Annual review of entomology, 36*(1), 431-457.

Byrne, D. N., Hendrix, D. L., & Williams III, L. H. (2003). Presence of trehalulose and other oligosaccharides in hemipteran honeydew, particularly Aleyrodidae. *Physiological Entomology, 28*(2), 144-149.

Byrne, D. N., & Miller, W. B. (1990). Carbohydrate and amino acid composition of phloem sap and honeydew produced by Bemisia tabaci. *Journal of Insect Physiology, 36*(6), 433-439.

Chen, B.-H., Wang, C.-C., Hou, Y.-H., Mao, Y.-C., & Yang, Y.-S. (2015). Mechanism of sulfotransferase pharmacogenetics in altered xenobiotic metabolism. *Expert opinion on drug metabolism & toxicology, 11*(7), 1053-1071.

Chen, W., Hasegawa, D. K., Kaur, N., Kliot, A., Pinheiro, P. V., Luan, J., . . . Sun, H. (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 biology, 14*(1), 1-15.

Cheung, P., Roberts, C., & Perkins Jr, H. (1980). Implications of disaccharides in sticky-cotton processing: honeydew contamination. *Textile Research Journal, 50*(1), 55-59.

Coley, P. D., Bryant, J. P., & Chapin, F. S. (1985). Resource availability and plant antiherbivore defense. *Science, 230*(4728), 895-899.

Conn, E. E. (1980). Cyanogenic compounds. *Annual Review of Plant Physiology, 31*(1), 433-451.

Cooke, R. D. (1978). An enzymatic assay for the total cyanide content of cassava (Manihot esculenta Crantz). *Journal of the Science of Food and Agriculture, 29*(4), 345-352.

Cooper-Driver, G., & Swain, T. (1976). Cyanogenic polymorphism in bracken in relation to herbivore predation. *Nature, 260*(5552), 604-604.

Cott, H. B. (1940). Adaptive coloration in animals.

Cristofoletti, P. T., Ribeiro, A. F., Deraison, C., Rahbé, Y., & Terra, W. R. (2003). Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid Acyrthosiphon pisum. *Journal of Insect Physiology, 49*(1), 11-24.

Croteau, R., Kutchan, T. M., & Lewis, N. G. (2000). Natural products (secondary metabolites). *Biochemistry and molecular biology of plants, 24*, 1250-1319.

Czapek, F. (1921). Spezielle Biochemie, Biochemie der Pflanzen, vol. 3, G. *Fischer Jena*, 369.

Darby, F., Heenan, M., & Smith, J. (1966). The absence of glucuronide conjugates from 1-naphthol dosed flies and grass grubs; Detection of 1-naphthylphosphate. *Life Sciences, 5*(16), 1499-1502.

Darwin, C. (1859). *The origin of species. 6th* (Vol. 570): John Murray, London.

Dauterman, W. C. (1983). Role of hydrolases and glutathione S-transferases in insecticide resistance *Pest resistance to pesticides* (pp. 229-247): Springer.

Davidson, E. W., Segura, B. J., Steele, T., & Hendrix, D. L. (1994). Microorganisms influence the composition of honeydew produced by the silverleaf whitefly, Bemisia argentifolii. *Journal of Insect Physiology, 40*(12), 1069-1076.

Dinsdale, A., Cook, L., Riginos, C., Buckley, Y., & De Barro, P. (2010). Refined global analysis of Bemisia tabaci (Hemiptera: Sternorrhyncha: Aleyrodoidea: Aleyrodidae) mitochondrial cytochrome oxidase 1 to identify species level genetic boundaries. *Annals of the Entomological Society of America, 103*(2), 196-208.

Douglas, A. (1993). The nutritional quality of phloem sap utilized by natural aphid populations. *Ecological Entomology, 18*(1), 31-38.

Douglas, A. (1998). Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria Buchnera. *Annual review of entomology, 43*(1), 17-37.

Douglas, A. (2006). Phloem-sap feeding by animals: problems and solutions. *Journal of Experimental Botany, 57*(4), 747-754.

Du, L., Lykkesfeldt, J., Olsen, C. E., & Halkier, B. A. (1995). Involvement of cytochrome P450 in oxime production in glucosinolate biosynthesis as demonstrated by an in vitro microsomal enzyme system isolated from jasmonic acid-induced seedlings of Sinapis alba L. *Proceedings of the National Academy of Sciences, 92*(26), 12505-12509.

Duus, F. (1979). ComprehensiveOrganic Chemistry·, Thiocarbonyl Compounds. *Pergamon Press: Oxford, England, 3*, 376.

Eakteiman, G., Moses-Koch, R., Moshitzky, P., Mestre-Rincon, N., Vassao, D. G., Luck, K., . . . Morin, S. (2018). Targeting detoxification genes by phloem-mediated RNAi: a new approach for controlling phloem-feeding insect pests. *Insect Biochemistry and Molecular Biology, 100*, 10-21. doi:10.1016/j.ibmb.2018.05.008

Ehrlich, P. R., & Raven, P. H. (1964). Butterflies and plants: a study in coevolution. *Evolution*, 586-608.

Esch, T. H., & Tjallingii, W. (1990). Fine structure of aphid stylets in plant tissue *Insects-Plants' 89* (pp. 475-476).

Eto, M., Kishimot.K, Matsumura, K., Ohshita, N., & Oshima, Y. (1966). Studies on Saligenin Cyclic Phosphorus Esters with Insecticidal Activity .9. Derivatives of Phosphonic and Phosphonothionic Acids. *Agricultural and Biological Chemistry, 30*(2), 181-+.

Falk, K. L., & Gershenzon, J. (2007). The desert locust, Schistocerca gregaria, detoxifies the glucosinolates of Schouwia purpurea by desulfation. *Journal of Chemical Ecology, 33*(8), 1542-1555.

Fan, T. W.-M., & Conn, E. E. (1985). Isolation and characterization of two cyanogenic β-glucosidases from flax seeds. *Archives of Biochemistry and Biophysics, 243*(2), 361-373.

Ferreira, C., Torres, B. B., & Terra, W. R. (1998). Substrate specificities of midgut β-glycosidases from insects of different orders. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, 119*(1), 219-225.

Firdaus, S., van Heusden, A. W., Hidayati, N., Supena, E. D. J., Mumm, R., de Vos, R. C., . . . Vosman, B. (2013). Identification and QTL mapping of whitefly resistance components in Solanum galapagense. *Theoretical and applied genetics, 126*(6), 1487-1501.

Gachon, C. M., Langlois-Meurinne, M., Henry, Y., & Saindrenan, P. (2005). Transcriptional co-regulation of secondary metabolism enzymes in Arabidopsis: functional and evolutionary implications. *Plant molecular biology, 58*(2), 229-245.

Gatehouse, J. A. (2002). Plant resistance towards insect herbivores: a dynamic interaction. *New phytologist, 156*(2), 145-169.

Gennadius, P. (1889). Disease of tobacco plantations in the Trikonia. *The aleurodid of tobacco. Ellenike Georgia, 5*, 1-3.

Geu-Flores, F., Nielsen, M. T., Nafisi, M., Møldrup, M. E., Olsen, C. E., Motawia, M. S., & Halkier, B. A. (2009). Glucosinolate engineering identifies a γ-glutamyl peptidase. *Nature Chemical Biology, 5*(8), 575-577.

Gibson, G. G., & Skett, P. (2013). *Introduction to drug metabolism*: Springer.

Gidda, S., & Varin, L. (2006). Biochemical and molecular characterization of flavonoid 7-sulfotransferase from Arabidopsis thaliana. *Plant Physiology and Biochemistry, 44*(11-12), 628-636.

Giordanengo, P., Brunissen, L., Rusterucci, C., Vincent, C., van Bel, A., Dinant, S., . . . Bonnemain, J.-L. (2010). Compatible plant-aphid interactions: how aphids manipulate plant responses. *Comptes rendus biologies, 333*(6-7), 516-523.

Goławska, S., & Łukasik, I. (2012). Antifeedant activity of luteolin and genistein against the pea aphid, Acyrthosiphon pisum. *Journal of Pest Science, 85*(4), 443-450.

Grant, D. (1991). Detoxification pathways in the liver *Journal of inherited metabolic disease* (pp. 421-430): Springer.

Grimaldi, D., & Engel, M. S. (2005). *Evolution of the Insects*: Cambridge University Press.

Guinard, L. (1890). *Recherches sur la localisation des principes actifs des Crucifères*: J. Mirsch.

Halkier, B. A., & Du, L. (1997). The biosynthesis of glucosinolates. *Trends in Plant Science, 2*(11), 425-431.

Halkier, B. A., & Gershenzon, J. (2006). Biology and biochemistry of glucosinolates. *Annu. Rev. Plant Biol., 57*, 303-333.

Hansen, A. K., & Moran, N. A. (2011). Aphid genome expression reveals host–symbiont cooperation in the production of amino acids. *Proceedings of the National Academy of Sciences, 108*(7), 2849-2854.

Hao, P., Liu, C., Wang, Y., Chen, R., Tang, M., Du, B., . . . He, G. (2008). Herbivore-induced callose deposition on the sieve plates of rice: an important mechanism for host resistance. *Plant Physiology, 146*(4), 1810-1820.

Haritatos, E., Keller, F., & Turgeon, R. (1996). Raffinose oligosaccharide concentrations measured in individual cell and tissue types in Cucumis melo L. leaves: implications for phloem loading. *Planta, 198*(4), 614-622.

Hayashi, H., & Chino, M. (1990). Chemical composition of phloem sap from the uppermost internode of the rice plant. *Plant and Cell Physiology, 31*(2), 247-251.

Heil, M. (2009). Damaged-self recognition in plant herbivore defence. *Trends in Plant Science, 14*(7), 356-363.

Hendrix, D. L., & Salvucci, M. E. (2001). Isobemisiose: an unusual trisaccharide abundant in the silverleaf whitefly, Bemisia argentifolii. *Journal of Insect Physiology, 47*(4-5), 423-432.

Hendrix, D. L., & Wei, Y.-a. (1994). Bemisiose: an unusual trisaccharide in Bemisia honeydew.

Hendrix, D. L., Wei, Y.-a., & Leggett, J. E. (1992). Homopteran honeydew sugar composition is determined by both the insect and plant species. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry, 101*(1-2), 23-27.

Hodgson, E. (1985). Microsomal monooxygenases. Comp insect physiol, biochem: Pergamon press Oxford, UK.

Hudes, G. R., Greenberg, R., Krigel, R. L., Fox, S., Scher, R., Litwin, S., . . . Comis, R. (1992). Phase II study of estramustine and vinblastine, two microtubule inhibitors, in hormone-refractory prostate cancer. *Journal of clinical oncology, 10*(11), 1754-1761.

Ioannides, C. (2002). *Enzyme systems that metabolise drugs and other xenobiotics*: J. Wiley.

Ishaaya, I. (1993). Insect detoxifying enzymes: their importance in pesticide synergism and resistance. *Archives of Insect Biochemistry and Physiology, 22*(1-2), 263-276.

Ishaaya, I., Ascher, K. R. S., & Casida, J. E. (1983). Pyrethroid Synergism by Esterase Inhibition in Spodoptera-Littoralis (Boisduval) Larvae. *Crop Protection, 2*(3), 335-343.

Ishaaya, I., & Casida, J. E. (1981). Pyrethroid Esterase(S) May Contribute to Natural Pyrethroid Tolerance of Larvae of the Common Green Lacewing (Neuroptera, Chrysopidae). *Environmental Entomology, 10*(5), 681-684.

Jacobson, A. L., Duffy, S., & Sseruwagi, P. (2018). Whitefly-transmitted viruses threatening cassava production in Africa. *Current Opinion in Virology, 33*, 167-176.

Jensen, R. A. (1976). Enzyme recruitment in evolution of new function. *Annual review of microbiology, 30*(1), 409-425.

Jing, X., White, T. A., Luan, J., Jiao, C., Fei, Z., & Douglas, A. E. (2016). Evolutionary conservation of candidate osmoregulation genes in plant phloem sap-feeding insects. *Insect Molecular Biology, 25*(3), 251-258.

Jones, P., & Vogt, T. (2001). Glycosyltransferases in secondary plant metabolism: tranquilizers and stimulant controllers. *Planta, 213*(2), 164-174.

Jørgensen, K., Morant, A. V., Morant, M., Jensen, N. B., Olsen, C. E., Kannangara, R., . . . Bak, S. (2011). Biosynthesis of the cyanogenic glucosides linamarin and lotaustralin in cassava: isolation, biochemical characterization, and expression pattern of CYP71E7, the oxime-metabolizing cytochrome P450 enzyme. *Plant Physiology, 155*(1), 282-292.

Kannangara, R., Motawia, M. S., Hansen, N. K., Paquette, S. M., Olsen, C. E., Møller, B. L., & Jørgensen, K. (2011). Characterization and expression profile of two UDP-glucosyltransferases, UGT85K4 and UGT85K5, catalyzing the last step in cyanogenic glucoside biosynthesis in cassava. *The Plant Journal, 68*(2), 287-301.

Kim, J. H., & Jander, G. (2007). Myzus persicae (green peach aphid) feeding on Arabidopsis induces the formation of a deterrent indole glucosinolate. *The Plant Journal, 49*(6), 1008-1019.

Kim, J. H., Lee, B. W., Schroeder, F. C., & Jander, G. (2008). Identification of indole glucosinolate breakdown products with antifeedant effects on Myzus persicae (green peach aphid). *The Plant Journal, 54*(6), 1015-1026.

Koroleva, O. A., Davies, A., Deeken, R., Thorpe, M. R., Tomos, A. D., & Hedrich, R. (2000). Identification of a new glucosinolate-rich cell type in Arabidopsis flower stalk. *Plant Physiology, 124*(2), 599-608.

Kossel, A. (1891). Ueber die chemische Zusammensetzung der Zelle. *Du Bois-Reymond's Archiv/Arch Anat Physiol Physiol Abt*, 181-186.

Lattanzio, V., Arpaia, S., Cardinali, A., Di Venere, D., & Linsalata, V. (2000). Role of endogenous flavonoids in resistance mechanism of Vigna to aphids. *Journal of Agricultural and Food Chemistry, 48*(11), 5316-5320.

Lee, S. S., & Scott, J. G. (1989). Microsomal cytochrome P450 monooxygenases in the house fly (Musca domestica L.): biochemical changes associated with pyrethroid resistance and phenobarbital induction. *Pesticide Biochemistry and Physiology, 35*(1), 1-10.

Liska, D. J. (1998). The detoxification enzyme systems. *Altern Med Rev, 3*(3), 187-198.

Luna, E., Pastor, V., Robert, J., Flors, V., Mauch-Mani, B., & Ton, J. (2011). Callose deposition: a multifaceted plant defense response. *Molecular Plant-Microbe Interactions, 24*(2), 183-193.

Lüthy, B., & Matile, P. (1984). The mustard oil bomb: rectified analysis of the subcellular organisation of the myrosinase system. *Biochemie und Physiologie der Pflanzen, 179*(1-2), 5-12.

Malka, O., Shekhov, A., Reichelt, M., Gershenzon, J., Vassão, D. G., & Morin, S. (2016). Glucosinolate desulfation by the phloem-feeding insect Bemisia tabaci. *Journal of Chemical Ecology, 42*(3), 230-235.

Markovich, O., Kafle, D., Elbaz, M., Malitsky, S., Aharoni, A., Schwarzkopf, A., . . . Morin, S. (2013). Arabidopsis thaliana plants with different levels of aliphatic-and indolyl-glucosinolates affect host selection and performance of Bemisia tabaci. *Journal of Chemical Ecology, 39*(11-12), 1361-1372.

Martin, J. H., Mifsud, D., & Rapisarda, C. (2000). The whiteflies (Hemiptera: Aleyrodidae) of Europe and the Mediterranean basin. *Bulletin of entomological research, 90*(5), 407-448.

McMahon, J. M., White, W. L., & Sayre, R. T. (1995). Cyanogenesis in cassava (Manihot esculenta Crantz). *Journal of Experimental Botany, 46*(7), 731-741.

Medina-Ortega, K. J., & Walker, G. P. (2015). Faba bean forisomes can function in defence against generalist aphids. *Plant, cell & environment, 38*(6), 1167-1177.

Meyer, U. A. (1996). Overview of enzymes of drug metabolism. *Journal of pharmacokinetics and biopharmaceutics, 24*(5), 449-459.

Mikkelsen, M. D., Naur, P., & Halkier, B. A. (2004). Arabidopsis mutants in the C–S lyase of glucosinolate biosynthesis establish a critical role for indole-3-acetaldoxime in auxin homeostasis. *The Plant Journal, 37*(5), 770-777.

Mithöfer, A., & Boland, W. (2012). Plant defense against herbivores: chemical aspects. *Annual review of plant biology, 63*, 431-450.

Mitra, P. S., Basu, N. K., Basu, M., Chakraborty, S., Saha, T., & Owens, I. S. (2011). Regulated phosphorylation of a major UDP-glucuronosyltransferase isozyme by tyrosine kinases dictates endogenous substrate selection for detoxification. *Journal of Biological Chemistry, 286*(2), 1639-1648.

Mkpong, O. E., Yan, H., Chism, G., & Sayre, R. T. (1990). Purification, characterization, and localization of linamarase in cassava. *Plant Physiology, 93*(1), 176-181.

Moore, B. (1967). Hydrogen cyanide in the defensive secretions of larval Paropsini (Coleoptera: Chrysomelidae). *Australian Journal of Entomology, 6*(1), 36-38.

Morkunas, I., Woźniak, A., Formela, M., Marczak, Ł., Narożna, D., Borowiak-Sobkowiak, B., . . . Grimm, B. (2016). Pea aphid infestation induces changes in flavonoids, antioxidative defence, soluble sugars and sugar transporter expression in leaves of pea seedlings. *Protoplasma, 253*(4), 1063-1079.

Moulis, C., Joucla, G., Harrison, D., Fabre, E., Potocki-Veronese, G., Monsan, P., & Remaud-Simeon, M. (2006). Understanding the polymerization mechanism of glycoside-hydrolase family 70 glucansucrases. *Journal of Biological Chemistry, 281*(42), 31254-31267.

Naegeli, C. v. (1861). Über die Siebröhren von Cucurbita. SB bayer. *Akad. Wiss, 1*, 212-238.

Nahrstedt, A., & Davis, R. (1983). Occurrence, variation and biosynthesis of the cyanogenic glucosides linamarin and lotaustralin in species of the Heliconiini (Insecta: Lepidoptera). *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry, 75*(1), 65-73.

Nartey, F. (1968). Studies on cassava, Manihot utilissima Pohl—I. Cyanogenesis: the biosynthesis of linamarin and lotaustralin in etiolated seedlings. *Phytochemistry, 7*(8), 1307-1312.

Navas-Castillo, J., Fiallo-Olivé, E., & Sánchez-Campos, S. (2011). Emerging virus diseases transmitted by whiteflies. *Annual review of phytopathology, 49*, 219-248.

Nikolov, Z. L., Meagher, M. M., & Reilly, P. J. (1989). Kinetics, equilibria, and modeling of the formation of oligosaccharides from D-glucose with Aspergillus niger glucoamylases I and II. *Biotechnology and bioengineering, 34*(5), 694-704.

Nintemann, S. J., Hunziker, P., Andersen, T. G., Schulz, A., Burow, M., & Halkier, B. A. (2018). Localization of the glucosinolate biosynthetic enzymes reveals distinct spatial patterns for the biosynthesis of indole and aliphatic glucosinolates. *Physiologia Plantarum, 163*(2), 138-154. doi:10.1111/ppl.12672

Oliveira, M. R. V., Henneberry, T. J., & Anderson, P. (2001). History, current status, and collaborative research projects for *Bemisia tabaci*. *Crop Protection, 20*(9), 709-723. doi:Doi 10.1016/S0261-2194(01)00108-9

Olsen, L. R., Gabel-Jensen, C., Nielsen, P. A., Hansen, S. H., & Badolo, L. (2014). Identification of a Functional Homolog of the Mammalian CYP3A4 in Locusts. *Drug Metabolism and Disposition, 42*(7), 1153-1162.

Olsen, L. R., Gabel-Jensen, C., Wubshet, S. G., Kongstad, K. T., Janfelt, C., Badolo, L., & Hansen, S. H. (2016). Characterization of midazolam metabolism in locusts: the role of a CYP3A4-like enzyme in the formation of 1'-OH and 4-OH midazolam. *Xenobiotica, 46*(2), 99-107.

Opitz, S. E., Jensen, S. R., & Müller, C. (2010). Sequestration of glucosinolates and iridoid glucosides in sawfly species of the genus Athalia and their role in defense against ants. *Journal of Chemical Ecology, 36*(2), 148-157.

Palmer, I. S., El Olson, O., Halverson, A. W., Miller, R., & Smith, C. (1980). Isolation of factors in linseed oil meal protective against chronic selenosis in rats. *The Journal of nutrition, 110*(1), 145-150.

Peche, K. (1913). Mikrochemischer Nachweis des Myrosins. *Ber. Deut. Bot. Ges, 31*, 458-462.

Peiser, G. D., Wang, T.-T., Hoffman, N. E., Yang, S. F., Liu, H.-w., & Walsh, C. T. (1984). Formation of cyanide from carbon 1 of 1-aminocyclopropane-1-carboxylic acid during its conversion to ethylene. *Proceedings of the National Academy of Sciences, 81*(10), 3059-3063.

Pentzold, S., Zagrobelny, M., Rook, F., & Bak, S. (2014). How insects overcome two-component plant chemical defence: plant ß-glucosidases as the main target for herbivore adaptation. *Biological Reviews, 89*(3), 531-551. doi:10.1111/brv.12066

Pfalz, M., Vogel, H., & Kroymann, J. (2009). The gene controlling the indole glucosinolate modifier1 quantitative trait locus alters indole glucosinolate structures and aphid resistance in Arabidopsis. *The Plant Cell, 21*(3), 985-999.

Piotrowski, M., Schemenewitz, A., Lopukhina, A., Müller, A., Janowitz, T., Weiler, E. W., & Oecking, C. (2004). Desulfoglucosinolate sulfotransferases from Arabidopsis thaliana catalyze the final step in the biosynthesis of the glucosinolate core structure. *Journal of Biological Chemistry, 279*(49), 50717-50725.

Price, D., Karley, A., Ashford, D., Isaacs, H., Pownall, M., Wilkinson, H., . . . Douglas, A. (2007). Molecular characterisation of a candidate gut sucrase in the pea aphid, Acyrthosiphon pisum. *Insect Biochemistry and Molecular Biology, 37*(4), 307-317.

Purnapatre, K., Khattar, S. K., & Saini, K. S. (2008). Cytochrome P450s in the development of target-based anticancer drugs. *Cancer letters, 259*(1), 1-15.

Quaintance, A. L., & Banks, N. (1900). Contributions toward a monograph of the American Aleurodidae.

Rask, L., Andréasson, E., Ekbom, B., Eriksson, S., Pontoppidan, B., & Meijer, J. (2000). Myrosinase: gene family evolution and herbivore defense in Brassicaceae. *Plant molecular biology, 42*(1), 93-114.

Ratzka, A., Vogel, H., Kliebenstein, D. J., Mitchell-Olds, T., & Kroymann, J. (2002). Disarming the mustard oil bomb. *Proceedings of the National Academy of Sciences, 99*(17), 11223-11228.

Raven, J. A. (1983). Phytophages of xylem and phloem: a comparison of animal and plant sap-feeders *Advances in Ecological Research* (Vol. 13, pp. 135-234): Elsevier.

Rhimi, M., Haser, R., & Aghajari, N. (2008). Bacterial sucrose isomerases: properties and structural studies. *Biologia, 63*(6), 1020.

Robyt, J. F., Yoon, S.-H., & Mukerjea, R. (2008). Dextransucrase and the mechanism for dextran biosynthesis. *Carbohydrate Research, 343*(18), 3039-3048.

Roth, J. A. (1986). Sulfoconjugation: role in neurotransmitter and secretory protein activity. *Trends in Pharmacological Sciences, 7*, 404-407.

Saheed, S., Larsson, K. A., Delp, G., Botha, C., Jonsson, L., & Bradley, G. (2007). Wound callose synthesis in response to Russian wheat aphid and bird cherry-oat aphid feeding on barley cv Clipper. *South African Journal of Botany, 73*(2), 310-310.

Salvucci, M. E. (2003). Distinct sucrose isomerases catalyze trehalulose synthesis in whiteflies, Bemisia argentifolii, and Erwinia rhapontici. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, 135*(2), 385-395.

Salvucci, M. E., Wolfe, G. R., & Hendrix, D. L. (1997). Effect of sucrose concentration on carbohydrate metabolism in Bemisia argentifolii: biochemical mechanism and physiological role for trehalulose synthesis in the silverleaf whitefly. *Journal of Insect Physiology, 43*(5), 457-464.

Samuels, M., & Howe, C. (1970). Vinblastine in the management of testicular cancer. *Cancer, 25*(5), 1009-1017.

Scanlan, J. L., Gledhill-Smith, R. S., Battlay, P., & Robin, C. (2020). Genomic and transcriptomic analyses in Drosophila suggest that the ecdysteroid kinase-like (EcKL) gene family encodes the 'detoxification-by-phosphorylation'enzymes of insects. *Insect Biochemistry and Molecular Biology*, 103429.

Selmar, D., Lieberei, R., & Biehl, B. (1988). Mobilization and Utilization of Cyanogenic Glycosides - the Linustatin Pathway. *Plant Physiology, 86*(3), 711-716. doi:DOI 10.1104/pp.86.3.711

Selmar, D., Lieberei, R., & Biehl, B. (1988). Mobilization and utilization of cyanogenic glycosides: the linustatin pathway. *Plant Physiology, 86*(3), 711-716.

Selmar, D., Lieberei, R., Biehl, B., Nahrstedt, A., Schmidtmann, V., & Wray, V. (1987). Occurrence of the cyanogen linustatin in Hevea brasiliensis. *Phytochemistry, 26*(8), 2400-2401.

Sharma, S., Kooner, R., & Arora, R. (2017). Insect pests and crop losses *Breeding insect resistant crops for sustainable agriculture* (pp. 45-66): Springer.

Smith Jr, C. R., Weisleder, D., Miller, R. W., Palmer, I. S., & Olson, O. E. (1980). Linustatin and neolinustatin: cyanogenic glycosides of linseed meal that protect animals against selenium toxicity. *The Journal of Organic Chemistry, 45*(3), 507-510.

Soler, R., Badenes-Pérez, F. R., Broekgaarden, C., Zheng, S. J., David, A., Boland, W., & Dicke, M. (2012). Plant-mediated facilitation between a leaf-feeding and a phloem-feeding insect in a brassicaceous plant: from insect performance to gene transcription. *Functional Ecology, 26*(1), 156-166.

Sparks, T. C., Hahn, D. R., & Garizi, N. V. (2017). Natural products, their derivatives, mimics and synthetic equivalents: role in agrochemical discovery. *Pest Management Science, 73*(4), 700-715.

Srivastava, V., & Thakre, R. P. (2000). Synergism and antagonism among the fungi grown in honeydew secretion on leaf surface of Nagpur mandarin. *Indian Phytopathology, 53*(2), 193-196.

Srivastava, V. K., & Tuteja, N. (2014). Calcium powered phloem protein of SEO gene family "Forisome" functions in wound sealing and act as biomimetic smart materials. *Plant signaling & behavior, 9*(9), e29438.

Stauber, E. J., Kuczka, P., Van Ohlen, M., Vogt, B., Janowitz, T., Piotrowski, M., . . . Wittstock, U. (2012). Turning the 'mustard oil bomb'into a 'cyanide bomb': aromatic glucosinolate metabolism in a specialist insect herbivore. *Plos One, 7*(4), e35545.

Szwergold, B. S., Kappler, F., & Brown, T. R. (1990). Identification of Fructose 3-Phosphate in the Lens of Diabetic Rats. *Science, 247*(4941), 451-454.

Szwergold, B. S., Kappler, F., Brown, T. R., Pfeffer, P., & Osman, S. F. (1989). Identification of D-Sorbitol 3-Phosphate in the Normal and Diabetic Mammalian Lens. *Journal of Biological Chemistry, 264*(16), 9278-9282.

Tang, T., Zhao, C., Feng, X., Liu, X., & Qiu, L. (2012). Knockdown of several components of cytochrome P450 enzyme systems by RNA interference enhances the susceptibility of Helicoverpa armigera to fenvalerate. *Pest Management Science, 68*(11), 1501-1511.

Tjallingii, W., & Esch, T. H. (1993). Fine structure of aphid stylet routes in plant tissues in correlation with EPG signals. *Physiological Entomology, 18*(3), 317-328.

Tripathi, M., & Mishra, A. (2007). Glucosinolates in animal nutrition: A review. *Animal feed science and technology, 132*(1-2), 1-27.

Tzin, V., Yang, X. W., Jing, X. F., Zhang, K., Jander, G., & Douglas, A. E. (2015). RNA interference against gut osmoregulatory genes in phloem-feeding insects. *Journal of Insect Physiology, 79*, 105-112. doi:10.1016/j.jinsphys.2015.06.006

Ünligil, U. M., & Rini, J. M. (2000). Glycosyltransferase structure and mechanism. *Current opinion in structural biology, 10*(5), 510-517.

van Ohlen, M., Herfurth, A.-M., Kerbstadt, H., & Wittstock, U. (2016). Cyanide detoxification in an insect herbivore: Molecular identification of β-cyanoalanine synthases from Pieris rapae. *Insect Biochemistry and Molecular Biology, 70*, 99-110.

Vassão, D. G., Wielsch, N., Gomes, A. M. d. M. M., Gebauer-Jung, S., Hupfer, Y., Svatoš, A., & Gershenzon, J. (2018). Plant defensive β-glucosidases resist digestion and sustain activity in the gut of a lepidopteran herbivore. *Frontiers in Plant Science, 9*, 1389.

Visser, T. J., Kaptein, E., Glatt, H., Bartsch, I., Hagen, M., & Coughtrie, M. W. (1998). Characterization of thyroid hormone sulfotransferases. *Chemico-biological interactions, 109*(1-3), 279-291.

Walling, L. L. (2008). Avoiding effective defenses: strategies employed by phloem-feeding insects. *Plant Physiology, 146*(3), 859-866.

Wang, X. W., Zhao, Q. Y., Luan, J. B., Wang, Y. J., Yan, G. H., & Liu, S. S. (2012). Analysis of a native whitefly transcriptome and its sequence divergence with two invasive whitefly species. *Bmc Genomics, 13*. doi:Artn 529

10.1186/1471-2164-13-529

Wang, Y., Spitz, M. R., Tsou, A. M.-H., Zhang, K., Makan, N., & Wu, X. (2002). Sulfotransferase (SULT) 1A1 polymorphism as a predisposition factor for lung cancer: a case-control analysis. *Lung Cancer, 35*(2), 137-142.

Wei, Y.-a., Hendrix, D. L., & Nieman, R. (1996). Isolation of a novel tetrasaccharide, bemisiotetrose, and glycine betaine from silverleaf whitefly honeydew. *Journal of Agricultural and Food Chemistry, 44*(10), 3214-3218.

Wei, Y. A., Hendrix, D. L., & Nieman, R. (1997). Diglucomelezitose, a novel pentasaccharide in silverleaf whitefly honeydew. *Journal of Agricultural and Food Chemistry, 45*(9), 3481-3486. doi:DOI 10.1021/jf970228e

White, W. L., & Sayre, R. T. (1992). *Partial purification and characterization of hydroxynitrile lyase from cassava.* Paper presented at the Roca, William M.; Thro, Ann Marie (eds.). International Scientific Meeting Cassava Biotechnology Network (1, 1992, Cartagena de Indias, Colombia). Proceedings.

Wilkinson, C., & Brattsten, L. (1972). Microsomal drug metabolizing enzymes in insects. *Drug metabolism reviews, 1*(1), 153-227.

Wilkinson, T., Ashford, D., Pritchard, J., & Douglas, A. (1997). Honeydew sugars and osmoregulation in the pea aphid Acyrthosiphon pisum. *Journal of Experimental Biology, 200*(15), 2137-2143.

Wilkinson, T., Koga, R., & Fukatsu, T. (2007). Role of host nutrition in symbiont regulation: impact of dietary nitrogen on proliferation of obligate and facultative bacterial endosymbionts of the pea aphid Acyrthosiphon pisum. *Applied and environmental microbiology, 73*(4), 1362-1366.

Will, T., Tjallingii, W. F., Thönnessen, A., & van Bel, A. J. (2007). Molecular sabotage of plant defense by aphid saliva. *Proceedings of the National Academy of Sciences, 104*(25), 10536-10541.

Williams, R. T. (1959). *Detoxication mechanisms: the metabolism and detoxication of drugs, toxic substances, and other organic compounds*: Wiley.

Winde, I., & Wittstock, U. (2011). Insect herbivore counteradaptations to the plant glucosinolate–myrosinase system. *Phytochemistry, 72*(13), 1566-1575.

Winter, J. M., & Tang, Y. (2012). Synthetic biological approaches to natural product biosynthesis. *Current opinion in biotechnology, 23*(5), 736-743.

Wittstock, U., Agerbirk, N., Stauber, E. J., Olsen, C. E., Hippler, M., Mitchell-Olds, T., . . . Vogel, H. (2004). Successful herbivore attack due to metabolic diversion of a plant chemical defense. *Proceedings of the National Academy of Sciences, 101*(14), 4859-4864.

Wouters, F. C., Reichelt, M., Glauser, G., Bauer, E., Erb, M., Gershenzon, J., & Vassao, D. G. (2014). Reglucosylation of the benzoxazinoid DIMBOA with inversion of stereochemical configuration is a detoxification strategy in Lepidopteran herbivores. *Angewandte Chemie-International Edition, 53*(42), 11320-11324. doi:10.1002/anie.201406643

Xu, J., De Barro, P., & Liu, S. (2010). Reproductive incompatibility among genetic groups of Bemisia tabaci supports the proposition that the whitefly is a cryptic species complex. *Bulletin of entomological research, 100*(3), 359.

Zagrobelny, M., Bak, S., & Møller, B. L. (2008). Cyanogenesis in plants and arthropods. *Phytochemistry, 69*(7), 1457-1468.

Zhang, X., Dong, J., Wu, H., Zhang, H., Zhang, J., & Ma, E. (2019). Knockdown of cytochrome P450 CYP6 family genes increases susceptibility to carbamates and pyrethroids in the migratory locust, Locusta migratoria. *Chemosphere, 223*, 48-57.

Zubair, M., Khan, M. Z., Rauf, I., Raza, A., Shah, A. H., Hassan, I., . . . Mansoor, S. (2020). Artificial micro RNA (amiRNA)-mediated resistance against whitefly (Bemisia tabaci) targeting three genes. *Crop Protection, 137*, 105308.

Acknowledgements

Firstly, and it goes without saying, that I am extremely grateful to **Jonathan Gershenzon**, and **Daniel Vassão** for not only selecting me for this project, but supporting me throughout the entirety of my thesis. We have certainly shared many memories together, from my first email to **Jonathan** asking whether it would be rude if I applied to this program having not completed my Bachelors yet, to dancing together in Africa at the inception meeting of the Cassava Whitefly Project. You both have been an inspiration for me in so many ways and have further helped to develop my scientific skills and given me confidence in my scientific inquiries. Even more than that, you have both been a great support in dealing with the difficulties that have arisen during my thesis and I have always felt very well backed throughout the many trying days. Thank you **Daniel** for also for giving me the freedom to explore avenues that truly interested me, and providing guidance in the pursuit of my own research endeavours. The opportunity to perform my PhD thesis in this lab has been one of absolute privilege, in terms of both opportunity for personal growth, and scientific growth.

I would also like to thank **Michael Reichelt**, for embodying the knowledgeable and passionate teacher in analytics that everyone is lost without. Your experience and mentorship have allowed me to become extremely confident in operation of various analytical platforms and your door was always open for questions and advice. I am sure you know that you are a pillar of the department and it is always a pleasure to work with, and learn from you!

Angela Schneider. You have been invaluable to me and helpful in every obstacle that I have come across while adjusting to German culture or navigating my way through countless forms. When you could, you would accompany me to the immigration office, and we might even go for lunch after, which was always a great comfort and a joy to have your assistance. Thank you so much for all that you do around the department.

I would also like to thank **Katrin Luck** and **Bettina Raguschke** for their input and assistance in the performance of various molecular tasks from basic cloning when my hands were seemingly unworthy to amplify a gene, to sequencing my successful (and unsuccessful) transformations.

Osnat Malka and **Shai Morin**…although we have certainly had our differences and there was a storm in our collaboration's seas, you were an integral part of my coming to Jena and pursuing my graduate work here on whiteflies, so I thank you for that!

Martin Niebergall. What can I say? When it comes to anything and I mean anything technical that is troubling myself, or any one of my fellow colleagues, you are always ready to swoop in and save the day. You have made so many things less stressful, simply by involving yourself and putting any worries to rest. Thank you so much for your steadfast and endless support.

Thank you to my **office mates** (both past and present) who have provided me with useful scientific discussion and daily conversation which has been undoubtedly helpful when the days would drag. Furthermore, I would like to extend my thanks and gratitude to the entire **biochemistry department** at the MPI-CE, for providing a welcoming and warm environment to work and grow. Thank you to the current members who I still see in the hallways everyday, and those that have past through and gone on to the next challenge in life. A special thanks goes to those that have assisted me in my scientific and thesis edits (**Felix Feistel, Bhawana Israni**).

Deep thanks to **Andrew O'Donnell**, **Anton Shekov**, **Benjamin Bartels**, **Wiebke Haegar**, **Joeren Carstens** for being the warmest of friends throughout my thesis. Whether we were enjoying a beer or six, dancing, crawling through dungeons, or watching a movie after a tiring week, your impact on my life is significant, and will always be near and dear to me. For everyone that I have not mentioned by name here and have contributed in some way to my doctoral studies or life, thank you very much.

I would like to thank my mother, **Pamela Easson**. You have instilled in me a deep sense for the natural world, which has gradually taken root in my life and become the tree from which I derive all of my confidence and structure. You have supported me wholly; my entire life and have always had endless confidence in my abilities as a person, and in the end as an academic. I would not be half the man I am today if it was not for you, and your support will never be taken for granted. I love you and thank you for everything I have.

Finally, I would like to thank my love **Melanie Werlich**. Although it took some time for us to find each other despite working in the same department for so long, the wait

was well worth it. You make everyday a joy and inspire me to be a better person, enhancing my life in almost every way imaginable. I have never been happier before finding you and I look forward to the rest of life's adventures, with you by my side.

Eigenständigkeitserklärung

Hiermit erkläre ich, dass mir die geltende Promotionsordnung (§ 5 PromO) der Biologisch-Pharmazeutischen Fakultät der Friedrich-Schiller-Universität Jena bekannt ist. Die vorliegende Arbeit wurde selbständig und ohne unzulässige Hilfesmittel oder Benutzung anderer als der angegebenen Hilfsmittel angefertigt. Übernommene Inhalte aus anderen Quellen und von anderen Personen, die in dieser Arbeit Verwendung fanden oder auf welche direkt Bezug genommen wird, sind als solche eindeutig kenntlich gemacht.

Ich versichere, dass weder ein Promotionsberater in Anspruch genommen wurde, noch das Dritte unmittelbar noch mittelbar von mir geldwerte Leistungen für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die vorgelegte Arbeit wurde zu keinem früheren Zeitpunkt, weder im Inland noch im Ausland in gleicher oder ähnlicher Form bei einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens eingereicht.

Jena, 23.11.2020	Michael L.A.E. Easson
(Ort, Datum)	(Unterschrift der Verfasserin)

Curriculum Vitae

Michael L.A.E. Easson
Max Planck Institute for Chemical Ecology (Beutenberg Campus)
Hans-Knoell-Strasse 8
Jena, Thüringen, Germany D-07745
++49 (0)3641 57 0 Ext. 1330
measson@ice.mpg.de

Education

Bachelor of Science, April 2015	Biochemistry with Honours with completion of Thesis, Brock University, St.Catharines Ontario
Doctoral Thesis, Present	Thesis in Biochemistry currently enrolled, Max Planck Institute for Chemical Ecology, Jena Germany

Research Experience

Doctoral Thesis (Dr. rer. nat. of Biochemistry), Department of Biochemistry, Max Planck Institute for Chemical Ecology, Jena Germany (May 2015- Present)

Project: Sucrase-transglucosidases and the interplay between detoxification and osmoregulation in phloem feeding insects. Results of this work were published in a peer reviewed journal (*Nature Chemical Biology*), with two further papers in preparation.

Supervisor: Jonathan Gershenzon and Severin Sasso

Bachelor of Science Biochemistry Thesis, Department of Biological Sciences, Brock University, St. Catharines ON (September 2014- April 2015)

Project: Identification of two epidermally expressed genes involved in catharanthine and tabersonine biosynthesis in *Catharanthus roseus*. Results of this work were published in a peer reviewed journal (*PNAS)*

Supervisor: Dr. Vincenzo De Luca

NSERC Research Assistant, Department of Biological Sciences, Brock University, St.Catharines ON (May 2014-August 2014)

Project: Completion of the vindoline pathway in *Catharanthus roseus* with expression in yeast. Results of this work were published in a peer reviewed journal (*PNAS*)

Supervisor: Dr. Vincenzo De Luca

NSERC Research Assistant, Department of Biological Sciences, Brock University, St.Catharines ON (May 2013-August 2013)

Project: Mass screening of EMS mutagenized *Catharnthus roseus* plants for differences in secondary metabolic accumulation by thin layer chromatography. Resulted in the identification of alkaloid knockout plants as well as numerous other metabolic mutants. Results of this work were published in multiple peer reviewed journal (*Phytochemistry and Planta*)
Supervisor: Dr. Vincenzo De Luca

Research Volunteer, Department of Chemistry, Brock University, St. Catharines ON (February 2013-April 2013)
Project: Learning of more advanced techniques in organic synthesis under the supervision of a PhD Student.
Supervisor: Dr. Thomas Hudlicky

Research Profile

Scientific Skills

Analytical Chemistry

Extensive experience with HP/UPLC-MS/UV separation and analysis of complex mixtures in targeted and untargeted approaches, compound optimization and purification, as well as advanced techniques of separation such as Trapped Ion-Mobility Spectrometry (TIMS) analysis. I also have a basic understanding of NMR operation and its usage in structural elucidation of low-molecular weight compounds.

Molecular Biology

RNA- & DNA-isolation and purification, PCR, Real Time Quantitative PCR, Plasmid Cloning, DNA Sequence Analysis, DNA Gel Electrophoresis, SDS-Page-Electrophoresis and Western Blots, protein expression and purification in multiple systems (yeast, bacteria and insects cells), with accompanying enzyme kinetic analysis.

Chemistry

Experience in compound derivitization, and chemical transformation of natural products for structural elucidation or more efficient purification.

Languages

English (mother tongue)
German (A1.1 Advanced)

Computer Skills

High level of expertise in Microsoft Office (Word, Excel, PowerPoint), Adobe products (Illustrator, Photoshop), GraphPad, multiple mass spectrometry spectral visualization softwares.

Other

Drivers license (G)

Professional Development

May 2020	Metabolic engineering & Evolution of diversity, Lecture in: MINI-LECTURE SERIES on Natural Products Biochemistry, speaker: Dr. Sarah E. O'Connor, Prof. Jonathan Gershenzon
May 2020	Major chemical reactions in biosynthesis, Lecture in: MINI-LECTURE SERIES on Natural Products Biochemistry, speaker: Dr. Sarah E. O'Connor
May 2020	Plant terpenoids & Iridoids in your Life!, Lecture in: MINI-LECTURE SERIES on Natural Products Biochemistry, speaker: Dr. Sarah E. O'Connor, Prof. Jonathan Gershenzon
May 2020	Nitrogen-containing compounds in plants: Alkaloids and others, Lecture in: MINI-LECTURE SERIES on Natural Products Biochemistry, speaker: Dr. Sarah E. O'Connor, Prof. Jonathan Gershenzon
Feb 2020	The basics of light and fluorescence microscopy, speaker: Dr. Veit Grabe
Feb 2019	Excursion to Jena BioScience, speaker: Nadine Müller-Dittmann
Nov 2018	Guided Tour Bayer Crop Science, speaker: Dr. Klaus Raming
Oct 2016	RNA-Seq Workshop, speaker: Dr. Claudia Voelckel, Dr. Andrew Heidel

Refereed Journal Publications

Easson, M.L., Gershenzon, J., Vassão, D.G. (2021). Isomerization and oligomerization of dietary disaccharides by Bemisia tabaci transglucosidases. (In preparation).

Easson, M.L., Malka, O., Paetz, C., Hojná, A., Reichelt, M., Stein, B., Colvin, J., Winter, S., Morin, S., Gershenzon, J., Vassão, D.G. (2020). Cassava cyanogenic glucosides are detoxified via multiple pathways in the pest Bemisia tabaci. (In preparation).

Malka, O*., **Easson, M. L***., Paetz, C., Götz, M., Reichelt, M., Stein, B., ... & Mondaca, L. L. (2020). Glucosylation prevents plant defense activation in phloem-feeding insects. *Nature Chemical Biology*, 1-7.

Koerte, S., Keesey, I. W., **Easson, M. L.**, Gershenzon, J., Hansson, B. S., & Knaden, M. (2020). Variable dependency on associated yeast communities influences host range in Drosophila species. *Oikos*.

Kidd, T., **Easson, M. L**., Qu, Y., & De Luca, V. (2019). Inter-organ transport of secologanin allows assembly of monoterpenoid indole alkaloids in a Catharanthus roseus mutant. *Phytochemistry*, *159*, 119-126.

Qu, Y., **Easson, M. L**., Simionescu, R., Hajicek, J., Thamm, A. M., Salim, V., & De Luca, V. (2018). Solution of the multistep pathway for assembly of corynanthean, strychnos, iboga, and aspidosperma monoterpenoid indole alkaloids from 19E-geissoschizine. *Proceedings of the National Academy of Sciences*, *115*(12), 3180-3185.

Edge, A., Qu, Y., **Easson, M. L**., Thamm, A. M., Kim, K. H., & De Luca, V. (2018). A tabersonine 3-reductase Catharanthus roseus mutant accumulates vindoline pathway intermediates. *Planta*, *247*(1), 155-169.

Qu, Y., **Easson, M. L**., Froese, J., Simionescu, R., Hudlicky, T., & De Luca, V. (2015). Completion of the seven-step pathway from tabersonine to the anticancer drug precursor vindoline and its assembly in yeast. *Proceedings of the National Academy of Sciences*, 201501821.

* Denotes equal contribution and thus shared first author position

Professional Presentations

Oral Presentations

Easson M. (2019). A spoonful of sugar … makes the toxins go down: Novel polyglucosylation for detoxification of two-component defenses in phloem-feeding insects. Talk presented at Institute Symposium, Max Planck Institute for Chemical Ecology, Jena, DE

Easson M. (2018). The sweet life: Mechanistic parallels between osmoregulation and detoxification in Bemisia tabaci. Talk presented at 17th IMPRS Symposium, International Max Planck Research School, Dornburg, DE

Poster Presentations

Easson M., Malka O., Reichelt M., Paetz C., Stanisic A., Winter S., Morin S., Giddings Vassão D., Gershenzon J. (2018). When Life Gives you Sugar: Metabolic Fates of Activated Defenses in Phloem Feeding Insects. Poster presented at Institute Symposium, MPI für Chemische Ökologie, Jena, DE.

Easson M. (2017). When life gives you sugar: Metabolic fates of phytoanticipins in phloem feeding insects. Poster presented at 2017 ISCE, APACE, International Society of Chemical Ecology, Kyoto, JP.

Jeschke V., Shekhov A., **Easson M.**, Christoff Wouters F., Luck K., Seibel E., Secinti S., Manivannan A., Sun R., Pandit S.S., Gershenzon J., Giddings Vassão D. (2016). Detoxification and Mode of Action of Plant Defenses in Insect Herbivores. Poster presented at SAB Meeting 2016, MPI for Chemical Ecology, Jena, DE.

Easson, M., Malka, O., Stanisic, M., Van Brunschot, S., Reichelt, M., Paetz, C., Morin, S., Seal, S., Colvin, J., Gershenzon, J., Vassão, D.G. (2016). Biochemical Mechanisms for Whitefly Utilization of Cassava: Detoxification of Cyanogenic Glycosides. Poster presented at Cassava Whitefly Project meeting in Lilongwe, Malawi.

Easson, M., Malka, O., Reichelt, M., Morin, S., Gershenzon, J., Vassão, D.G. (2015). Metabolism of cassava compounds by *B. tabaci* species. Poster presented at cassava whitefly project meeting in Greewich, United Kingdom.

Fellowships, Scholarships and Awards

2018	IMPRS travel award for best talk at the 17th IMPRS Symposium
2016	Award for best poster presentation at Cassava Whitefly Project meeting in Lilongwe, Malawi.
2011, 2012, 2013, 2014	Brock University Entrance Bursary
2014	Natural Science and Engineering Research Council of Canada Undergraduate Research Award
2012, 2014	Deans Honour List
2012, 2014	Academic-All-Canadian Athlete
2013	Natural Science and Engineering Research Council Undergraduate Research Award
2011	Brock University Academic Enhanced Entrance Bursary

References

Dr. Jonathan Gershenzon
Department of Biochemistry
Max-Planck Institute for Chemical Ecology
Jena, Germany, 07745
Phone: +49 (0)3641 57 1300
Email: gershenzon@ice.mpg.de

Dr. Daniel Giddings Vassão
Department of Biochemistry
Max-Planck Institute for Chemical Ecology
Jena, Germany, 07745
Phone: +49 (0)3641 57 1333
Email: vassao@ice.mpg.de

Dr. Vincenzo De Luca, **Professor**
Department of Biological Sciences
Cairns Health and Bioscience Research Centre
Brock University
St. Catharines, ON Canada L2S 3A1
Phone: +1 (905) 688-5550 ext. 4554
Email: vdeluca@brocku.ca