TOMATO FLAVOR MOLECULES:
A STORY OF GUAIACOL AND GLYCOSYLATION

By

MELISSA HAMNER MAGEROY

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2011

Then you will understand what is right and just and fair—every good path.
For wisdom will enter your heart, and knowledge will be pleasant to your soul.
Discretion will protect you, and understanding will guard you.
~Proverbs 2:9-11

ACKNOWLEDGMENTS

I thank my parents and sisters for their love, support, and encouragement throughout life. I also thank my grandparents for instilling in me a love for learning and constant support and love. I thank my husband, Jon Mageroy, for opening my eyes to so many things, and his loving support.

I also want to acknowledge Eric Schmelz at the United States Department of Agriculture (USDA) for his help on the purification of petunia glycosides and Alisdair Fernie and Takayuki Tohge for their assistance in analyzing tomato glycosides. I want to thank my undergraduate advisor, Dr. James Shinkle, for first giving me the opportunity to do plant research and directing me to graduate school. I want to thank my lab members: Dawn Bies, Peter Bliss, Abbye Floystad, Charles Goulet, Yusuke Kamiyoshihara, Mark Taylor, and Denise Tieman for all of their help, support, and friendship. I especially want to thank Charles Goulet and Denise Tieman for all their guidance, assistance with experiments, and helpful conversations. I also want to thank my advisor Harry Klee for sharing his knowledge and scientific wisdom with me. I also thank the members of my committee: Kevin Folta and Andrew Hanson, who both allowed me to do rotations in their lab which contributed to publications, and Jon Stewart.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AADC	Aromatic L-amino acid decarboxylases
ADH	Alcohol dehydrogenase
AdoMet	*S*-adenosyl-*L*-methionine
CAPS	Cleaved amplified polymorphic sequence
CCD	Carotenoid cleavage dioxygenase
CCRC	Complex Carbohydrate Research Center
EST	Expressed sequence tag
FISH	Fluorescence *in situ* Hybridization
GC-MS	Gas chromatography mass spectroscopy
HPL	Hydroperoxide lyase
IL	Introgression line
LC-MS	Liquid chromatography mass spectroscopy
LOX	Lipoxygenase
MES	2-(N-morpholino)ethanesulfonic acid
MW	Molecular weight
OMT	*O*-methyltansferase
RFLP	Restriction fragment length polymorphism
PCR	Polymerase chain reaction
PG	Polygalacturonase
PSPG	Plant Secondary Product Glycosyltransferase Motif
QTL	Quantitative trait locus
UGT	UDP-dependent glycosyltransferases

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

TOMATO FLAVOR MOLECULES: A STORY OF GUAIACOL AND GLYCOSYLATION

By

MELISSA HAMNER MAGEROY

December 2011

Chair: Harry J. Klee
Major: Plant Molecular and Cellular Biology

O-Methyltransferases (OMT) are important enzymes responsible for synthesis of many small molecules, including lignin monomers, flavonoids, alkaloids, and aroma compounds. One such compound is guaiacol, a volatile molecule with a smoky aroma that contributes to tomato flavor. Little is known about the pathway and regulation of guaiacol synthesis. One possible synthetic route is via catechol methylation. We identified a tomato *O*-methyltransferase (CTOMT1) with homology to a *Nicotiana tabacum* catechol OMT. *CTOMT1* was cloned from *Solanum lycopersicum* cv. M82 and expressed in *E. coli*. Recombinant CTOMT1 enzyme preferentially methylated catechol, producing guaiacol. To validate the *in vivo* function of CTOMT1, gene expression was decreased or increased in transgenic *S. lycopersicum* plants. Suppression of *CTOMT1* resulted in significantly reduced fruit guaiacol emissions. *CTOMT1* overexpression resulted in slightly increased fruit guaiacol production, suggesting that catechol availability might limit guaiacol production. To test this hypothesis, wild type (WT) and *CTOMT1* overexpressing tomato pericarp discs were supplied with exogenously applied catechol. Guaiacol production increased in both WT and transgenic fruit discs, although to a much greater extent in *CTOMT1* overexpressing discs. Finally, we identified two

introgression lines, 10-1 and 10-1-1, with increased guaiacol and higher *CTOMT1* expression. Taken together, these data led us to conclude that CTOMT1 is a catechol-*O*-methyltransferase that produces guaiacol in tomato fruit.

Many of the small volatile molecules, including guaiacol, that contribute to tomato flavor are glycosylated. A large family of glycosyltransferases (UGTs) is known to glycosylate a broad group of secondary metabolites, including flavor compounds. Little is known about UGTs or the function of these glycosides in tomato. One of the largest pools of glycosylated flavor volatiles is 2-phenylethanol glycoside. In order to identify tomato UGTs that glycosylate 2-phenylethanol, we first characterized 2-phenylethanol glycoside. We found that the 2-phenylethanol aglycone was attached to a polysaccharide and perhaps a malonyl glucose moiety. To screen potential candidate UGTs, we co-expressed the 2-phenylethanol biosynthetic pathway with candidate UGTs through transient expression in *Nicotiana benthamiana*. One potential 2-phenylethanol UGT was identified, SGN-U578227. UGT candidates were also overexpressed and suppressed down in tomato fruit. However, no changes to volatiles in the 2-phenylethanol synthesis pathway were found.

CHAPTER 1
TOMATO FLAVOR

Uncovering Lost Flavor

Domestication of tomato (*Solanum lycopersicum*) by Native Americans most likely began in Central America before the arrival of Europeans (Tanksley, 2004). By the end of the 19th century, a wide range of cultivars, that are today generally referred to as 'heirloom varieties' were available. These varieties were diverse in color, size, and flavor (Bai and Lindhout, 2007). Today, most fresh market tomatoes are bland and exhibit 'green' flavor characteristics while lacking 'floral/fruit' ones (Baldwin *et al*., 2008).

This loss in flavor profile diversity is in part due to the economic pressure on breeders to select for characteristics that reduce production costs, such as increased yield, larger fruit size, and longer shelf-life. These traits tend to be negatively correlated with good flavor (Bai and Lindhout, 2007; Klee, 2010). Improving or even maintaining tomato flavor through breeding has proven to be a daunting task as flavor is a complex trait that is comprised of a mixture of sugars, acids, and volatiles (Baldwin et al., 2008; Tieman et al., 2006a). From a sensory perspective, the interaction and balance of these three components are important for creating good tomato flavor as flavor perception is a summation of both taste and smell. Sweetness has been shown to enhance the perception of volatiles with floral/fruity notes, while sourness heightens the perception greens notes (Baldwin et al., 1998). Although attempts to improve taste by increasing sugars and acids have shown some success (Jones and Scott, 1983), it has been difficult to overcome the correlation between small fruit size and increased sugar content (Klee, 2010).

Breeding for increased volatile production is an even greater challenge as volatiles are synthesized from multiple precursors, including amino acids, fatty acids, and carotenoids (Goff and Klee, 2006). Also, a large number of quantitative trait loci (QTLs) affecting their synthesis have been identified (Tieman *et al*., 2006a; Mathieu *et al*., 2008). To further complicate matters, flavor is also affected by environmental conditions, harvest maturity, and postharvest handling (Baldwin, 2002; Petro-Turza, 1986). In order to overcome these challenges, we must employ new tools in molecular breeding and methods for gene discovery (Klee, 2010). As we strive to recover lost flavor, we also have a great opportunity to explore gene regulation, enzyme function, and secondary metabolism.

Peeling Apart Flavor Volatiles

The first description of tomato fruit alcohols and acetaldehydes was made in 1934 (Gustafson). For the next twenty years, there was a continued interest in describing and identifying components that contribute to tomato aroma, but not until 1968 was there any real quantitative analysis of tomato volatiles (Johnson *et al*., 1968). Quantitative methods to analyze fresh tomato flavor volatiles were further developed by R.G. Buttery (Buttery *et al*., 1987; Buttery *et al*., 1988; Buttery *et al.* 1989). Over 400 different volatiles have been identified from tomato fruit (Buttery *et al.*, 1989; Petro-Turza, 1986). However, only about 30 of these volatiles are present at perceivable levels to humans (Table 1-1) (Tieman *et al*., 2006a). Approximately, 100 QTLs that impact volatiles and their precursors have been identified. Generally, genes associated with these QTLs encode regulatory factors controlling volatile biosynthetic pathways or enzymes in these pathways (Klee, 2010).

Volatile Biosynthesis

Tomato flavor volatiles are synthesized from diverse classes of molecules including: fatty acids, amino acids, and carotenoids (Tieman *et al*., 2006a). Many aroma volatiles increase during fruit ripening (Baldwin *et al*., 1991). The increase is thought to be due to changes in substrate availability and enzyme accumulation (Baldwin *et al*., 2000). For example, the synthesis of some volatile precursors, such as carotenoids, is known to be ripening regulated and occurs upon chromoplast differentiation (Bramley, 2002). Regulation of synthesis of enzymes involved in volatile synthesis is not well understood. However, comparisons between ripening mutants and their controls indicates that the ripening process induces changes in both mRNA and protein accumulation (Biggs *et al*., 1985). Although isolating volatile synthetic enzymes has been challenging due to the genetic redundancy and broad substrate specificity, the role for several important enzymes have been discovered (Klee, 2010).

Fatty Acid Derived Volatiles

Fatty Acid derived volatiles are characterized as having “tomato,” “green,” or “grassy,” odors (Goff and Klee, 2006). Examples of fatty acid derived volatiles are hexenal and *cis*-3-hexenol. Synthesis of these volatiles begins with the oxidation of C_{18} compounds, such as linoleic acid and linolenic acid, to form fatty acid hydroperoxides. This reaction is catalyzed by non-heme iron-containing dioxygenases named lipoxygenases (LOX). Although there are at least five LOX genes expressed in ripe fruit, so far only *LOXC* has been shown to have a significant impact on the production of C_6 volatiles in tomato fruit (Chen *et al.*, 2004).

Hydroperoxides formed by LOX can be further modified by hydroperoxide lyase (HPL). HPL cleaves at the hydroperoxide containing carbon to form an aldehyde and an

oxoacid (Riley *et al.*, 1996). Resulting aldehydes can be further reduced to form their corresponding alcohols (Matsui, 2006). In tomato, an alcohol dehydrogenase (ADH) has been shown to catalyze the synthesis of these C_6-alcohols (Speirs *et al.*, 1998)

Amino Acid Derived Volatiles

Amino acid precursors of aroma molecules include alanine, valine, leucine, isoleucine, and phenylalanine (Baldwin *et al*., 2000). Although the pathways are not well established, it is thought that branched chain amino acids contribute to the formation of volatiles such as 3-methylbutanal/ol and 2-methylbutanal/ol. Both the amino acids and their α-keto acids could serve as building blocks for volatile synthesis. It has been shown that application of both amino acids and α-keto acids to fruit pericarp disc stimulates synthesis of volatiles, although application of the keto acids leads to significantly higher rates of volatiles synthesis (Klee, 2010; Kochevenko *et al*., manuscript under review). A similar result has also been observed in *Cucumis melo* L. fruit (Gonda *et al*., 2010).

Much more is known about phenylalanine derived volatiles, which include phenylacetaldehyde, 2-phenylethanol, and 1-nitro-2-phenethane. These volatiles are described as fruity/floral (Tieman *et al.,* 2006b).The first and major flux-controlling step of the pathway to synthesis of these volatiles is catalyzed by a family of genes called aromatic L-amino acid decarboxylases (AADC). These enzymes convert phenylalanine to phenethylamine (Tieman *et al*., 2006b). In the 2-phenylethanol biosynthetic pathway, the final step from phenylacetaldehyde to 2-phenylethanol is catalyzed by a small family of aldehyde reductases (Tieman *et al*., 2007).

Carotenoid Derived Volatiles

Volatiles derived from carotenoids include: β-ionone, β-damascenone, geranlyacetone, and pseudoionone. These volatiles are also characterized as fruity/floral (Klee, 2010). Although these volatiles tend to be present at low levels in tomato fruit, they can have a great impact on flavor due to their low odor thresholds (Baldwin *et al*., 2000). Carotenoid cleavage dioxygenases (CCD) have been shown to be important in the formation of geranylacetone and β-ionone (Simkin *et al*., 2004). These enzymes are capable of cleaving carotenoids at the 5, 6, 7, 8, or 9, 10 positions to produce aldehydes and ketones (Vogel *et al*., 2008).

Continuing the Search

Although many volatile pathways have been identified, many are still unknown. Additionally, the understanding of how these pathways are regulated is still elemental (Klee, 2010). The scope of this work is to identify the biosynthetic pathway and enzymes that synthesize guaiacol. Also explored is a molecule modification process that may serve to regulate the pool of free volatiles.

Table 1-1. Volatiles that contribute to tomato aroma.

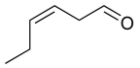
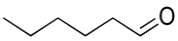
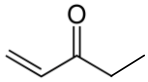
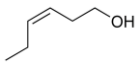
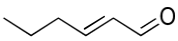
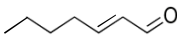
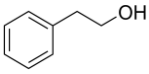
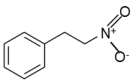
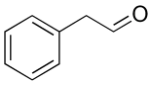
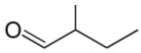
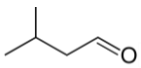
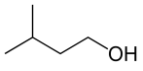
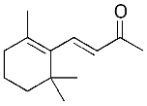
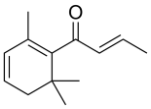
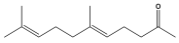
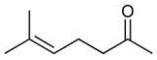
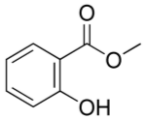
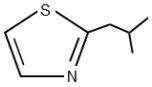
	Structure	Conc. (ppb)	Odor threshold (ppb)	Precursor	Description
cis-3-Hexenal		12000	0.25	Fatty acid	green, grassy
Hexanal		3000	5	Fatty acid	green, grassy
1-Penten-3-one		520	1	Fatty acid	green, citrus
cis-3-Hexenol		150	70	Fatty acid	green, leafy
trans-2-Hexenal		60	17	Fatty acid	green, grassy
trans-2-Heptenal		60	13	Fatty acid	green, grassy
2-Phenylethanol		1900	750	Phe	floral, rose
1-Nitro-2-phenylethane		17	2	Phe	floral, sweet
Phenylacetaldehyde		15	4	Phe	floral, honey
2+3-Methylbutanal		27	1	Ile	fruity, cocoa
			0.2	Leu	

Table 1-1. (Continued)

	Structure	Conc. (ppb)	Odor threshold (ppb)	Precursor	Description
3-Methylbutanol		380	120	Phe	fruity, wine
β-ionone		4	0.007	Carotenoid	floral, fruity
β-damascenone		1	0.002	Carotenoid	fruity, sweet
Geranylacetone		57	60	Carotenoid	sweet, citrus
6-Methyl-5-hepten-2-one		130	2000	Carotenoid	fruity, citrus
Methyl salicylate		48	40	Phe and/or chorismate	Winter-green
2-Isobutylthiazole		36	3.5	Unknown	musty

Adapted from Baldwin *et al*., 2000 and Goff and Klee *et al*., 2007

CHAPTER 2
GUAIACOL SYNTHIESIS: METHYLATION OF CATECHOL BY AN *O*-METHYLTRANSFERASE

Overview

The Flavor Molecule Guaiacol

Guaiacol (2-methoxyphenol) is found in many processed food products such as wine, roasted coffee, tea, cocoa, and food additives, like liquid smoke (Serra Bonvehí and Ventura Coll, 1998; Dorfner *et al*., 2003; Guillen *et al*., 1995; Hayasaka *et al*., 2010; Kumazawa and Masuda, 2002). Guaiacol is not commonly found in fresh fruits and vegetables, but is an important contributor to tomato flavor. Common flavor descriptors of guaiacol are medicinal or smoky (Álvarez-Rodríguez *et al*., 2003). Guaiacol has been described as an undesirable compound in many fruits, based on its medicinal-like aroma (Zierler *et al*., 2004; Zanor *et al*., 2009). However, guaiacol has not been well correlated with either liking or disliking in tomato fruit. Therefore, we were interested in identifying genes responsible for its synthesis to permit guaiacol content alterations . Little is known about how guaiacol is synthesized in plants. Based upon its structure, we deduced that it could be synthesized by a methylation of catechol.

Small Molecule Methyltransferases

Methylation of small molecules is catalyzed by methyltransferases that transfer a methyl group from *S*-adenosyl-*L*-methionine (AdoMet) to an acceptor molecule. Usually methyltransferases that methylate hydroxyl or carboxyl groups are known as *O*-methyltransferases (OMTs), but some methylate nitrogen and sulfur groups on small molecules (Noel *et al*., 2003). Many plant OMTs with important functions in phenylpropanoid biosynthesis have been identified. These enzymes synthesize secondary metabolites such as lignin, flavonoids and aroma molecules (Gang *et al*.,

2002). These aroma molecules include several volatiles that are important contributors to *Rosa chinensis* (China Rose), *Ocimum basilicum* (Basil), and *Clarkia breweri* (Fairy Fans) fragrance (Gang *et al*., 2002; Gang *et al*., 2005; Wang *et al*., 1997). Synthesis of the tomato flavor volatile methyl salicylate has also been shown to require enzymatic activity of an OMT. In this biosynthetic pathway, the carboxyl group of salicylic acid is methylated to produce methyl salicylate (Tieman *et al*., 2010).

Catechol: Discovery and Synthesis

We deduce that catechol is the direct precursor of guaiacol, based upon structurally similarities. Catechol (1,2-dihydroxybenzene) was first isolated by H. Reinsch in 1893 from *Mimosa catechu* by distilling “catechin”. Catechol is known to be found in fruits and vegetables. The browning observed in cut apples is in part due to the oxidation of catechol to benzoquinone (Matheis, 1983; Zheng *et al*., 2010).

In microorganisms, catechol is known to be synthesized from phenol, benzoic acid, salicylic acid, or 2,3-dihydroxybenzoic acid (Evans *et al*., 1951; Katagiri *et al*., 1962; Subba Rao *et al*., 1967). One well-characterized bacterial gene that converts salicylic acid to catechol is *nahG*, a salicylic acid hydroxylase (Gaffney *et al*., 1993; Yamamoto *et al*., 1965). Many studies in transgenic plants have demonstrated a great reduction in the salicylic acid content of plants expressing *nahG* (van Wees and Glazebrook, 2003; Boss *et al*., 2010). There has been little exploration of catechol synthesis in plants. One study, in *Tecoma stuns,* suggests that catechol is synthesized from anthranilic acid, while another, in *Gaultheria adenothrix,* suggests that it is synthesized from salicylic acid (Ellis and Towers, 1969).

Synthesis of Guaiacol

Guaiacol can be synthesized from catechol by the addition of a methyl group. Methylation of catechol is catalyzed by orthodiphenol-*O*-methyltransferases (Pellegrini *et al*., 1993). OMTs with activity on various orthodiphenolic substrates, including catechol, have previously been characterized in *Fragaria* x *ananassa, N. tabacum* and *O. basilicum* (Collendavelloo *et al*., 1981; Maury *et al*., 1999; Gang *et al*., 2002; Wein *et al*., 2002), although a direct *in vivo* role for any enzyme involved in guaiacol synthesis has not been demonstrated.

Investigation of Guaiacol Synthesis in Tomato

Here we describe both *in vitro* and *in vivo* approaches to identify a tomato catechol OMT. The catalytic activity of candidate OMTs on catechol was first determined using recombinant protein expressed in *E. coli*. This activity was then confirmed *in vivo* by transgene expression in tomato fruit.

Results

Identification of a Catechol O-Methyltransferase (CTOMT1) from *S. lycopersicum*

Potential *S. lycopersicum* catechol OMT candidates were selected by identification of coding sequences with similarity to previously characterized small molecule OMTs (Figure 2-1). Five candidate genes were selected, SGN-U582403, SGN-U565623, SGN-U319245, SGN-U575022, and SGN-U321686. Full length cDNAs were synthesized from *S. lycopersicum* cv. M82 ripe fruit RNA. Candidate genes were cloned into an expression vector in *E. coli*. Initial screens were performed by adding catechol directly to bacterial cultures expressing recombinant protein and measuring guaiacol production. Only SGN-U582403 converted catechol to guaiacol.

Specificity and Specific Activity

To further test the specificity of SGN-U582403 for catechol, the activity of SGN-U582403 on substrates with similar structure to catechol was measured. Recombinant enzyme purified from *E. coli* was incubated with [^{14}C]AdoMet and the following potential substrates: catechol, guaiacol, salicylic acid, benzoic acid, orcinol, caffeic acid, protocatechuic aldehyde, 2,5-dimethyl-4-methyoxy-3(2H)-furanone, and pyrogallol. SGN-U582403 had relatively high activity with catechol as a substrate, much lower activity on protocatechuic aldehyde and slight activity on orcinol, caffeic acid, and pyrogallol (Table 2-1).

The specific activity of SGN-U582403 with catechol as a substrate was also measured. Purified enzyme was incubated with excess [^{14}C]AdoMet and various concentrations of catechol. The enzyme was determined to have a K_m of 8.36 ± 1.78 µM and K_{cat} value of 9.67 ± 2.42 s^{-1} (Figure 2-2). Guaiacol was confirmed as the product by GC-MS. Based on these results, the gene encoding the SGN-U582403 protein was renamed CTOMT1.

Overexpression of CTOMT1 *in planta*

To further test the function of CTOMT1 *in planta,* a full-length *CTOMT1* cDNA was cloned into pHK1001 (Figure 2-3) for constitutive overexpression. The construct was transformed into *S. lycopersicum* cv. Flora-Dade. Seventeen independent lines were initially screened for transgene expression. Based on this screen, the four best overexpressing lines were further analyzed for *CTOMT1* mRNA levels (Figure 2-4A) and guaiacol synthesis in ripe fruits (Figure 2-4B). Guaiacol production was significantly increased in two of the four lines overexpressing *CTOMT1*. However, the increased guaiacol was not proportional to the increased RNA levels; up to a 26-fold increase in

transcript resulted in only a two-fold increase in guaiacol. A similar lack of correlation between *CTOMT1* transcript abundance and guaiacol production was observed throughout development of non-transgenic Flora-Dade fruit. Although *CTOMT1* expression was the highest in immature green fruit, guaiacol production was much higher in ripe fruit (Figure 2-5).

Suppression of CTOMT1 *in planta*

CTOMT1 was also cloned into pK2WG7 (Figure 2-3) (Karimi *et al*., 2002) for antisense knock-down and transformed into cv. Flora-Dade. Twenty-five lines were initially screened for supression of *CTOMT1* RNA using leaf tissue. The four lines with greatest RNA reduction were further screened for *CTOMT1* mRNA levels in ripe fruit (Figure 2-4C). Volatile emissions were also determined (Figure 2-4D). The guaiacol levels were significantly reduced in all four antisense lines, confirming the role of CTOMT1 in guaiacol synthesis.

Catechol Feeding of Fruit Pericarp Disc

While antisense lines in which *CTOMT1* levels were greatly reduced produced significantly less guaiacol than controls, over-expression of *CTOMT1* had much less effect on guaiacol levels. These results suggested that while CTOMT1 is essential for guaiacol synthesis, it may not be rate-limiting under normal circumstances. Rather, synthesis of catechol might limit the production of guaiacol in *CTOMT1*-overexpressing plants. We tested this hypothesis by feeding catechol to fruit pericarp discs of Flora-Dade (WT) and *CTOMT1*-overexpressing lines. Volatiles were collected after incubation for 4 h. Both WT and *CTOMT1* discs produced more guaiacol when supplied with exogenous catechol. However, while WT catechol-fed discs exhibited a 36-fold increase in guaiacol synthesis, *CTOMT1* discs exhibited a 52-fold increase (Table 2-2; Figure 2-

6). These results indicate that while CTOMT1 catalyzes the conversion of catechol to guaiacol, the availability of catechol likely limits guaiacol synthesis in fruit tissue.

Use of *nahG* Transgenic Plants for *in vivo* Analysis of Increase Catechol

We also observed that transgenic fruit overexpressing the salicylate hydrolase gene, *nahG*, produce 38-fold more guaiacol than the non-transgenic cv. Ailsa Craig control (data not shown). Historically, *nahG* plants have been used in pathogen response studies because much of the salicylic acid pool is converted to catechol, which accumulates (Gaffney *et al*., 1993; Van Wees and Glazebrook, 2003). *nahG* plants provide an *in vivo* confirmation that, when catechol levels are increased guaiacol production is also increased.

Discussion

We identified potential candidates by screening for catechol methylation with tomato homologs of previously characterized orthodiphenol OMTs. Of the five candidate *S. lycopersicum* proteins that were screened for catechol methylation activity, only CTOMT1 was capable of converting catechol to guaiacol, suggesting that this enzyme is solely responsible for guaiacol synthesis *in vivo*. The closest homolog of this protein in sequence databases (81% identity) is an enzyme with *in vitro* catechol-OMT activity from *N. tabacum* (Collendavelloo *et al.,* 1981; Maury *et al.,* 1999; Pellegrini *et al.,* 1993). The *N. tabacum CTOMT* gene is highly inducible by pathogen infection (Pellegrini *et al.,* 1993). *In vivo* effects on catechol and guaiacol pools have not been reported.

The activity of CTOMT1 on catechol was confirmed by recombinant enzyme assays. The K_m and K_{cat} values were similar to those of other characterized diphenol-*O*-methyltransferases (www.brenda-enzymes.org). While CTOMT1 preferentially methylates catechol, it does have some activity on protocatechuic aldehyde, oricinol ,

caffeic acid, and pyrogallol. All of these molecules have a similar basic structure of a benzene ring with at least two hydroxyl groups. CTOMT1 was unable to methylate molecules lacking two hydroxyl groups, indicating that the diphenol structure is important for substrate recognition. While plant OMTs usually have a high degree of selectivity, a few are promiscuous and catalyze methylation of structurally related compounds (Lam *et al.,* 2007, Wein *et al., 2002*). However, CTOMT1 exhibited a strong preference for catechol over other tested diphenol compounds.

In order to test if that CTOMT1 is a catechol OMT *in vivo*, its expression was increased or reduced in transgenic tomato plants. Suppression of the endogenous gene significantly reduced guaiacol emission, indicating that CTOMT1 is the major, if not only, enzyme responsible for guaiacol synthesis. Although there were high levels of *CTOMT1* expression in overexpressing plants, there were not correspondingly large increases in guaiacol production. It is probable that catechol levels limit guaiacol production, as normal endogenous levels of catechol must be low and high levels have been shown to be toxic to plants (Morse *et al.,* 2007; Van Wees and Glazebrook, 2003). This hypothesis is further supported by the fact that in Flora-Dade, *CTOMT1* expression decreases with ripening, while guaiacol production increases.

When we tested the hypothesis that catechol is limiting guaiacol production by supplying fruit pericarp discs with exogenous catechol, we were able to significantly increase guaiacol emission. Both WT and *CTOMT1*-overexpressing discs produced more guaiacol when supplied with non-limiting catechol. However, the increase in guaiacol synthesis was much greater in *CTOMT1*-overexpressing discs than in non-transgenic controls. These results indicate that under certain circumstances, *CTOMT1*

expression influences the rate of guaiacol synthesis but there must be a level of control that precedes catechol. This conclusion is further supported by the elevated guaiacol levels in *nahG* overexpressing plants, which have high catechol levels.

Materials and Methods

Phylogenetic Tree of Small Molecule Methyltransferases

S. lycopersicum OMT candidates were identified by a TBLASTN search of the sol genomic network Lycopersicon combined (tomato) unigene database using *O. basilicum* chavicol OMT and eugenol OMT amino acid sequences (Q93WU3; Q93WU2). Other similar proteins were identified by conducting a BLASTP search of the NCBI non-redundant protein sequences using candidate CTOMTs. Twenty three OMT amino acid sequences were used to generate a protein alignment. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The bootstrap test (1000 replicates) was used to calculate the percentage of replicate trees in which the associated taxa clustered together (Felsenstein, 1985). The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965). The analysis involved 25 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 306 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al., 2011).

CTOMT1 *in vitro* Expression and Purification

SGN-U582403 (*CTOMT1*) was PCR-amplified from *S. lycopersicum* and *S. pennellii* fruit cDNA. The products were cloned into pENTR/D/TOPO vectors and sequenced (CHUL Research Center, http://www.sequences.crchul.ulaval.ca). The coding regions were then cloned into vector pET160 by recombination and transformed into *E. coli* BL21-DE3 (Invitrogen, http://www.invitrogen.com) for inducible protein

expression. Bacteria were precultured for 16 h at 37ºC in Luria-Bertani broth containing 50 µg/mL carbenicillin and the culture was used to inoculate 100 mL of the same medium. Cells were grown at 24ºC to an OD_{600} of 0.5. Protein expression was induced by adding isopropyl-1-ß-D-thiogalactoside to the medium at a final concentration of 0.1 mM. Induced cultures continued growing at 25ºC for 16 h.

Cells were harvested by centrifugation (10 min, 4 420 g) and resuspended in 6 mL of lysis buffer (1x Phosphate-buffered saline (PBS)), lysozyme, 10% v/v glycerol, and Bacterial Protease Inhibitor Cocktail [Sigma, http://www.sigmaaldrich.com/]) and lysed with sonication. Protein was purified using Ni-Talon® (Clonetech, http://www.clontech.com/) affinity chromatography. The column was washed with 1X PBS containing 5 mM imidazole. Imidazole concentration was increased to 150 mM in the elution buffer. Protein levels were quantified using Bradford Reagent and Bovine Serum Albumin as a standard (BioRad, http://www.bio-rad.com/). Protein was stored in 16% glycerol at -80ºC.

Enzymatic Assay

For relative activity assays, 2.7 µg purified enzyme was assayed at 25ºC in a 100 µL reaction containing 50 mM Tris–HCl, pH 7.5, 100 mM KCl, 2.8 mM β-mercaptoethanol, 15 µM substrate, 10 mM AdoMet, 0.4 mM [methyl-^{14}C] AdoMet (specific activity 50.4 mCi mmol^{-1}; Amersham). Substrates were diluted in 50% ethanol. Assays were done in triplicate, including boiled enzyme controls. After 30 min at 25ºC the reactions were stopped by adding equal volumes of hexanes. The methylated substrate was extracted on a vortex mixer for 15 sec and centrifuging (5 min, 13 200 g). 50 µL of the organic layer was counted for 10 min in 3 ml Ready Gel Scintillation Fluid (Beckman Coulter, http://www.beckmancoulter.com). Counts for the boiled enzyme

controls were subtracted from the sample counts, and activity for catechol was normalized to 100%. For specific activity on catechol, procedures were the same as above, except for substrate concentrations. Concentrations used were: 0, 1, 5, 15, 25, and 50 µM. The conversion of catechol to guaiacol was validated by repeating the experiment using only unlabeled AdoMet and analyzing organic layer by GC-MS.

Production of Transgenic Plants

The full-length open reading frame of *CTOMT1* was cloned into a vector, pHK1001, containing the constitutive FMV 35S promoter (Richins *et al*., 1987) followed by the *nos* 3´ terminator, for overexpression. *S. lycopersicum* cv. Flora-Dade cotyledons were transformed by Agrobacterium-mediated transformation (McCormick *et al*., 1986) with the kanamycin selectable marker, NPTII. Antisense constructions were made by cloning a full length *CTOMT1* into pK2WG7 (Karimi *et al*., 2002). Antisense constructs were made by the Plant Transformation Core Research Facility at the University of Nebraska (http://unlcms.unl.edu/biotech/plant-transformation).

Volatile Collection

Volatiles were collected from tomato fruits according to Tieman *et al*. (2006a). One hundred grams of fruit was chopped and placed in a glass tube. Each tube was fitted with a rubber stopper. One end of the tube was attached to the volatile collection apparatus which allows for the regulation of air flow over the column. The other end was attached to a column containing SuperQ resin. Air was passed over the samples and volatiles were collected on a SuperQ Resin for 1 h. Five µL of nonyl acetate were added to each column as an internal control of column recovery. Volatiles were eluted off the column with methylene chloride and run on a GC/MS and GC for analysis as described in Tieman *et al*. (2006a).

Gene Expression Analysis

Tomato fruit was chopped and quickly frozen in liquid nitrogen. Samples were stored at -80º until further use. RNA was extracted using Plant RNeasy kit (Qiagen, http://www.qiagen.com). Possible genomic DNA contamination was removed by on column DNaseI treatment for 15 min at room temperature. Quantitative PCR was performed with StepOnePlus™ Real-Time PCR System using total amount of 325 ng total RNA, Taqman® 1-step kit (Applied Biosystems, http://www.appliedbiosystems.com), 500 nm forward and reverse primer. A total reaction volume of 25 µl was used. A standard curve was generated using pENTR-OMT1 ranging from 10^5 to 10^{10} copies per 5 µL.

***CTOMT1* Expression and Guaiacol Quantification through Flora-Dade Ripening**

RNA was extracted and volatiles were collected from Flora-Dade fruit at the following stages: immature green, mature, turning, and red ripe. RNA extraction and volatile collection were performed as above described.

Catechol Feeding of Fruit Pericarp Disc

Tomato discs were cut from pericarp tissue of ripe Flora-Dade and *CTOMT1*-overexpressing fruit using a size 10 borer. One hundred discs were used for each sample treatment. Discs were placed in Petri dishes and an “X” was cut in the top of each with a razor blade. 10 µl of either water or 1 M catechol dissolved in water were pipetted into each disc. Covers were placed on petri dishes and discs were left to incubate for 4 h. Discs were then placed in glass tubes and volatiles were extracted as previously described. Guaiacol was quantified by GC/MS using a guaiacol standard curve.

Table 2-1. Relative activity of CTOMT1 on substrates with similar structure to catechol.

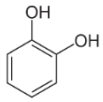
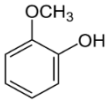
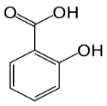
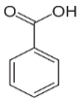
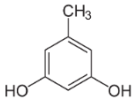
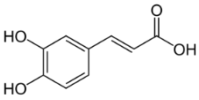
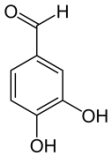
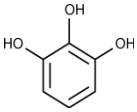
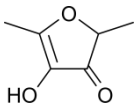
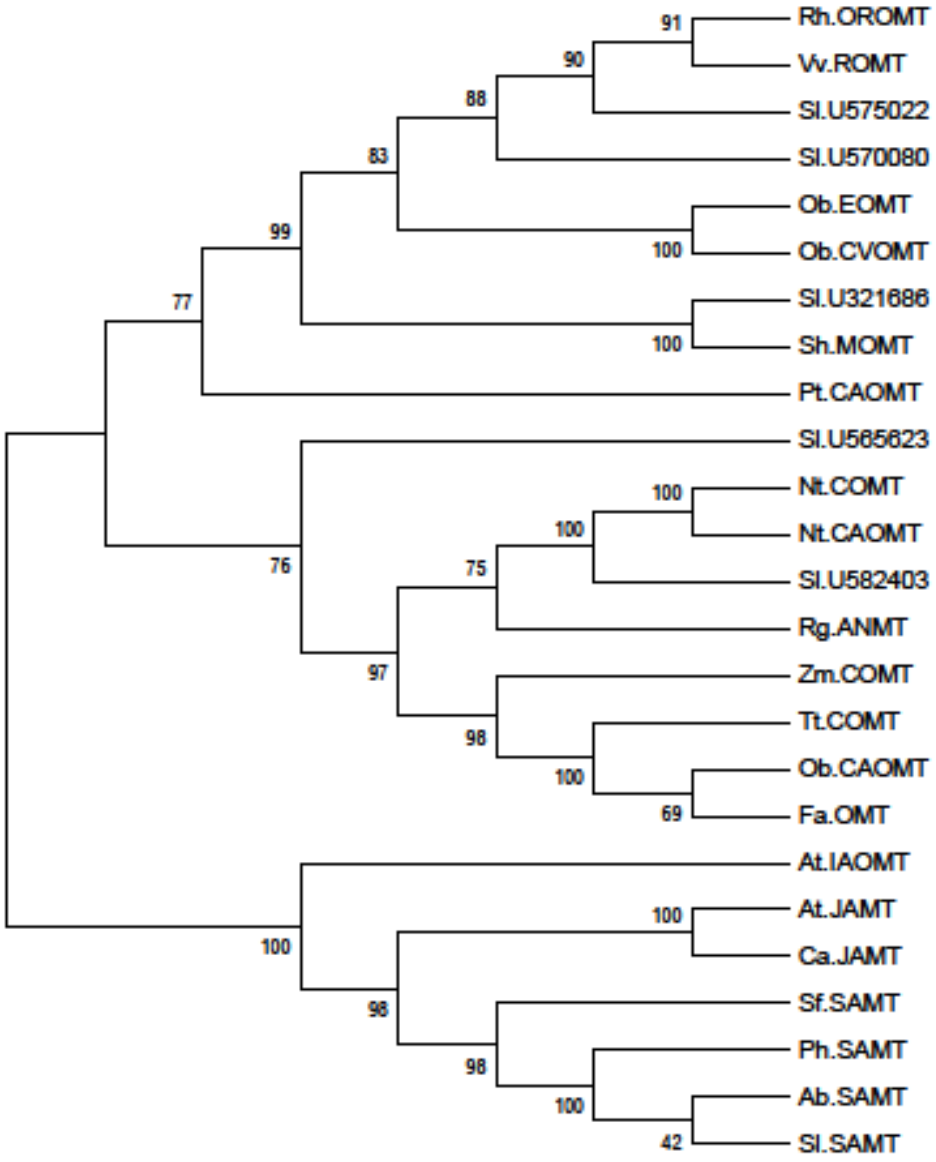
Substrate	Structure	SlCTOMT1 activity
Catechol		100%
Guaiacol		0%
Salicylic acid		0%
Benzoic acid		0%
Oricinol		9%
Caffeic acid		2%
Protocatechuic aldehyde		17%
Pyrogallol		4%
2,5-dimethyl-4-methoxy-3(2H)-furanone		0%

Table 2-2. Catechol feeding of tomato discs.

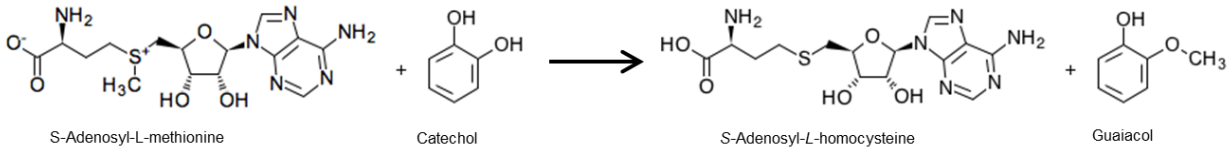
	Guaiacol emitted (nmol $disc^{-1}$ h^{-1})	
	Control	Catechol
Flora-Dade	0.04 ± 0.01	0.36 ± 0.04
OE1683	0.54 ± 0.30	4.53 ± 0.83

Data are means ± SE.

Figure 2-1. Identification of potential *S. lycopersicum* CTOMTs. Potential tomato CTOMT candidates were identified by finding coding sequences with similar to known OMTs. Protein alignment and a Neighbor-Joining tree were done using MEGA5 (Tamura *et al*., 2011). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. ANMT= anthranilate N-methyltransferase CAOMT= caffeic acid *O*-methyltransferase; CTOMT= catechol *O*-methyltransferase; CVOMT= chavicol *O*-methyltransferase; EOMT= eugenol *O*-methyltransferase ; IAOMT= indole-3-acetate *O*-methyltransferase; JAMT= jasmonic acid carboxyl methyltransferase; MOMT= myricetin *O*-methyltransferase; OROMT= orcinol *O*-methyltransferase; ROMT= reserveratrol *O*-methyltransferase; SAMT= salicylic acid carboxyl methyltransferase. Ab.SAMT is *Atropa belladonna* (BAB39396). At.IAOMT and At.JAMT are *Arabidopsis thaliana* (Q9FLN8; AAG23343). Ca.JAMT is *Capsicum annuum* (ABB02661).Fa.OMT is *Fragaria* x *ananassa* (AAF28353).Nt.CAOMT and Nt.CTOMT are *Nicotiana tabacum* (AAL91506; CAA50561). Ob.CAOMT, Ob.CVOMT and Ob.EOMT are *Ocimum basilicum* (Q9XGV9; Q93WU3; Q93WU2). Ph.SAMT is *Petunia hybrida* (AAO45013). Pt.CAOMT is *Pinus taeda* (AAC49708). Rh.OROMT is *Rosa hybrida* (AAM23004). Sf.SAMT is *Stephanotis floribunda* (CAC33768). Sh.MOMT is *Solanum habrochaites* (ADZ76434). Rg.ANMT is *Ruta graveolens* (A9X7L0). SL.U575022, SL.U570080, SL.U321686, SL.U582403, and SL.U565623 are *S. lycopersicum* unigenes. Tt.CTOMT is *Thalictrum tuberosum* (AAD29844). Vv.ROMT is *Vitis vinifera* (CAQ76879). ZM.CTOMT is *Zea mays* (NP_001106047).


Rh.OROMT
Vv.ROMT
Sl.U575022
Sl.U570080
Ob.EOMT
Ob.CVOMT
Sl.U321686
Sh.MOMT
Pt.CAOMT
Sl.U565623
Nt.COMT
Nt.CAOMT
Sl.U582403
Rg.ANIMT
Zm.COMT
Tt.COMT
Ob.CAOMT
Fa.OMT
At.IAOMT
At.JAMT
Ca.JAMT
Sf.SAMT
Ph.SAMT
Ab.SAMT
Sl.SAMT
91
90
88
83
99
100
77
100
100
100
75
76
97
98
100
69
100
100
98
98
100
42

A



B

CTOMT1	K_m (µM)	K_{cat} (s^{-1})
S. lycopersicum	8.36±1.78	9.67±2.42

Figure 2-2. Enzyme activity of CTOMT1. (A) The predicted pathway for the synthesis of guaiacol from catechol. This is the reaction that CTOMT1 is thought to catalyze. (B) The specific activity and turnover rate of CTOMT1. Values were determined using non-linear regression. These values were similar to values found for previously characterized diphenol OMTs as listed on BRENDA (www.brenda-enzymes.org/).

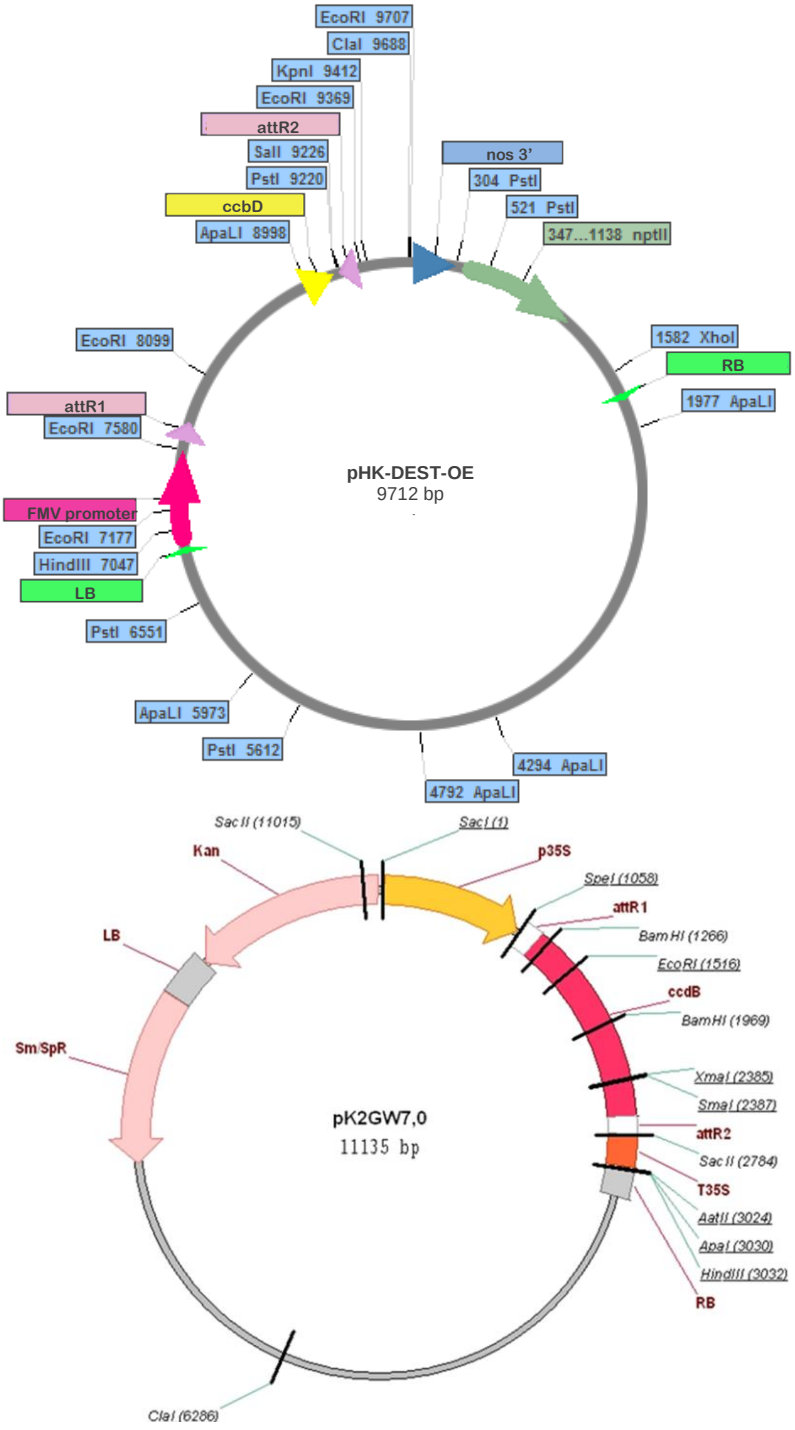


Figure 2-3. Vectors for transgene expression. The pHK-DEST-OE plasmid map was created using SerialCloner 2-1 and the pK2GW7 plasmid is based on sequence data from Karimi *et al*., 2002.

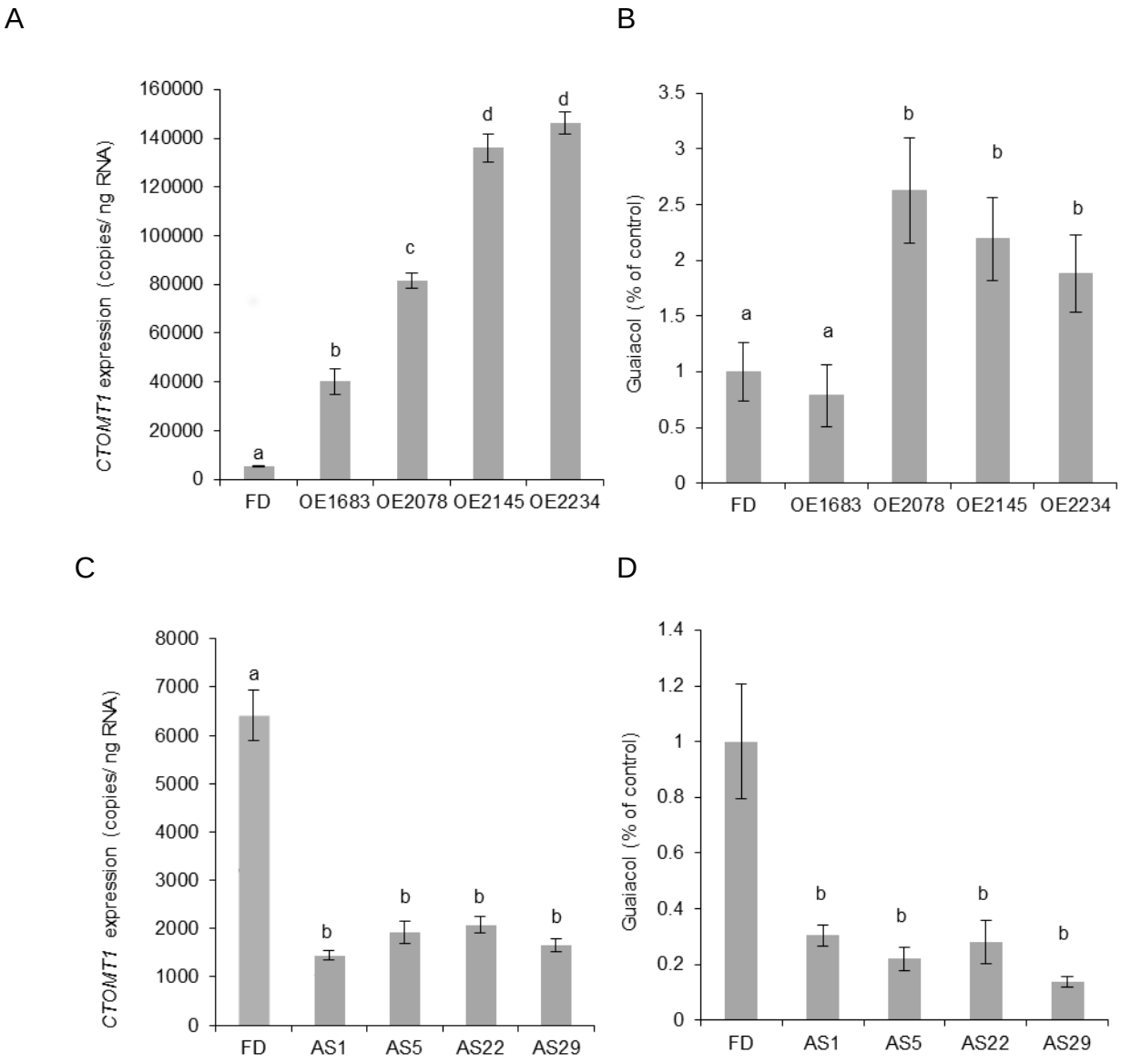


Figure 2-4. Gene expression and guaiacol levels in transgenic plants. (A)*CTOMT1* mRNA levels for four lines overexpressing (OE) *CTOMT1*. (B) Guaiacol emission for four OE *CTOMT1* lines as a percentage of control (FD). (C) Suppression of endogenous mRNA levels for four antisense (AS) lines. (D) Guaiacol emission for four AS *CTOMT1* lines as a percentage of control (FD). Total mRNA and volatiles were extracted from ripe fruit. Error bars represent standard error of the mean. Tukey's HSD was used to determine significant differences (P <0.05). Statistical groups are indicated by use of different letters adjacent to bars.

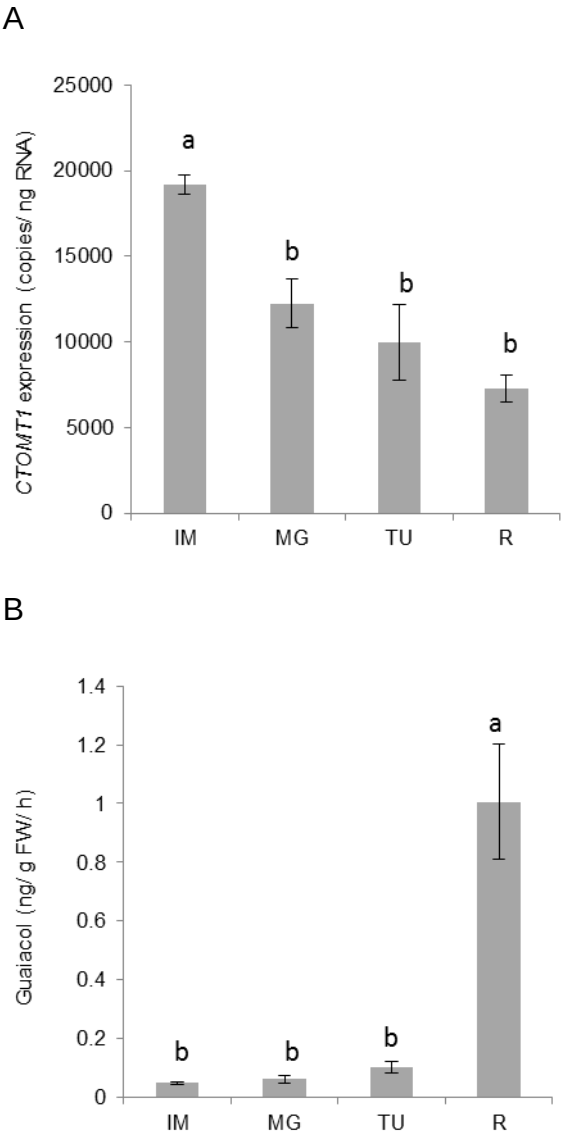


Figure 2-5. *CTOMT1* expression and guaiacol production through fruit development. (A) *CTOMT1* RNA was measured in immature green (IM), mature green (MG), turning (TU), and ripe (F) Flora-Dade fruit. (B) Fruit guaiacol levels. Error bars represent standard error of the mean. Tukey's HSD was used to determine significant differences ($P < 0.05$). Statistical groups are indicated by letters.

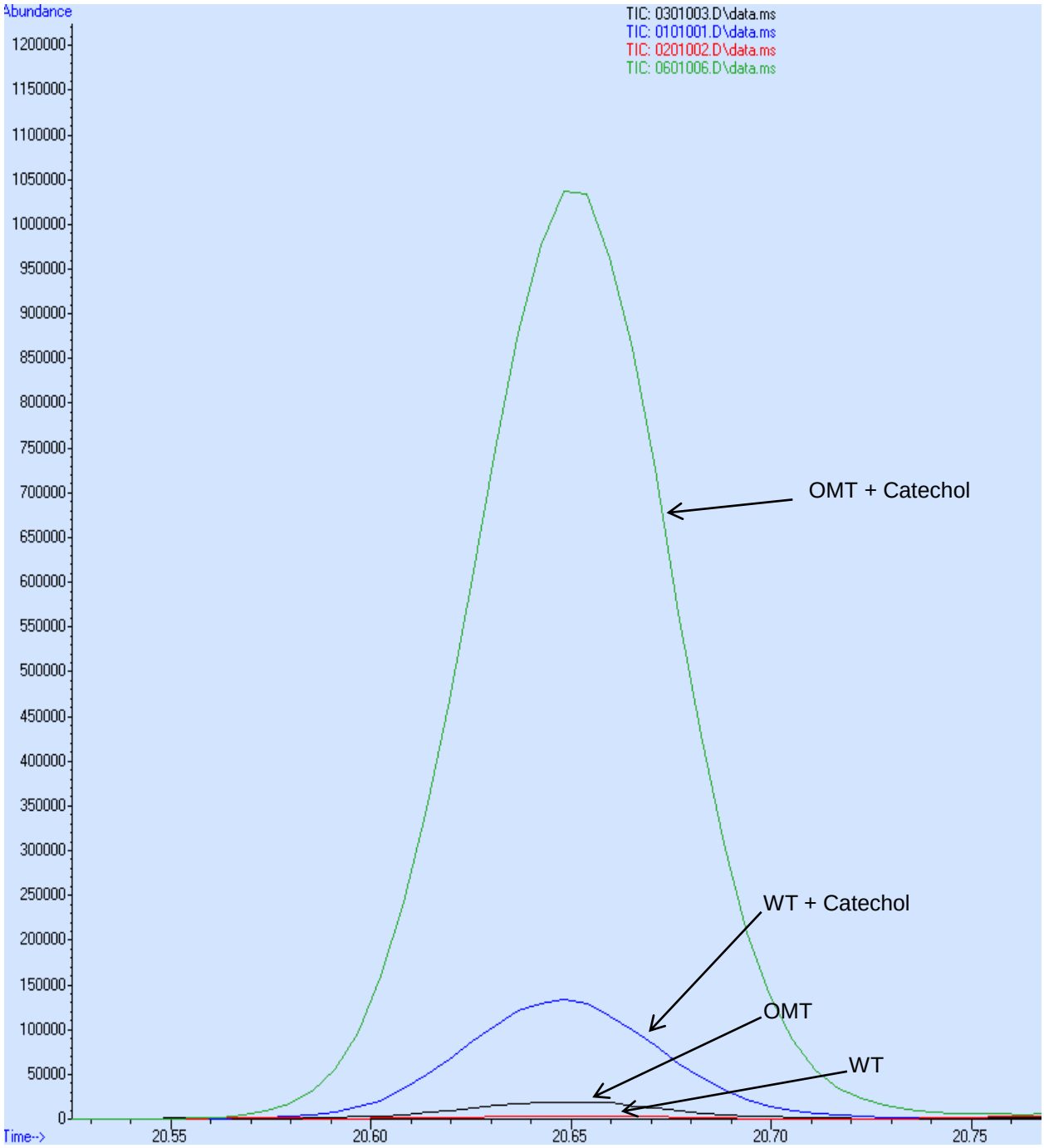


Figure 2-6. Guaiacol spectrum of catechol fed disc. This spectrum shows the guaiacol peak as measured by GC-MS. As shown, the *CTOMT1* overexpressing disc (—) produces slightly more guaiacol than the WT (—) when just fed water. However, when supplied with non-limiting catechol, *CTOMT1* overexpressing disc (—) produce much more guaiacol than catechol fed WT disc (—).

CHAPTER 3
QUANTITATIVE TRAIT LOCUS ASSOCIATED WITH INCREASED GUAIACOL

Overview

Looking for Genetic Variation

There are now 13 recognized species of cultivated tomato, *S. lycopersicum*, and its wild relatives. Formerly, these species were under the segregated genus *Lycopersicon*, but now have been grouped into the genus *Solanum*, along with potato and eggplant (Peralta *et al*., 2006). Much diversity exists between and within species. Although there is much phenotypic variation within *S. lycopersicum* accessions, it is low in genetic diversity as compared to its wild relatives, as measured by allozymes and restriction fragment length polymorphisms (RFLPs) (Miller and Tanksley, 1990; Rick and Yoder, 1988).

One explanation for this lower diversity is the breeding system that *S. lycopersicum* utilizes. Tomato and its wild relatives can be separated into two types of breeding systems, those that are self-incompatible and those that are self-compatible. Self-incompatible species show more variation that all self-compatible species (Miller and Tanksley, 1990; Rick and Yoder, 1988).

The primary cause of the lack of genetic variation among *S. lycopersicum* accessions is domestication. The modern tomato has gone through three major genetic bottlenecks: early domestication by the native peoples of Central America, transport of a limited germplasm to Europe, and intensive breeding during the development of tomato as a commercial crop after World War II (Bai and Lindhout, 2007; Klee, 2010).

Enhancing Diversity

Good QTL mapping requires high variation and marker coverage (Rick and Yoder, 1988). Outcrossing tomato with its wild relatives integrates novel quantitative variation which can enhance the discovery of QTLs (Eshed and Zamir, 1994). The green-fruited *S. pennellii* is a good choice as the wild parent in such a cross because it produces fruit after controlled self-pollination, the offspring of *S. lycopersicum* × *S. pennellii* are fertile, and there is a wealth of allozymic differences between the two species (Rick, 1960; Rick and Yoder, 1988).

An previously described introgression population of 75 *S. lycopersicum* lines, each containing a single homozygous *S. pennellii* chromosomal segment in an otherwise *S. lycopersicum* genome, was created by initially crossing *S. lycopersicum* cv. M82 as the female parent with *S. pennellii* as the male parent. The F1 hybrid of this cross was then backcrossed to the M82 parent and repeatedly selfed. Altogether the lines provide complete representation of the wild species genome and are nearly isogenic to M82 (Eshed and Zamir, 1994; Eshed and Zamir, 1995).

Identifying Quantitative Trait Loci

An initial analysis of the introgression population, revealed 23 QTLs for total soluble solids and 18 QTLs for fruit size (Eshed and Zamir, 1995). This same population was later used to identify QTLs affecting fruit flavor volatile emission. After collecting data from introgression lines (ILs) grown in multiple locations over five different seasons, 25 different loci were identified affecting at least one of 23 different volatiles. Additionally, the ILs have been used in numerous other ways, such as metabolic profiling and identification of loci affecting salinity stress, demonstrating their value as a tool in the discovery of gene trait associations (Frary *et al*., 2011; Schauer *et al*., 2006).

Results

Finding a Guaiacol QTL

In analysis of ILs by Tieman *et al.* (2006a), increased guaiacol levels were reported in IL10-1. Further analysis of the complete data set indicated the presence of a guaiacol QTL on the overlapping IL 10-1-1 (fdr = 0.000220175) near the top of chromosome 10. Using the *S. lycopersicum* genome sequence database (http://solgenomics.net/), a cleaved amplified polymorphic sequence (CAPS) marker was developed (details in Materials and Methods) to map the position of *CTOMT1* (Figure 3-1). The gene was located within the *S. pennellii* segment of IL 10-1 but not 10-1-1. IL 10-1 contains approximately 60 million base pairs of *S. pennellii* DNA. IL 10-1-1 is a much smaller segment contained within IL10-1 and adjacent to the position of *CTOMT1*. To further elaborate the nature of the QTL, guaiacol was collected from ripe IL10-1, IL10-1-1, and M82 fruits (Figure 3-2A). Both IL 10-1 and IL 10-1-1 had elevated guaiacol production relative to M82. However, IL 10-1-1 produced significantly more guaiacol than IL 10-1. Both ILs also produced significantly more methylsalicylate as well (Figure 3-3).

Comparing Activity of CTOMT1 Orthologs from *S. lycopersicum* and *S. pennellii*

In order to determine the underlying cause for the increased guaiacol production in these ILs, the orthologous *S. pennellii CTOMT1* was cloned and sequenced. Differences in the promoter and the coding sequences were found (Appendix A). Some changes in the coding sequence resulted in amino acid differences (Figure 3-3). However, when the specific activity of recombinant *S. pennellii CTOMT1* was tested there was no significant difference the specific activity between the *S. pennellii* and the M82 enzymes (Table 3-1).

Gene Expression Analysis

CTOMT1 mRNA levels were also measured from ripe IL10-1, IL10-1-1, and M82 fruits (Figure 3-2B). RNA expression analysis indicated that *CTOMT1* RNA levels were significantly increased in both ILs. Analysis of these lines seems to indicate that increased guaiacol is associated with increased *CTOMT1* expression.

Catechol and Salicylic Acid Quantification

In order better understand how the catechol synthesis pathway is affected in ILs 10-1 and 10-1-1, catechol and salicylic acid were quantified from ripe fruit. Catechol and salicylic acid were extracted from ground tissue and then derivatized and quantified using GC-MS (Figure 3-5). A trend of lower catechol was found in the ILs. This same trend was also seen for salicylic acid concentrations.

Discussion

Through screening of an introgression population we were able to identify a QTL associated with higher guaiacol production near the top of chromosome 10. *CTOMT1* was mapped to the region covered by IL 10-1, with higher guaiacol and higher *CTOMT1* expression than M82. However, another IL, 10-1-1, adjacent to but not including the *CTOMT1* gene, also synthesized significantly more guaiacol and had higher *CTOMT1* expression. Although in these lines increased guaiacol production seems associated with increased expression of *CTOMT1,* we know that this is not always the case, as exemplified in *CTOMT1* overexpressing transgenics. Something other than just increased gene expression is contributing to increased guaiacol levels. There must be a genetic element within the *S. pennellii*-derived 10-1-1 segment that directs higher guaiacol production. Recently, it has been shown that many tomato QTLs are in *trans* to structural genes encoding enzymes that contribute to the phenotype. Steinhauser *et al*.

(2011) found that 17 out of 27 robust QTLs that they mapped were located *trans* to the structural gene that encodes the corresponding enzyme activity.

The most probable explanation for the increase of guaiacol in IL10-1 and IL 10-1-1 is the existence of a *trans*-acting regulatory element contained in IL 10-1-1. The overexpression of a single transcription factor can increase the production of a whole pathway of enzymes and metabolites (Dal Cin *et al*., 2011). Perhaps a transcription factor that upregulates the guaiacol synthesis pathway is located on IL 10-1-1. Preliminary results indicate that both methylsalicylate and eugenol are also increased in these ILs, while catechol and salicylic acid levels decrease. The precursor for methylsalicylate is salicylic acid (Tieman *et al*., 2010). Salicylic acid has also been proposed as the precursor for catechol in *G. adenothrix* (Ellis and Towers, 1969). Salicylic acid and eugenol have cinnamic acid as a common precursor (Koeduka *et* al., 2006; Métraux, 2002). If this entire pathway is upregulated we would expect to see increased gene expression of other genes in the pathway and increased metabolite levels. This hypothesis could be tested using global expression analysis of the ILs compared to M82 and more indepth metabolite analysis.

Another possible cause for the observed phenotypes of IL10-1 and IL10-1-1 is that an enzyme that converts salicylic acid to catechol is located in 10-1-1. As previously discussed, catechol availability limits guaiacol production. If there is an enzyme on IL10-1-1 that increases the production of catechol, then *CTOMT1* expression may also be stimulated. As catechol is toxic to plants, it is expected that its concentration would be tightly regulated; catechol may act as regulator of *CTOMT1* expression in order to keep its levels low. BLAST analysis of the 10-1-1 segment against known enzymes that

produce catechol did not reveal any potential candidates; however, this analysis was not exhaustive.

A third possibility is that a chromosomal rearrangement has occurred. Chromosomal rearrangements can cause suppression of recombination (Verlaan *et al*., 2011). This in combination with low marker coverage can result in discrepancies between genetic and physical maps (Liharska *et al*., 1996). Fluorescence *in situ* Hybridization (FISH) is a useful technique for discovering chromosomal rearrangements and correctly mapping introgressed genes from wild species (Verlaan *et al*., 2011). Also, chromosome architecture and loci interactions are not well understood. It is know that transcription enhancers can be located some distance away from the core promoter. DNA looping facilitates the interaction of the enhancer with the promoter (Kagey *et al*., 2010). Perhaps there is such an element located in 10-1-1 that is interacting with the promoter of *CTOMT1*.

It is also possible that there are two guaiacol QTLs represented by these ILs. However, this is not very likely as *CTOMT1* does not seem to be a QTL. This conclusion is supported by the result that overexpression of the gene alone is not sufficient to increase guaiacol production. To fully rule out *CTOMT1* as a QTL by itself, we would need a recombinant that only included the region above 10-1-1. This is very difficult to achieve as *CTOMT1* is near the end of the chromosome. It is possible that there could be a *CTOMT1* allele that acts as a QTL for decreased guaiacol. Further analysis of heirloom varieties and wild species could help identify allozymes with decreased catechol methylation activity.

Our understanding of chromosome structure and gene regulation only scratches the surface of the depth of complexity that exists. This lack of knowledge makes sorting out complex traits difficult. Hopefully, as technologies and techniques develop, we will be able to unravel more of these regulatory networks.

Materials and Methods

Mapping and Volatile Analysis of *CTOMT1*

A guaiacol QTL on the overlapping ILs 10-1 and 10-1-1 was identified as previously described (Tieman *et al*., 2006a). Guaiacol contents for each of five seasons as well as summary values for the combined seasons are available at http://ted.bti.cornell.edu/cgi-bin/TFGD/metabolite/metabolite_info.cgi?ID=M0000025. To determine the map position of *CTOMT1,* a marker that distinguishes between the *S. lycopersicum* and *S. pennellii* alleles was developed. The following primers were used: F- ATTAATGCTTTCCTGTCGAACC and R- ACCTCCAACATCAACCAAAGTT. The product size was 3.7 kb. Amplification products were digested with DdeI (New England Biolabs, www.neb.com). Genomic sequence alignments of *S. lycopersicum* and *S. pennellii* were done with ClustalW using genomic sequence data provided by A.R. Fernie (personal communication). Protein purification and enzyme activity assays were performed as described in Chapter 2. Volatiles were collected and RNA was quantified from ripe M82, IL10-1, and IL10-1-1 fruit as described in Chapter 2.

Catechol and Salicylic Acid Quantification

Catechol was extracted from ripe M82, IL10-1-, and IL10-1-1 fruits ($n \geq 3$). After grinding tissue in liquid nitrogen, 3 g were measured and catechol was extracted with 3 ml of acetonitrile. As an internal control, 500 ng of 4-nitrophenol were added to each sample. Samples were vortexed and centrifuged for 10 min at 25 000 g. The

supernatant was placed in a glass vial and dried under nitrogen gas. A standard curve was also made using catechol, salicylic acid, and 4-nitrophenol using each standard at the following quantities: 0, 10 ng, 50 ng, 100 ng, 500 ng, 1000 ng, 5000 ng, 10000 ng. Samples were resuspended in 200 µl of anhydrous acetonitrile. For derivatization, 100 µl of the resuspended sample was placed in a new vial with 100 µl of *N*-methyl-*N*-trimethylsilyltrifluoroacetamide (MSTFA) with 1% trimethylchlorosilane (TMCS) (Thermo Scientific, www.thermoscientific.com). Reaction vials were placed at 70˚c for 1 h. Samples were then analyzed using GC-MS on a Agilent 5975 GC/MSD (www.chem.agilent.com) (He carrier gas; 0.7 ml min^{-1}; splitless injector 250˚c, injection volume 2 µl) with a Agilent DB-5ms column ((5%-Phenyl)-methylpolysiloxane; 30 m long, 250 µm i.d., 1 µm film thickness). The temperature was programmed from 100˚c (4 min hold) at 8˚c min^{-1} to 300˚c. Source and quadrupole temperatures were 230˚c and 150˚c respectively. Ions selected for detection were as follows: catechol- 136,166,239 (Figure 3-6); salicylic acid- 135, 193, 267; 4-nitrophenol- 150, 196, 211. Compounds were identified by retention times with standards and specific ions.

Table 3-1. Comparison of enzyme activity between CTOMT1 orthologs from *S. lycopersicum* and *S. pennellii*.

CTOMT1	K_m (μM)	K_{cat} (s^{-1})
S. lycopersicum	8.36±1.78	9.67±2.42
S. pennellii	13.14±3.93	10.87±4.39

Data are means ± SE.

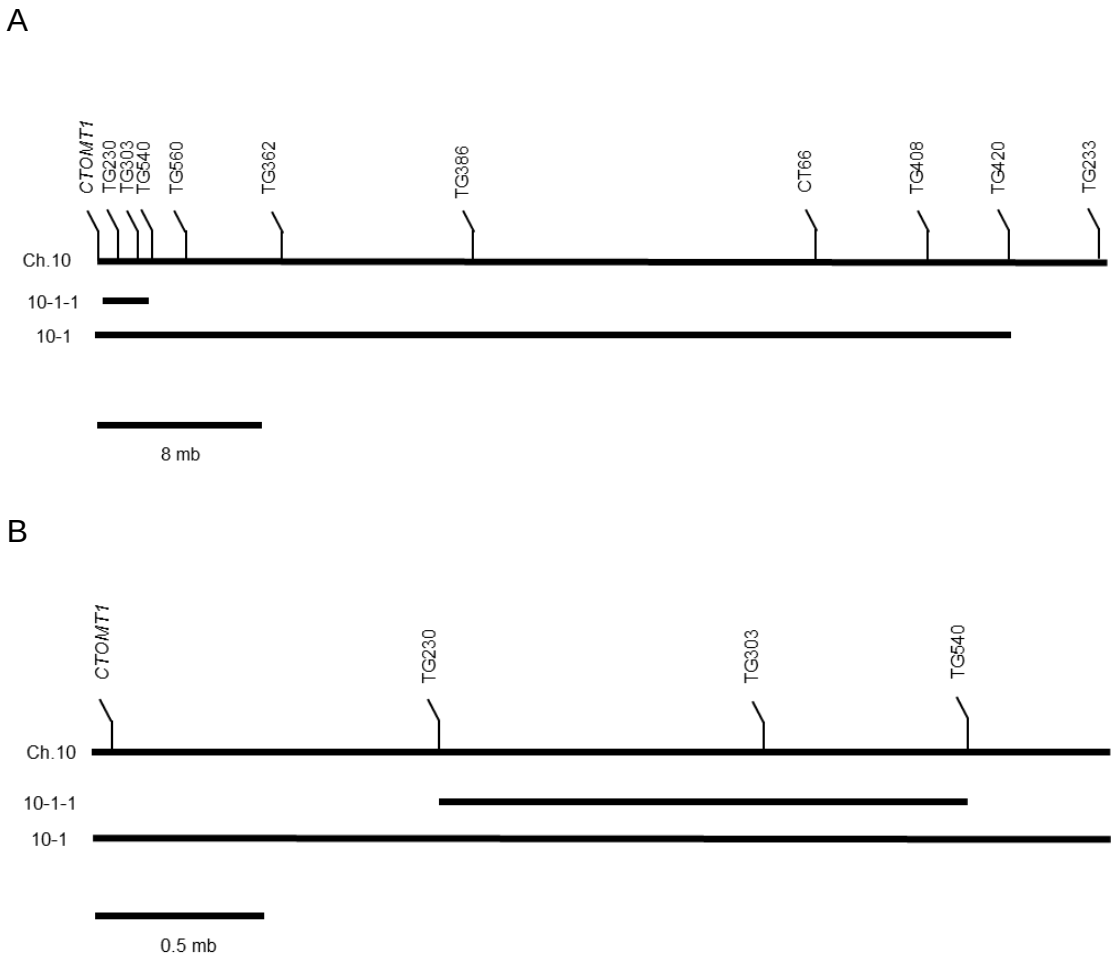


Figure 3-1. Chromosome 10 map with positions of ILs 10-1 and 10-1-1. (A) Genetic markers are shown above the black line. Positions of *S. pennellii* segments in ILs 10-1 and 10-1-1 were originally determined by Eshed and Zamir (1995). Fine mapping of *CTOMT1* was done by developing a new marker for the gene. The *S. pennellii* allele of *CTOMT1* was present on 10-1, but not 10-1-1. Physical distances were determined using chromosomal sequence data from sol genomic network (www.solgenomics.net). (B) An expansion of the region encompassing 10-1-1 is shown. Marker sequences are available on the sol genomic network. See Materials and Methods for the sequence of *CTOMT1* marker primers.

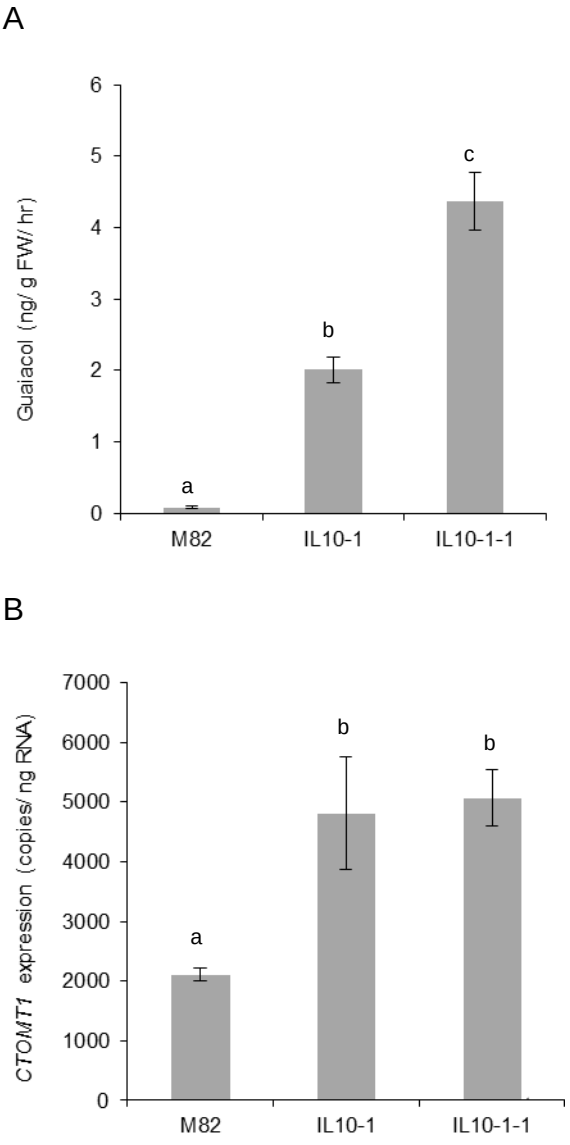


Figure 3-2. CTOMT1 expression and guaiacol emission from ILs. (A) Increased guaiacol in fruit from ILs 10-1 and 10-1-1. (B) Increased mRNA levels of *CTOMT1* in ILs 10-1 and 10-1-1. Error bars represent standard error of the mean. Tukey's HSD was used to determine significant differences (P <0.05). Statistical groups are indicated by letters.

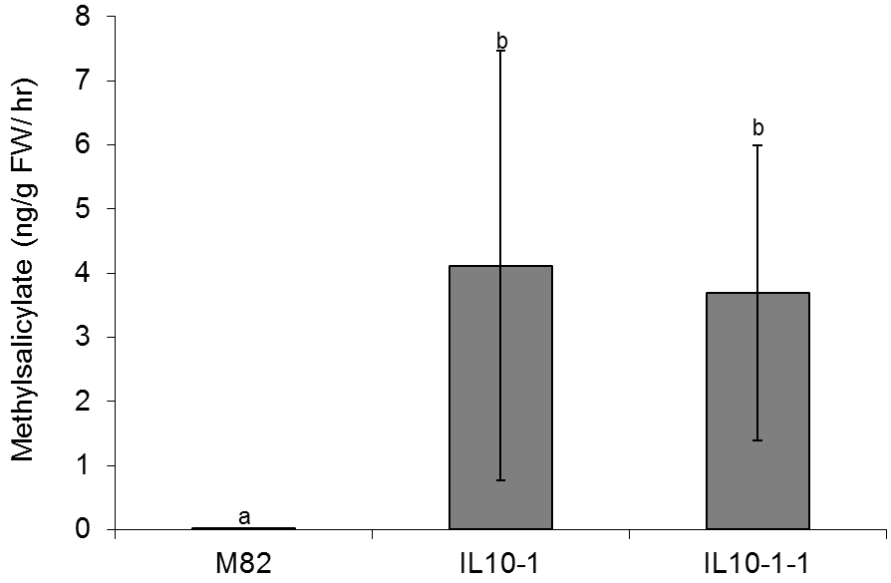


Figure 3-3. Methyl salicylate levels in M82 and ILs 10-1 and 10-1-1. Increased methyl salicylate levels in fruit from ILs 10-1 and 10-1-1. Error bars represent standard error of the mean. Tukey's HSD was used to determine significant differences (P <0.05). Statistical groups are indicated by letters.

```
S.lycopersicum   MGSTANIQLATQSEDEERNCTYAMQLLSSSVLPFVLHSTIQLDVFDILAKDKAATKLSAL
S.pennellii      MGSTANIQLPTQSENEERNCTYAMQLLSSSVLPFVLHSTIQLDVFEILAKDKAATKLSAL
                 *********.****:******************************:**************

S.lycopersicum   EIVSHMPNCKNPDAATMLDRMLYVLASYSLLDCSVVEEGNGVTERRYGLSRVGKFFVRDE
S.pennellii      EIVSHMPNCKNPDAATMLDRMLYVLASYSLLDCTVVEEGNGVTERRYGLSRVGKFFVRDE
                 ********************************:************************

S.lycopersicum   DGASMGPLLALLQDKVFINSWFELKDAVLEGGVPFDRVHGVHAFEYPKLDPKFNDVFNQA
S.pennellii      DGASMGPLLALLQDKAFINSWFELKDAVLEGGVPFDRVHGVHAFEYPKLDPKFNDVFNQA
                 ***************.****************************************

S.lycopersicum   MINHTTVVMKRILENYKGFENLKTLVDVGGGLGVNLKMITSKYPTIKGTNFDLPHVVQHA
S.pennellii      MINHTTVVMKRILENYKGFENLKTLVDVGGGLGVNLKMITSKYPTIKGTNFDLPHVVQHA
                 ************************************************************

S.lycopersicum   PSYPGVDHVGGDMFESVPQGDAIFMKWILHDWSDGHCLKLLKNCHKALPDNGKVIVVEAN
S.pennellii      TSYPGVDHVGGDMFESVPQGDAIFMKWILHDWSDGHCLKLLKNCHKALPDNGKVIVVEAN
                 .************************************************************

S.lycopersicum   LPVKPDTDTTVVGVSQCDLIMMAQNPGGKERSEQEFRALASEAGFKGVNLICCVCNFWVM
S.pennellii      LPVKPDTDTTVVGVSQCDLIMMAQNPGGKERSEQEFRALASEAGFKGVNLICCVCNFWVM
                 ************************************************************

S.lycopersicum   EFYK
S.pennellii      EFYK
                 ****
```

Figure 3-4. Amino acid alignment of *S. lycopersicum* and *S. pennellii* CTOMT1. Amino acid sequences were determined by translating the coding region of *CTOMT1*. An asterisk (*) indicates a fully conserved residue. A colon (:) indicates conservation between groups of strongly similar properties. A period (.) indicates conservation between groups of weakly similar properties. Protein alignment was performed with ClustalW.

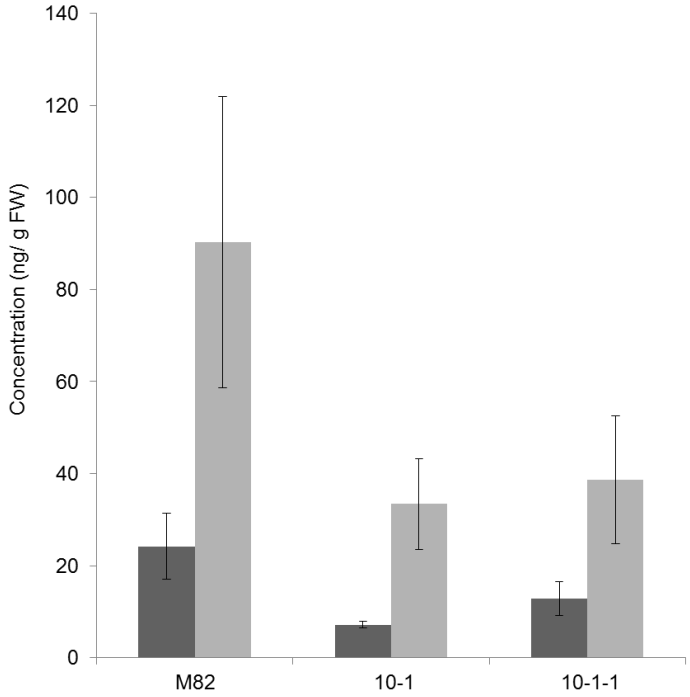
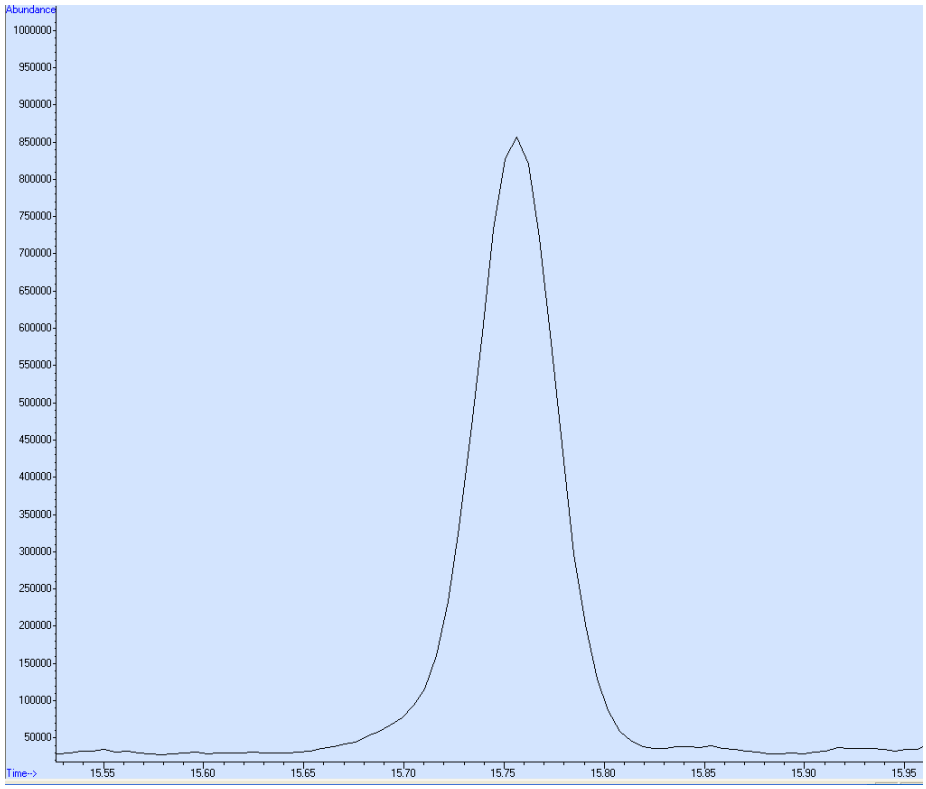


Figure 3-5. Catechol and salicylic acid quantification. Catechol (■) and salicylic acid (■) were quantified from fruit ILs 10-1 and 10-1-1 and M82. Error bars represent standard error of the mean.

A



B

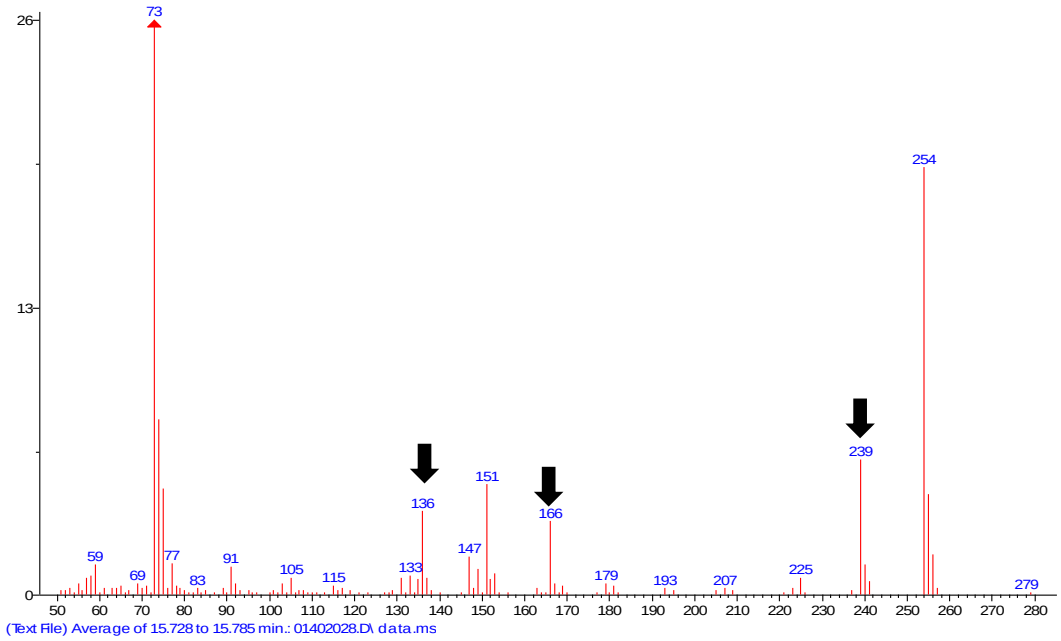


Figure 3-6. Spectra of silylated catechol. (A) Peak of silylated catechol 500 ng standard. (B) Mass spectrum of derivatized catechol. Ions 239, 166, and 136 were used to identify catechol in tomato fruit samples.

CHAPTER 4
GLYCOSYLATION OF FLAVOR MOLECULES

Overview

Small Molecule Glycosides

Many classes of plant secondary metabolites are glycosylated. These classes include hormones, betalains, phenylpropanoids, terpenoids, steroids, flavonoids, coumarins, glucosinates, cyanogenic glucosides, and cyclic hydroxamic acids (Bowles *et al*., 2006). Glycosylation increases solubility and stability and decreases chemical reactivity. For some molecules, such as *p*-aminobenzoate, glycosylation promotes membrane-transport and storage in the vacuole (Brazier-Hicks *et al*., 2007; Eudes *et al*., 2008; Lim *et al*., 2002; Lim *et al*., 2004; Paquette *et al*., 2003). For other molecules, like saponins and quercetin, glycosylation is necessary for bioactivity (Cartwright *et al*., 2008; Osbourn *et al*., 2003).

Family 1 Glycosyltransferases

Glycosyltransferases are ubiquitous across all kingdoms of life (Campbell *et al*., 1997). Small lipophilic molecules are glycosylated by the Family 1 glycosyltransferases. Most glycosyltransferases in this family require UDP-activated sugars and are, therefore, called UDP-dependent glycosyltransferases (UGTs) (Osmani *et al*., 2009).

Four plant UGTs have been crystallized. Although the primary sequence may be quite variable, the secondary structure is quite conserved (Shao *et al*., 2005). The secondary structure consists of two Rossmann fold domains connected by an interdomain linker. The two domains form a deep, narrow pocket for substrate binding that is made accessible via movement of the flexible linker (Bowles *et al*., 2005; Osmani *et al*., 2009).

The C-terminal domain contains a highly conserved region called the Plant Secondary Product Glycosyltransferase motif (PSPG) (Figure 4-1) (Jones and Vogt, 2001; Lim and Bowles, 2004; Paquette *et al*., 2003; Vogt and Jones, 2000). This motif is important for recognition and binding of the UDP-sugar. UDP-glucose is the most commonly utilized sugar; however, UDP-galactose, UDP-xylose, UDP-mannose and UDP-glucuronic acid are preferred by some UGTs (Kohara *et al*., 2007; Masada *et al*., 2007; Weis *et al*., 2008).

The UGT N-terminal domains are much more variable than the C-terminal domains. The N-terminal domain is believed to be important for interaction with the acceptor molecule (Shao *et al*., 2005). N-terminal amino acid sequence homology is not an indicator of substrate preference. A pair of UGTs with 20% homology in this domain may have activity on the same substrate while another pair with 80% homology may recognize different substrates (Osmani *et al*., 2009).

Studies of the phylogeny and activities of *Arabidopsis* UGTs have provided important information about UGT structure and function. 120 UGTs have been identified in the Arabidopsis genome (Paquette *et al*., 2003). These enzymes have been placed into 14 phylogenetic groups (Li *et al*., 2001; Ross *et al*., 2001; Osmani *et al*., 2009). Enzymes within a group tend to show activity on the same types of substrates (Lim *et al*., 2003).

Aromatic Small Molecule Glycosides

Many of the small molecules important to tomato flavor are glycosylated (Table 4-1) (Birtiç *et al*., 2009; Buttery *et al*., 1990; Marlatt *et al*., 1992; Ortiz-Serrano and Gil, 2007). Glycosylation of these molecules prevents their volatilization and, therefore, contribution to tomato flavor. One of the largest pools of glycosidically-bound tomato

volatiles has been observed for the flavor molecule 2-phenylethanol (Birtiç *et al*., 2009; Buttery *et al*., 1990; Marlatt *et al*., 1992). IL8-2-1 from the *Solanum lycopersicum* X *Solanum pennellii* introgression population has been characterized as a very high 2-phenylethanol producing line (Tadmor *et al*., 2002; Tieman *et al*., 2006b). This line also produces elevated levels of a 2-phenylethanol glycoside (Tieman, unpublished). It can be estimated that roughly two-thirds of the total 2-phenylethanol pool is contained in the glycosylated form.

2-Phenylethanol Glycosylation

Much of what is known about the structure and accumulation of 2-phenylethanol glycoside comes from studies on roses. These studies indicated that 2-phenylethanol is bound to either a mono- or disaccharide composed of hexose, hexose-hexose, or hexose-pentose sugars (Watanabe *et al*., 2001). The concentration of 2-phenylethanol glycoside increased and decreased in a diurnal pattern antithetical to the concentration of free 2-phenylethanol, suggesting that the glycoside serves as a storage molecule for timed release of the aglycone (Hayashi *et al*., 2004; Picone *et al*., 2004). This pattern is not what has been observed in fruits. Concentrations of glycosides were observed to increase with fruit ripening (Birtiç *et al*., 2009; Groyne *et al*., 1999). Enzymatic release of aromatic aglycones from these glycosides has been demonstrated to be important in increasing the flavor of wine (Cabrita *et al*., 2006; Ugliano and Molo, 2008).

Results

Isolation and Characterization of 2-Phenylethanol Glycoside

In order to develop a protocol for the isolation and characterization of 2-phenylethanol glycosides, *Petunia hybrida* ‘Mitchell Diploid’ flowers were chosen for glycoside analysis. Petunia was selected because it is also a member of the

Solanaceae family and a prolific producer of 2-phenylethanol. A crude extract was made from 400 flowers and fractionated using flash chromatography. Further purification was done using LC-MS. Fractions with mass spectra that indicated the presence of both di- and tri-saccharides bound to 2-phenylethanol were collected. The fraction with the highest purity was predicted to contain a 2-phenylethanol tri-saccharide. This fraction was dried and sent to the Complex Carbohydrate Research Center (CCRC) at the University of Georgia (www.ccrc.uga.edu) for further analysis. Reports from CCRC indicated the purified glycoside contains primarily terminal and 1, 2-linked glucose (Figure 4-2).

To determine the 2-phenylethanol glycoside structure from tomato, a crude extraction was made from ripe fruit of IL8-2-1 and its M82 parent. These extracts were sent for analysis to Dr. Alisdair Fernie at the Max Planck Institute for Molecular Plant Physiology. LC-MS data showed a peak with MW 387 that was highly abundant in IL8-2-1, but absent in M82 (Figure 4-3). One possible structure assignment for this product is 2-phenylethyl-6'-O-malonyl-ß-D-glucopyranoside. This compound has been sent to the RIKEN Institute for NMR structural confirmation.

Selection of Tomato UGT Candidates and Cloning

Known *Arabidopsis* family 1 UGTs were used to search the sol genomic network Lycopersicon combined (tomato) unigene database for UGT sequences (Figure 4-4). Over 100 probable family 1 UGTs were found. In order to narrow the list of candidates, only unigenes with a high number of fruit ESTs were chosen. Six candidates were selected: SGN-U578221, SGN-U578227, SGN-U565076, SGN-U584032, SGN-U576693, and SGN-U571691 (Table 4-2). Full length cDNAs were synthesized from *S. lycopersicum* cv. M82 ripe fruit RNA. Candidate genes were cloned vectors for

expression in *E. coli*. Although bacteria were grown in minimal media at low temperatures, the protein was completely insoluble. Attempts to increase expression levels and solubility were made by adding plasmids coding for rare codon tRNAs and protein chaperones to expression bacteria. However, results were not improved.

The next strategy for characterizing candidate UGTs was to use transient expression in *Nicotiana benthamiana*. For transient expression, candidate genes were cloned into the binary vector, pHK1001, and transformed into *Agrobacterium tumefaciens*. The *Agrobacterium* containing various pHK1001 constructs were infiltrated into *N.benthamiana* leaves along with p19, a silencing suppressor, and PAAS, a phenylacetaldehyde synthase from *Petunia hybrida,* (Kaminaga *et al*., 2006; Voinnet *et al*., 2003). After 5 days, the leaves were chopped and volatiles were collected. Expression of p19 with PAAS produce leaves with high levels of 2-phenylethanol. Only one of the UGT candidates, SGN-U578227, was observed to lower 2-phenylethanol when expressed together with p19 and PAAS (Figure 4-5).

Transgenic Expression of UGT Candidates in Tomato

Candidate gene pHK1001 overexpression (pHK_{OE}) constructs were also transformed into tomato. SGN-U578227 was also cloned in anitsense orientation (pHK_{AS}) for RNA suppression. As we were interested in identifying a UGT that glycosylated 2-phenylethanol, *Agrobacterium* was used to transform IL8-2-1 cotyledons for transgene expression. The numbers of independent transgenic lines for each construct were as follows: 6 pHK_{OE}- 578221, 60 pHK_{OE}-565076, 7 pHK_{OE} -584032, 14 pHK_{OE}-576693, 2 pHK_{OE}-571691, 68 pHK_{OE}-578227, and 6 pHK_{AS}-578227. Due to time and resources constraints only 26 of the pHK_{OE}-578227 and 28 of the pHK_{OE}-565076 were tested for transgene expression using leaf tissue. Volatiles were measured for ripe

fruit of 24 of the pHK_{OE}-578227 plants and all the plants of the other constructs except pHK_{OE}-565076. None of the candidate genes impacted volatiles in the 2-phenylethanol pathway, including phenylacetaldehyde and 1-nitro-phenylethane, or other observed volatiles.

During the production of these transgenic plants, SGN-U565076 was found to be upregulated by the overexpression of the transcription factor ORDORANT1 (ODO1) (Dal Cin *et al*., 2011). ODORANT1 is a MYB transcription factor from *Petunia hybrid* important for the regulation of volatile benzenoid synthesis (Verdonk *et al*., 2005). *ODORANT1* was also overexpressed in tomato plants to gain understanding of the phenylalanine metabolic pathway. Although no changes in volatile emissions were detected in the tomato plants, there were significant increases in phenylpropanoid glucosides. Because SGN-U565076 (formerly SGN-U217248) was the only glucosyltransferase found to be upregulated, it was proposed as the likely candidate for phenylpropanoid glucosylation (Dal cin *et al*., 2011). For this reason, ripe fruits from 11 of the best SNG-U565076 overexpressing plants were collected, lyophylized, and sent to for analysis to Dr. Alisdair Fernie at the Max Planck Institute for Molecular Plant Physiology for metabolite analysis.

Discussion

Analysis of glycosides from both petunia and tomato showed that the 2-phenylethanol glycosides are decorated with multiple sugars. This is similar to what has been observed in other studies on flower and tomato glycosides, although the precise structure of tomato glycosides remains unknown (Tikunov *et al*., 2010; Watanabe *et al*., 2001). Based upon the product of MW 387 found in the IL8-2-1 enriched peaks, one possible glycoside present is 2-phenylethyl-6’-O-malonyl-ß-D-glucopyranoside. 6’-O-

malonyl-ß-D-glucopyranosides have been shown to be conjugated with flavor molecules in raspberry, strawberry, guava, and papaya (Withopf *et al*., 1997).

The high number of UGT unigenes found in the tomato database is not surprising as 120 UGTs have been identified in the Arabidopsis genome (Paquette *et al*., 2003). This is part of the challenge of working with UGTs; they are numerous and redundant. Each enzyme is believed to recognize multiple substrates and multiple enzymes recognize a single substrate.

Another challenge of working with UGTs is that when recombinant proteins are produced, they are often in the insoluble fractions and sometimes only slightly present in the soluble fraction (Personal communication, Eng-Kiat “Jack” Lim). Expression of these proteins in yeast versus *E. coli* often does not improve yield (Personal communication, Wilfred Schwab). Development of an *in planta* expression system for screening and characterization of UGTs may be a way to overcome this challenge. Attempts to solubilize protein using urea could also be made, but this can be problematic as it usually result in improper protein refolding.

One attempt to develop an *in planta* screen for a 2-phenylethanol UGT was done by expressing candidate UGTs with PAAS from *Petunia hybrida*. This bifunctional enzyme catalyzes the decarboxylation and oxidation of phenylalanine to form phenylacetaldehyde (Kaminaga *et al*., 2006). Enzymes that reduce phenylacetaldehyde to 2-phenylethanol appear to be present in other members of the Solanceae family, petunia and tomato (Tieman *et al*., 2007). Fortunately, there was also an endogenous enzyme in *N. benthamiana* that performed this last step, allowing us to reconstruct the 2-phenylethanol pathway in a plant that is easily transformable by only adding PAAS.

Using this approach, a candidate UGT SGN-U578227 that decreased 2-phenylethanol pools was identified.

However, when SGN-U578227 was overexpressed or suppressed down in tomato fruit, no changes in volatiles that are part of the 2-phenylethanol synthetic pathway were observed. The overexpression of the other candidate UGTs was also not effective in changing emitted volatile levels. Altering the endogenous expression of a single UGT may not be able to affect 2-phenylethanol emission from fruit if the activity of another UGT compensates for the gain or loss of activity. This scenario is highly probable as UGTs are known to be nonspecific and redundant.

It is also possible that none of these UGTs glycosylate 2-phenylethanol in tomato. However, due to the high level of expression of these genes in tomato fruit it is likely that they have an important role there. Many other metabolites in the phenylpropanoid pathway are known to be glycosylated, such as flavonoids, anthocyanins, and lignin. Further analysis of these metabolites is needed in order to assess if there are any effects of increasing the expression of these UGTs.

Materials and Methods

Glycoside Extraction from *Petunia hybrida* Flowers

Petunia glycosides were extracted from 400 flowers with ethyl acetate according to the methods of Oka et al. (1999). Extracts were dried under N_2 gas and resuspended in methanol. Soluble material was transferred to a new vial and dried under N_2 gas. Resulting residue was resuspended in 30% methanol and soluble material was dried. Residue was resuspended in 30% methanol and separated by flash chromatography on a C18 column. Resulting fractions with detectable UV absorption at 280 nm were analyzed for the presence of 2-phenylethanol glycoside by hydrolysis of the glycoside

and measurement of free 2-phenylethanol by GC-MS (see below). Fractions containing 2-phenylethanol were further purified by LC-MS. Mass spectra seemed to indicate the presence of both di- and trisaccharides bound to 2-phenylethanol. Peaks with the highest purity of the hypothesized 2-phenylethanol trisaccharide were collected and dried. The resulting powder (approximately 500 ng) was sent to the Complex Carbohydrate Research Center (CCRC) at the University of Georgia for further analysis.

Glycoside Extraction from Tomato Fruits

Tomato glycosides were extracted from 1 kg of ripe M82 and IL8-2-1 fruits by pureeing fruit in a blender and then centrifuging the puree at 9000 g for 10 min. The supernant was run over a Sep-Pak® C18 column (Waters, www.waters.com) and eluted with methanol. The elutant was dried, lyophilized, and shipped to the Max Planck Institute in Golm, Germany. Crude extracts were further purified by LC-MS.

Phylogenetic Tree of UGTs

S. lycopersicum UGT candidates were identified by a TBLASTN search of the sol genomic network Lycopersicon combined (tomato) unigene database using previously characterized *Arabidopsis* UGT sequences (www.p450.kvl.dk). The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The bootstrap test (1000 replicates) was used to calculate the percentage of replicate trees in which the associated taxa clustered together (Felsenstein, 1985). The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965). The analysis involved 42 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 285 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura *et al.,* 2011).

Cloning and Protein Expression in *E. coli*

UGT candidate genes, SGN-U578221, SGN-U578227, SGN-U565076, SGN-U584032, SGN-U576693, and SGN-U571691, were PCR-amplified from *S. lycopersicum* fruit cDNA. The products were cloned into pENTR/D/TOPO vectors and sequenced (CHUL Research Center, http://www.sequences.crchul.ulaval.ca). The coding regions were then cloned into vectors pDEST15 and pDEST17 by recombination and transformed into *E. coli* BL21-DE3 (Invitrogen, http://www.invitrogen.com) or Rosetta™(DE3)pLysS Competent Cells (EMD, http://www.emdchemicals.com) for inducible protein expression. Bacteria were precultured for 16 h at 37ºC in Luria-Bertani broth containing 50 µg/mL carbenicillin and the culture was used to inoculate 1L of M9 minimal medium. Cells were grown at 24ºC to an OD_{600} of 0.5. Protein expression was induced by adding isopropyl-1-ß-D-thiogalactoside to the medium at a final concentration of 0.1 mM. Induced cultures continued growing at 20ºC for 16 h. See chapter 3 for protein 6Xhis-tag protein purification steps.

For GST-tag protein purification, cells were centrifugation (10 min, 4 420 g) and resuspended in PBS and Bacterial Protease Inhibitor Cocktail (Sigma, http://www.sigmaaldrich.com). Cells were lysed with sonication. Proteins were purified from cell lysates using gravity flow with Glutathione-Superflow Resin (Clontech, http://www.clontech.com) according to the manufacturer's instructions.

Transient Expression in *N.benthamiana*

Bacteria carrying pHK1001 constructs (see Chapter 2), p19 (Voinnet *et al*., 2003), and PAAS (Kaminaga *et al*., 2006) were precultured for 2 days at 28ºC in Luria-Bertani broth containing 50 µg/mL spectinomycin and the cultures were used to inoculate 100 mL of the same medium. Cells were grown at overnight at 28 ºC. Cultures were pelleted

by centrifugation (10 min, 6 000 g). The pellet was resuspended in 200 mL MES buffer [10 mM $MgCl_2$; 10 mM MES, pH 5.8]. Three mL syringes were used to infiltrate the leaves. Three different treatments were used: p19 alone, p19 with PAAS, and p19 with PAAS and pHK1001 constructs. Five leaves of five plants were used per treatment. After 5 days, leaves were finely chopped. Six grams of sample was placed in a glass tube. Volatile collection procedure was as described in chapter 3, except that volatiles were collected from 2 h.

Transgenic Expression in Tomato

The full-length open reading frame of the UGT candidate genes were cloned into a vector, pHK1001, containing the constitutive FMV 35S promoter (Richins *et al*., 1987) followed by the *nos* 3´ terminator, in both the sense and antisense directions. *S. lycopersicum* IL8-2-1 cotyledons were transformed by Agrobacterium-mediated transformation (McCormick *et al*., 1986) with the kanamycin selectable marker, NPTII.

RNA Extraction and Volatiles Collection

RNA was extracted from leaf tissue as described in Chapter 2. RNA was quantified using Power SYBR® Green RNA-to-CT™ *1-Step* Kit (Invitrogen, www.invitrogen.com). A total reaction volume of 20 µL and 100 ng of RNA was used per reaction. Quantitative PCR was performed with StepOnePlus™ Real-Time PCR System. Volatiles were collected as described in Chapter 2.

Table 4-1. Comparison of free and bound volatiles from ripe cv. Moneymaker tomatoes

Volatile	Free compound (µg/L)	Glycoside (µg/L)
Hexanal	1827.0	402.9
3-Methylbutanol	623.6	7240.0
trans-2-Hexenal	122.4	192.8
cis-3-Hexenol	3121.0	23.4
2-Isobutylthiazole	12.0	Tr
Methyl salicylate	1.3	2.5
Guaiacol	3240.0	18.6
2-Phenylethanol	67.0	539.8
β-Ionone	2.0	Tr

Adapted from Ortiz-Serrano and Gil, 2007

Table 4-2. EST counts for UGT candidate unigenes

Unigene	Total ESTs	ESTs from fruit	Nonspecified
SGN-U578221	135	55	22
SGN-U578227	93	83	2
SGN-U565076	21	12	7
SGN-U584032	77	64	9
SGN-U576693	30	24	2
SGN-U571691	7	1	4

EST numbers from http://solgenomics.net/

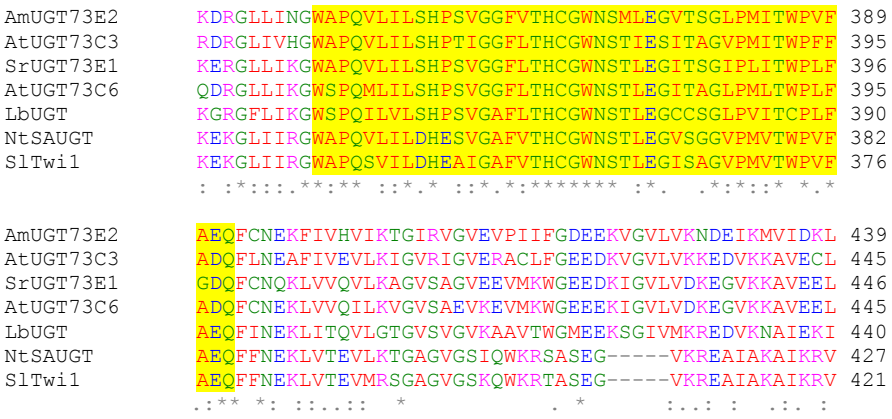


Figure 4-1. Amino acid alignment of PSPG region. Amino acid sequences from UGTs were aligned using ClustalW. The highlighted region shows the 44 amino acids of the PSPG region. An asterisk (*) indicates a fully conserved residue. A colon (:) indicates conservation between groups of strongly similar properties. A period (.) indicates conservation between groups of weakly similar properties. AmUGT73E1 is *Antirrhinum majus* (BAG16513). AtUGT73C3 is *Arabidopsis thaliana* (NP_181216). SrUGT73E1 is *Stevia rebaudiana* (AAR06917). AtUGT73C6 is *Arabidopsis thaliana* (NP_181217). LbUGT is *Lycium barbarum* (BAG80555).NtSAUGT is *Nicotiana tabacum* (AAB36652). SlTwi1 is *Solanum lycopersicum* (CAA59450).

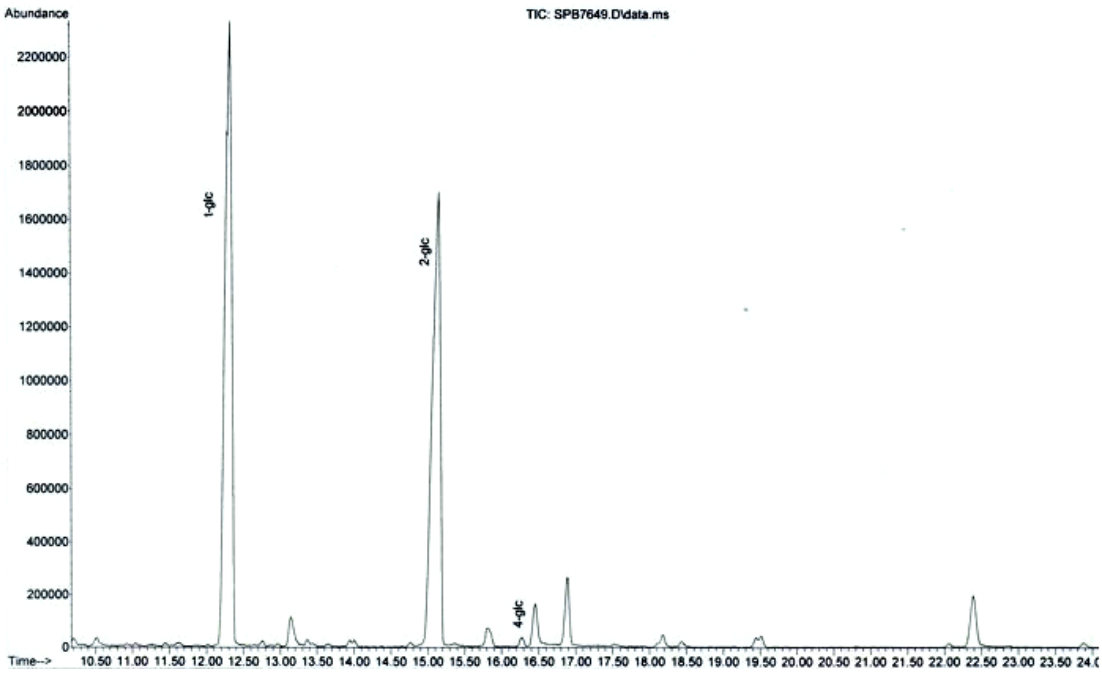


Figure 4-2. Isolated glycosides from Petunia. LC-MS analysis of purified putative 2-phenylethanol glycoside was performed at the CCRC by hydrolysis of the glycoside bond.

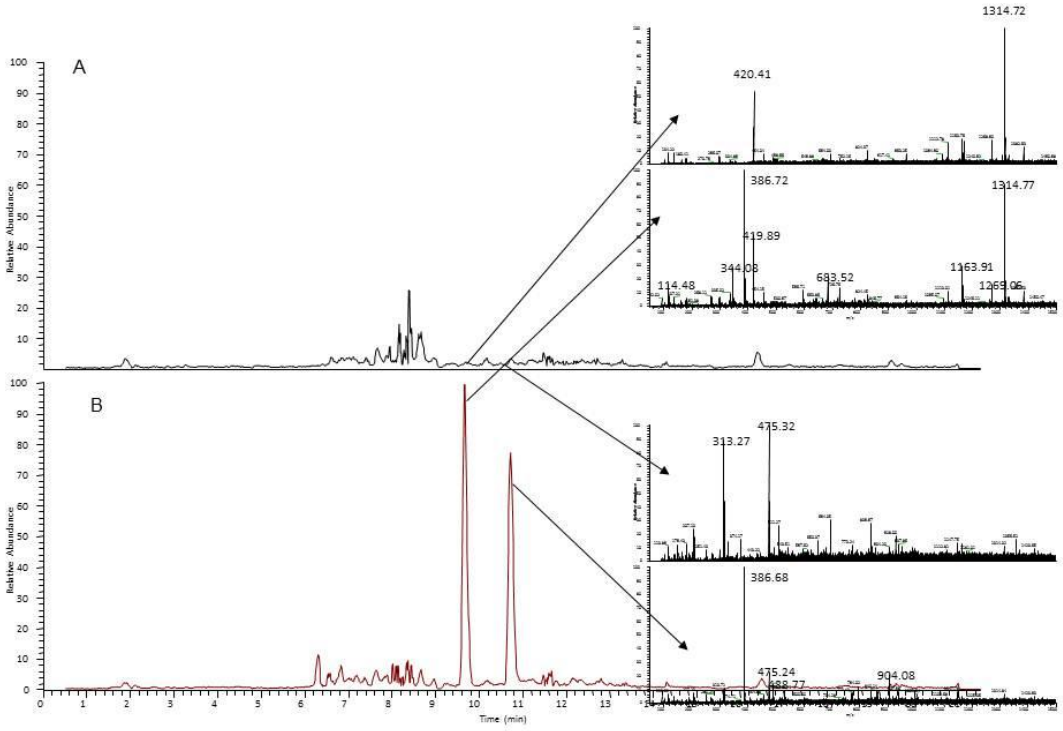
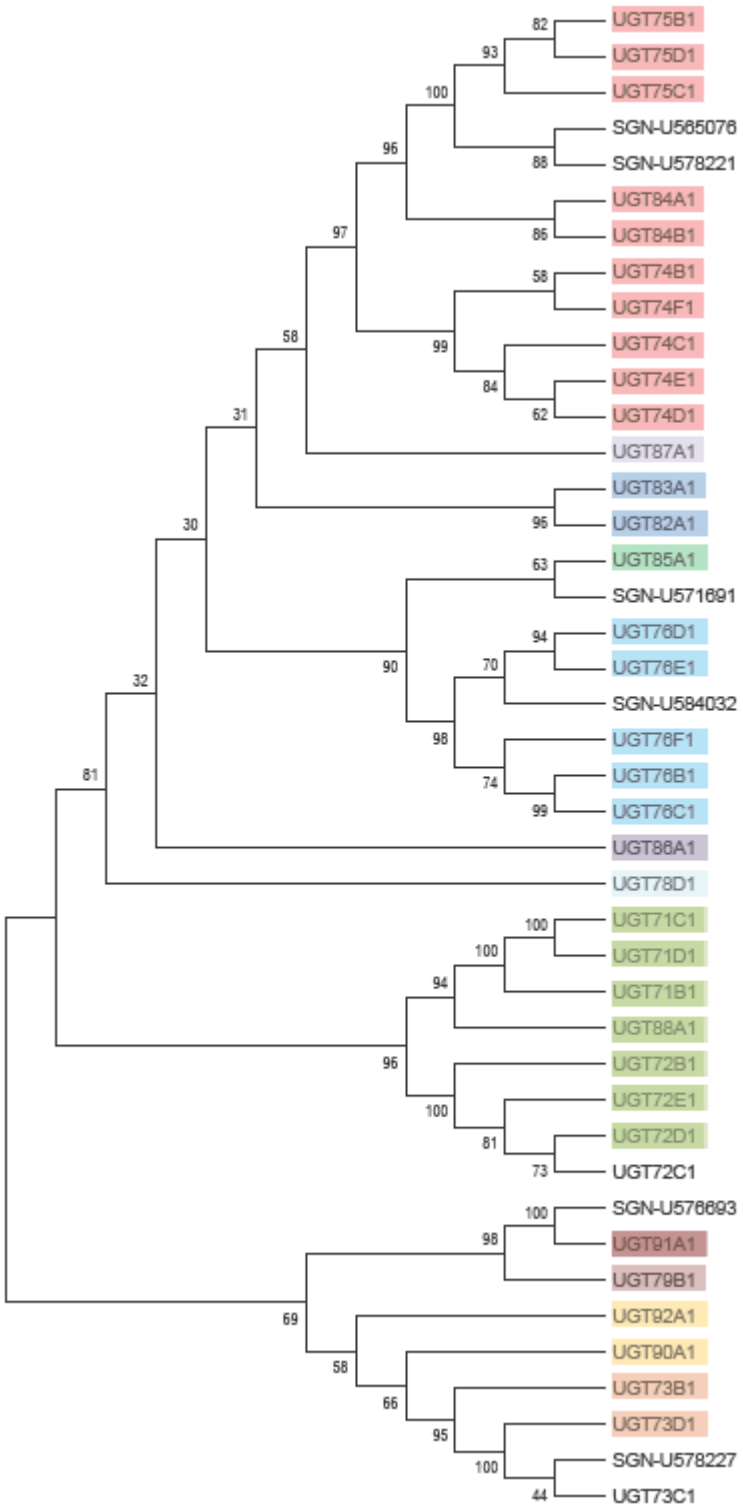


Figure 4-3. Comparison of M82 and IL8-2-1 glycosides. Glycosides were extracted from ripe (A) M82 and (B) IL8-2-1 fruits. Crude extracts were analyzed in the lab of A. R. Fernie at the Max Planck Institute by T. Tohge using LC-MS. Two highly abundant peaks were observed in the IL8-2-1 by the M82 spectrum. Fractionation of these peaks (shown in inset) revealed a unique product of MW 387 in IL 8-2-1.

Figure 4-4. Homology of tomato UGT candidates with known UGTs from *Arabidopsis*. Amino acid sequences of the 8 tomato UGT candidates (SGN-U) and representatives of different *Arabidopsis* UGT families (www.p450.kvl.dk) were aligned using ClustalW. A Neighbor-Joining tree was constructed using MEGA5 (Tamura *et al*., 2011). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. Colored boxes represent *Arabidopsis* groups as defined by Li *et al*., 2001 and are as follows: ■ group A, ■ group C,■ group D, ■ group E, ■ group F, ■ group G, ■ group H, ■ group I, ■ group J, ■ group K, ■ group L.


UGT75B1
UGT75D1
UGT75C1
SGN-U565076
SGN-U578221
UGT84A1
UGT84B1
UGT74B1
UGT74F1
UGT74C1
UGT74E1
UGT74D1
UGT87A1
UGT83A1
UGT82A1
UGT85A1
SGN-U571691
UGT76D1
UGT76E1
SGN-U584032
UGT76F1
UGT76B1
UGT76C1
UGT86A1
UGT78D1
UGT71C1
UGT71D1
UGT71B1
UGT88A1
UGT72B1
UGT72E1
UGT72D1
UGT72C1
SGN-U576693
UGT91A1
UGT79B1
UGT92A1
UGT90A1
UGT73B1
UGT73D1
SGN-U578227
UGT73C1

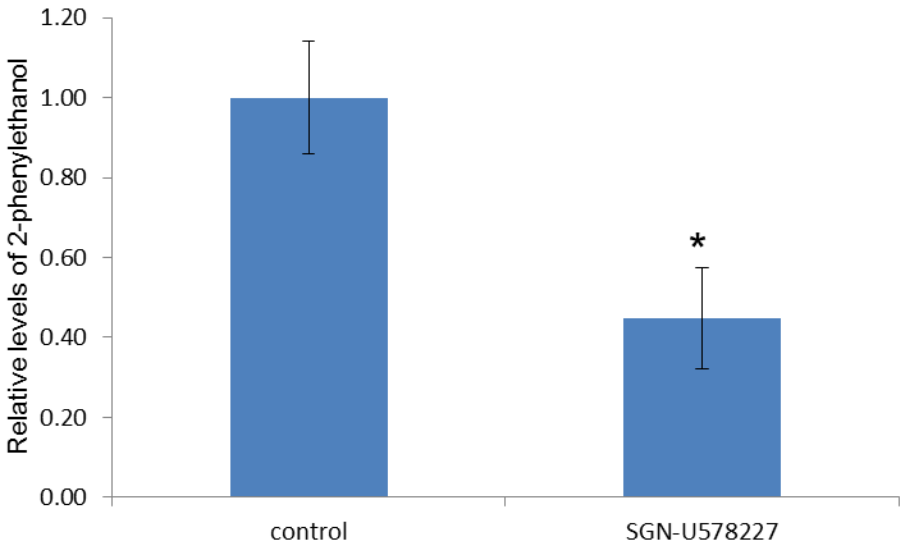


Figure 4-5. Relative levels of 2-phenylethanol from transient expression in *N. benthamiana*. Volatiles were collected from *N. benthamiana* leaves 5 days after inoculation with p19 and PAAS (control) or p19, PAAS, and pHK-SGN-U578227. 2-phenylethanol levels were normalized to the control. Error bars represent standard error of the mean. The asterisk (*) indicates a significant difference ($p<0.05$) by Student's t-test.

CHAPTER 5
CONCLUSIONS

In the course of this work, a catechol OMT, CTOMT1, capable of synthesizing guaiacol from catechol in tomato fruits has been identified. *In vitro* biochemical analysis showed an affinity for catechol similar to other previously characterized catechol OMTs. Altering expression of the *CTOMT1* gene can significantly affect the levels of guaiacol synthesis in tomato fruit. However, under some circumstances, steps leading up to catechol synthesis can limit the ability of the fruit to synthesize guaiacol. The ability of CTOMT1 to increase guaiacol production when catechol is not limiting was demonstrated by feeding *CTOMT1* overexpressing fruit pericarp discs with excess amounts of catechol. Since reduced expression of *CTOMT1* results in reduced guaiacol synthesis, it should be possible to obtain fruits with significantly reduced guaiacol synthesis by a variety of transgenic and non-transgenic techniques. This might be done by looking for natural variation in heirloom varieties or, as we show here, suppressing down *CTOMT1* expression.

Understanding glycosyltransferase activity and regulation is of great importance in our ability to redirect volatiles from glycosylation and to create a more flavorful tomato or, as in the case of guaiacol, remove undesirable volatiles by glycosylating them. However, this endeavor presents many challenges as glycosyltransferases lack single substrate specificity and putative glycosyltransferases are numerous in the tomato genome. Proper characterization of glycosides will help to determine what types of reaction and enzymes are required for flavor molecule glycosylation. Additionally, development of new techniques for screening UGT candidates will help to overcome the challenge expressing enzymes in non-plant systems.

APPENDIX
PROMOTER AND GENOMIC SEQUENCE COMPARISON

Below is shown the promoter and genomic sequence comparison of *CTOMT1* between *S. lycopersicum* cv. M82, *S. lycopersicum* cv. Heinz1706, and *S. pennellii*. Sequence alignment was performed with ClustalW. Bases in red are coding sequence.

```
M82             TATATATCAATTTTCAACATGATA-------TAAAGATAGGCATTTGGAGAGAAATTTTG
Heinz1706       TATATATCAATTTTCAACATGATA-------TAAAGATAGGCATTTGGAGAGAAATTTTG
S.pennellii     TATATATCAATTTTCAACATAAGAATTCAGATAAAGATAGGCATTTGGAGAGAAATTTTG
                ******************** * *       ****************************

M82             ATAAAACCCGATGTTGTACAAATTTTAAGGGGTTTAGTTGTAGAATATGAAATACTAGTA
Heinz1706       ATAAAACCCGATGTTGTACAAATTTTAAGGGGTTTAGTTGTAGAATATGAAATACTAGTA
S.pennellii     ATAAAACCCGATTTTGTACAAATTTTAAGGG-TTCAGTTATAGCATATGAAATACTAGTA
                ************ ****************** ** **** *** ****************

M82             ATCTGTAGCATGAAATAAGAAGGCATATAGGTATCGGGTGGCATTATTTTATGAGCCCAC
Heinz1706       ATCTGTAGCATGAAATAAGAAGGCATATAGGTATCGGGTGGCATTATTTTATGAGCCCAC
S.pennellii     ATCTGTAGCATGAAATAAGAAGGCATATAGGTATCGGGTGCCATTATTTTATGAGCCCAC
                ************************************** ******************

M82             TTAGCCATTAACTTTCAAAATAAAATGGAACATACTTGGGCCAGAACTCAGCGAATATGG
Heinz1706       TTAGCCATTAACTTTCAAAATAAAATGGAACATACTTGGGCCAGAACTCAGCGAATATGG
S.pennellii     TTAGCCATTAA-----------AAATGGAACATAATTGGGCCAGACCTCAGCGAATATGG
                ***********           ************ ********** **************

M82             GCTTAGCAATTGCATATGGACCTCACTGT-----------TAGGTTCCCTATATAACTAG
Heinz1706       GCTTAGCAATTGCATATGGACCTCACTGT-----------TAGGTTCCCTATATAACTAG
S.pennellii     GCTTAATAATTGCATATGGACCTTACTGTGCAACTTATGCTAGGTTCCCTATATAACTAG
                *****  **************** *****           ********************

M82             TTCACATAATACTTATTGTTACAAGGAGCTGAATTTGTAACAAGGTCATATATATATATA
Heinz1706       TTCACATAATACTTATTGTTACAAGGAGCTGAATTTGTAACAAGGTCATATATATATATA
S.pennellii     TTCACATAATACTTTTTGTTACAAGGAGCTGAATTTGTAACAAGGTCATATATATATA--
                ************** ******************************************

M82             TATACTGTACTAACATTATAAAGATAAGCGTGCAAAAACCAATAAACGAATAATGGTTAC
Heinz1706       TATACTGTACTAACATTATAAAGATAAGCGTGCAAAAACCAATAAACGAATAATGGTTAC
S.pennellii     --CACTGTACTAACATTATAAAGATAAGTGTGCAAAAACCAACAAACGGATAATGGTTAC
                ************************ ************* ***** ***********

M82             TTCGGATTAAGTCAAGGAAATACTGAAATCAAGGAGTTTGATTTTAAGAAAGAACTTTAC
Heinz1706       TTCGGATTAAGTCAAGGAAATACTGAAATCAAGGAGTTTGATTTTAAGAAAGAACTTTAC
S.pennellii     TTCGGATTAAGTCCAGGAAATACTAAAATCAAGGAGTTTGATTTTAAGAAAGAACTTTAC
                ************* ********** ***********************************

M82             TTGATGTAGAATTTTAAATCAAGAGAGTTTTGAGT-------------------------
Heinz1706       TTGATGTAGAATTTTAAATCAAGAGAGTTTTGAGT-------------------------
S.pennellii     TTTATGTAGAATTTTAAATCAAGAGAGTTTTGAGTTGAGGTGGAGTTTGAAGTGAAAGAA
                ** *******************************

M82             -AACACTCTGAAGAGTTGTGCTTGAGAGTCACTCAGAACAAGGTGTGCACTCACAGAGCA
Heinz1706       -AACACTCTGAAGAGTTGTGCTTGAGAGTCACTCAGAACAAGGTGTGCACTCACAGAGCA
S.pennellii     TAACTCTCTAAAGAGTTGTGCTTGAGAGTCACTCAGAACAAGGTGTGCACTCACAGAGCA
                *** **** ************************************************

M82             AAAACCAATTGGCTTCGCCAATGTTGTTTGACTATTGAAGGAACACATTGAAGAATCAGG
```

```
Heinz1706       AAAACCAATTGGCTTCGCCAATGTTGTTTGACTATTGAAGGAACACATTGAAGAATCAGG
S.pennellii     AAAACCAATTGGCTTCGCCAATGTTGTTTGACTATTGAAGGAACACATTGAAGAATCAGG
                ************************************************************

M82             TCCTAATGCAACTACAAGTTTTAGCCTTCATGTGTTCATTTGAGTTGTAATATTAATGCA
Heinz1706       TCCTAATGCAACTACAAGTTTTAGCCTTCATGTGTTCATTTGAGTTGTAATATTAATGCA
S.pennellii     TCCTAATGCAACTACAAGTTTTAGCGTTCATGTGTTCATG-GAGTTGTAATATTAATGCA
                ************************ *************  ******************

M82             AATCTTAATTTTAGTGTGTTATTGGATTATAATCTTCAAGTTGGGATAACTTAAAGATTT
Heinz1706       AATCTTAATTTTAGTGTGTTATTGGATTATAATCTTCAAGTTGGGATAACTTAAAGATTT
S.pennellii     A-TCTTAATTCTTGTGTGTTATTGGATTATAATCTTCAAGTTGGGATAACTTAAAGATTT
                * ******** * ********************************************

M82             GAGGACATATATCTTGAGAGGTTTGTGAGTTGTTAAGAATTAGAGTTCATAATTTTGTGG
Heinz1706       GAGGACATATATCTTGAGAGGTTTGTGAGTTGTTAAGAATTAGAGTTCATAATTTTGTGG
S.pennellii     GAGGACATATATCTTGAGAGGTTTGTGAGTTGTTAAGAATTAGAGTTCATAATTTTGTAT
                ************************************************************

M82             TTTAGAATTGTTAAGAATTAGAGTTCATAATGTCTTGTTGAAAGGCTCATTGTGCTTTAG
Heinz1706       TTTAGAATTGTTAAGAATTAGAGTTCATAATGTCTTGTTGAAAGGCTCATTGTGCTTTAG
S.pennellii     ATTACA------------TAGA-TTCATAATGTCTTGTTGAAAGGCTCATTGTGGATTAG
                *** *            **** ********************************  ****

M82             AAAAGTTGTGGTTAAATGTTGTAGATGTACATGTGATTTTT---------GTGAACTGGA
Heinz1706       AAAAGTTGTGGTTAAATGTTGTAGATGTACATGTGATTTTT---------GTGAACTGGA
S.pennellii     AAAAGTTGTGGTTAAATGTTGTAGATGTACATGTGATTTTTATGACTTTTGTGAGCTGGA
                ****************************************         **** *****

M82             TATTTTTACATAAAAATAATGTAGTGTTGATACCATTTTGTTAAAGCACATTAGCCAAGA
Heinz1706       TATTTTTACATAAAAATAATGTAGTGTTGATACCATTTTGTTAAAGCACATTAGCCAAGA
S.pennellii     TATTTTTACATAAAAATAATGTAGTGTTTATACTATCTTGTTAAAGCACATTAGCCAAGA
                **************************** **** ** ************************

M82             TACTAGTTAATGGGACTATAAGTAGCGTCACGAAATTCTTTTGGTGGAATTCGGTTTAGA
Heinz1706       TACTAGTTAATGGGACTATAAGTAGCGTCACGAAATTCTTTTGGTGGAATTCGGTTTAGA
S.pennellii     TACTAGTTAATGGGACTATAAGTAGCGTCACGAAATTCTTTTGGTGGAATTCGGTTTAGA
                ************************************************************

M82             ATTAATAGAAGTTTAGCATTAATAGGATTATTTCAAGTGACAAGAGAATCATTCAAAATG
Heinz1706       ATTAATAGAAGTTTAGCATTAATAGGATTATTTCAAGTGACAAGAGAATCATTCAAAATG
S.pennellii     ATTAATAGAAGTTTAGCATTAATAGGATTATTTAGAGTAACAAGAGAATCATTCAAAATG
                *********************************  *** **********************

M82             GTAAACATCACTTACTTAGGAAGCTAAAAGAAAACCTTAAAGTAGGTTATCTTTCCATCT
Heinz1706       GTAAACATCACTTACTTAGGAAGCTAAAAGAAAACCTTAAAGTAGGTTATCTTTCCATCT
S.pennellii     GTAAACATCACTTACTTAGGAAGCTAAAAAAACAA-TTAAAGTAGGT-ATCTTTCCATCT
                **************************** ** *  ************ ************

M82             AGAAAGTGATTGAAAAAATAAATCTAGTAATGTTAGGTGCAACAACTTTAAATCATCACA
Heinz1706       AGAAAGTGATTGAAAAAATAAATCTAGTAATGTTAGGTGCAACAACTTTAAATCATCACA
S.pennellii     AGAAAGTGATTGAAAAA-CAAATCTAGTAATGTTAGGTGCAACAACTTTGAATCATCACA
                *****************  ****************************** **********

M82             GAGAAAAAGTGGACTGAAAAAGAATAATCATAAAGGAGATTTCATGATTTGATATATACA
Heinz1706       GAGAAAAAGTGGACTGAAAAAGAATAATCATAAAGGAGATTTCATGATTTGATATATACA
S.pennellii     AAGAAAAAGTGGACTGAAAAAGAATAATCATAAAGGAGATTTCATGATTTGATATATACA
                ************************************************************

M82             TACATAT--------------------TTATTTTTTGTTATACCAAACAAAGATTTTATT
Heinz1706       TACATAT--------------------TTATTTTTTGTTATACCAAACAAAGATTTTATT
S.pennellii     TATATATACATACATATTTATTTTATTTTATTTTTTGTTATACCAAACAAAGATTTTATT
                ** ****                    *********************************
```

```
M82             TATTTATTGTTTTTAAAAAAAAGAAAAATCTCTAGTTGAAGACTCTTTCTTGCAAATTTC
Heinz1706       TATTTATTGTTTTTAAAAAAAAGAAAAATCTCTAGTTGAAGACTCTTTCTTGCAAATTTC
S.pennellii     TATTTATTGTTTTTAAAATGAAGAAAAATCTCTAGTTGAAGACTCTTTCTTGCAAATTTC
                ******************  ****************************************

M82             AACAAGCAACGTATCAAGGTAAAAAAAAAAAAAATAACACATGTAATGTATCTTCATATGT
Heinz1706       AACAAGCAACGTATCAAGGTAAAAAAAAAAAAAATAACACATGTAATGTATCTTCATATGT
S.pennellii     AACAAGTAACTTATCAAGGTAAAAAAAA-----TAACACATGTAATGTATCTTCATATGT
                ****** *** *****************     ****************************

M82             CATCATTAAATAGAAGGGGTTAGTTAGAATTTAGTAATACATTCAAAAAAAACTAACAGC
Heinz1706       CATCATTAAATAGAAGGGGTTAGTTAGAATTTAGTAATACATTCAAAAAAAACTAACAGC
S.pennellii     CATCATTAAATAGAAGGGGTTAGTTAGAATTTAGTAATACATTCAAAAAAAAAACTCACC
                **************************************************    ** *

M82             AATTCAATGTATCTTCATATATTAATGTGGTCATATCAACCTTGAACATATTAAACAATA
Heinz1706       AATTCAATGTATCTTCATATATTAATGTGGTCATATCAACCTTGAACATATTAAACAATA
S.pennellii     CATTCAATGTATCTTCATATATTAATGTGGTCATATTATCTTTGAACATATTAAACAATA
                ********************************** * * ******************

M82             TAATAGAGAAATAAAATTTGTAAATATCGATATTCTACTTCAACTAGACAATTACATTGT
Heinz1706       TAATAGAGAAATAAAATTTGTAAATATCGATATTCTACTTCAACTAGACAATTACATTGT
S.pennellii     TAATAGAAAAATAAAATTTGTAAATGTCGATATTCTACTTCAAGTAGACAATTACATTGT
                ******* ***************** ***************** ****************

M82             TTGTATTCACAATTTTGATAAAGTAATGAGAAGTAAATTAATAGAATACAATAGGAATTT
Heinz1706       TTGTATTCACAATTTTGATAAAGTAATGAGAAGTAAATTAATAGAATACAATAGGAATTT
S.pennellii     TTGTATTCACAATTTTGATAAAGTAATGAGAAGTAAATTAATAGAATACAATAGGAATTT
                ************************************************************

M82             GTATATCCATCGTTAAAAGTCAAGAGATAAAACAAACTTT-ATGTATTTAATTATCTAAG
Heinz1706       GTATATCCATCGTTAAAAGTCAAGAGATAAAACAAACTTT-ATGTATTTAATTATCTAAG
S.pennellii     GTATATCCATCGTTAAAAGTCAAGAGATAAAACAAACTTTTATGTATTTAATTATCTTAG
                ************************************ ***************** **

M82             AGTCAATTAACTAATTGTATGTTAATATGATGGTTAGGTGAAGAAAACATGTTATAGTAA
Heinz1706       AGTCAATTAACTAATTGTATGTTAATATGATGGTTAGGTGAAGAAAACATGTTATAGTAA
S.pennellii     AGTCAATTAACTAATTGTATGTTAATATGATGGTTAGGTGAAGAAAACATGTTATAGTAA
                ************************************************************

M82             TATTGTATGAGGAAAATATGAAGAAAATGACTGAATTCTCTTGTTCAGTAAAGCAGACAG
Heinz1706       TATTGTATGAGGAAAATATGAAGAAAATGACTGAATTCTCTTGTTCAGTAAAGCAGACAG
S.pennellii     TATTATATGAGGAAAATATGAAGAAAATGACTGAATTCTCTTGTTCAGTAAAGCAGACAG
                **** ********************************************************

M82             CCAATCACATGTTAAGTGGCCTACTCTCCACTTTTTT-AGTGGACCTTATGCTTCACTAA
Heinz1706       CCAATCACATGTTAAGTGGCCTACTCTCCACTTTTTT-AGTGGACCTTATGCTTCACTAA
S.pennellii     CAAATCACATGTTAAGTGGCCTACTCTCCATTTTTTTTAGTGGACCTTGTGCTTCACTAA
                * ************************** ****** ********** ***********

M82             CTTTT-TTTTTTTTACCAAAAGCAATAATTTTTAATCCAAACAGTAAACAAAAAAAAAAA
Heinz1706       CTTTT-TTTTTTTTACCAAAAGCAATAATTTTTAATCCAAACAGTAAACAAAAAAAAAAA
S.pennellii     CTTATATTTTTTTTACCAAAAGCAATAATTTTTAATCCAAACAGTAAACACAAAAAGAAA
                *** * ******************************************* ***** ***

M82             A-CATACCACCAACTCACATATACAGGAAGTAACTGTGCACAATGGAAGAAGGAAATGGA
Heinz1706       --CATACCACCAACTCACATATACAGGAAGTAACTGTGCACAATGGAAGAAGGAAATGGA
S.pennellii     CACATACCACCAACTCACATATACAGGAAGTAACTGTGCACAATGGAAGAAGGAAAAGGA
                ************************************************** ***

M82             GCGATCCACTGCTGCTTCGAGATGTTATTATTACAATTTTCAGATTGAACTGAATATACT
Heinz1706       GCGATCCACTGCTGCTTCGAGATGTTATTATTACAATTTTCAGATTGAACTGAATATACT
```

```
S.pennellii     GCGATCCACTGCTACTTCGCGATGTTATTATTAGAATTTTCAGATTGAACTGAATATACT
                ************* ***** ************* **************************

M82             GCTTTCAAGTCATGAACGTGAGATAAAATAATAATATTAGGCAGATAGAGGGAGTGATAT
Heinz1706       GCTTTCAAGTCATGAACGTGAGATAAAATAATAATATTAGGCAGATAGAGGGAGTGATAT
S.pennellii     GCTTTCAAGTCATGAACGTGAGATAAAAAAATAATATTAGGCAGATAGAGAGAGTGATAT
                *************************** ******************** *********

M82             ATACTTCATTAGTCTCCATTTATATAATTATTTTTCTTTTTCATCAGTAAACAAAAAAAA
Heinz1706       ATACTTCATTAGTCTCCATTTATATAATTATTTTTCTTTTTCATCAGTAAACAAAAAAAA
S.pennellii     ATACTTCATTGATCTC----------------------TTTAATCAGTAA-CAAAAAAGA
                **********  ****                     *** ******** ******* *

M82             AGAAAATATTTTTATATTTAATAACAAATTAATTTTTAAATAAATCAGAACAGATAGAAT
Heinz1706       AGAAAATATTTTTATATTTAATAACAAATTAATTTTTAAATAAATCAGAACAGATAGAAT
S.pennellii     ATAAAACATTTCTATATTTAATAACAAATTAATTTTTAAATATATCAGAACAGATAGAAT
                * **** **** ****************************** *****************

M82             GCCACTATGCAATTGAAAAAGAACAAAAAACGAATGAAAAGCAGACGCATTACTAATATT
Heinz1706       GCCACTATGCAATTGAAAAAGAACAAAAAACGAATGAAAAGCAGACGCATTACTAATATT
S.pennellii     GCCACTATGCAATTGAAAAAGAACAAAAAACGAATGAAAAGCAGACGCATTACTAATATT
                ************************************************************

M82             CCCACCAAGAAATCAATTATGACCAATCTTTGACAAAACAACAATTCTTGGTTTGATATT
Heinz1706       CCCACCAAGAAATCAATTATGACCAATCTTTGACAAAACAACAATTCTTGGTTTGATATT
S.pennellii     CCCACCAAGAAATCAATTATGACCAATCTTTGACAAAACAACAATTCTTGGTTTGATGTT
                ************************************************************ **

M82             TATAAAAGGGTAGTCTAACCCCATTATACATCATCTTGAGGCCTAACAAAACACTCCAAG
Heinz1706       TATAAAAGGGTAGTCTAACCCCATTATACATCATCTTGAGGCCTAACAAAACACTCCAAG
S.pennellii     TATAAAAGGGTACTCTAACCCCATTATACATCATCTTGAGGCCTAACAAAACACTCCAAG
                ************ ********************************************

M82             CAGCAAAAATAACATTTTCTTGTTCATCTCTAAGTTCTTTTTAGCTATGGGATCGACAGC
Heinz1706       CAGCAAAAATAACATTTTCTTGTTCATCTCTAAGTTCTTTTTAGCTATGGGATCGACAGC
S.pennellii     CAGCAAAAATAACATTTTCTTGTTCATCTCTAAGTTCTTTTTAGCTATGGGATCGACAGC
                ************************************************************

M82             AAATATCCAGTTAGCAACACAATCGGAAGACGAAGAGCGTAATTGCACGTACGCCATGCA
Heinz1706       AAATATCCAGTTAGCAACACAATCGGAAGACGAAGAGCGTAATTGCACGTACGCCATGCA
S.pennellii     AAATATCCAGTTACCAACACAATCGGAAAACGAAGAGCGTAATTGCACGTACGCCATGCA
                ************* ************** ********************************

M82             ACTACTCTCATCGTCAGTGCTTCCCTTCGTTTTGCACTCAACTATCCAATTGGATGTTTT
Heinz1706       ACTACTCTCATCGTCAGTGCTTCCCTTCGTTTTGCACTCAACTATCCAATTGGATGTTTT
S.pennellii     ACTACTCTCATCGTCAGTGCTTCCCTTCGTTTTGCACTCAACTATCCAATTGGATGTTTT
                ************************************************************

M82             TGACATACTCGCAAAAGATAAAGCCGCCACTAAACTATCTGCTTTAGAAATTGTGTCTCA
Heinz1706       TGACATACTCGCAAAAGATAAAGCCGCCACTAAACTATCTGCTTTAGAAATTGTGTCTCA
S.pennellii     TGAGATACTCGCAAAAGATAAAGCCGCCACTAAACTATCTGCTTTAGAAATTGTGTCTCA
                *** ********************************************************

M82             CATGCCTAACTGTAAGAACCCTGATGCCGCTACCATGCTAGACCGGATGCTTTATGTCCT
Heinz1706       CATGCCTAACTGTAAGAACCCTGATGCCGCTACCATGCTAGACCGGATGCTTTATGTCCT
S.pennellii     CATGCCTAACTGTAAGAACCCTGATGCCGCTACCATGCTAGACCGGATGCTTTATGTCCT
                ************************************************************

M82             AGCTAGTTATTCTTTACTCGATTGCTCGGTTGTTGAAGAGGGAAATGGGGTGACCGAAAG
Heinz1706       AGCTAGTTATTCTTTACTCGATTGCTCGGTTGTTGAAGAGGGAAATGGGGTGACCGAAAG
S.pennellii     AGCTAGTTATTCTTTACTCGATTGTACTGTTGTTGAAGAGGGAAATGGGGTGACCGAAAG
                ************************  * ********************************
```

```
M82             GCGCTATGGTCTGTCACGAGTGGGGAAATTTTTTGTACGTGATGAAGATGGTGCATCCAT
Heinz1706       GCGCTATGGTCTGTCACGAGTGGGGAAATTTTTTGTACGTGATGAAGATGGTGCATCCAT
S.pennellii     GCGCTATGGTCTGTCACGAGTGGGGAAATTTTTTGTACGTGATGAAGATGGTGCATCCAT
                ************************************************************

M82             GGGACCATTGTTGGCTTTGCTTCAAGATAAAGTATTCATTAACAGCTGGTCAGTTTTCTC
Heinz1706       GGGACCATTGTTGGCTTTGCTTCAAGATAAAGTATTCATTAACAGCTGGTCAGTTTTCTC
S.pennellii     GGGACCATTGTTGGCTTTGCTTCAAGATAAAGCATTCATTAACAGCTGGTCAGTTTTCTC
                ******************************** ***************************

M82             TTTTTACTGCAGCAATCTTTCTTTTTAACCAAACTTTTATCATGTCAATTGTATGTGGTC
Heinz1706       TTTTTACTGCAGCAATCTTTCTTTTTAACCAAACTTTTATCATGTCAATTGTATGTGGTC
S.pennellii     TTTTTACTGCAGCAATCTTTCTTTTTAACCAAACTTTTATCATGTCAATTGTATGTGGTC
                ************************************************************

M82             ATCCTAGTATAACCTAACAAATTGAGTATATATTAGAGATTTTCTCACAATATAAGTGAG
Heinz1706       ATCCTAGTATAACCTAACAAATTGAGTATATATTAGAGATTTTCTCACAATATAAGTGAG
S.pennellii     ATCCTAGTATAACCTAACAAATTGAGTATATATTAGAGATTTTCTCACAATATATGTGAG
                ********************************************************* *****

M82             TCAGAGTCAGGTGGATATATCATGCAAAGTTGAAGACCCTTTTTTGATCCCTTCATTATA
Heinz1706       TCAGAGTCAGGTGGATATATCATGCAAAGTTGAAGACCCTTTTTTGATCCCTTCATTATA
S.pennellii     TCAGAGTCAGGTGGATATATCATGCAAAGTTGAAGACCCTTTTTTGATCCCTTCATTATA
                ************************************************************

M82             TTCTTAATATACAAAACATGTATCTTTGCTGGCTATTATATTAGGGCGGCC-AAATAGAT
Heinz1706       TTCTTAATATACAAAACATGTATCTTTGCTGGCTATTATATTAGGGCGGCC-AAATAGAT
S.pennellii     TTCTTAATATACAAAACATGTATCTTTGCTGGCTATTATATTAGGGCGGCCCAAATAGAT
                ************************************************* ********

M82             AATTATTCCTATATATTACTTCATGAAGGAATCTCAGAATATTAATGCTTTCCTGTCGAA
Heinz1706       AATTATTCCTATATATTACTTCATGAAGGAATCTCAGAATATTAATGCTTTCCTGTCGAA
S.pennellii     AATTATTCCTATATATTACTTCATGAAGGAATCTCAGAATATTAATGCTTTCCTGTCGAA
                ************************************************************

M82             CCATCTGGTATCCAAAACTCACTAGGCCGACCAATTAAAATCCATGATGCATAGGACCTA
Heinz1706       CCATCTGGTATCCAAAACTCACTAGGCCGACCAATTAAAATCCATGATGCATAGGACCTA
S.pennellii     CCATCCGGTATCCAAAACTCACTAGGCCGACCAATTCAAATCCATGATGCATAGGACCTA
                ***** ****************************** ***********************

M82             TGACAGAGTGAATGAGTCTATTTCCTAGCTCGAATCAAAGATTTCTGATCAAGTGTAAAG
Heinz1706       TGACAGAGTGAATGAGTCTATTTCCTAGCTCGAATCAAAGATTTCTGATCAAGTGTAAAG
S.pennellii     TTACAGAGTGAATGAGTCTATTTCCTAGCTCGAATCAAAGATTTCTGATCAAGTGTGAAG
                * ********************************************************* ***

M82             TGATGTGATCATGAGACTAATGGAATTTGTAAGTTAATTACAGTTATCATGTTAACAAAT
Heinz1706       TGATGTGATCATGAGACTAATGGAATTTGTAAGTTAATTACAGTTATCATGTTAACAAAT
S.pennellii     TGATGTGATCATGAGACTAATGGAATTTGTAAGTTAATTACAATTATCATGTTAACAAAT
                ****************************************** *****************

M82             ACATCAACTGGTTCAAGTTAGCATATAAATTGCTAACAGAATG----------------T
Heinz1706       ACATCAACTGGTTCAAGTTAGCATATAAATTGCTAACAGAATG----------------T
S.pennellii     ACATCAACTGGTTCAAGTTAGCATATAAATTGCTAAGAGAATACTTTTGCATGAGCCTAT
                ********************************** *****                 *

M82             GTCCACTCAATTGCCAAAGATCAAGGGTACACTATAATTTCAAGAAATTGTTGGATAGTT
Heinz1706       GTCCACTCAATTGCCAAAGATCAAGGGTACACTATAATTTCAAGAAATTGTTGGATAGTT
S.pennellii     GTCCACTCAACTGCCAAAGATCAAGGGTACACTGTAATTTCGAGAAATTATTGGATAGTT
                ********** ********************** ******* ******* **********

M82             AGAGTACGTATGTTATCAGACTCATACTCGTGAAATTACACTGAATATATTGTC-----T
Heinz1706       AGAGTACGTATGTTATCAGACTCATACTCGTGAAATTACACTGAATATATTGTC-----T
S.pennellii     AGGGTACGTATGTTATCAGACTCATACTCGTGAAATTACACTGAATATATTGTTATTAAT
```

```
                ** ************************************************      *

M82             GCTGGATACTTGGGAATTAATTGCTTCCAGATGAGACTGAGGCGTAATTATAGTAGTTGT
Heinz1706       GCTGGATACTTGGGAATTAATTGCTTCCAGATGAGACTGAGGCGTAATTATAGTAGTTGT
S.pennellii     GCTGGATACTTGGGAATTAATTGCTTCCAGATGAGACTGAGGCGTAATTATAGTAGTTGT
                ************************************************************

M82             ATTTCTGACTCTCTCTATCTAATTTAAATTACAGGTTTGAACTAAAAGATGCAGTACTTG
Heinz1706       ATTTCTGACTCTCTCTATCTAATTTAAATTACAGGTTTGAACTAAAAGATGCAGTACTTG
S.pennellii     ATTTCTGACTCTCTCTATCTAATTTAAATTACAGGTTTGAACTAAAAGATGCAGTACTTG
                ************************************************************

M82             AAGGTGGAGTTCCATTTGACAGGGTGCATGGTGTACATGCATTTGAATATCCAAAATTGG
Heinz1706       AAGGTGGAGTTCCATTTGACAGGGTGCATGGTGTACATGCATTTGAATATCCAAAATTGG
S.pennellii     AAGGTGGAGTTCCATTTGACAGGGTGCATGGTGTACATGCATTTGAATATCCAAAATTGG
                ************************************************************

M82             ACCCAAAGTTCAATGATGTTTTCAACCAGGCAATGATAAACCACACAACTGTTGTCATGA
Heinz1706       ACCCAAAGTTCAATGATGTTTTCAACCAGGCAATGATAAACCACACAACTGTTGTCATGA
S.pennellii     ACCCAAAGTTCAATGATGTTTTCAACCAGGCAATGATCAACCACACAACTGTTGTCATGA
                ********************************** **********************

M82             AAAGAATACTTGAAAATTACAAAGGTTTTGAGAATCTCAAAACTTTGGTTGATGTTGGAG
Heinz1706       AAAGAATACTTGAAAATTACAAAGGTTTTGAGAATCTCAAAACTTTGGTTGATGTTGGAG
S.pennellii     AAAGAATACTTGAAAATTACAAAGGTTTTGAGAATCTCAAAACTTTGGTTGATGTTGGAG
                ************************************************************

M82             GTGGTCTTGGTGTTAATCTCAAGATGATTACATCTAAATACCCCACAATTAAGGGCACTA
Heinz1706       GTGGTCTTGGTGTTAATCTCAAGATGATTACATCTAAATACCCCACAATTAAGGGCACTA
S.pennellii     GTGGTCTTGGAGTTAATCTCAAGATGATTACATCTAAATACCCCACAATTAAGGGCACTA
                ********** **************************************************

M82             ATTTTGATTTGCCTCATGTTGTTCAACATGCACCTTCCTATCCTGGTACCTTAATTCTTG
Heinz1706       ATTTTGATTTGCCTCATGTTGTTCAACATGCACCTTCCTATCCTGGTACCTTAATTCTTG
S.pennellii     ATTTTGATTTGCCTCATGTTGTTCAACATGCAACTTCCTATCCTGGTACCTTAATTCTTG
                ******************************** ****************************

M82             TTTTATTGTTCACTTTGATACTTTGTTTCAATGTTAGAGATTTATACTTTGTTTCAATGT
Heinz1706       TTTTATTGTTCACTTTGATACTTTGTTTCAATGTTAGAGATTTATACTTTGTTTCAATGT
S.pennellii     TTTTATTGTTCAATTTGATACTTTGTTTCA-------------------------ATGT
                ************ ***************                           ****

M82             TAGAGATTTAAATTACAATTCATTGGATTGTTTTGTTTGCAAACAAGTTATACAGAGATT
Heinz1706       TAGAGATTTAAATTACAATTCATTGGATTGTTTTGTTTGCAAACAAGTTATACAGAGATT
S.pennellii     TAGAGATTTAAATTACAATTCATTGGATTGTTTTGTTTGCAAACAAGTTATGCAGAGATT
                ************************************************* ********

M82             ATAATACGAGGTTTAAAATAATAACGAGATTCTTTAATCGATAGATTTCTAAAATGGTAG
Heinz1706       ATAATACGAGGTTTAAAATAATAACGAGATTCTTTAATCGATAGATTTCTAAAATGGTAG
S.pennellii     ATAATACGAGGTTTAAAATAATAACGAGATTCTTTGATCG--------------------
                ********************************** ****

M82             CTCTCAATTTCCTAACATGAACTGAATTTGTCTTAATAAATATTGCAGGGGTGGATCATG
Heinz1706       CTCTCAATTTCCTAACATGAACTGAATTTGTCTTAATAAATATTGCAGGGGTGGATCATG
S.pennellii     ----------------------------------------------CAGGGGTGGATCATG
                                                               ***************

M82             TTGGGGGAGATATGTTTGAAAGTGTTCCACAAGGAGATGCTATTTTTTATGAAGGTAATGT
Heinz1706       TTGGGGGAGATATGTTTGAAAGTGTTCCACAAGGAGATGCTATTTTTTATGAAGGTAATGT
S.pennellii     TTGGGGGAGATATGTTTGAAAGTGTTCCACAAGGAGATGCTATTTTTTATGAAGGTAATGT
                ************************************************************

M82             CCAAATCTTTAGCAGAGGCTGTATGTATGTACTGTGCATATATTTGGCTTACATGTCGAA
```

```
Heinz1706       CCAAATCTTTAGCAGAGGCTGTATGTATGTACTGTGCATATATTTGGCTTACATGTCGAA
S.pennellii     CCAAATCTTTAGCAGAGGCAATATGTATGTACTGTGCATATATTTGGCTTACATGTCGAA
                ******************  ****************************************

M82             AGTCTTCTTTAATTTCTTAGATTTTGTGTTCAGTCAAACAAACTTTATTTTGTTCCTCAC
Heinz1706       AGTCTTCTTTAATTTCTTAGATTTTGTGTTCAGTCAAACAAACTTTATTTTGTTCCTCAC
S.pennellii     AGTCTTCTTTAATTT-TTAGATTTTGTGTTCAGTCAAACAAACTTTATTTTGTTCCTCAC
                *************** ****************************************

M82             ATAACCGATGCGAGTTATGTAACGCTTCTTTTTGTTTCACAAATTAGCGGACCTAAATTC
Heinz1706       ATAACCGATGCGAGTTATGTAACGCTTCTTTTTGTTTCACAAATTAGCGGACCTAAATTC
S.pennellii     ATAACCGATGCGAGTTATGTAACGCTTCTTTTTGTTTCACAAATTAGCGGACCTAAATTC
                ************************************************************

M82             AATACTTTTGGGTTCACAAACTTTGGGTTGACCGATTTATGAAATAAAAAAGAAGTCGCT
Heinz1706       AATACTTTTGGGTTCACAAACTTTGGGTTGACCGATTTATGAAATAAAAAAGAAGTCGCT
S.pennellii     AATACTTCTGGGTTCACAAACTTTGGGTTGACCAATTTATGAAATAAAAAAGAAGTCGCG
                ******* ************************ ************************

M82             CACAACTTGTGTCAGCTGGAACTTAACTACTTGCATAGTCTTGCATTCCTGTTCTTCACC
Heinz1706       CACAACTTGTGTCAGCTGGAACTTAACTACTTGCATAGTCTTGCATTCCTGTTCTTCACC
S.pennellii     CACAACTTGTGTCAGCTGGAACTTAACTACTTGCATAGTCTTGCATTCCTGTTCTTCACC
                ************************************************************

M82             AATAGTATCTATAACCTATGATTAATAAGGGCATTCTGTGTTTATTGATGAAAAGTGGAT
Heinz1706       AATAGTATCTATAACCTATGATTAATAAGGGCATTCTGTGTTTATTGATGAAAAGTGGAT
S.pennellii     AATAGTATCTATAACCTATGATTAATAAGGACATTCTGTGTTTATTGATGAAAAGTGGAT
                ****************************** ******************************

M82             CCTTCATGACTGGAGTGATGGTCACTGCCTCAAATTGCTGAAGAACTGTCATAAGGCTCT
Heinz1706       CCTTCATGACTGGAGTGATGGTCACTGCCTCAAATTGCTGAAGAACTGTCATAAGGCTCT
S.pennellii     CCTTCATGACTGGAGTGATGGTCACTGCCTCAAATTGCTGAAGAACTGCCATAAGGCTCT
                ************************************************ ***********

M82             ACCGGACAACGGAAAGGTGATTGTTGTGGAGGCCAATCTACCAGTGAAACCTGATACTGA
Heinz1706       ACCGGACAACGGAAAGGTGATTGTTGTGGAGGCCAATCTACCAGTGAAACCTGATACTGA
S.pennellii     ACCGGACAACGGAAAGGTGATTGTTGTGGAGGCCAATCTACCAGTGAAACCTGATACTGA
                ************************************************************

M82             TACCACAGTGGTTGGAGTTTCACAATGTGATTTGATCATGATGGCTCAGAATCCCGGAGG
Heinz1706       TACCACAGTGGTTGGAGTTTCACAATGTGATTTGATCATGATGGCTCAGAATCCCGGAGG
S.pennellii     TACCACAGTGGTTGGAGTTTCACAATGTGATTTGATCATGATGGCTCAGAATCCGGGAGG
                ********************************************************* *****

M82             TAAAGAGCGTTCTGAACAGGAGTTTCGGGCATTGGCAAGTGAAGCTGGATTCAAAGGTGT
Heinz1706       TAAAGAGCGTTCTGAACAGGAGTTTCGGGCATTGGCAAGTGAAGCTGGATTCAAAGGTGT
S.pennellii     CAAAGAGCGTTCTGAACAGGAGTTTCGGGCATTGGCAAGTGAAGCTGGATTCAAAGGTGT
                ************************************************************

M82             TAACCTAATATGTTGTGTCTGTAATTTTTGGGTCATGGAATTTTACAAGTAGATTTCCAC
Heinz1706       TAACCTAATATGTTGTGTCTGTAATTTTTGGGTCATGGAATTTTACAAGTAGATTTCCAC
S.pennellii     TAACCTAATATGTTGTGTCTGTAATTTTTGGGTCATGGAATTTTACAAGTAGATTTCCAC
                ************************************************************

M82             AACCTACTTCGCTCTTATGATTATGTATTTTCGTGGCACTCTGGGACTGGAATTTATAAA
Heinz1706       AACCTACTTCGCTCTTATGATTATGTATTTTCGTGGCACTCTGGGACTGGAATTTATAAA
S.pennellii     AACCTACTTCGCTCTTATGATTATGTACTTTCGTGGCACTCTGGGACTGGAATTTATAAA
                ************************** *******************************

M82             CTAGCCCAGCTTGAATGTTTGACGTTGATTCCTAATAATATTTATATTACTACTTGTTTG
Heinz1706       CTAGCCCAGCTTGAATGTTTGACGTTGATTCCTAATAATATTTATATTACTACTTGTTTG
S.pennellii     TTAGCCCAGCTTGAATGTTTGACGTTGATTCCTAATAATATTTATATTACTACTTGTTTG
                ************************************************************
```

```
M82             TTTCTCTAGTTTGAGAGGATGTCATTAACTCATTGTAACTTCTGTCTTAATAATATTTAT
Heinz1706       TTTCTCTAGTTTGAGAGGATGTCATTAACTCATTGTAACTTCTGTCTTAATAATATTTAT
S.pennellii     TTTCTCTAGTTTGAGAGGATGTCAT--------TGTAACTTCTGTCTTAATAATATTTAT
                ************************        ****************************

M82             ATATTCCTCTGTTCCATTTGATATGATGCCTTCCTTTTTAGTTTTCAGAAAAAAGAATGA
Heinz1706       ATATTCCTCTGTTCCATTTGATATGATGCCTTCCTTTTTAGTTTTCAGAAAAAAGAATGA
S.pennellii     ATATTCCTCTGTTCCATTTGATATGATGCCTTCCTTTTTAGTTTTCAGCAAAA-GAATGA
                ********************************************** **** ******

M82             ACCCAAACATACGTAACCCGTCCAATCCGCCCAGAATTTTAAGGGTTGGGCTCAAGATAA
Heinz1706       ACCCAAACATACGTAACCCGTCCAATCCGCCCAGAATTTTAAGGGTTGGGCTCAAGATAA
S.pennellii     ACCCAAACATACGTAACCCGTCCAATCCGCCCAGAATTTTAAGGGTTGGGCTCAAAATAA
                ********************************************************* ****

M82             TTTGAAATGGGTTCAATCTCAACCCATTCAAGCAAAGAGAATTCTCAATTGAGCCCAATT
Heinz1706       TTTGAAATGGGTTCAATCTCAACCCATTCAAGCAAAGAGAATTCTCAATTGAGCCCAATT
S.pennellii     TTTGAATTGGGTTCAATCTCAACCCATTCAAGCAA-GAAAATTCTCAATTGAGCCCAATT
                ****** **************************** ** **********************

M82             CAATCTCCAATTTCAACCCGTTTTAAAAAAATTTATTAAGATATGTTCCTATATTGAAAG
Heinz1706       CAATCTCCAATTTCAACCCGTTTTAAAAAAATTTATTAAGATATGTTCCTATATTGAAAG
S.pennellii     CAATCTCCAATTTCAACCCGTTTTAAATTTTTTTATTAAGATATGTTCCTATATTGAAAG
                *************************    ****************************

M82             TATGAATTATTATCTATTTAACATCTTTTAGAATTTATCTATCAATTTGTTACTTTTTTA
Heinz1706       TATGAATTATTATCTATTTAACATCTTTTAGAATTTATCTATCAATTTGTTACTTTTTTA
S.pennellii     TATGAGTTATTATATATTTAACATCTCTCGGGATTTATCTATCAATTTGTTATTTTTTTA
                ***** ******* ************ *  * ******************** *******

M82             ACAAAAAATTCTTGAGCCGAAATTCAAATTGTGATTATAAAAGTTATATATCAATATGTT
Heinz1706       ACAAAAAATTCTTGAGCCGAAATTCAAATTGTGATTATAAAAGTTATATATCAATATGTT
S.pennellii     ACAAAAAATTCTTGAGTCGAAATTCAAATTGTGATTATAAAAGTTATATATCAATATGTT
                **************** ****************************************

M82             AAATTATTGAGATTAATCGGATCAAATTGGGTAGGTCAA-GACCAACCCCGTTTTTTAGC
Heinz1706       AAATTATTGAGATTAATCGGATCAAATTGGGTAGGTCAA-GACCAACCCCGTTTTTTAGC
S.pennellii     AAATTATTGTGATTAATCGGGTCAAATTGGGCAGGTCAAAGACCAAACCCGTTTTTTAGC
                ********* ********** ********** ******* ****** *************

M82             CCATTTGA---------------ACCCAAAGTAAACTTGGGCGGGTCGAGACCCAACCCA
Heinz1706       CCATTTGA---------------ACCCAAAGTAAACTTGGGCGGGTCGAGACCCAACCCA
S.pennellii     CCATTTGAGCCCAACCCATTTGAACCCAAAGTAAACTTGGGCGGGTCGAGACCCAACCCA
                ********               ************************************

M82             ATTTCTATTCAACCCATTGTAATATTTTAAATTTCAACC--ACCCGCCCATTTGACACCC
Heinz1706       ATTTCTATTCAACCCATTGTAATATTTTAAATTTCAACC--ACCCGCCCATTTGACACCC
S.pennellii     ATTTCTATTCAACCCATTGTAATATTTTAAATTTCAACCCAACCCGCCCATTTGACACCC
                ************************************  ******************

M82             CTAATTATTATTTTTATTTTCATATTTCCTTTTTCAAACTGCTTTGGGGTGCTTTAGGAA
Heinz1706       CTAATTATTATTTTTATTTTCATATTTCCTTTTTCAAACTGCTTTGGGGTGCTTTAGGAA
S.pennellii     CTAATTATTATTTTTATTTTCATATTTCCTTTTTCAAACTGCTTTGGGATGCTTTAGGAA
                ********************************************** ***********

M82             ACCACACTTTGTCTCTACGAGGTAGGAATAAGGTCTATGTACACTCTACCCTACCCAGAC
Heinz1706       ACCACACTTTGTCTCTACGAGGTAGGAATAAGGTCTATGTACACTCTACCCTACCCAGAC
S.pennellii     ACCACACTTTGTCTCTACGAAGTAGGAATAAGGTTTATGTACAATCTACCCTACCCAGAC
                ******************* ************* ******** ***************

M82             TATACTTGTGAGATTACACTGGATATGCATCCAGTTGTTGTTGTTGGGTTCTAGACTCTA
Heinz1706       TATACTTGTGAGATTACACTGGATATGCATCCAGTTGTTGTTGTTGGGTTCTAGACTCTA
```

```
S.pennellii    TACACTTGTGAAATTAC-------------------GTTATTGTTGGGTTCTAGACTCTA
               ** ******** *****                  *** ********************

M82            ATCTTTTCAAGTTACTAGGAGTAACTTGTACAAATTCAAATCAACTTTTGTAACAAACAT
Heinz1706      ATCTTTTCAAGTTACTAGGAGTAACTTGTACAAATTCAAATCAACTTTTGTAACAAACAT
S.pennellii    ATCTTTTCAAGTTACTAGGAGTAACTTGTACAAATTCAA-TCAACTTTTGTAACAATCAT
               ************************************ *************** ***

M82            GGAGTTTGAGCCAAAGATACTGGCTTTAGCCGAGCCCATACCTCCCTAGTCCCTCCACCC
Heinz1706      GGAGTTTGAGCCAAAGATACTGGCTTTAGCCGAGCCCATACCTCCCTAGTCCCTCCACCC
S.pennellii    GGAATTTGAGCCAAAGATACTGGCTTTAGCCGAGCCCATACCTCCCTAGTCCCTCCACCC
               *** ************************************************************

M82            CTACTAGGATGAGGCATTGCCTCTTCACGATTTGGATGTATAGCTATTGGACTATATAAC
Heinz1706      CTACTAGGATGAGGCATTGCCTCTTCACGATTTGGATGTATAGCTATTGGACTATATAAC
S.pennellii    CTACTAGGATGAGGCATTGCCTCTTCACGATTTGGATGTATAGCTATTGGACTATATAAC
               ************************************************************

M82            CATAGT-------AACATGTTTTATTGCACAAGTTCTTTTAAGCCATTGAATTAGCAAAG
Heinz1706      CATAGT-------AACATGTTTTATTGCACAAGTTCTTTTAAGCCATTGAATTAGCAAAG
S.pennellii    CATAGTCCATAGTAACATGTTTTATTGCACAAGTTCTTTTAAGCCATTGAATTAGCAAAG
               ******       *********************************************

M82            ATATGGATTTACTTGAAAGCATTTGATATACATTAACTTCCAACTGCTAATGAGAACATA
Heinz1706      ATATGGATTTACTTGAAAGCATTTGATATACATTAACTTCCAACTGCTAATGAGAACATA
S.pennellii    ATATGGATTTACTTGAAAGCATTTGATATACATTAACTTCCAACTGCTAACGAGAACATA
               ************************************************* *********

M82            TTGAAGGTGAGGAAATGAAAAGACAATATACAGATAAGCACATATATAGACATAGTTCAG
Heinz1706      TTGAAGGTGAGGAAATGAAAAGACAATATACAGATAAGCACATATATAGACATAGTTCAG
S.pennellii    TTGAAGGTGAGGAAATGAAAAGACAATATACAGATAAGCACATATATAGACATAGTTCAG
               ************************************************************

M82            TTGGGTTTTATTCTGTTAGAATAAAAAGACAAAAGATCGAAGCAGAGTTTACATTTTGAA
Heinz1706      TTGGGTTTTATTCTGTTAGAATAAAAAGACAAAAGATCGAAGCAGAGTTTACATTTTGAA
S.pennellii    TTGG-TTTTATTCTGTTAGAATAAAAAGACAAAAGATCGAAGCAGAGTTTACATTTTGAA
               **** ********************************************************

M82            GAGCAAAGCTGCAAGATTGCTCAACTGAAATCTATTTTGACCATGTCTCTGCAGCAGCAT
Heinz1706      GAGCAAAGCTGCAAGATTGCTCAACTGAAATCTATTTTGACCATGTCTCTGCAGCAGCAT
S.pennellii    GAGCAAAGCTGCAAGATTGCTCAACTGAAATCTATTTTGACCATGTCTCTGCAGCAGCAT
               ************************************************************

M82            CGGACTATGTTTCCATTTAGCTGCTCCCAAGAATATCCTTGTACAATTCCTTGATTTTTG
Heinz1706      CGGACTATGTTTCCATTTAGCTGCTCCCAAGAATATCCTTGTACAATTCCTTGATTTTTG
S.pennellii    CGGACTATGTTTCCATTTAGCTGCTCCCAAGAATATCCTTGTATAATTCCTTGATTTTTG
               ****************************************** ***************

M82            CTATCAAAGCTTCTCTGTTAGCAGCAAAAGGTAATAAGAAGAGAGGTAAGCTAGGATGAA
Heinz1706      CTATCAAAGCTTCTCTGTTAGCAGCAAAAGGTAATAAGAAGAGAGGTAAGCTAGGATGAA
S.pennellii    CTATCAAAGCTTCTCTGTTAGCAGCAAAAGGTAATAAGAAGAGAGGTAAGCTAGGATGAA
               ************************************************************

M82            CACACAAAGGTATGAATAATAAACTTAACTTCCACTAGTTCATATACAAAGGAACGGAAA
Heinz1706      CACACAAAGGTATGAATAATAAACTTAACTTCCACTAGTTCATATACAAAGGAACGGAAA
S.pennellii    CACACAAAGGTATGAATAATAAACTTAACTTCCACTAGTTCATATACAAAGGAACGGAAA
               ************************************************************

M82            TAACCTGTCAGGTTTGAAGATCAAGTTCTTGTAGGCAAACCACTGCAGAGGAAGGGAAGA
Heinz1706      TAACCTGTCAGGTTTGAAGATCAAGTTCTTGTAGGCAAACCACTGCAGAGGAAGGGAAGA
S.pennellii    TAACCTGTCAGGTTTGAAGATCAAGTTCTTGTAGGCAAACCACTGCAGAGGAAGGGAAGA
               ************************************************************
```

```
M82             TCCCCCAAAAACGTGTAAATGAAGTCAAGATAACATGGTAATCGATTATATAGTTCAAAG
Heinz1706       TCCCCCAAAAACGTGTAAATGAAGTCAAGATAACATGGTAATCGATTATATAGTTCAAAG
S.pennellii     TCCCCCAAAAACGTGTAAATGAAGTCAAGATAACATGGTAATCGATTATATAGTTCAAAG
                ************************************************************

M82             TTTAACCAAACAAGCATTGATGAAGCCTGATGCTAATGCCTATGCAATATGGTTCAAAGA
Heinz1706       TTTAACCAAACAAGCATTGATGAAGCCTGATGCTAATGCCTATGCAATATGGTTCAAAGA
S.pennellii     TTTAACCAAACAAGCATTGATGAAGCCTGATGCTAATGCCTATGCAATATGGTTCAAAGA
                ************************************************************

M82             AAGGATTTAACTTAAGTATAACGTTTATTTTTTACCCTATCAGTGTAATTATTGGTTATA
Heinz1706       AAGGATTTAACTTAAGTATAACGTTTATTTTTTACCCTATCAGTGTAATTATTGGTTATA
S.pennellii     AAGGATTTAACTTAAGTATAACGTTTATTTTTTACCCTATCAGTGTAATTATTGGTTATA
                ************************************************************

M82             GCATTCAGGTTACAACATACAGAGTAGTGGCTAAGAGTGAAAATATTTCAACTTACACCG
Heinz1706       GCATTCAGGTTACAACATACAGAGTAGTGGCTAAGAGTGAAAATATTTCAACTTACACCG
S.pennellii     GCATTCAGGTTACAACATACAGAGTAGTGGTTAAGAGTGAAAATATTTCAACTTACACCG
                ***************************** ****************************

M82             GTGTAACCAATTCCTACAAGCTCAAGAACACCAGGAATCAAAGGAAGCCTGTCAATTGCC
Heinz1706       GTGTAACCAATTCCTACAAGCTCAAGAACACCAGGAATCAAAGGAAGCCTGTCAATTGCC
S.pennellii     GTGTAACCAATTCCTACAAGCTCAAGAACACCAGGAATCAAAGGAAGCCTGTCAATTGCC
                ************************************************************

M82             TGCCACAATTAAGCAAGTCCCTTTGTCAAAGTTTGTTATAGTTTGCAACACCTGAACTAA
Heinz1706       TGCCACAATTAAGCAAGTCCCTTTGTCAAAGTTTGTTATAGTTTGCAACACCTGAACTAA
S.pennellii     TGCCACAATTAAGCAAGTCCCTTTGTCAA-GTTTGTTATAGTTTGCAACACCTGAACAAA
                ***************************** ************************** **

M82             AATGGAACTAAAAAA-GTCTTAAGGTAGCAATTAGTAGCAGTA
Heinz1706       AATGGAACTAAAAAA-GTCTTAAGGTAGCAATTAGTAGCAGTA
S.pennellii     AATGGAACTAAAAAAAGTCTTAAGGTAGCAATTAGTAGCAGTA
                *************** ************************
```

LIST OF REFERENCES

Álvarez-Rodríguez, M.L., Belloch, C., Villa, M., Uruburu, F., Larriba, G. and Coque, J. R. (2003) Degradation of vanillic acid and production of guaiacol by microorganisms isolated from cork samples. *FEMS Microbiol. Lett.,* **220**(1), 49-55.

Bai, Y. and Lindhout, P. (2007) Domestication and breeding of tomatoes: What have we gained and what can we gain in the future? *Ann. Bot.,* **100**(5), 1085-1094.

Baldwin, E.A. (2002) Fruit flavor, volatile metabolism, and consumer perceptions. In (M. Knee, ed) Boca Raton, FL: CRC Press.

Baldwin, E.A., Nisperos-Carriedo, M. O. and Moshonas, M. G. (1991) Quantitative analysis of flavor and other volatiles and for certain constituents of two tomato cultivars during ripening. *J. Am. Soc. Hort. Sci.,* **116**(2), 265-269.

Baldwin, E.A., Scott, J.W., Einstein, M.A., Malundo, T.M.M., Carr, B.T., Shewfelt, R.L. and Tandon, K.S. (1998) Relationship between sensory and instrumental analysis for tomato flavor. *J. Am.Soc. Hort. Sci.,* **123**,906–15.

Baldwin, E.A., Scott, J. W., Shewmaker, C. K. and Schuch, W. (2000) Flavor trivia and tomato aroma: Biochemistry and possible mechanisms for control of important aroma components. *HortScience,* **35**(6), 1013-1022.

Baldwin, E.A., Goodner, K. and Plotto, A. (2008) Interaction of volatiles, sugars, and acids on perception of tomato aroma and flavor descriptors. *J. Food Sci.,* **73**(6), S294-S307.

Biggs, M.S., Harriman, R.W., and Handa, A.K. (1986) Changes in gene expression during tomato fruit ripening. *Plant Physiol*., **81**, 395-403.

Birtić, S., Ginies, C., Causse, M., Renard, C. M. G. C. and Page, D. (2009) Changes in volatiles and glycosides during fruit maturation of two contrasted tomato (*Solanum lycopersicum*) lines. *J. Agric. Food Chem.,* **57**(2), 591-598.

Boss, W.F., Sederoff, H. W., Im, Y. J., Moran, N., Grunden, A. M. and Perera, I. Y. (2010) Basal signaling regulates plant growth and development. *Plant Physiol.,* **154**(2), 439-443.

Bowles, D., Lim, E., Poppenberger, B. and Vaistij, F. E. (2006) Glyosyltransferase of lipophilic small molecules. *Annu. Rev. Plant Biol.,* **57**(1), 567-597.

Bowles, D., Isayenkova, J., Lim, E. and Poppenberger, B. (2005) Glycosyltransferases: Managers of small molecules. *Curr. Opin. Plant Biol.,* **8**(3), 254-263.

Bramley, P.M. (2002) Regulation of carotenoid formation during tomato fruit ripening and development. *J. Exp. Bot*., **53**(377), 2107-2113.

Brazier-Hicks, M., Offen, W. A., Gershater, M. C., Revett, T. J., Lim, E., Bowles, D. J., Davies, G. J. and Edwards, R. (2007) Characterization and engineering of the bifunctional *N*- and *O*-glucosyltransferase involved in xenobiotic metabolism in plants. *PNAS,* **104**(51), 20238-20243.

Buttery, R.G., Teranishi, R. and Ling, L. C. (1987) Fresh tomato aroma volatiles: A quantitative study. *J. Agric. Food Chem.,* **35**(4), 540-544.

Buttery, R.G., Takeoka, G., Teranishi, R. and Ling, L. C. (1990) Tomato aroma components: Identification of glycoside hydrolysis volatiles. *J. Agric. Food Chem.,* **38**(11), 2050-2053.

Buttery, R.G., Teranishi, R., Ling, L. C., Flath, R. A. and Stern, D. J. (1988) Quantitative studies on origins of fresh tomato aroma volatiles. *J. Agric. Food Chem.,* **36**(6), 1247-1250.

Buttery, R., Teranishi, R., Flath, R. and Ling, L. (1989) Fresh tomato volatiles - composition and sensory studies. *ACS Symp. Ser.,* **388**, 213-222.

Cabrita, M.J., Freitas, A. M. C., Laureano, O. and Stefano, R. D. (2006) Glycosidic aroma compounds of some Portuguese grape cultivars. *J. Sci. Food Agric.,* **86**(6), 922-931.

Campbell, J., Davies, G., Bulone, V. and Henrissat, B. (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.,* **326**(3), 929-942.

Cartwright, A.M., Lim, E., Kleanthous, C. and Bowles, D. J. (2008) A kinetic analysis of regiospecific glucosylation by two glycosyltransferases of *Arabidopsis thaliana*: Domain swapping to introduce new activities *J. Biol. Chem.,* **283**(23), 15724-15731.

Chen, G., Hackett, R., Walker, D., Taylor, A., Lin, Z. and Grierson, D. (2004) Identification of a specific isoform of tomato lipoxygenase (TomloxC) involved in the generation of fatty acid-derived flavor compounds. *Plant Physiol.,* **136**(1), 2641-2651.

Collendavelloo, J., Legrand, M., Geoffroy, P., Barthelemy, J. and Fritig, B. (1981) Purification and properties of the three *O*-diphenol-*O*-methyltransferases of tobacco leaves. *Phytochemistry,* **20**(4), 611-616.

Dal Cin, V., Tieman, D. M., Tohge, T., McQuinn, R., de Vos, R. C. H., Osorio, S., Schmelz, E. A., Taylor, M. G., Smits-Kroon, M. T., Schuurink, R. C., Haring, M. A., Giovannoni, J., Fernie, A. R. and Klee, H. J. (2011) Identification of genes in the phenylalanine metabolic pathway by ectopic expression of a MYB transcription factor in tomato fruit. *Plant Cell,* **23**(7), 2738-2753.

Dorfner, R., Ferge, T., Kettrup, A., Zimmermann, R. and Yeretzian, C. (2003) Real-time monitoring of 4-vinylguaiacol, guaiacol, and phenol during coffee roasting by resonant laser ionization time-of-flight mass spectrometry. *J. Agric. Food Chem.,* **51**(19), 5768-5773.

Ellis, B.E. and Towers, G. H. N. (1969) The biogenesis of catechol in *Gaultheria*. *Phytochemistry,* **8**(8), 1415-1417.

Eshed, Y. and Zamir, D. (1995) An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics,* **141**(3), 1147-1162.

Eshed, Y. and Zamir, D. (1994) A genomic library of *Lycopersicon pennellii* in *Lycopersicon esculentum* - a tool for fine mapping of genes. *Euphytica,* **79**(3), 175-179.

Eudes, A., Bozzo, G. G., Waller, J. C., Naponelli, V., Lim, E., Bowles, D. J., Gregory, J. F.,III and Hanson, A. D. (2008) Metabolism of the folate precursor p-aminobenzoate in plants: Glucose ester formation and vacuolar storage *J. Biol. Chem.,* **283**(22), 15451-15459.

Evans, W.C., Smith, B. S. W., Linstead, R. P. and Elvidge, J. A. (1951) Chemistry of the oxidative metabolism of certain aromatic compounds by micro-organisms. *Nature,* **168**(4279), 772-775.

Felsenstein J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, **39**, 783-791.

Frary, A., Keles, D., Pinar, H., Gol, D. and Doganlar, S. (2011) NaCl tolerance in *Lycopersicon pennellii* introgression lines: QTL related to physiological responses. *Biol. Plant.,* **55**(3), 461-468.

Gaffney, T., Friedrich, L., Vernooij, B., Negrotto, D., Nye, G., Uknes, S., Ward, E., Kessmann, H. and Ryals, J. (1993) Requirement of salicylic acid for the induction of systemic acquired resistance. *Science,* **261**(5122), 754-756.

Gang, D.R. (2005) Evolution of Flavors and Scents. *Annu. Rev. Plant Biol.,* **56**(1), 301-325.

Gang, D.R., Lavid, N., Zubieta, C., Chen, F., Beuerle, T., Lewinsohn, E., Noel, J. P. and Pichersky, E. (2002) Characterization of phenylpropene *O*-methyltransferases from sweet basil: Facile change of substrate specificity and convergent evolution within a plant *O*-methyltransferase family. *Plant Cell,* **14**(2), 505-519.

Goff, S.A. and Klee, H. J. (2006) Plant volatile compounds: Sensory cues for health and nutritional value? *Science,* **311**(5762), 815-819.

Gonda, I., Bar, E., Portnoy, V., Lev, S., Burger, J., Schaffer, A. A., Tadmor, Y., Gepstein, S., Giovannoni, J. J., Katzir, N. and Lewinsohn, E. (2010) Branched-chain and aromatic amino acid catabolism into aroma volatiles in *Cucumis melo* L. fruit. *J. Exp. Biol.,* **61**(4), 1111-1123.

Groyne, J., Lognay, G. and Marlier, M. (1999) Accumulation of glycosidically bound compounds in *Fragaria x ananassa* cv. elsanta fruits at various developmental stages. *Biotechnol. Agron. Soc.,* **3**, 5-9.

Guillen, M.D., Manzanos, M. J. and Zabala, L. (1995) Study of a commercial liquid smoke flavoring by means of gas Chromatography/Mass spectrometry and fourier transform infrared spectroscopy. *J. Agric. Food Chem.,* **43**(2), 463-468.

Gustafson, F.G. (1934) Production of alcohol and acetaldehyde by tomatoes. *Plant Physiol.,* **9**(2), 359-367.

Hayasaka, Y., Baldock, G. A., Pardon, K. H., Jeffery, D. W. and Herderich, M. J. (2010) Investigation into the formation of guaiacol conjugates in berries and leaves of grapevine *Vitis vinifera* L. cv. cabernet sauvignon using stable isotope tracers combined with HPLC-MS and MS/MS analysis. *J. Agric. Food Chem.,* **58**(4), 2076-2081.

Hayashi, S., Yagi, K., Ishikawa, T., Kawasaki, M., Asai, T., Picone, J., Turnbull, C., Hiratake, J., Sakata, K., Takada, M., Ogawa, K. and Watanabe, N. (2004) Emission of 2-phenylethanol from its β-glucopyranoside and the biogenesis of these compounds from [$_2$H^8] -phenylalanine in rose flowers. *Tetrahedron,* **60**(33), 7005-7013.

Jean-Pierre, M. (2002) Recent breakthroughs in the study of salicylic acid biosynthesis. *Trends Plant Sci*., **7**(8), 332-334.

Johnson, J.H., Gould, W. A., Badenhop, A. and Johnson, R. M. (1968) Quantitative comparison of isoamylol, pentanol, and 3-hexen-1-ol in tomato juice: varietal and harvest differences and processing effects. *J. Agric. Food Chem.,* **16**(2), 255-258.

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

Jones, R.J. and Scott, S. J. (1983) Improvement of tomato flavor by genetically increasing sugar and acid content. *Euphytica,* **32**, 845-855.

Kagey, M.H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., van Berkum, N. L., Ebmeier, C. C., Goossens, J., Rahl, P. B., Levine, S. S., Taatjes, D. J., Dekker, J. and Young, R. A. (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature,* **467**(7314), 430-435.

Kaminaga, Y., Schnepp, J., Peel, G., Kish, C. M., Ben-Nissan, G., Weiss, D., Orlova, I., Lavie, O., Rhodes, D., Wood, K., Porterfield, D. M., Cooper, A. J. L., Schloss, J. V., Pichersky, E., Vainstein, A. and Dudareva, N. (2006) Plant phenylacetaldehyde synthase is a bifunctional homotetrameric enzyme that catalyzes phenylalanine decarboxylation and oxidation. *J. Biol.Chem.,* **281**(33), 23357-23366.

Karimi, M., Inzé, D. and Depicker, A. (2002) GATEWAY™ vectors for agrobacterium-mediated plant transformation. *Trends Plant Sci.,* **7**(5), 193-195.

Katagiri, M., Yamamoto, S. and Hayaishi, O. (1962) Flavin adenine dinucleotide requirement for the enzymic hydroxvlation and decarboxyla-tion of salicylic acid. *J. Biol. Chem.,* **237**(7), 2413-2414.

Klee, H.J. (2010) Improving the flavor of fresh fruits: Genomics, biochemistry, and biotechnology. *New Phytol.,* **187**(1), 44-56.

Kochevenko, A., Araújo, W.L., Maloney, G.S., Tieman, D.M., Do, P.T., Taylor, M.G., Klee, H.J. and Fernie, A.R. (Under review) Catabolism of branched chain amino acids supports respiration but not volatile synthesis in tomato fruits. *Mol. Plant*.

Koeduka, T., Fridman, E., Gang, D. R., Vassão, D. G., Jackson, B. L., Kish, C. M., Orlova, I., Spassova, S. M., Lewis, N. G., Noel, J. P., Baiga, T. J., Dudareva, N. and Pichersky, E. (2006) Eugenol and isoeugenol, characteristic aromatic constituents of spices, are biosynthesized via reduction of a coniferyl alcohol ester. *PNAS,* **103**(26), 10128-10133.

Kohara, A., Nakajima, C., Yoshida, S. and Muranaka, T. (2007) Characterization and engineering of glycosyltransferases responsible for steroid saponin biosynthesis in solanaceous plants. *Phytochemistry,* **68**(4), 478-486.

Kumazawa, K. and Masuda, H. (2002) Identification of potent odorants in different green tea varieties using flavor dilution technique. *J. Agric. Food Chem.,* **50**(20), 5660-5663.

Lam, K., Ibrahim, R., Behdad, B. and Dayanandana, S. (2007) Structure, function, and evolution of plant *O*-methyltransferases. *Genome,* **50**(11), 1001-1013.

Li, Y., Baldauf, S., Lim, E. K. and Bowles, D. J. (2001) Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. *J. Biol. Chem.,* **276**(6), 4338-4343.

Liharska, T., Wordragen, M. v., Kammen, A. v., Zabel, P. and Koornneef, M. (1996) Tomato chromosome 6: Effect of alien chromosomal segments on recombinant frequencies. *Genome,* **39**(3), 485-491.

Lim, E.K. and Bowles, D. J. (2004) A class of plant glycosyltransferases involved in cellular homeostasis. *EMBO J.,* **23**, 2915-2922.

Lim, E., Doucet, C. J., Li, Y., Elias, L., Worrall, D., Spencer, S. P., Ross, J. and Bowles, D. J. (2002) The activity of *Arabidopsis* glycosyltransferases toward salicylic acid, 4-hydroxybenzoic acid, and other benzoates. *J. Biol. Chem.,* **277**(1), 586-592.

Lim, E., Baldauf, S., Li, Y., Elias, L., Worrall, D., Spencer, S. P., Jackson, R. G., Taguchi, G., Ross, J. and Bowles, D. J. (2003) Evolution of substrate recognition across a multigene family of glycosyltransferases in *Arabidopsis*. *Glycobiology,* **13**(3), 139-145.

Marlatt, C., Ho, C. T. and Chien, M. (1992) Studies of aroma constituents bound as glycosides in tomato. *J. Agric. Food Chem.,* **40**(2), 249-252.

Masada, S., Terasaka, K. and Mizukami, H. (2007) A single amino acid in the PSPG-box plays an important role in the catalytic function of CaUGT2 (curcumin glucosyltransferase), a group D family 1 glucosyltransferase from *Catharanthus roseus*. *FEBS Lett.,* **581**(14), 2605-2610.

Matheis, G. (1983) Enzymatic browning of foods - quantitative relationships between browning and food constituents. *Z. Lebensm. Unters. Forsch.,* **176**(6), 454-462.

Mathieu, S., Dal Cin, V., Fei, Z., Li, H., Bliss, P., Taylor, M. G., Klee, H. J. and Tieman, D. M. (2009) Flavour compounds in tomato fruits: Identification of loci and potential pathways affecting volatile composition. *J. Exp. Bot.,* **60**(1), 325-337.

Matsui, K. (2006) Green leaf volatiles: Hydroperoxide lyase pathway of oxylipin metabolism. *Curr. Opin. Plant Biol.,* **9**(3), 274-280.

Maury, S., Geoffroy, P. and Legrand, M. (1999) Tobacco O-methyltransferases involved in phenylpropanoid metabolism. the different caffeoyl-coenzyme A/5-hydroxyferuloyl-coenzyme A 3/5-O-methyltransferase and caffeic Acid/5-hydroxyferulic acid 3/5-O-methyltransferase classes have distinct substrate specificities and expression patterns. *Plant Physiol.,* **121**(1), 215-224.

McCormick, S., Niedermeyer, J., Fry, J., Barnason, A., Horsch, R. and Fraley, R. (1986) Leaf disc transformation of cultivated tomato (*L. esculentum*) using *Agrobacterium tumefaciens*. *Plant Cell Reports,* **5**(2), 81-84.

Miller, J.C. and Tanksley, S. D. (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus *Lycopersicon*. *TAG,* **80**(4), 437-448.

Morse, A.M., Tschaplinski, T. J., Dervinis, C., Pijut, P. M., Schmelz, E. A., Day, W. and Davis, J. M. (2007) Salicylate and catechol levels are maintained in *nahG* transgenic poplar. *Phytochemistry,* **68**(15), 2043-2052.

Noel, J.P., Dixon, R. A., Pichersky, E., Zubieta, C. and Ferrer, J. (2003) Chapter two structural, functional, and evolutionary basis for methylation of plant small molecules. *Recent Adv. Phytochem.* **37**, 37–58.

Oka, N., Ohishi, H., Hatano, T., Hornberger, M., Sakata, K. and Watanabe, N. (1999) Aroma evolution during flower opening in *Rosa damascena* mill. *Z. Naturforsch. C.,* **54**(12), 889-895.

Ortiz-Serrano, P. and Gil, J. V. (2007) Quantitation of free and glycosidically bound volatiles in and effect of glycosidase addition on three tomato varieties (*Solanum lycopersicum* L.). *J. Agric. Food Chem.,* **55**(22), 9170-9176.

Osbourn, A.E. (2003) Saponins in cereals. *Phytochemistry,* **62**(1), 1-4.

Osmani, S.A., Bak, S. and Møller, B. L. (2009) Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. *Phytochemistry,* **70**(3), 325-347.

Paquette, S., Møller, B. L. and Bak, S. (2003) On the origin of family 1 plant glycosyltransferases. *Phytochemistry,* **62**(3), 399-413.

Pellegrini, L., Geoffroy, P., Fritig, B. and Legrand, M. (1993) Molecular cloning and expression of a new class of ortho-diphenol-*O*-methyltransferases induced in tobacco (*Nicotiana tabacum* L.) leaves by infection or elicitor treatment. *Plant Physiology,* **103**(2), 509-517.

Peralta, I.E., Knapp, S. and Spooner, D. M. (2006) Nomenclature for wild and cultivated tomatoes. *Report of the Tomato Genetics Cooperative,* **56**, 6-12.

Petro-Turza, M. (1986) Flavor of tomato and tomato products. *Food Reviews International,* **2**(3), 309-351.

Picone, J.M., Clery, R. A., Watanabe, N., MacTavish, H. S. and Turnbull, C. G. N. (2004) Rhythmic emission of floral volatiles from *Rosa damascena* semperflorens cv.‘Quatre saisons’. *Planta,* **219**(3), 468-478.

Richins, R.D., Scholthof, H. B. and Shepherd, R. J. (1987) Sequence of figwort mosaic virus DNA (caulimovirus group). *Nucleic Acids Res.,* **15**(20), 8451-8466.

Rick, C.M. (1960) Hybridization between *Lycopersicon esculentum* and *Solanum pennellii* - phylogenetic and cytogenetic significance. *PNAS,* **46**(1), 78-82.

Rick, C. and Yoder, J. (1988) Classical and molecular-genetics of tomato - highlights and perspectives. *Annu. Rev. Genet.,* **22**, 281-300.

Riley, J.C.M., Willemot, C. and Thompson, J. E. (1996) Lipoxygenase and hydroperoxide lyase activities in ripening tomato fruit. *Postharvest Biol. Technol.,* **7**(1-2), 97-107.

Ross, J., Li, Y., Lim, E. and Bowles, D. J. (2001) Higher plant glycosyltransferases. *Genome Biol.,* **2**(2), reviews3004.1–3004.6.

Saitou, N. and Nei, M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. and Evol*., **4**, 406-425.

Schauer, N., Semel, Y., Roessner, U., Gur, A., Balbo, I., Carrari, F., Pleban, T., Perez-Melis, A., Bruedigam, C., Kopka, J., Willmitzer, L., Zamir, D. and Fernie, A. R. (2006) Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat. Biotechnol.,* **24**(4), 447-454.

Serra Bonvehí, J. and Ventura Coll, F. (1998) Evaluation of smoky taste in cocoa powder. *J. Agric. Food Chem.,* **46**(2), 620-624.

Shao, H., He, X., Achnine, L., Blount, J. W., Dixon, R. A. and Wang, X. (2005) Crystal structures of a multifunctional Triterpene/Flavonoid glycosyltransferase from *Medicago truncatula*. *Plant Cell,* **17**(11), 3141-3154.

Simkin, A.J., Schwartz, S. H., Auldridge, M., Taylor, M. G. and Klee, H. J. (2004) The tomato *carotenoid cleavage dioxygenase 1* genes contribute to the formation of the flavor volatiles *β*-ionone, pseudoionone, and geranylacetone. *Plant J.,* **40**(6), 882-892.

Speirs, J., Lee, E., Holt, K., Yong-Duk, K., Scott, N. S., Loveys, B. and Schuch, W. (1998) Genetic manipulation of alcohol dehydrogenase levels in ripening tomato fruit affects the balance of some flavor aldehydes and alcohols. *Plant Physiol.,* **117**(3), 1047-1058.

Steinhauser, M., Steinhauser, D., Gibon, Y., Bolger, M., Arrivault, S., Usadel, B., Zamir, D., Fernie, A. R. and Stitt, M. (2011) Identification of enzyme activity quantitative trait loci in a *Solanum lycopersicum* x *Solanum pennellii* introgression line population. *Plant Physiol.*, [Epub ahead of print].

Subba Rao, P.V., Moore, K. and Towers, G. H. N. (1967) The conversion of tryptophan to 2,3-dihydroxybenzoic acid and catechol by *Aspergillus niger*. *Biochem. Biophys. Res. Commun.,* **28**(6), 1008-1012.

Tadmor, Y., Fridman, E., Gur, A., Larkov, O., Lastochkin, E., Ravid, U., Zamir, D. and Lewinsohn, E. (2002) Identification of malodorous, a wild species allele affecting tomato aroma that was selected against during domestication. *J. Agric. Food Chem.,* **50**(7), 2005-2009.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. and Evol.,* **28** (10), 2731-2739.

Tanksley, S.D. (2004) The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *Plant Cell,* **16**, S181-S189.

Tieman, D.M., Loucas, H. M., Kim, J. Y., Clark, D. G. and Klee, H. J. (2007) Tomato phenylacetaldehyde reductases catalyze the last step in the synthesis of the aroma volatile 2-phenylethanol. *Phytochemistry,* **68**(21), 2660-2669.

Tieman, D.M., Taylor, M., Schauer, N., Fernie, A. R., Hanson, A. D. and Klee, H. J. (2006b) Tomato aromatic amino acid decarboxylases participate in synthesis of the flavor volatiles 2-phenylethanol and 2-phenylacetaldehyde. *PNAS,* **103**(21), 8287-8292.

Tieman, D.M., Zeigler, M., Schmelz, E. A., Taylor, M. G., Bliss, P., Kirst, M. and Klee, H. J. (2006a) Identification of loci affecting flavour volatile emissions in tomato fruits. *J. Exp. Bot.,* **57**(4), 887-896.

Tieman, D., Zeigler, M., Schmelz, E., Taylor, M. G., Rushing, S., Jones, J. B. and Klee, H. J. (2010) Functional analysis of a tomato salicylic acid methyl transferase and its role in synthesis of the flavor volatile methyl salicylate. *Plant J.,* **62**(1), 113-123.

Tikunov, Y.M., de Vos, R. C. H., González Paramás, A. M., Hall, R. D. and Bovy, A. G. (2010) A role for differential glycoconjugation in the emission of phenylpropanoid volatiles from tomato fruit discovered using a metabolic data fusion approach. *Plant Physiol.,* **152**(1), 55-70.

Ugliano, M. and Moio, L. (2008) Free and hydrolytically released volatile compounds of vitis vinifera L. cv. fiano grapes as odour-active constituents of fiano wine. *Anal. Chim. Acta,* **621**(1), 79-85.

Van Wees, S.C.M. and Glazebrook, J. (2003) Loss of non-host resistance of *Arabidopsis nahG* to *Pseudomonas syringae* pv. phaseolicola is due to degradation products of salicylic acid. *Plant J.,* **33**(4), 733-742.

Verdonk, J., Haring, M., van Tunen, A. and Schuurink, R. (2005) ODORANT1 regulates fragrance biosynthesis in petunia flowers. *Plant Cell,* **17**(5), 1612-1624.

Verlaan, M.G., Szinay, D., Hutton, S. F., de Jong, H., Kormelink, R., Visser, R. G., Scott, J. W. and Bai, Y. (2011) Chromosomal rearrangements between tomato and *Solanum chilense* hamper mapping and breeding of the TYLCV resistance gene ty-1. *Plant J.,* [Epub ahead of print]

Vogel, J.T., Tan, B. C., McCarty, D. R. and Klee, H. J. (2008) The carotenoid cleavage dioxygenase 1 enzyme has broad substrate specificity, cleaving multiple carotenoids at two different bond positions. *J. Biol. Chem.,* **283**, 11364-11373.

Vogt, T. and Jones, P. (2000) Glycosyltransferases in plant natural product synthesis: Characterization of a supergene family. *Trends Plant Sci.,* **5**(9), 380-386.

Voinnet, O., Rivas, S., Mestre, P. and Baulcombe, D. (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J.,* **33**(5), 949-956.

Wang, J., Dudareva, N., Bhakta, S., Raguso, R. A. and Pichersky, E. (1997) Floral scent production in *Clarkia breweri* (onagraceae) II. localization and developmental modulation of the enzyme S-adenosyl-L-methionine:(iso)eugenol O-methyltransferase and phenylpropanoid emission. *Plant Physiol.,* **114**(1), 213-221.

Watanabe, S., Hashimoto, I., Hayashi, K., Yagi, K., Asai, T., Knapp, H., Straubinger, M., Winterhalter, P. and Watanabe, N. (2001) Isolation and identification of 2-phenylethyl disaccharide glycosides and mono glycosides from rose flowers, and their potential role in scent formation. *Biosci. Biotechnol. Biochem.,* **65**(2), 442-445.

Wein, M., Lavid, N., Lunkenbein, S., Lewinsohn, E., Schwab, W. and Kaldenhoff, R. (2002) Isolation, cloning and expression of a multifunctional *O*-methyltransferase capable of forming 2,5-dimethyl-4-methoxy-3(2H)-furanone, one of the key aroma compounds in strawberry fruits. *Plant J.,* **31**(6), 755-765.

Weis, M., Lim, E., Bruce, N. C. and Bowles, D. J. (2008) Engineering and kinetic characterisation of two glucosyltransferases from *Arabidopsis thaliana*. *Biochimie,* **90**(5), 830-834.

Withopf, B., Richling, E., Roscher, R., Schwab, W. and Schreier, P. (1997) Sensitive and selective screening for 6'-O-malonylated glucoconjugates in plants. *J. Agric. Food Chem.,* **45**(3), 907-911.

Yamamoto, S., Katagiri, M., Maeno, H. and Hayaishi, O. (1965) Salicylate hydroxylase, a monooxygenase requiring flavin adenine dinucleotide. *J.Biol. Chem.,* **240**(8), 3408-3413.

Zanor, M.I., Rambla, J., Chaïb, J., Steppa, A., Medina, A., Granell, A., Fernie, A. R. and Causse, M. (2009) Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents. *J. Exp. Bot.,* **60**(7), 2139-2154.

Zheng, L.T., Ryu, G., Kwon, B., Lee, W. and Suk, K. (2008) Anti-inflammatory effects of catechols in lipopolysaccharide-stimulated microglia cells: Inhibition of microglial neurotoxicity. *Eur. J. Pharmacol.,* **588**(1), 106-113.

Zierler, B., Siegmund, B. and Pfannhauser, W. (2004) Determination of off-flavour compounds in apple juice caused by microorganisms using headspace solid phase microextraction–gas chromatography–mass spectrometry. *Anal. Chim. Acta,* **520**(1-2), 3-11.

Zuckerkandl, E. and Pauling, L. (1965) Evolutionary divergence and convergence in proteins. Edited in *Evolving Genes and Proteins* by V. Bryson and H.J. Vogel, pp. 97-166. Academic Press, New York.

BIOGRAPHICAL SKETCH

Melissa Hamner Mageroy grew up in San Antonio, Texas. She attended Trinity University in San Antonio, where she majored in biology. During her time at Trinity she took part in undergraduate research and was award a Summer Undergraduate Research Fellowship from the American Society of Plant Biology. This experience sparked her interest in plant biology and her desire to pursue a PhD in the field. After graduating from Trinity in 2007, she entered the Plant Molecular and Cellular Biology Graduate program at the University of Florida. Since that time, she has been studying enzymes that are important in the synthesis and regulation of tomato flavor molecules in the lab of Dr. Harry J. Klee. After graduating, she will be moving to Vancouver, British Columbia to begin a post-doctoral position in the lab of Joerg Bohlmann.