THE BIOCHEMISTRY AND GENETICS OF PLANT SURFACE LIPID BIOSYNTHESIS AND ASSEMBLY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Trevor Howard Yeats

January 2012

THE BIOCHEMISTRY AND GENETICS OF PLANT SURFACE LIPID BIOSYNTHESIS AND ASSEMBLY

Trevor Howard Yeats, Ph. D.

Cornell University 2012

The waxy cuticle that covers the aerial epidermis of plants provides protection from desiccation and environmental stresses and was likely a critical adaptation that allowed plants to colonize land. The cuticle consists of various organic solvent soluble waxes that are embedded within, and layered upon, a non-soluble polyester matrix of cutin. Despite the fundamental importance of the cuticle, relatively little is known about its synthesis and metabolism. Tomato (*Solanum lycopersicum*) is an emerging model species for studying cuticle biology is, as it has extensive genetic resources and a fruit cuticle that is substantial and easily isolated. This work describes the results of three studies that leverage the attributes of tomato as a model for cuticle biology. First, a survey of the morphology and chemistry of the fruit cuticles in 6 wild relatives of tomato is described. While these species are separated by less than 7 million years of evolution, they are endemic to diverse environments, and substantial differences in cuticle morphology and composition were observed. The genetic basis of one of these traits was further dissected by chemical analysis of a *S. habrochaites* introgression line population. The second study used a proteomic approach to identify proteins associated with the cuticle of expanding tomato fruits. This study identified several candidates for the extracellular transport and assembly of cuticle precursors. One of

these was independently identified by forward genetics of the *cutin deficient 1* (*cd1*) mutant and shown to be required for accumulation of the cutin polymer. The *CD1* gene encodes an extracellular GDSL-motif lipase/hydrolase family protein, and the final study describes the biochemical characterization of the CD1 protein and the *cd1* mutant. In extracts from the *cd1* mutant, but not the wild type, the 2-monoacylglycerol (2-MAG) derivative of the major cutin monomer of tomato was identified by GC-MS. *In vitro* biochemical characterization of CD1 demonstrated acyltransferase activity with a preference for a 2-MAG acyl donor. Collectively, these results suggest that CD1 catalyzes extracellular cutin polymerization via successive transesterification of 2-MAG derivatives of cutin monomers, thus indicating for the first time the monomeric precursor and enzymatic mechanism of cutin polymerization.

BIOGRAPHICAL SKETCH

Trevor Yeats was born in 1982 in Santa Clara, California and grew up in Santa Rosa, California. He completed a bachelor of science degree in Biochemistry at the University of British Columbia that included honors thesis work on cuticular triterpenoid biosynthesis in the laboratory of Reinhard Jetter. He entered into a doctoral program in Plant Biology at Cornell University in 2005, where he inadvertently continued studying the plant cuticle in the Rose lab.

"I'm not bullshitting you… back me up on this bullshit!"

- Dr. Shamoni Maheshwari

Dedicated to my family, friends and Marie, for love and support in the past, present and future.

ACKNOWLEDGMENTS

In addition to those that helped me with each project and are thus acknowledged in the appropriate chapters, I would like to specifically thank a number of people that have had an ongoing role in helping me be a happy and productive graduate student.

First I would like to thank my advisor, Joss Rose, who has provided me with an incredible range of opportunities and flexibility as a graduate student. As well, I would like to thank the rest of my committee: Jim Giovannoni and Karl Niklas for feedback, discussion and help with all of these projects. I would also like to express my gratitude to Tadhg Begley, who is no longer at Cornell, but had many helpful ideas while he was on my committee. I also learned a lot from him through my participation in the NIH Chemistry Biology Interface training grant program, which he directed at the time.

I have been lucky enough to work with some of the kindest and most helpful people over the last six years, and it has been a joy to share a lab and office with everybody. Specific thanks are due to those in the group with an interest in cuticle biology, including Greg Buda, Yonghua He, Tal Isaacson, Gloria Lopez-Casado, Laetitia Martin and Antonio Matas. Many helpful discussions and occasional journal clubs helped form and refine a lot of the ideas, experiments and discussion in this dissertation. One former labmate that I'd like to specifically acknowledge is Bree Urbanowicz, who taught me most of what I know about protein biochemistry.

Outside of the lab, most of my time at Cornell has been deeply enriched by the great friends I made while living at Stewart Little Coop. A large number of my

housemates were scientists, and even biologists, but I don't really remember talking all that much about science. On the other hand, I do remember talking a lot about science with my friends in the Barley Legal String Band, to the great frustration of our one member that was not a plant scientist. All together, I have had the best friends and the happiest times in grad school, and I really miss everybody that has graduated and moved on before me.

Finally I want to thank Marie, who has been very supportive when experiments have failed, papers have been rejected and science hasn't worked the way that it's supposed to. But most of all, I am lucky to have a partner that understands why the Champagne tastes so sweet when the science does work.

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH	v
DEDICATION	vi
ACKNOWLEDGMENTS	vii
TABLE OF CONTENTS	ix
CHAPTER 1. Introduction	1
CHAPTER 2. The fruit cuticles of wild tomato species exhibit architectural and chemical diversity, providing a new model for studying the evolution of cuticle function	33
CHAPTER 3. Mining the surface proteome of tomato (*Solanum lycopersicum*) fruit for proteins associated with cuticle biogenesis	75
CHAPTER 4. Enzymatic synthesis of the plant biopolyester cutin from hydroxyacylglycerol precursors	119
CHAPTER 5. Conclusions and future directions	153
APPENDIX 1. Supplementary material for Chapter 3	160
APPENDIX 2. Related publications	182

CHAPTER 1

Introduction

The plant cuticle is an extracellular hydrophobic layer that covers the aerial epidermis of all land plants, providing protection against desiccation and external environmental stresses and helping to determine organ boundaries (Nawrath, 2006). The cuticle is composed of two main components, a lipid polyester, known as cutin, and an assortment of organic solvent soluble lipids that are collectively referred to as waxes. An additional, non-esterified polymeric lipid, cutan, is also observed in some species. Waxes can occur embedded with the cutin matrix as intracuticular wax, or they can be layered on the cutin surface as epicuticular wax (ECW) films or crystals. The entire cuticular complex is contiguous with the polysaccharide cell wall, and all three components, cutin, wax, and cell wall polysaccharides occur in overlapping gradients to form a complex composite structure (Figure 1.1).

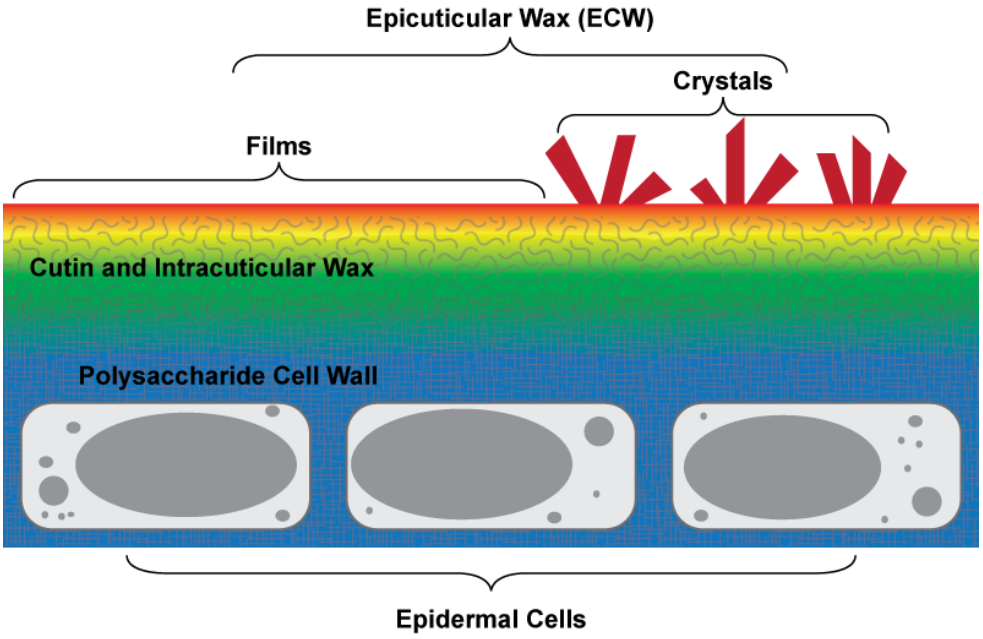


Figure 1.1. The architecture of the epidermis of aerial plant organs.

Functions of the cuticle

As the primary interface between the plant and its environment, the cuticle has many roles in protecting the plant, the most notable of which is acting as a barrier to the loss of water and apoplastic polar metabolites (Riederer and Schreiber, 2001). The hydrophobicity of the cuticle also limits its wettability, which helps to limit the growth of pathogens on the plant surface (Reina-Pinto and Yephremov, 2009) and the cuticle serves as a physical barrier to penetration by pathogens. Accordingly, many bacteria and fungi secrete cutinase enzymes in order to digest the cutin matrix (Kolattukudy, 2001). Such damage and disassembly of the cuticle does not go unnoticed by the plant, and innate immune responses are activated by the disassembly and permeabilization of the cuticle (L'haridon *et al.*, 2011).

Several functions have been specifically proposed for the ECW crystals that are found on many cuticle surfaces. When present, ECW crystals are seen by the naked eye as a dull waxy bloom on leaves, stems or fruits (e.g. the surface of blueberries and grapes) and can have elaborate microscopic structures that can impart unique hydrophobic surface properties to the cuticular surface. In many species, a highly textured hydrophobic surface can cause water to bead and wash away particulate contamination on the surface of the plant (Barthlott and Neinhuis, 1997). ECW crystals can disperse excessive UV radiation that can be damaging to the plant (Pfündel *et al.*, 2006) and can also contribute to plant-insect interactions. For example, they are the basis of the slippery surface of the pitcher plant (*Nepenthes spp.*) that prevents the escape of trapped insects (Gorb *et al.*, 2005).

Many mutants that are defective in cuticle biosynthesis also exhibit aberrant organ fusions, suggesting that the cuticle plays an important role in determining organ boundaries during development. Two mechanisms for this have been discussed: either the cuticle directly establishes a physical barrier between adjacent polysaccharide cell walls to prevent their merging, or the diffusion of developmental signals (e.g. hormones) is limited by the nascent cuticle (Nawrath, 2006). During anisotropic growth of plant organs, the cuticle may also play an important mechanical role in regulating growth as, together with the rest of the outer periclinal cell wall, it is under more tensile strength than the rest of the organ (Kutschera and Niklas, 2007).

Structure and composition of the cuticle

The cutin matrix

The cutin polymer is a polyester composed of hydroxy fatty acids, dicarboxylic acids and glycerol. The relative proportions of these components vary substantially with species and ontogeny, and despite the well characterized composition of monomers, the macromolecular structure of the polymer is less well known (Pollard *et al.*, 2008).

Cutin monomers

The major cutin monomers can broadly be divided among four groups based on their structure and biosynthesis, as depicted in Figure 1.2. In many species, a single group or compound predominates, while other species show a mixture of different monomers. The C16 group of monomers are derived from ω- and midchain

hydroxylation of hexadecanoic acid (**1**), yielding 16-hydroxhexadecanoic acid (**2**) or 10,16-dihydroxyhexadecanoic acid (**3**) and its positional isomers. The C18 group of monomers are derived from a combination of ω-hydroxylation, epoxidation and epoxide hydrolysis of *cis*-9-octadecenoic acid (**4**) to yield 18-hydroxy-*cis*-9-octadecenoic acid (**5**), 18-hydroxy-9-epoxyoctadecanoic acid (**6**) and 9,10,18-trihydroxyoctadecanoic acid (**7**) (Kolattukudy, 2001). The analogous compounds with an additional *cis*-12 unsaturation are also common. More recently, a third type of cutin rich in dicarboxylic acids was identified from the leaves and stems of arabidopsis (*Arabidopsis thaliana*; Bonaventure *et al.*, 2004; Franke *et al.*, 2005). The major monomer of this group is derived from oxidation of octadeca-*cis*-9, *cis*-12-dieneoate (**8**) to give octadeca-*cis*-6, *cis*-9-diene-1,18-dioate[1] (**9**). Lesser amounts of the dicarboxylic acid homolog of (**1**) and (**4**) are also found in this cutin. Additionally, although it is not often quantified for technical reasons due to its high polarity, glycerol (**10**) is a substantial component of many cutins (Graca *et al.*, 2002). It has been detected in arabidopsis leaf cutin (Franke *et al.*, 2005), and it is assumed to be an essential component of such cutin polymers (discussed below).

[1] This compound is named according to IUPAC conventions: although it is directly derived from ω-oxidation of octadeca-*cis*-9, *cis*-12-dieneoate, the numbering of the carbon chain is reversed so as to have the lowest possible numbering for the double bond position.

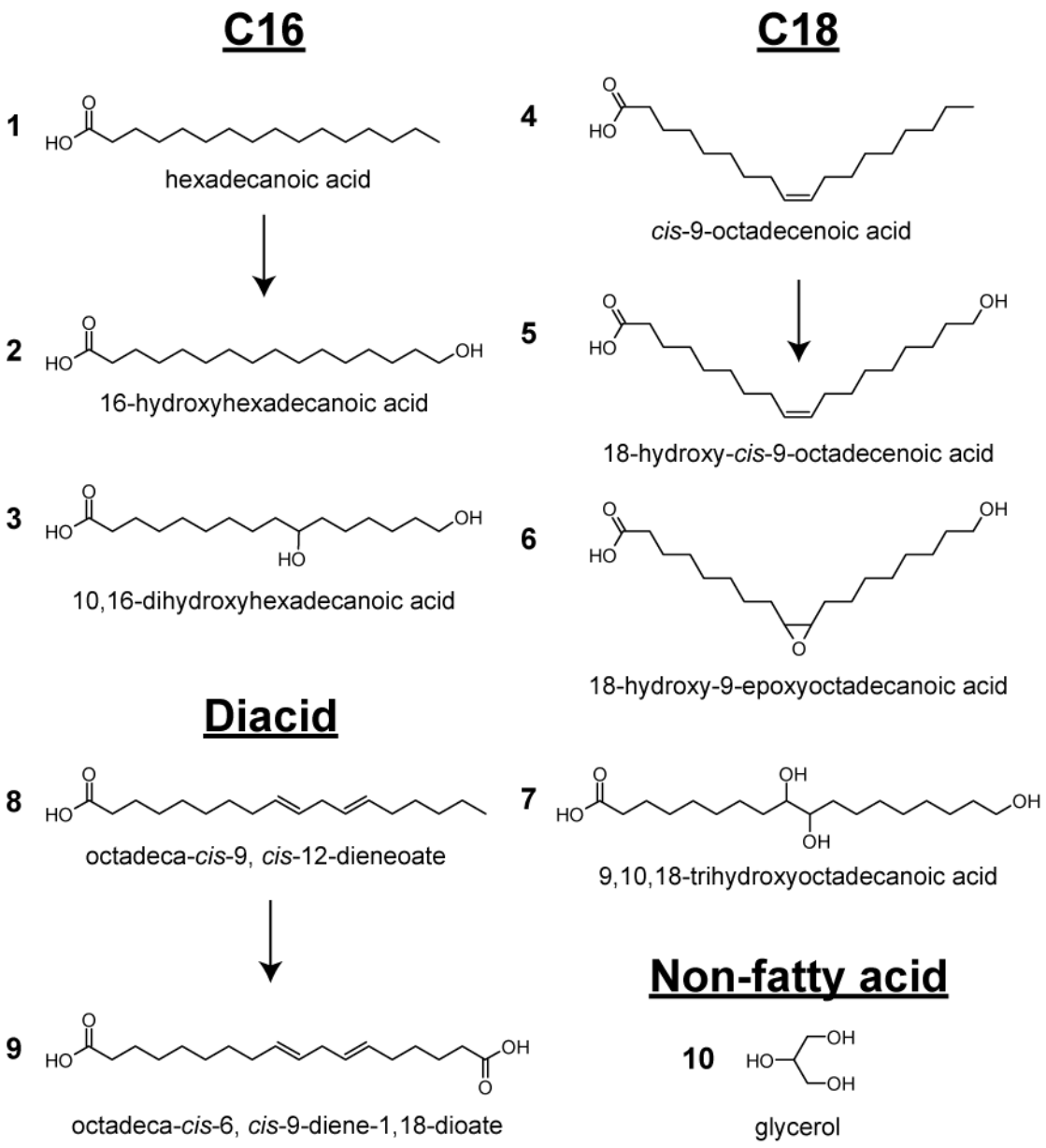


Figure 1.2. The four groups of cutin monomers and the fatty acids from which they are derived.

Structure of cutin polymers

Based on known cutin monomers, three polyester structural motifs can be envisioned (Pollard *et al.,* 2008). The simplest is a linear polymer consisting of either a homopolymer of ω-hydroxy acid (Figure 1.3a) or a copolymer of dicarboxylic acids and glycerol (Figure 1.3b). A highly branched, dendritic structure occurs when both midchain- and ω-hydroxy groups of a dihydroxy acid are highly esterified (Figure 1.3c), or a similar structure can occur with the dicarboxylic acid-glycerol copolymer (Figure 1.3d). Finally, a truly crosslinked polyester is only possible if the monomers include dicarboxylic acids and polyhydroxy compounds. For the dicarboxylic acid-glycerol copolymer, such a crosslinked structure is depicted (Figure 1.3e). Note that for the dicarboxylic acid structures, a polyester homopolymer cannot occur and so in such a case glycerol, or another polyhydroxy compound, is required regardless of the degree of branching or crosslinking.

For the common case of a cutin that is rich in dihydroxy fatty acids, assuming random linkage of the monomers, the structure is likely a mixture of a linear and branched polymer (Figure 1.3a, c). True crosslinking between polymeric chains would require dicarboxylic acids, which are only minor constituents in most cutins. Despite extensive surveys of the chemical composition of cutin from many species and organs (Kolattukudy, 2001), arabidopsis leaves and stems are the only known cutins composed primarily of dicarboxylic acids. Thus, these cutins may be highly atypical and exceptionally crosslinked, or they could adopt a structure that is closely analogous to typical cutins with a dimer of glycerol and dicarboxylic acid substituting for a monomer of dihydroxy fatty acid (Figure 1.3b, d).

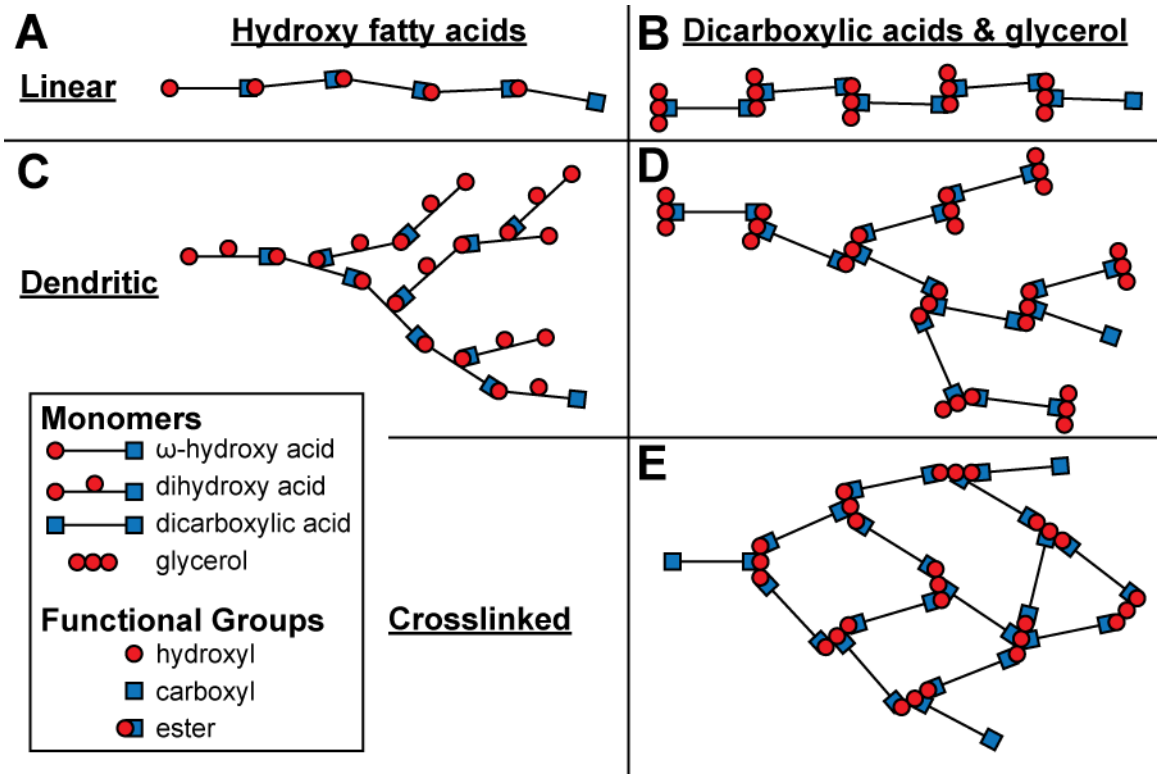


Figure 1.3. Theoretical structural motifs of cutin polymers. This figure is adapted from Pollard *et al.,* 2008. A. Linear polymer of ω-hydroxy acids. B. Linear copolymer of dicarboxylic acid and glycerol. C. Dendritic polymer of dihydroxy acids. D. Dendritic copolymer of dicarboxylic acid and glycerol. E. Crosslinked copolymer of glycerol and dicarboxylic acid. Inset is a key to the symbols used. For representative structures, see Figure 1.2.

There have been several attempts to experimentally determine the polymeric structure of cutin. One approach is to label the free hydroxyls of intact cutin followed by conventional depolymerization and characterization of the monomers (Kolattukudy, 2001). This approach applied to the cutins of *Solanum lycopersicum, Rose canina* and *Ribes nigrum*, showed that nearly all of the primary (ω) hydroxyl

groups and approximately half of the secondary (midchain) hydroxyl groups are esterified (Deas and Holloway, 1976; Kolattukudy, 1976). Another approach is the characterization of oligomers produced by partial depolymerization of the polymer by enzymatic or chemical methods; however, the products of these treatments are difficult to characterize because they are often heterogenous and of low abundance. Additionally, primary esters are less sterically hindered and are thus more readily cleaved by most treatments and so the identification of a given bond in an oligomer structure confirms its presence in the original polymer, but not its quantitative abundance. This further limits the value of this approach.

Partial depolymerization of tomato cutin by methanolysis resulted in the production of oligomers that were predominantly composed of secondary esters (Graca and Lamosa, 2010). In two experiments where fruit cutin from lime (*Citrus aurantifolia*) was partially depolymerized using two treatments with opposite regiospecificity, oligomers with the complementary linkages were identified. Specifically, enzymatic hydrolysis with pancreatic lipase, an enzyme that specifically cleaves esters of primary alcohols, yielded oligomers with mostly secondary ester linkages (Ray and Stark, 1998). Meanwhile, treatment of lime fruit cutin with iodotrimethylsilane, an ester deprotecting reagent with specificity for secondary esters, yielded a suite of oligomers with mostly primary ester linkages (Ray *et al.*, 1998). More recent studies using hydrofluoric acid to produce oligomers of lime cutin have confirmed the identity of these linkages and have led to the putative identification of a cutin oligomer linked to a disaccharide, suggesting covalent linkage of the cutin polymer to cell wall polysaccharides (Tian *et al.*, 2008). Taken together, these results

support a model for cutin structure that is a mixture of linear and dendritic domains, with the function and average size of each domain remaining to be defined. Whether these domains are defined entirely by the macroscopic composition of monomers, or are more tightly controlled at the microscopic level during ontogeny, remains a matter of speculation.

A final enigmatic aspect of cutin structure is the presence of a second lipidic polymer, cutan, that is recalcitrant to treatments that break the ester bonds of cutin. Cutan is rich in ether and C-C bonds, but its structure is otherwise unknown (Pollard *et al.*, 2008). The occurrence of cutan varies greatly between species, although a broad taxonomic survey suggest that its presence is restricted to relatively few extant species (Gupta *et al.*, 2006).

Waxes

While the cuticle is built on the structural skeleton defined by the cutin polymer, many of the cuticular functions are associated with the waxes that fill and coat this polymer. The cuticular waxes consist of a collection of very long chain (C20-C34) acyl lipids that commonly occur in homologous series of varying chain lengths, with either even or odd chain lengths predominating, depending on the compound class. These include acids, aldehydes, alcohols, alkanes, ketones and esters (Figure 1.4). Esters typically have longer chain lengths, on account of their dimeric nature (C30-C60). In some species, unsaturated and bifunctional acyl compounds are also observed, as are a variety of non-aliphatic lipids, including triterpenoids and flavonoids (Jetter *et al.*, 2006).

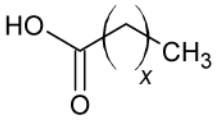
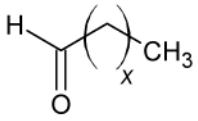
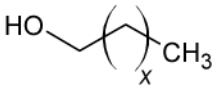
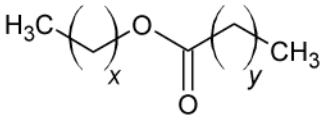
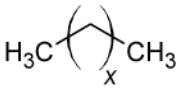
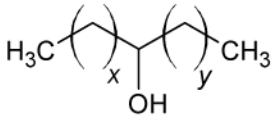
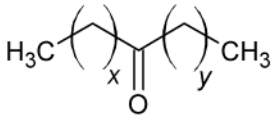
Compound class	Structure	Common chain lengths
Acids		Even
Aldehydes		Even
Primary alcohols		Even
Esters		Even
Alkanes		Odd
Secondary alcohols		Odd
Ketones		Odd

Figure 1.4. Common classes of acyl lipids in cuticular waxes.

The spatial distribution of different wax compounds in the cuticle is not homogenous. In general, alcohols and cyclic compounds, such as triterpenoids, are more abundant in the intracuticular fraction while alkanes and fatty acids accumulate at higher levels in the epicuticular layer, although when present, elaborate epicuticular

wax crystal morphologies are usually associated with polar waxes (Buschhaus and Jetter, 2011). Within the cuticle, waxes are believed to associate spontaneously into both crystalline and amorphous zones that can be observed by spectroscopic and X-ray diffraction techniques (Jetter *et al.*, 2006).

Molecular biology and biochemistry of cuticle biosynthesis

Cutin biosynthesis

Upon the discovery of the atypical cutin composition of arabidopsis stems and leaves, it was thought that this species might be of limited value as a model for studying cutin biosynthesis (Pollard *et al.*, 2008). However, the floral organs of arabidopsis are rich in the typical C16 dihydroxyhexadecanoic acid cutin monomer, thus allowing genetic dissection of a more common cutin (Li-Beisson *et al.*, 2009). The intracellular steps of monomer biosynthesis are now partially known at the molecular level as a result of studies of arabidopsis mutants, although the relative order of some specific steps in the pathway and details of monomer trafficking and polymerization are currently unknown. The pathway described below is depicted in Figure 1.5.

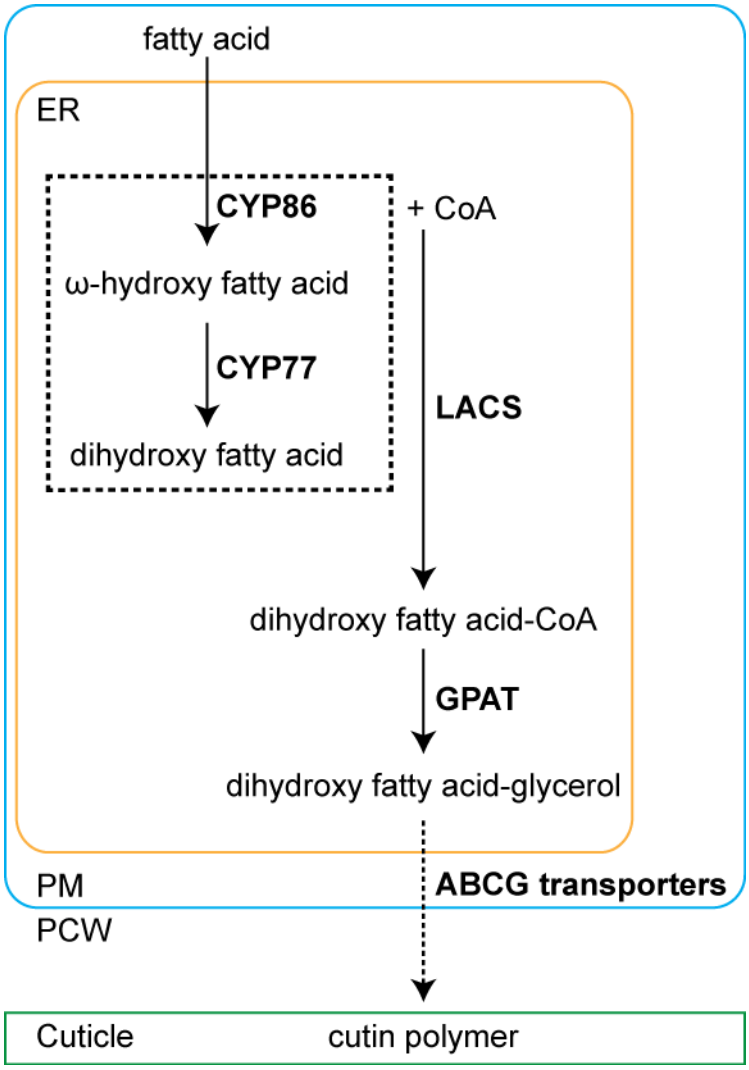


Figure 1.5. The cutin biosynthetic pathway of C16 cutin monomers in arabidopsis. The names of the proteins or protein classes known to mediate some of the steps are shown. Although the order of the two CYP steps in the dashed box is known, it is not known whether their *in vivo* subtrates are free acids or CoA derivatives. ER, Endoplasmic Reticulum; PM, Plasma Membrane; PCW, Polysaccharide Cell Wall; CoA , Coenzyme A. Enzyme abbreviations are given in the text.

The current molecular model of C16 cutin biosynthesis begins with fatty acid synthesis in the plastid. The next three steps occur in the endoplasmic reticulum and consist of ω- and midchain hydroxylation and synthesis of an acyl-CoA intermediate. The relative order of these steps is not known, although it has been shown that the ω-hydroxylation precedes the midchain hydroxylation and that the final product of these steps is most likely a dihydroxyhexadecanoic acid-CoA ester (Li-Beisson *et al.*, 2009). The enzymes responsible for these steps are encoded by genes with the common phenotype of a reduction in the amount of cutin polymer. The ω-hydroxylase is encoded by members of the CYP86 subfamily of cytochrome P450s (*CYP86A4* in arabidopsis flowers; Li-Beisson *et al.*, 2009), while the midchain hydroxylase is encoded by the CYP77 subfamily (*CYP77A2* in arabidopsis flowers; Li-Beisson *et al.*, 2009). The acyltransferases that synthesize acyl-CoA are encoded by the long chain acyl-CoA synthase (LACS) family, which consists of nine members in arabidopsis, and both *LACS1* and *LACS2* appear to be responsible for C16 cutin monomer biosynthesis (Lu *et al.*, 2009).

The hydroxyacyl-CoA product of these initial steps is a likely substrate for glycerol-3-phosphate acyltransferase (GPAT). In arabidopsis, C16 cutin biosynthesis is dependent on *GPAT6* (Li-Beisson *et al.*, 2009). *In vitro* characterization of this enzyme identified two unique features of this type of plant GPAT enzyme: these enzymes specifically acylate the *sn-2* position of glycerol-3-phosphate and they contain an additional phosphatase activity (Yang *et al.*, 2010). Thus, the current model

of C16 biosynthesis has 2-mono(10,16-dihydroxyhexadecanoyl)glycerol (2-MHG) as the final known precursor of the cutin polymer (Figure 1.6).

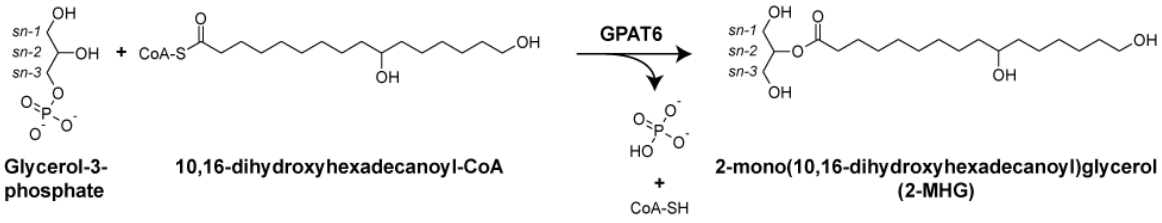


Figure 1.6. Proposed activity of GPAT6-type enzymes. Although only dicarboxylic acid-CoA substrates were tested *in vitro* (Yang *et al.*, 2010), the mutant phenotype of *gpat6* suggests that 10,16-dihydroxyhexadecanoic acid would be the primary GPAT6 substrate (Li-Beisson *et al.*, 2009), forming 2-MHG as shown.

Any additional intermediates before the cutin polymer is formed are unknown, as is the mechanism of cutin polymerization. The molecular basis of trafficking of cutin precursors from the ER to the cutin polymer has been partially resolved since several plasma membrane localized ATP binding cassette G transporters (ABCGs) have been identified that are required for cutin biosynthesis and presumably export of cutin precursors. In arabidopsis, these include *ABGC11* (Bird *et al.*, 2007), *ABCG13* (Panikashvili *et al.*, 2011) and *ABCG32* (Bessire *et al.*, 2011). The mechanism of apoplastic transport of cutin precursor lipids across the polar polysaccharide cell wall is unknown. Apoplastic lipid-transfer proteins (LTPs) have been proposed to facilitate their solubility, although any evidence for their role in cutin precursor transport is indirect (Yeats and Rose, 2008). Depending on the nature of the precursor of polymer biosynthesis, it may be soluble in its own right. For example, the four hydroxyl groups of 2-MHG could impart sufficient polarity to make this compound water soluble

(Pollard *et al.*, 2008). Finally, the mechanism of polymerization is unknown. The extracellular hydrolase-related enzyme BDG (Kurdyukov *et al.*, 2006) and proteins of the GDSL-motif lipase/hydrolase family (Park *et al.*, 2010; Reina *et al.*, 2007) have been proposed as cutin synthases, but evidence for these roles is currently indirect and incomplete. Conversely, it has also been proposed that cutin polymerization may be a spontaneous non-enzymatic process (Dominguez *et al.*, 2010).

Wax biosynthesis

Arabidopsis has been an excellent model for determining the molecular biology of wax biosynthesis. Unlike the cutin, the wax of arabidopsis stems and leaves is fairly typical and contains all of the major categories of acyl lipids. Moreover, the presence of ECW crystals on stems, imparting a glaucous appearance in the wild type, has provided an easy screen for wax deficient mutants. Such mutants exhibit a glossy stem phenotype and are termed *eceriferum* (*cer*; Koornneef *et al.*, 1989). It is through molecular analyses of these and other wax mutants that an increasingly complete pathway for acyl wax biosynthesis has been established (Figure 1.6).

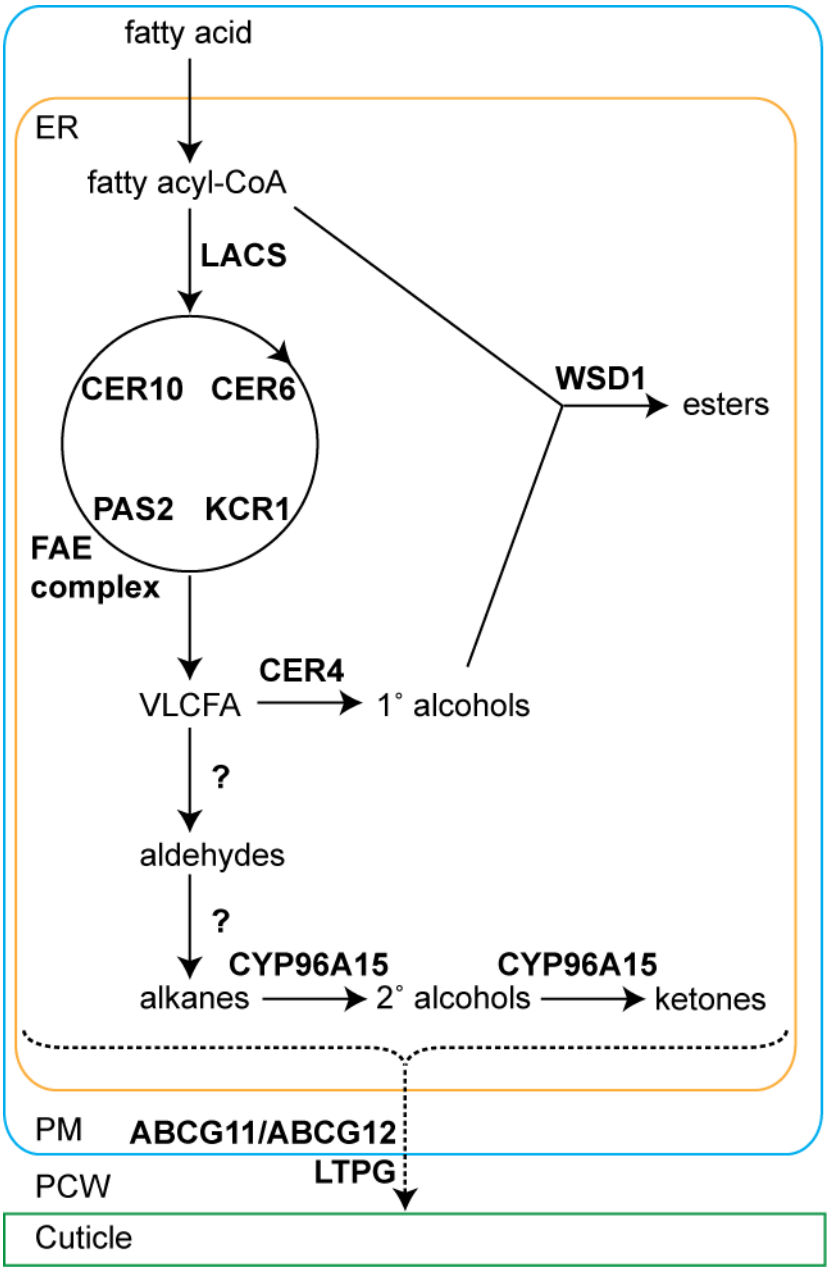


Figure 1.6. The biosynthesis of acyl wax compounds in arabidopsis. The names of the proteins or protein classes known to mediate some of the steps are shown. Enzyme abbreviations are given in the text, other abbreviations are as in Figure 1.5.

Like cutin, wax biosynthesis begins with C16 or C18 fatty acid biosynthesis in the plastid. Following transfer of these free fatty acids to the ER, their corresponding CoA thioesters are synthesized by a long chain acyl-CoA synthase (LACS). In arabidopsis, the specific LACS isozymes responsible for this step is not known. The C16 acyl-CoA is then a substrate for the fatty acid elongase (FAE) complex. Through successive addition of two carbons per cycle derived from malonyl-CoA, the ultimate products of this complex are C20-C34 very long chain fatty acids (VLCFAs). The complex consists of four subunits, in the wax-synthesis pathway of arabidopsis these are CER6 (β-ketoacyl-CoA synthase), KCR1 (β-ketoacyl-CoA reductase), PAS2 (β-hydroxyacyl-CoA dehydratase) and CER10 (enoyl-CoA reductase; Kunst and Samuels, 2009). The elongation cycles can be terminated by a thioesterase to form free VLCFAs, or the VLCFA-CoA esters can undergo further modifications.

Primary alcohols can be produced from VLCFA-CoA by fatty acyl-coenzyme A reductase, an enzyme encoded by *CER4* in arabidopsis (Rowland *et al.*, 2006). The alcohol, along with fatty acyl-CoA, can serve as substrates for wax ester synthesis. The arabidopsis enzyme responsible for this is WSD1, an enzyme of the wax synthase/diacyl glycerol acyltransferase family (Li *et al.*, 2008).

A second branch of acyl wax biosynthesis leads to formation of aldehydes and ultimately alkanes. Curiously, in arabidopsis, LACS1, which is also required for C16 cutin monomer biosynthesis, appears to have an additional specificity for C30 VLCFA and is required for normal accumulation of downstream wax compounds (Lu *et al.*, 2009). Otherwise, the enzymes in this pathway are less clearly defined. First VLCFAs (possibly as CoA esters) are reduced to their corresponding aldehydes and are then

decarbonylated, producing an alkane with the loss of one carbon as carbon monoxide. The enzymes responsible for these steps have not been identified, although the aldehyde decarbonylase activity was demonstrated with crude extracts from *Pisum sativum* leaf epidermis (Cheesbrough and Kolattukudy, 1984; Schneider-Belhaddad and Kolattukudy, 2000). More recently, a cyanobacterial enzyme capable of decarbonylation of aldehydes and production of alkanes was identified and characterized (Schirmer *et al.*, 2010). The strongest candidate for a decarbonylase in higher plants from genetic studies is arabidopsis CER1. Although the sequence of this gene provides little insight into its enzymatic function, the *cer1* mutant accumulates aldehydes and is deficient in alkanes (Aarts *et al.*, 1995; Hannoufa *et al.*, 1993). Moreover, the protein is localized to the ER (Kamigaki *et al.*, 2009) and its overexpression leads to enhanced accumulation of alkanes in the cuticle (Bourdenx *et al.*, 2011). However, the biochemical activity of CER1 remains to be demonstrated. Alkanes can undergo further modification to form secondary alcohols and ketones. In arabidopsis, both of these oxidations are performed by the cytochrome P450 enzyme CYP96A15/MAH1 (Greer *et al.*, 2007).

After synthesis of the suite of wax compounds, they are exported from the ER, across the plasma membrane, through the polysaccharide cell wall and to the cuticular membrane. Most of these transport processes are poorly understood, although trafficking across the membrane has been shown to depend on ABCG transporters. In arabidopsis, two such genes have been identified, *CER5/ABCG12* (Pighin *et al.*, 2004) and *ABCG11* (Bird *et al.*, 2007). Both of these encode half transporters and, based on double mutant analysis and bimolecular fluorescent complementation (BiFC) analyses,

it has been suggested that a ABCG11/ABCG12 heterodimer is required for wax secretion (McFarlane *et al.*, 2010). Export of some wax compounds also appears to be facilitated by a GPI-anchored lipid-transfer protein, LTPG, which is bound to the extracellular side of the plasma membrane (Debono *et al.*, 2009). As with the extracellular transport of cutin monomers, nothing is known about how the hydrophobic waxes are transported across the hydrophilic environment of the polysaccharide cell wall to the cuticle: again, apoplastic LTPs have been proposed to play a role, but genetic or biochemical evidence is lacking (Yeats and Rose, 2008).

Tomato fruit as a model system

While current understanding of the molecular framework of cuticle biosynthesis has come mostly from studying arabidopsis, there are several shortcomings to its use as a model system for cuticle biology. As previously mentioned, with the exception of its floral organs, arabidopsis cutin is structurally unusual. Secondly, its cuticle is extremely thin and impossible to isolate intact (Franke *et al.*, 2005), and so water permeability measurements cannot be made with isolated cuticular membranes. When whole plant water loss is measured instead, the transpirational water loss is determined by substantial contributions from both the stomatal and the cuticular path (Kerstiens *et al.*, 2006).

Conversely, tomato fruit have a number of practical advantages with respect to studying the cuticle and present an attractive model system. First, the cuticle of tomato fruits is far more substantial than that of many plants and is easily isolated. Since the fruit is astomatous, cuticular water permeance can also be estimated with intact fruit

(Leide *et al.*, 2007) or isolated intact cuticular membranes (Schreiber *et al.*, 2006). Its chemical composition is relatively simple and well defined (Baker *et al.*, 1982) and a large germplasm collection including mutants, wild relatives and hundreds of cultivars is available. Importantly, the tomato genome has been sequenced and a range of associated genomics tools has been developed (www.solgenomics.net; Mueller *et al.*, 2005).

The tomato fruit is classified as a berry, reflecting its thick, fleshy pericarp and many seeds (Gillaspy *et al.*, 1993) and the major anatomical features of the mature tomato fruit are summarized in Figure 1.7a. The pericarp is defined by five primary tissue types (Figure 1.7b). The outer epidermis, which synthesizes the cuticle, is a single layer of non-photosynthetic cells. Upon maturity, the fruit cuticle is thick and can extend into the anticlinal and inner periclinal cell walls of the epidermis (Figure 1.7c). Beneath the epidermis are a few layers of collenchyma cells and under these are several layers of larger parenchyma cells. Defining the inner surface of the pericarp is another single-cell layer epidermis. Finally, the vascular bundles, containing the phloem and xylem, are found in the middle of the pericarp.

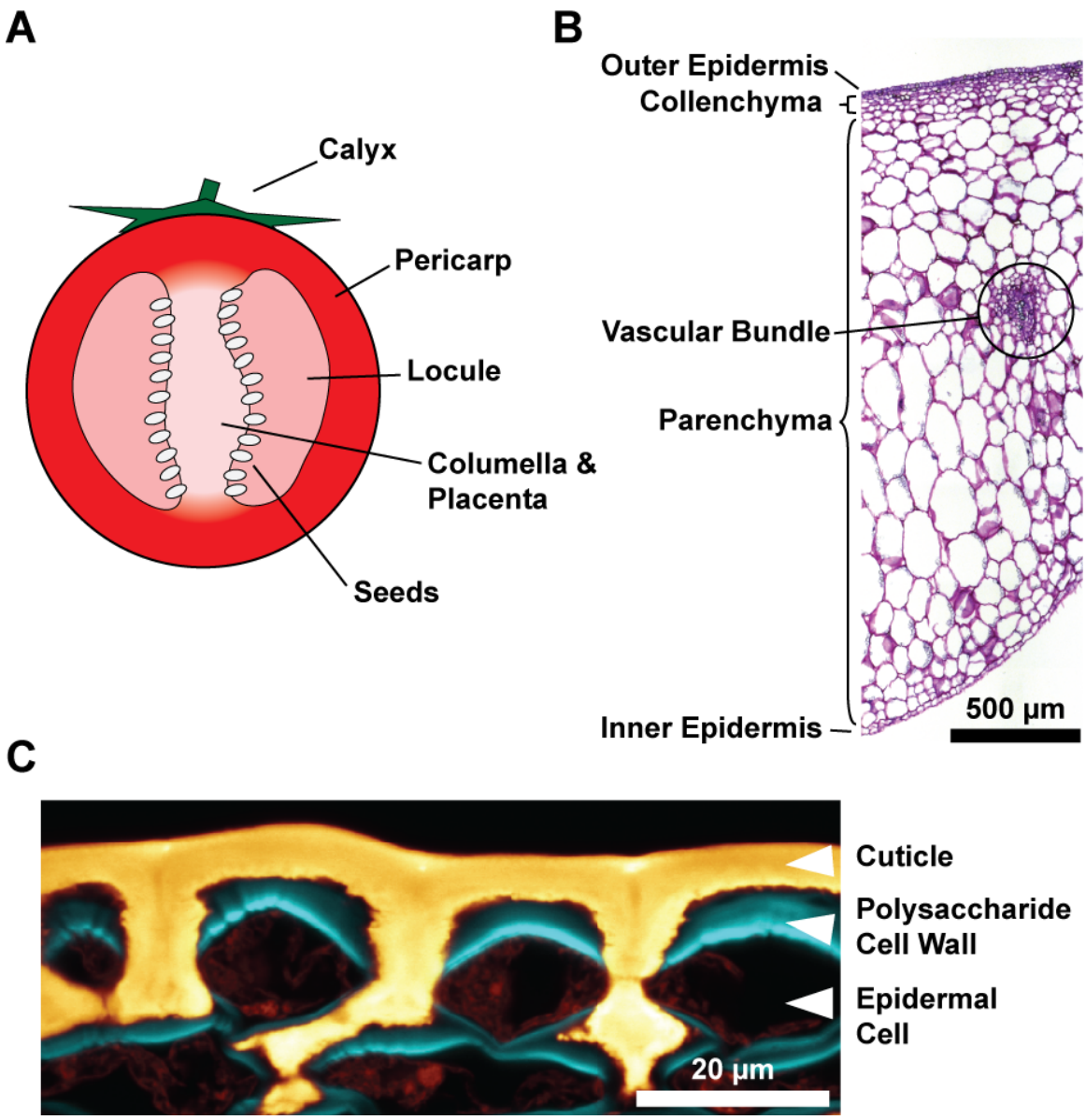


Figure 1.7. Tomato fruit anatomy. (a) Longitudinal schematic diagram of a mature tomato fruit. (b) Composite micrograph of an immature green tomato pericarp cross section showing the five principal tissues of the pericarp. (c) Confocal laser scanning micrograph of mature tomato outer epidermis. Lipids are stained with Auramine O (psuedocolored white and orange) and cellulose is stained with Calcofluor White (psuedocolored cyan).

The first few days of fruit growth and development are associated with cell division while the remaining growth phase results from cell expansion (Gillaspy *et al.*, 1993). During the course of fruit ontogeny, the surface area of the fruit, and thus the cuticle, can expand more than 500 times, and the cuticle coverage per unit of surface area can increase more than 10 fold (Baker *et al.*, 1982). Thus, early fruit development is accompanied by a rapid and substantial period of cuticle synthesis. Just as the high rate of cellulose synthesis by cotton (*Gossypium hirsutum*) fibers provided a model system that lead to the discovery of plant cellulose synthase (Pear *et al.*, 1996), presumably the genes and proteins required for of cuticle biosynthesis are highly expressed in the epidermis of the growing tomato fruit; a hypothesis that is addressed in this thesis.

The cutin of tomato fruit is composed primarily of 10,16-dihydroxyhexadecanoic acid, while its wax is primarily composed of very long chain *n*-alkanes, pentacyclic triterpenoids and the flavonoid naringenin-chalcone (Baker *et al.*, 1982). When the cuticle is isolated by incubation of pericarp samples with cellulases and pectinasse, a portion of the polysaccharides remain embedded in the cuticle rich residue that are relatively inaccessible to enzymatic hydrolysis. These intracuticular polysaccharides comprise about a third of the total cuticle mass and appear to contribute substantially to the cuticle stiffness (Lopez-Casado *et al.*, 2007).

Aside from its usefulness as a model for cuticle biology, applied research has also inspired several studies of the tomato fruit cuticle. For example, cultivar-specific fruit cracking has been associated with the cuticle biomechanical properties (Matas *et al.*, 2004). The enhanced shelf life, delayed softening and postharvest pathogen

resistance of the 'Delayed Fruit Deterioration' (DFD) tomato cultivar were also associated with the cuticle properties (Saladie *et al.*, 2007). Specifically, it has been proposed that the prolonged maintenance of DFD fruit texture results, at least in apart, from enhanced turgor maintenance, which in turn is likely due to decreased cuticular water permeability and increased mechanical support of the thicker cuticle of ripe DFD fruit. The enhanced level of cutin in ripe DFD fruit was also suggested to lead to a reduced susceptibility to postharvest infection by *Botrytis cinerea.* The correlation between resistance to *B. cinerea* and cutin amount was further extended in a study of three cutin deficient (*cd*) mutants, each with a greater than 95% reduction in the amount of cutin polymer in the fruit, and a corresponding increase in susceptibility to *B. cinerea* infection (Isaacson *et al.*, 2009).

Aims of the studies

The work described in this dissertation is focused on elucidating the poorly understood molecular mechanisms of cuticle biosynthesis using the experimental strengths of the tomato fruit model system. The first study focuses on exploiting the genetic diversity of tomato and its wild relatives in order to understand the genetic basis of the evolution of cuticle structure and function. The second study uses a proteomic approach to identify a suite of surface localized proteins that are expressed in expanding tomato fruit as a means of identifying candidate genes that may be involved in the extracellular steps of cuticle biosynthesis. The final study describes the biochemical characterization of one of these candidate proteins, a GDSL-motif lipase/hydrolase family protein that was independently discovered by forward genetics

to be encoded by the *CD1* gene. Finally, the appendices of this dissertation describe several collaborative projects that relate to cuticle biology, as well as a literature review on the topic of apoplastic lipid-transfer proteins.

References

Aarts, M.G., Keijzer, C.J., Stiekema, W.J. and Pereira, A. (1995) Molecular characterization of the *CER1* gene of Arabidopsis involved in epicuticular wax biosynthesis and pollen fertility. *Plant Cell* **7**, 2115–2127.

Baker, E.A., Bukovac, M.J. and Hunt, G.M. (1982) Composition of tomato fruit cuticle as related to fruit growth and development. In *The Plant Cuticle* (Cutler, D.F., Alvin, K.L. and Price, C.E., eds). London: Academic Press, pp. 33-44.

Barthlott, W. and Neinhuis, C. (1997) Purity of the sacred lotus, or escape from contamination in biological surfaces. *Planta* **202**, 1–8.

Bessire, M., Borel, S., Fabre, G. *et al.* (2011) A member of the PLEIOTROPIC DRUG RESISTANCE family of ATP binding cassette transporters is required for the formation of a functional cuticle in Arabidopsis. *Plant Cell* **23**, 1958–1970.

Bird, D., Beisson, F., Brigham, A., Shin, J., Greer, S., Jetter, R., Kunst, L., Wu, X., Yephremov, A. and Samuels, L. (2007) Characterization of Arabidopsis ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion. *Plant J.* **52**, 485–498.

Bonaventure, G., Beisson, F., Ohlrogge, J. and Pollard, M. (2004) Analysis of the aliphatic monomer composition of polyesters associated with Arabidopsis

epidermis: occurrence of octadeca-*cis*-6, *cis*-9-diene-1,18-dioate as the major component. *Plant J.* **40**, 920–930.

Bourdenx, B., Bernard, A., Domergue, F. *et al.* (2011) Overexpression of Arabidopsis *ECERIFERUM1* promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stresses. *Plant Physiol.* **156**, 29–45.

Buschhaus, C. and Jetter, R. (2011) Composition differences between epicuticular and intracuticular wax substructures: how do plants seal their epidermal surfaces? *J. Exp. Bot.* **62**, 841–853.

Cheesbrough, T.M. and Kolattukudy, P.E. (1984) Alkane biosynthesis by decarbonylation of aldehydes catalyzed by a particulate preparation from *Pisum sativum*. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6613–6617.

Deas, A.H.B. and Holloway, P.J. (1976) The intermolecular structure of some plant cutins. In *Lipids and Lipid Polymers in Higher Plants* (Tevini, M. and Lichtenthaler, H.K., eds). Berlin: Springer-Verlag, pp. 293-299.

Debono, A., Yeats, T.H., Rose, J.K., Bird, D., Jetter, R., Kunst, L. and Samuels, L. (2009) Arabidopsis LTPG is a glycosylphosphatidylinositol-anchored lipid transfer protein required for export of lipids to the plant surface. *Plant Cell* **21**, 1230–1238.

Dominguez, E., Heredia-Guerrero, J.A., Benitez, J.J. and Heredia, A. (2010) Self-assembly of supramolecular lipid nanoparticles in the formation of plant biopolyester cutin. *Mol. Biosyst.* **6**, 948–950.

Franke, R., Briesen, I., Wojciechowski, T., Faust, A., Yephremov, A., Nawrath, C. and Schreiber, L. (2005) Apoplastic polyesters in Arabidopsis surface tissues–a typical suberin and a particular cutin. *Phytochemistry* **66**, 2643–2658.

Gillaspy, G., Ben-David, H. and Gruissem, W. (1993) Fruits: A developmental perspective. *Plant Cell* **5**, 1439–1451.

Gorb, E., Haas, K., Henrich, A., Enders, S., Barbakadze, N. and Gorb, S. (2005) Composite structure of the crystalline epicuticular wax layer of the slippery zone in the pitchers of the carnivorous plant *Nepenthes alata* and its effect on insect attachment. *J. Exp. Biol.* **208**, 4651–4662.

Graca, J. and Lamosa, P. (2010) Linear and branched poly(omega-hydroxyacid) esters in plant cutins. *J. Agric. Food Chem.* **58**, 9666–9674.

Graca, J., Schreiber, L., Rodrigues, J. and Pereira, H. (2002) Glycerol and glyceryl esters of omega-hydroxyacids in cutins. *Phytochemistry* **61**, 205–215.

Greer, S., Wen, M., Bird, D., Wu, X., Samuels, L., Kunst, L. and Jetter, R. (2007) The cytochrome P450 enzyme CYP96A15 is the midchain alkane hydroxylase responsible for formation of secondary alcohols and ketones in stem cuticular wax of Arabidopsis. *Plant Physiol.* **145**, 653–667.

Gupta, N.S., Collinson, M.E., Briggs, D.E.G., Evershed, R.P. and Pancost, R.D. (2006) Reinvestigation of the occurrence of cutan in plants: implications for the leaf fossil record. *Paleobiology* **32**, 432–449.

Hannoufa, A., McNevin, J. and Lemieux, B. (1993) Epicuticular waxes of *eceriferum* mutants of *Arabidopsis thaliana*. *Phytochemistry* **33**, 851–855.

Isaacson, T., Kosma, D.K., Matas, A.J. *et al.* (2009) Cutin deficiency in the tomato fruit cuticle consistently affects resistance to microbial infection and biomechanical properties, but not transpirational water loss. *Plant J.* **60**, 363–377.

Jetter, R., Kunst, L. and Samuels, A.L. (2006) Composition of plant cuticular waxes. In *Biology of the Plant Cuticle* (Riederer, M. and Müller, C., eds). Oxford, UK: Blackwell, pp. 145-181.

Kamigaki, A., Kondo, M., Mano, S., Hayashi, M. and Nishimura, M. (2009) Suppression of peroxisome biogenesis factor 10 reduces cuticular wax accumulation by disrupting the ER network in *Arabidopsis thaliana*. *Plant Cell Physiol.* **50**, 2034–2046.

Kerstiens, G., Schreiber, L. and Lendzian, K.J. (2006) Quantification of cuticular permeability in genetically modified plants. *J. Exp. Bot.* **57**, 2547–2552.

Kolattukudy, P. (1976) Lipid polymers and associated phenols, their chemistry, biosynthesis, and role in pathogenesis. In *Recent Advances in Phytochemistry* (Loewus, F.A. and Runeckles, V.C., eds). New York: Plenum, pp. 185-246.

Kolattukudy, P.E. (2001) Polyesters in higher plants. *Adv. Biochem. Eng. Biotechnol.* **71**, 1–49.

Koornneef, M., Hanhart, C.J. and Thiel, F. (1989) A genetic and phenotypic description of *eceriferum* (*cer*) mutants in *Arabidopsis thaliana*. *J. Hered.* **80**, 118–122.

Kunst, L. and Samuels, L. (2009) Plant cuticles shine: advances in wax biosynthesis and export. *Curr. Opin. Plant Biol.* **12**, 721–727.

Kurdyukov, S., Faust, A., Nawrath, C. *et al.* (2006) The epidermis-specific extracellular BODYGUARD controls cuticle development and morphogenesis in Arabidopsis. *Plant Cell* **18**, 321–339.

Kutschera, U. and Niklas, K.J. (2007) The epidermal-growth-control theory of stem elongation: an old and a new perspective. *J. Plant Phys.* **164**, 1395–1409.

L'haridon, F., Besson-Bard, A., Binda, M. *et al.* (2011) A permeable cuticle is associated with the release of reactive oxygen species and induction of innate immunity. *PLoS Pathog.* **7**, e1002148.

Leide, J., Hildebrandt, U., Reussing, K., Riederer, M. and Vogg, G. (2007) The developmental pattern of tomato fruit wax accumulation and its impact on cuticular transpiration barrier properties: Effects of a deficiency in a beta-ketoacyl-coenzyme A synthase (LeCER6). *Plant Physiol.* **144**, 1667–1679.

Li-Beisson, Y., Pollard, M., Sauveplane, V., Pinot, F., Ohlrogge, J. and Beisson, F. (2009) Nanoridges that characterize the surface morphology of flowers require the synthesis of cutin polyester. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22008–22013.

Li, F., Wu, X., Lam, P., Bird, D., Zheng, H., Samuels, L., Jetter, R. and Kunst, L. (2008) Identification of the wax ester synthase/acyl-coenzyme A: diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. *Plant Physiol.* **148**, 97–107.

Lopez-Casado, G., Matas, A.J., Dominguez, E., Cuartero, J. and Heredia, A. (2007) Biomechanics of isolated tomato (*Solanum lycopersicum* L.) fruit cuticles: the role of the cutin matrix and polysaccharides. *J. Exp. Bot.* **58**, 3875–3883.

Lu, S., Song, T., Kosma, D.K., Parsons, E.P., Rowland, O. and Jenks, M.A. (2009) Arabidopsis *CER8* encodes LONG-CHAIN ACYL-COA SYNTHETASE 1 (LACS1) that has overlapping functions with LACS2 in plant wax and cutin synthesis. *Plant J.* **59**, 553–564.

Matas, A.J., Cobb, E.D., Bartsch, J.A., Paolillo, D.J. and Niklas, K.J. (2004) Biomechanics and anatomy of *Lycopersicon esculentum* fruit peels and enzyme-treated samples. *Am. J. Bot.* **91**, 352–360.

McFarlane, H.E., Shin, J.J., Bird, D.A. and Samuels, A.L. (2010) Arabidopsis ABCG transporters, which are required for export of diverse cuticular lipids, dimerize in different combinations. *Plant Cell* **22**, 3066–3075.

Mueller, L.A., Solow, T.H., Taylor, N. *et al.* (2005) The SOL Genomics Network: a comparative resource for Solanaceae biology and beyond. *Plant Physiol.* **138**, 1310–1317.

Nawrath, C. (2006) Unraveling the complex network of cuticular structure and function. *Curr. Opin. Plant Biol.* **9**, 281–287.

Panikashvili, D., Shi, J.X., Schreiber, L. and Aharoni, A. (2011) The Arabidopsis ABCG13 transporter is required for flower cuticle secretion and patterning of the petal epidermis. *New Phytol.* **190**, 113–124.

Park, J.J., Jin, P., Yoon, J., Yang, J.I., Jeong, H.J., Ranathunge, K., Schreiber, L., Franke, R., Lee, I.J. and An, G. (2010) Mutation in *Wilted Dwarf and Lethal 1* (*WDL1*) causes abnormal cuticle formation and rapid water loss in rice. *Plant Mol. Biol.* **74**, 91–103.

Pear, J.R., Kawagoe, Y., Schreckengost, W.E., Delmer, D.P. and Stalker, D.M. (1996) Higher plants contain homologs of the bacterial *celA* genes encoding the catalytic subunit of cellulose synthase. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12637–12642.

Pfündel, E.E., Agati, G. and Cerovic, Z.G. (2006) Optical properties of plant surfaces. In *Biology of the Plant Cuticle* (Riederer, M. and Müller, C., eds). Oxford, UK: Blackwell, pp. 216-249.

Pighin, J.A., Zheng, H., Balakshin, L.J., Goodman, I.P., Western, T.L., Jetter, R., Kunst, L. and Samuels, A.L. (2004) Plant cuticular lipid export requires an ABC transporter. *Science* **306**, 702–704.

Pollard, M., Beisson, F., Li, Y. and Ohlrogge, J.B. (2008) Building lipid barriers: biosynthesis of cutin and suberin. *Trends. Plant Sci.* **13**, 236–246.

Ray, A.K., Chen, Z.J. and Stark, R.E. (1998) Chemical depolymerization studies of the molecular architecture of lime fruit cuticle. *Phytochemistry* **49**, 65–70.

Ray, A.K. and Stark, R.E. (1998) Isolation and molecular structure of an oligomer produced enzymatically from the cuticle of lime fruit. *Phytochemistry* **48**, 1313–1320.

Reina-Pinto, J.J. and Yephremov, A. (2009) Surface lipids and plant defenses. *Plant Physiol. Biochem.* **47**, 540–549.

Reina, J.J., Guerrero, C. and Heredia, A. (2007) Isolation, characterization, and localization of *AgaSGNH* cDNA: a new SGNH-motif plant hydrolase specific to *Agave americana* L. leaf epidermis. *J. Exp. Bot.* **58**, 2717–2731.

Riederer, M. and Schreiber, L. (2001) Protecting against water loss: analysis of the barrier properties of plant cuticles. *J. Exp. Bot.* **52**, 2023–2032.

Rowland, O., Zheng, H., Hepworth, S.R., Lam, P., Jetter, R. and Kunst, L. (2006) *CER4* encodes an alcohol-forming fatty acyl-coenzyme A reductase involved in cuticular wax production in Arabidopsis. *Plant Physiol.* **142**, 866–877.

Saladie, M., Matas, A.J., Isaacson, T. *et al.* (2007) A reevaluation of the key factors that influence tomato fruit softening and integrity. *Plant Physiol.* **144**, 1012–1028.

Schirmer, A., Rude, M.A., Li, X., Popova, E. and del Cardayre, S.B. (2010) Microbial biosynthesis of alkanes. *Science* **329**, 559–562.

Schneider-Belhaddad, F. and Kolattukudy, P. (2000) Solubilization, partial purification, and characterization of a fatty aldehyde decarbonylase from a higher plant, Pisum sativum. *Arch. Biochem. Biophys.* **377**, 341–349.

Schreiber, L., Elshatshat, S., Koch, K., Lin, J. and Santrucek, J. (2006) AgCl precipitates in isolated cuticular membranes reduce rates of cuticular transpiration. *Planta* **223**, 283–290.

Tian, S., Fang, X., Wang, W., Yu, B., Cheng, X., Qiu, F., Mort, A.J. and Stark, R.E. (2008) Isolation and identification of oligomers from partial degradation of lime fruit cutin. *J. Agric. Food Chem.* **56**, 10318–10325.

Yang, W., Pollard, M., Li-Beisson, Y., Beisson, F., Feig, M. and Ohlrogge, J. (2010) A distinct type of glycerol-3-phosphate acyltransferase with sn-2 preference and phosphatase activity producing 2-monoacylglycerol. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 12040–12045.

Yeats, T.H. and Rose, J.K. (2008) The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). *Protein Sci.* **17**, 191–198.

CHAPTER 2

The fruit cuticles of wild tomato species exhibit architectural and chemical diversity, providing a new model for studying the evolution of cuticle function

Trevor H. Yeats[1], Gregory J. Buda[1], Zhonghua Wang[2], Noam Chehanovsky[3], Leonie C. Moyle[4], Reinhard Jetter[2,5], Arthur A. Schaffer[3] and Jocelyn K.C. Rose[1*]

[1] Department of Plant Biology, Cornell University, Ithaca, NY 14853, U.S.A.

[2] Department of Botany, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

[3] Institute of Field and Garden Crops, ARO, The Volcani Center, Bet Dagan 50250, Israel

[4] Department of Biology, Indiana University, Bloomington, IN, 47405 U.S.A.

[5] Department of Chemistry, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.

Summary

The cuticle covers the aerial epidermis of land plants and plays a primary role in water regulation and protection from external stresses. Remarkable species diversity in the structure and composition of its components, cutin and wax, have been catalogued, but few functional or genetic correlations have emerged. Tomato (*Solanum lycopersicum*) is part of a complex of closely related wild species endemic to the northern Andes and the Galapagos Islands (*Solanum* Sect. *Lycopersicon*). Although sharing an ancestor less than seven million years ago, these species are found in diverse environments and are subject to unique selective pressures. Furthermore, they are genetically tractable, since they can be crossed with *S. lycopersicum,* which has a sequenced genome. With the aim of evaluating the relationships between evolution, structure and function of the cuticle, we characterized the morphological and chemical diversity of fruit cuticles of seven species from *Solanum* Sect. *Lycopersicon*. Striking differences in cuticular architecture and quantities of cutin and waxes were observed, with wild species wax coverage exceeding that of *S. lycopersicum* by up to seven fold. Wax composition varied in the occurrence of wax esters and triterpenoid isomers. Using a *S. habrochaites* introgression line population, we mapped triterpenoid differences to a genomic region that includes two *S. lycopersicum* triterpene synthases. Based on known metabolic pathways for acyl wax compounds, hypotheses are discussed to explain the appearance of wax esters with atypical chain lengths. These results establish a model system for understanding the ecological and evolutionary functional genomics of plant cuticles.

Introduction

The plant cuticle is a waxy surface layer covering the primary aerial organs of all land plants. While it is central to limiting non-stomatal water loss, it is now clear that the cuticle plays a myriad of roles as the primary interface between the plant and its environment. It is an effective barrier against pests and pathogens (Reina-Pinto and Yephremov, 2009), shields the plant from excessive UV radiation (Pfündel *et al.*, 2006) and can act as a self-cleaning surface (Barthlott and Neinhuis, 1997). It also has a critical function in plant development, by establishing boundaries between nascent organs (Javelle *et al.*, 2011).

The hydrophobic cuticle is contiguous with the polysaccharide cell wall and consists primarily of a lipid polymer, cutin, and a variety of organic solvent-soluble compounds that are collectively termed waxes. Cutin is a polyester of ω- and midchain-substituted fatty acids and three main types exist, based on the predominant chain length of fatty acids and the nature of the substitutions. C16-type cutins are typically rich in dihydroxyhexadecanoic acid with midchain- and 16-hydroxy groups, while C18-type cutins are principally composed of 9,10-epoxy-18-hydroxyoctadecanoic acid or 9,10,18-trihydroxyoctadecanoic acid (Kolattukudy, 2001). More recently, analysis of arabidopsis (*Arabidopsis thaliana*) stem and leaf cutin identified a third type of cutin that is rich in C18 dicarboxylic acids (Bonaventure *et al.*, 2004; Franke *et al.*, 2005). Finally, although not typically detected for technical reasons, glycerol is found in varying abundance in the cutin polymers of many species (Graca *et al.*, 2002).

Waxes accumulate within the cutin matrix as intracuticular waxes, and they are also deposited on the outer surface of the cuticle as epicuticular crystals or films (Buschhaus and Jetter, 2011). Wax mixtures are typically more complex and variable than cutin, but broadly consist of homologous series of acyl lipids, derived from very-long-chain fatty acids (VLCFAs), and a variety of other lipophilic metabolites. The acyl lipids include alkanes, fatty acids, alcohols and wax esters, the occurrence and abundance of which vary between species and during ontogeny (Jenks and Ashworth, 1999). Examples of non-acyl wax compounds are pentacyclic triterpenoids, flavonoids and tocopherols, although their incidence is even more variable (Jetter *et al.*, 2006).

Fossil evidence suggests that evolution of a cuticle was one of the primary adaptations that allowed plants to colonize land, with both morphological and chemical evidence of the first cuticles dating to the late Silurian and early Devonian periods (Edwards, 1993; Niklas, 1980). Over the last four hundred million years substantial diversification in cuticle morphology and composition has occurred (Jeffree, 2006; Walton, 1990); however, attempts to correlate this variation with functional characteristics of the cuticle have had mixed success. On one hand, the self-cleaning of the plant surface by water beading, known as the lotus effect, was correlated with deposition of epicuticular wax crystals through a survey of ~ 10,000 diverse species (Barthlott and Neinhuis, 1997). On the other hand, the intuitive correlations between cuticle thickness or wax amount, and cuticular water permeability were not confirmed by a survey of the cuticles of 23 species (Riederer and Schreiber, 2001). However, it is important to note that such taxonomic comparisons have largely involved distantly related species, in the absence of genetic

resources, thus limiting molecular interpretation of cuticular diversity, while molecular models for cuticle biosynthetic pathways have mostly resulted from studies of mutants in model systems such as arabidopsis (Kunst and Samuels, 2009; Pollard *et al.*, 2008).

In order to resolve the complex interaction between cuticle structures, functions and evolution, a promising approach lies in the emerging discipline of ecological and evolutionary functional genomics (EEFG; Mitchell-Olds *et al.*, 2008). A prerequisite for this approach is a set of species with diverse ecological preferences, but with relatively recent evolutionary divergence, as well as access to the genomic tools of a model organism. For EEFG studies of the plant cuticle, an excellent model is presented by the wild relatives of tomato (*S. lycopersicum*). This group of approximately 14 species, *Solanum* Sect. *Lycopersicon*, evolved from a common ancestor less than seven million years ago and today is endemic to an array of environments in the northern Andes and Galapagos Islands (Peralta *et al.*, 2008). Furthermore, they can be readily crossed with *S. lycopersicum* and share a high degree of genomic synteny, allowing the sequenced genome of *S. lycopersicum* and its associated genomic resources to be utilized (Mueller *et al.*, 2005). A particularly useful resource for comparative genetic studies is the availability of several introgression line (IL) populations, consisting of discrete marker-defined homozygous segments of wild species chromosomes in a *S. lycopersicum* background (Prudent *et al.*, 2009; Monforte and Tanksley, 2000; Eshed and Zamir, 1995).

In addition to the advantages of *Solanum* Sect. *Lycopersicon* as an EEFG model system, tomato is an excellent experimental resource for cuticle biology in its own right. The tomato fruit cuticle is particularly substantial, enabling detailed

morphological analysis by light microscopy (Buda *et al.*, 2009), as well as easy isolation of a completely intact membrane, facilitating both biomechanical and chemical analysis. Additionally, the fruit epidermis is astomatous, greatly simplifying any tests of cuticular permeability. Finally, both the wax and cutin of tomato fruits have relatively simple compositions (Baker *et al.*, 1982).

To test the hypothesis that *Solanum* Sect. *Lycopersicon* represents a valuable EEFG model system for cuticle biology, we characterized the cuticular morphology and composition of seven of these species, which were selected to span the range of evolutionary history, geography and environment in *Solanum* Sect. *Lycopersicon* (Figure 2.1). While the cuticle may, at least in some plant species, be subject to dynamic adaptation to environmental stresses (Kosma *et al.*, 2009), for this study we first focused on establishing the static features of the cuticle by growing all tomato species in the same greenhouse conditions. We present the remarkable architectural and chemical diversity exhibited by cuticles of the wild relatives of tomato, and demonstrate the value of existing genetic resources to identify the molecular basis of the underlying compositional diversity.

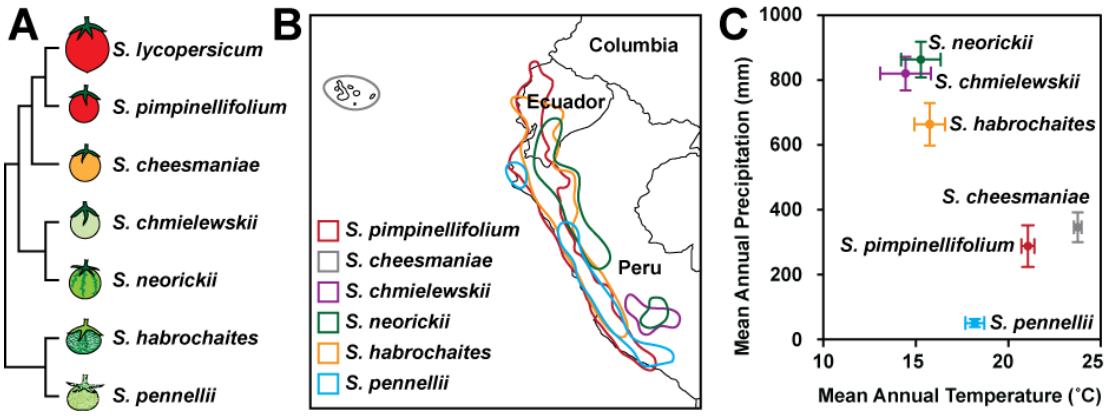


Figure 2.1. Phylogenetic and ecological context of the species in this study. (a) Phylogenetic tree of the seven *Solanum* species considered here, based on Rodriguez *et al.* 2009. (b) Geographical distribution of the six wild species based on the occurrence records of the Tomato Genetic Resource Center (TGRC, UC Davis, http://tgrc.ucdavis.edu). (c) Climate space plot of each species (mean and 95% CI) based on records for known geographical locations of each species (TGRC), using methods described in Nakazato *et al.* (2010).

Results

Two-dimensional Cuticle Morphology

In order to characterize the structural features of the fruit cuticles, isolated fruit pericarp cryosections were prepared and observed with differential interference contrast (DIC) light microscopy (Figure 2.2a-g). In *S. lycopersicum, S. pimpinellifolium* and *S. cheesmaniae* (Figure 2.2a-c, respectively), a flat surface topology was observed, with well developed anticlinal pegs (APs) between the epidermal cells. The APs of *S. pimpinellifolium* are particularly pronounced: the epidermal cells adopt a nearly conical shape and the anticlinal cell wall is nearly filled

by substantial but tapered pegs (Figure 2.2b). In the remaining species, a papillate surface topology was observed, generally leading to the incorporation of the AP structures into the continuous undulating organization of the cuticle (Figure 2.2d-g). *S. habrochaites* has an intermediate architecture, where the papillate morphology is less pronounced and APs are present, but only to half the depth of the epidermal cells (Figure 2.2f). The occurrence of subepidermal cuticular deposits (SD) correlates with the appearance of APs (Fig 2a-g), with *S. pimpinellifolium* exhibiting the greatest degree of cuticle accumulation in both of these locations (Figure 2.2b).

To complement the qualitative observations of cuticle structure, the cuticle thickness above each epidermal cell was quantified (Figure 2.2h). The two sister species *S. habrochaites* and *S. pennellii* had the thinnest cuticles, each with an average thickness of ~4 μm, while another set of sister species, *S. neorickii* and *S. chmielewskii,* had the thickest cuticles, with an average thickness of ~7 μm.

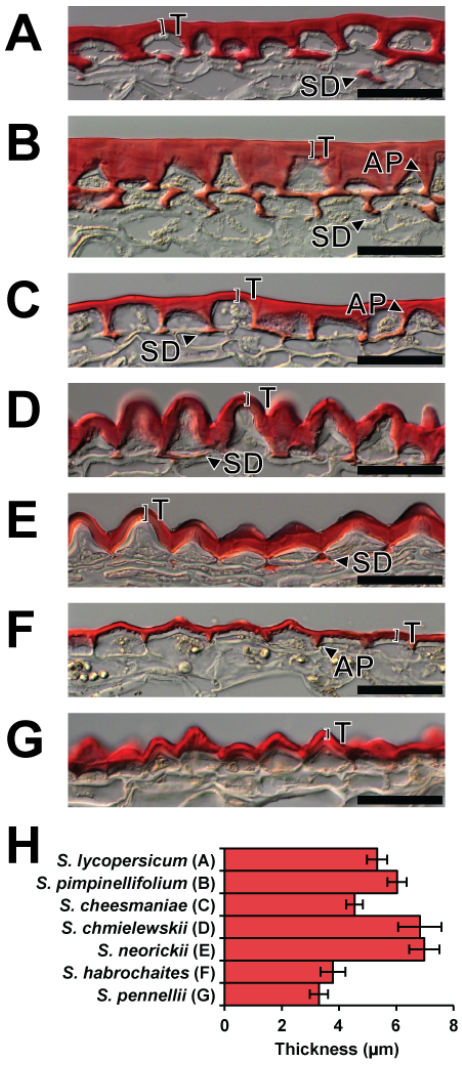


Figure 2.2. Two-dimensional cuticle morphology. (a-g) Light micrographs of sections of the fruit surface of tomato (*S. lycopersicum*) and six wild relatives, showing the cuticle stained with Oil red O. Scale bars = 50 µm. T, Thickness; SD, Subepidermal cuticular deposit; AP, Anticlinal peg. The species represented in each panel is shown in panel H of this figure. (h) Cuticle thickness measurements. Measurements were made above the center of each epidermal cell as indicated by bars labeled "T" in panels A-G. Error bars are S.E. for n=5.

Three-dimensional Cuticle Morphology

Since preliminary analysis by DIC microscopy showed considerable diversity in cuticle structure, cuticle morphology was further investigated for several of the species using three-dimensional confocal microscopy. Thick sections (30 μm) of fruit pericarp material were prepared by cryosectioning and stained with Auramine O. A series of optical sections were collected using a confocal scanning laser microscope and these were assembled into three-dimensional volume renderings. Using this technique, we were able to clearly observe a number of unique topological features that were either subtle or invisible by DIC microscopy.

In *S. pimpinellifolium* sections, the APs were divided down the middle by more lightly stained and presumably more polar anticlinal lamellae (AL) (Figure 2.3b,c). These clearly outlined the anticlinal boundary of each cell (Figure 2.3c). In species with papillate surface topology, this could be seen in striking relief (Figure 2.3d-k). Moreover, the internal surface topology (IST) was revealed. *S. habrochaites* has a particularly rough IST consisting of fused spherical globules of ~1 μm diameter (Figure 2.3g,h). In contrast, the IST of *S. neorickii* and *S. pennellii* were relatively smooth (Figure 2.3e,k).

Scanning electron microscopy (SEM) was used on a subset of species in order to confirm the structural features that were observed by light and confocal microscopy. While the papillate surface topology of *S. chmielewskii* and *S. habrochaites* was readily visible using SEM, the internal surface topology could not be observed, nor

was it possible to distinguish between the cuticular membrane and the polysaccharide cell wall (Figure 2.4).

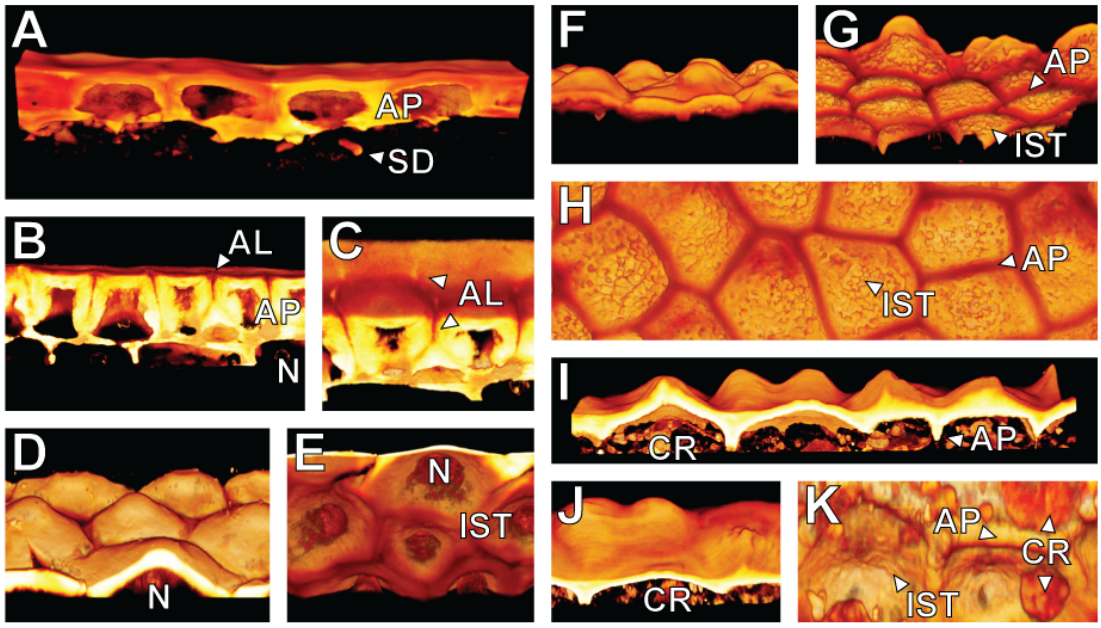


Figure 2.3. Three-dimensional cuticle morphology. Three-dimensional volume renderings constructed from confocal Z-stacks. (a) *S. lycopersicum.* (b-c) *S. pimpinellifolium*. (d-e) *S. neorickii.* (f-h) *S. habrochaites.* (i-k) *S. pennellii.* AP, Anticlinal peg; SD, Subepidermal cuticular deposit; AL, Anticlinal lamellae; N, Nucleus; IST, Internal surface topology; CR, Cellular remnants.

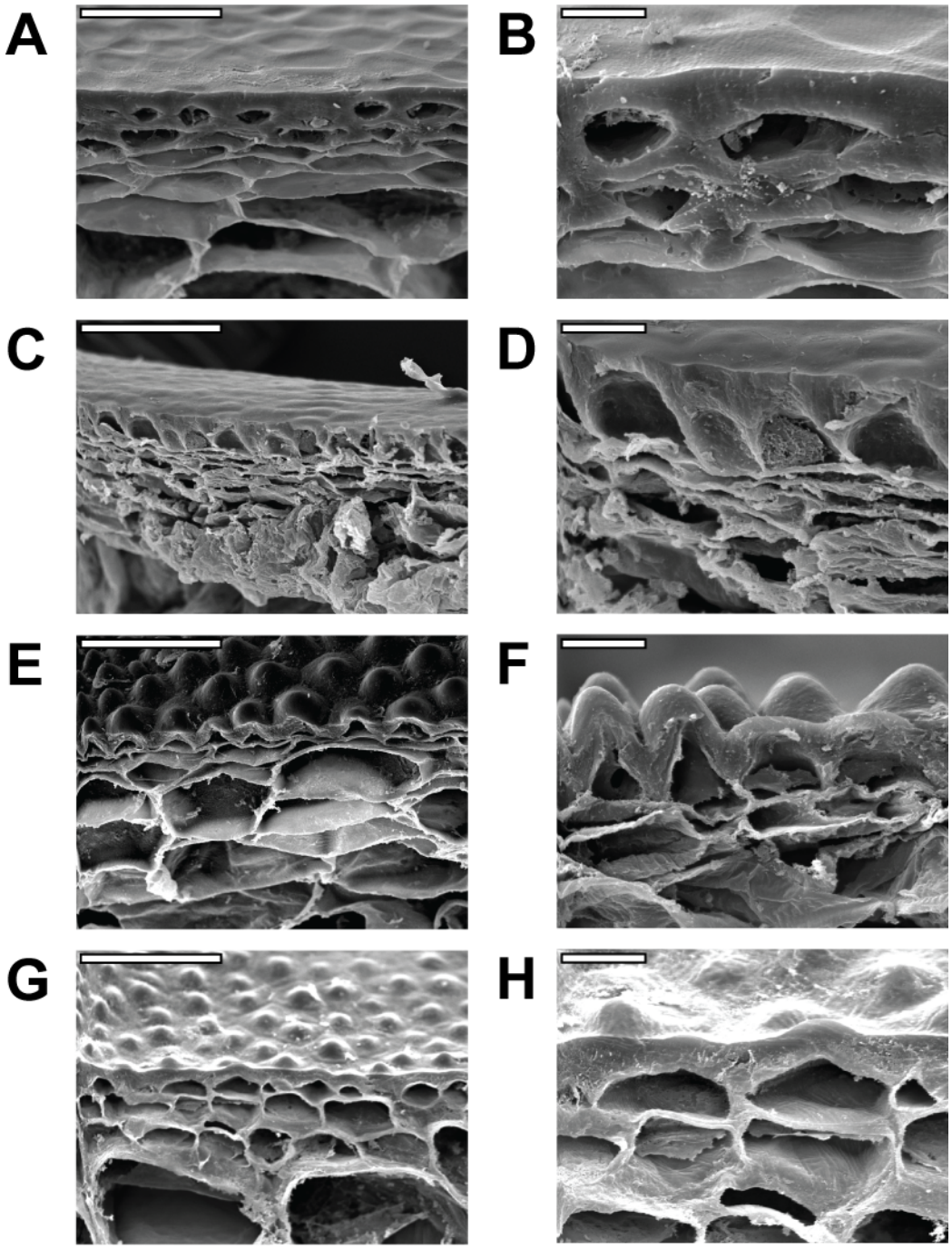


Figure 2.4. SEM micrographs of pericarp of selected *Solanum* species. (a-b). *S. lycopersicum*. (c-d). *S. pimpinellifolium.* (e-f). *S. chmielewskii.* (g-h). *S. habrochaites.* Scale bars are 100 µm for a, c, e and g and 20 µm for b, d, f and h.

Cutin chemical analysis

The predominant component of the tomato fruit cuticle is cutin, which typically exceeds the mass of wax by nearly 100-fold (Baker *et al.*, 1982). Thus, the major lipids stained in our microscopic analysis corresponded to the cutin polymer. To characterize the cutin in more detail, the fruit cuticular membranes from each species were enzymatically isolated, extracted of wax and depolymerized by methanolysis. The resulting methyl esters of cutin monomers were identified by GC-MS and quantified by GC-FID (Figure 2.5).

In terms of overall quantity of cutin identified, the chemical analysis generally confirmed the pattern based on microscopic thickness measurements (Figure 2.2h). However, a major exception is *S. pimpinellifolium*, which exhibits nearly twice the cutin load, while having nearly the same thickness as *S. lycopersicum*. This discrepancy is likely accounted for by the cutin in the thick APs of this species (Figure 2.2b, Figure 2.3b,c). Another notable result of comparing the microscopy images with the chemical analysis is the fact that, while *S. pennellii* and *S. habrochaites* both have similar cuticular thickness and architecture (Figure 2.2f-h), the cuticle of *S. habrochaites* is nearly three times as rich in cutin as *S. pennellii.*

In terms of cutin monomer composition, the most notable trend was the enhanced levels of 9,10,18-trihydroxyoctadecanoic acid in three species. While this monomer accounted for 6%, 10% and 6% of the cutin in *S. chmielewskii, S.*

habrochaites and *S. pennellii*, respectively, it corresponded to less than 3% of the cutin in the remaining species (Figure 2.5).

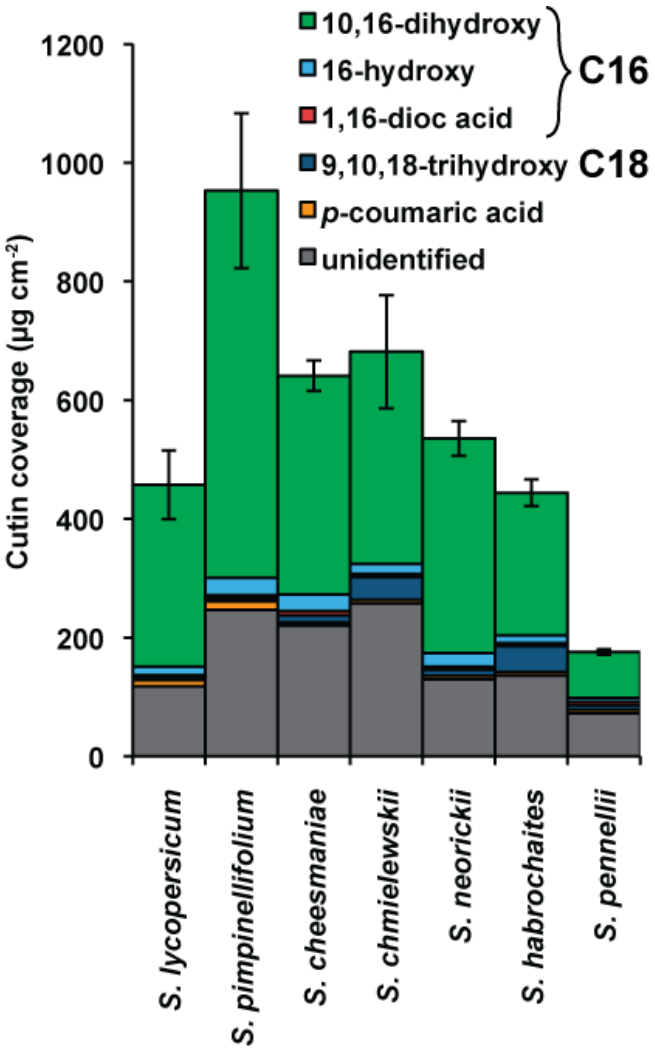


Figure 2.5. Cutin chemical composition. Monomer composition of cutin from fruit cuticles of various *Solanum* species. Error bars are S.E. for n=3.

Wax chemical analysis

Wax extracts from the isolated cuticles were analyzed by gas chromatography coupled to mass spectrometry and flame ionization detector (GC-MS and CG-FID). All wild species accumulated more wax than *S. lycopersicum*, although only three species (*S. chmielewskii, S. neorickii* and *S. pennellii*) show substantially greater wax levels (Figure 2.6a). The wax coverage of *S. neorickii* is particularly remarkable and exceeds that of *S. lycopersicum* by nearly seven-fold.

Considering the wax composition, alkanes represented the majority of identified compounds in all species (Figure 2.6a). Within the homologous series of *n*-alkanes, the typical predominance of odd chain lengths was observed in all species (Figure 2.6b), with the most common chain length being C31, except in *S. chmielewskii, S. neorickii* and *S. habrochaites*, where C29 was the most abundant (Figure 2.6b). In the same three species, alkyl esters were detected, accounting for 7-20% of the wax (Figure 2.6a). Alkyl esters typically show a chain length distribution favoring even chain lengths (Jetter *et al.*, 2006), but for the three species where alkyl esters were observed, relatively little even bias was seen (Figure 2.6c). To further investigate whether the unexpected abundance of odd-chain lengths was due to contributions from the fatty acid or alcohol moiety, wax extracts from *S. neorickii* were fractionated by thin layer chromatography (TLC) and the ester fraction was recovered and methanolyzed. GC analysis showed the esters to be composed of C16-C28 fatty acids that were mostly even in chain length, and C22-C30 alcohols of both even and odd chain lengths (Figure 2.7).

Across the species, the major non-aliphatic compounds observed were the pentacyclic triterpenoids, accounting for 1-35% of the total wax, although they could not be detected in *S. pennellii* (Figure 2.6a). The predominant isomers observed were α-amyrin, β-amyrin and δ-amyrin, with Ψ-taraxasterol and taraxasterol appearing at much lower levels. In *S. habrochaites*, β-amyrin was the only pentacyclic triterpenoid that was detected (Figure 2.6d).

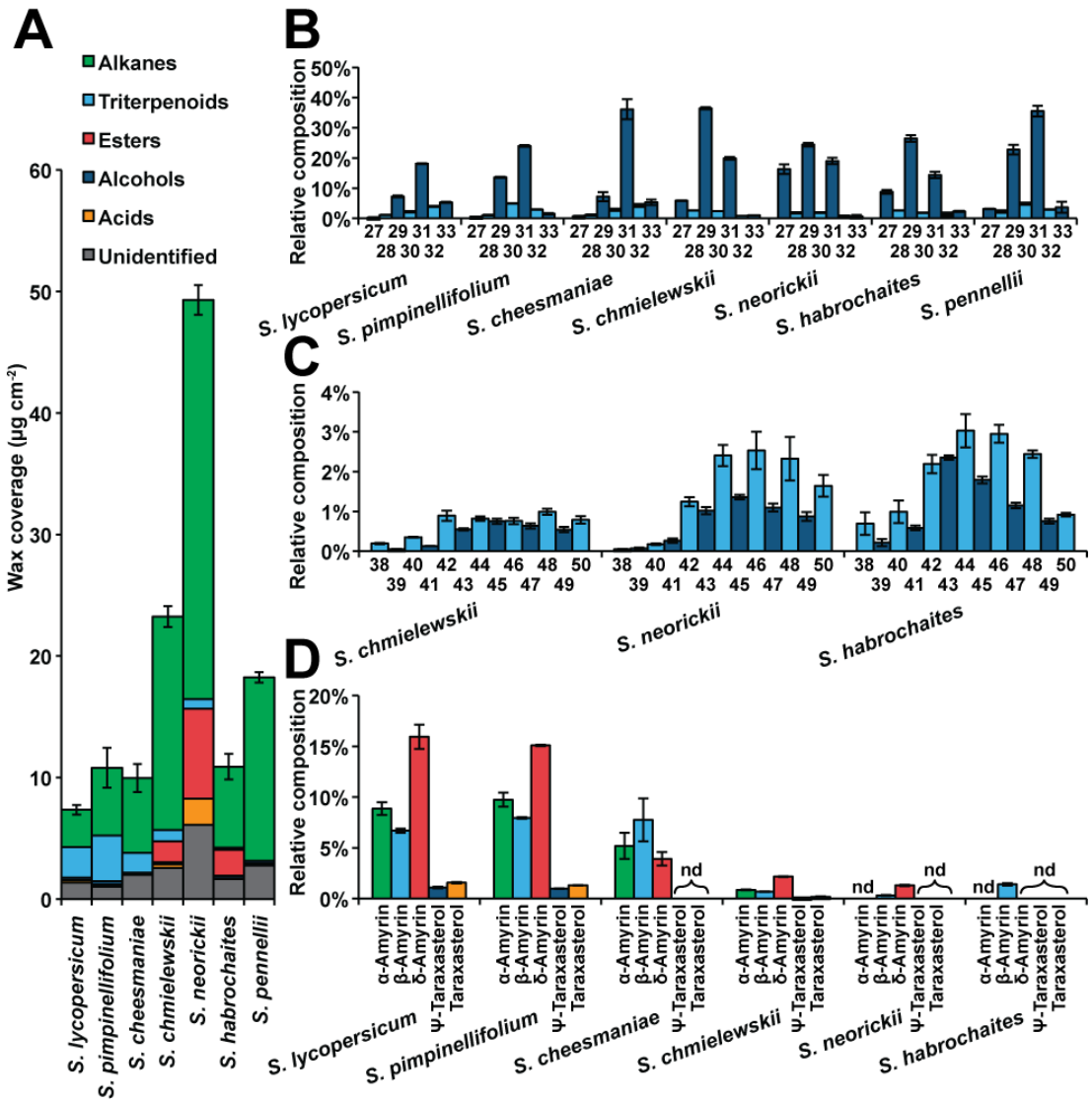


Figure 2.6. Wax chemical composition. (a) Total wax and compound class coverage for each species. (b) *n*-Alkane chain length distribution for each species. The number of carbons is indicated below. (c) Alkyl ester chain length distribution for the three species exhibiting this compound class. The number of carbons is indicated below. (d) Triterpenoid isomer distribution for the six species where triterpenoids were observed. nd, Not detected. Error bars are S.E. for n=3.

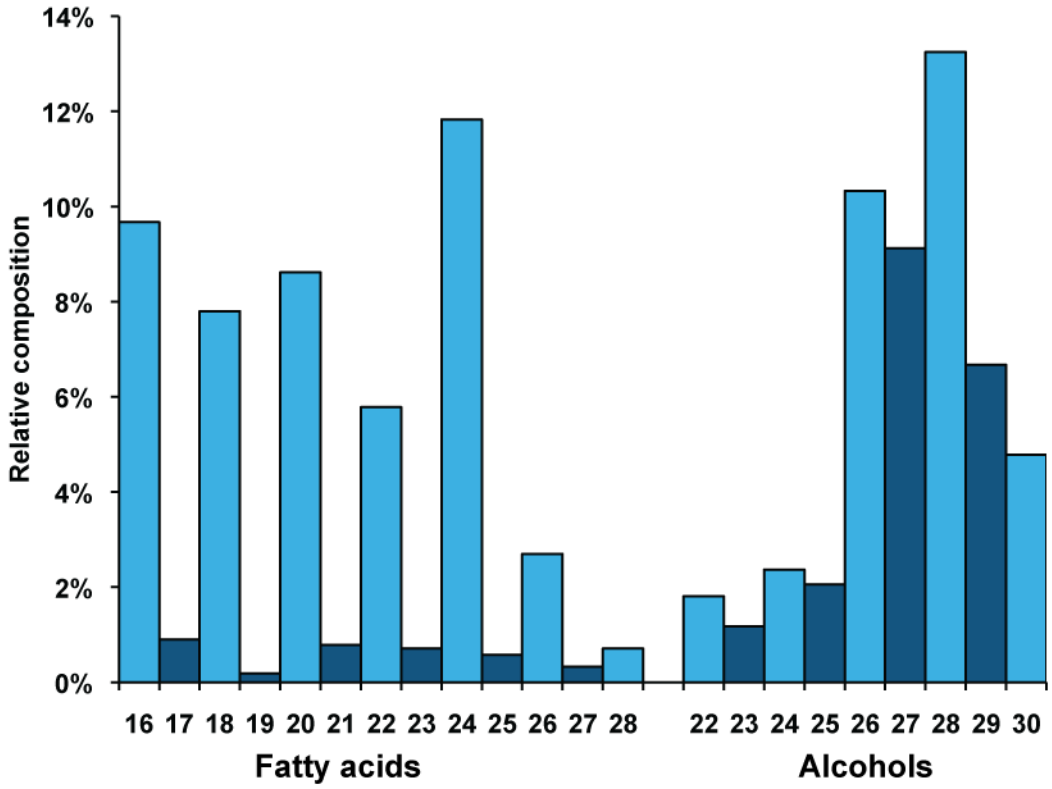


Figure 2.7. Chain length distribution of fatty acid and alcohol constituents of the wax ester fraction from *S. neorickii*. The wax esters were purified by TLC and methanolyzed with sodium methoxide and the resulting fatty acid methyl esters and alcohols (as TMS ethers) were analyzed by GC.

Genetic mapping of cuticle traits using IL populations

As a first demonstration of using ILs to genetically map cuticular traits, we focused on understanding the genetic basis for the absence of non-β-amyrin triterpenoid isomers in *S. habrochaites.* We isolated cuticular wax from a collection of 51 *S. habrochaites* ILs that collectively cover nearly the entire *S. habrochaites* genome (Monforte and Tanksley, 2000). GC analysis revealed that the wax extracts of

these lines all had a similar triterpenoid isomer composition to the *S. lycopersicum* parent with the exception of line LA3917, the triterpenoid profile of which consisted almost entirely of β-amyrin (Figure 2.8a). This line contains two unique DNA segments on chromosomes 1 and 12 that are not present in any of the other lines examined. The segment on the end of chromosome 12 includes the tandem loci *SlTTS1* and *SlTTS2* that encode the two triterpenoid synthases responsible for biosynthesis of the entire array of triterpenoid isomers that are observed in *S. lycopersicum* (Figure 2.8b). SlTTS1 specifically synthesizes β-amyrin, while SlTTS2 is a multifunctional synthase with δ-amyrin as its most abundant product (Figure 2.8c; Wang *et al.*, 2011).

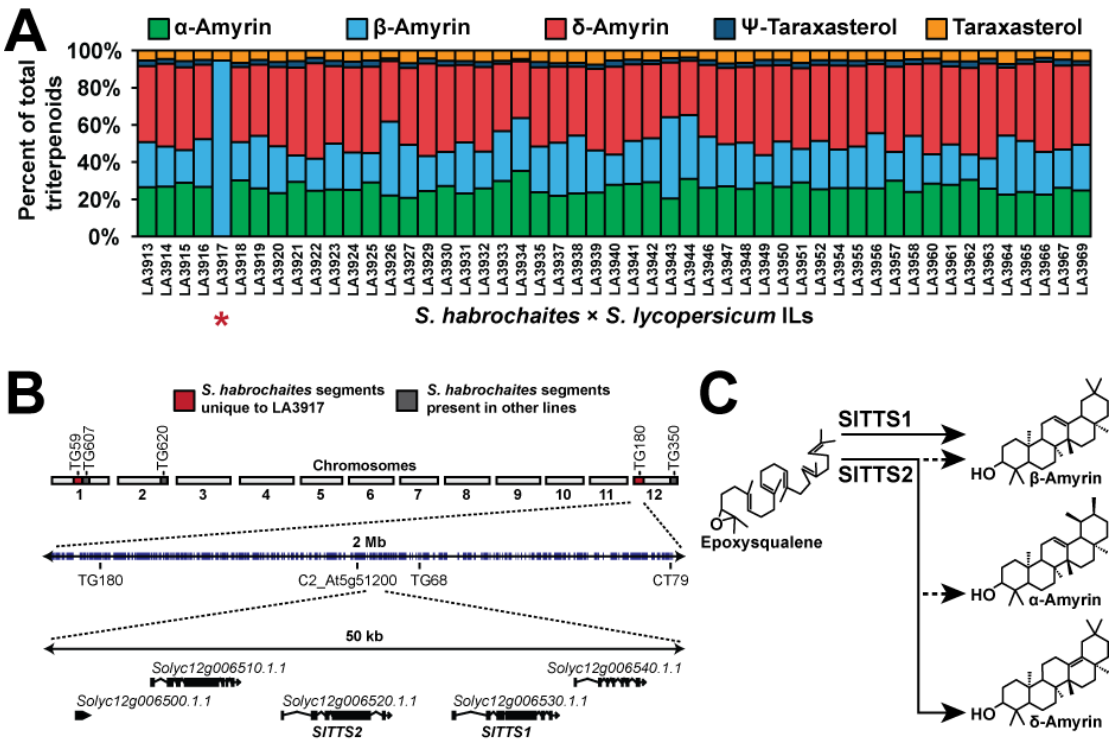


Figure 2.8. Genetic mapping of α- and δ-amyrin synthesis. (a) Relative contribution of triterpenoid isomers to the total cuticular triterpenoids for 51 *S. habrochaites* introgression lines in a *S. lycopersicum* background. LA3917 (indicated by *) had no detectable α- or δ-amyrin. (b) The genetic architecture of LA3917. The introgressed segments surrounding the markers TG59 and TG180 are uniquely represented in this line (shaded in red). The latter segment contains the tandem loci SlTTS1 and SlTTS2 that encode the two triterpene synthases of *S. lycopersicum*. (c) The isomer specificity of the two triterpene synthases of *S. lycopersicum*, SlTTS1 and SlTTS2. SlTTS1 synthesizes β-amyrin exclusively while SlTTS2 synthesizes primarily δ-amyrin and lesser amounts of α- and β-amyrin.

Discussion

Cuticle Morphology

The microscopic morphology of the fruit cuticle and the underlying epidermal cell layer was strikingly variable between the seven tomato species examined, despite relatively little variation in overall cuticle thickness (Figure 2.2). We observed no association between climatic factors (e.g. mean annual temperature and precipitation) and cuticular thickness (Figures 2.1 and 2.2), which is consistent with previous results showing no correlation between cuticle thickness and water permeability in a survey of taxonomically diverse species (Riederer and Schreiber, 2001) and a study of tomato mutants with varying degrees of fruit cutin deficiency (Isaacson *et al.*, 2009).

There is a general trend of a flat rather than undulating cuticle surface in species most closely related to *S. lycopersicum* (Figures 2.2 and 2.3), which would have a substantial effect on surface area and cuticular water conductance. For example, measurement of the surface following the two dimensional contour of the epidermis of *S. chmielewskii,* rather than the straight line assumed by approximating the surface area of a sphere, indicates that the effective surface area is ~50% greater. The effective cuticular transpiration, expressed as the flux of water across the epidermal surface, would thus be significantly affected by the cuticular topology on the cellular scale.

Previous studies of quantitative trait loci (QTLs) derived from *S. chmielewskii* identified several QTLs associated with increased or decreased fruit cuticular water permeability (Prudent *et al.*, 2009). The differences associated with these QTLs were

on the same order of magnitude as the ~50% difference in microscopic surface area and the ~200% difference in wax load that we observed in comparing *S. lycopersicum* and *S. chmielewskii* (Figures 2.2 and 2.6). It is likely that the variation in cuticular water permeability that was observed, considering macroscopic surface area, was due to either variation in cuticular morphology, or wax amount and composition. Along these lines, the previously described *Cwp1* genotype, containing a *S. habrochaites* allele of this gene that is expressed in the *S. lycopersicum* background, had enhanced cuticular water permeability that was associated not with altered cuticle chemistry, but rather with microfissures within the cuticle (Hovav *et al.*, 2007).

If variation in cuticular morphology were shaped by adaptive responses to environmental variation, one expectation might be that a lower surface area would be favored in species endemic to warm and dry environments. Indeed, this trend is observed with all the species except *S. pennellii*, which is endemic to regions with the lowest precipitation conditions of all the species considered, and yet features an undulating cuticle (Figures 2.1 and 2.2). Interestingly, *S. pennellii* frequently occurs in 'lomas formations'; mid-elevation fog-zone locations that have low precipitation but high ambient humidity (Dillon, 1989). Such a unique climatic condition might explain this anomalous observation. Clearly, many factors are important for determining cuticular water permeability. Wax composition and amount are often discussed as principal determinants of cuticular water permeability (Jenks and Ashworth, 1999) and pubescent surfaces, such as that of *S. habrochaites,* can also contribute to the transpirational barrier (Fernández *et al.*, 2011). However, from both a technical and

evolutionary perspective, the microscopic topology of the cuticle surface appears to be an important factor.

Cutin

The total amount of fruit cutin coverage varied from a low of 175 μg cm^{-2} in *S. pennellii* to a maximum of 950 μg cm^{-2} in *S. pimpinellifolium*. There was no association between climate factors and cutin amount, further suggesting the model that wax, and not cutin, is the major determinant of cuticular water permeability in tomato fruit (Isaacson *et al.*, 2009; Leide *et al.*, 2007). In terms of composition, there were few differences between the species, with the exception of the abundance of 9,10,18-trihydroxy octadecanoic acid, which accounts for 6-10% of the cutin in *S. habrochaites, S. pennellii* and *S. chmielewskii* but less than 3% of cutin in the other species. This monomer is widely distributed in plant cutins, and is one of the representative monomers of C18-type cutins (Kolattukudy, 2001). Radioactive feeding experiments with *cis-9*-octadecenoic acid have indicated that biosynthesis of 9,10,18-trihydroxy octadecanoic acid in *Spinacia oleracea* likely occurs by ω-hydroxylation of *cis-9*-octadecenoic acid followed by introduction of an epoxy group at the double bond by a cytochrome P450 epoxidase (Croteau and Kolattukudy, 1975a; Kolattukudy *et al.*, 1973). The resulting compound, 9,10-epoxy-18-hydroxyoctadecanoic acid is itself a common monomer of some cutin and suberin polymers, and for *Malus pumila* it was shown that opening of the epoxide ring to yield 9,10,18-trihydroxy octadecanoic acid likely occurs through the action of epoxide hydrolase enzymes (Croteau and Kolattukudy, 1975b). It can be assumed that the same enzymatic

reactions lead to formation of the trihydroxy fatty acid monomer found in tomato fruit cutin (Figure 2.9).

The genes encoding several of the cytochrome P450s involved in cutin biosynthesis have been identified by studying arabidopsi*s* mutants. The fatty acid ω-hydroxylase activity is encoded by genes of the CYP86 family (Pinot and Beisson, 2011), while a mid-chain hydroxylase is encoded by CYP77A6 (Li-Beisson *et al.*, 2009). Biochemical characterization of another member of the CYP77 family, CYP77A4, showed that, in addition to possessing a mid-chain hydroxylase activity, the enzyme also has epoxidase activity on unsaturated substrates (Sauveplane *et al.*, 2009).

Taken together, the current model for the synthesis of C18 mid-chain epoxy and trihydroxy cutin monomers involves CYP86 catalyzing ω-hydroxylation of an unsaturated substrate followed by epoxidation catalyzed by CYP77. In the species that we examined here, the increased prevalence of this pathway (as opposed to C16 monomer biosynthesis) could potentially occur through an increased C18 desaturase activity, or altered substrate specificity of the CYP86 ω-hydroxylase (Figure 2.9). Future work identifying the genetic basis of this shift to C18 monomers could be interesting as they have a greater potential for forming dendritic structures that may alter the cutin polymeric structure (Pollard *et al.*, 2008).

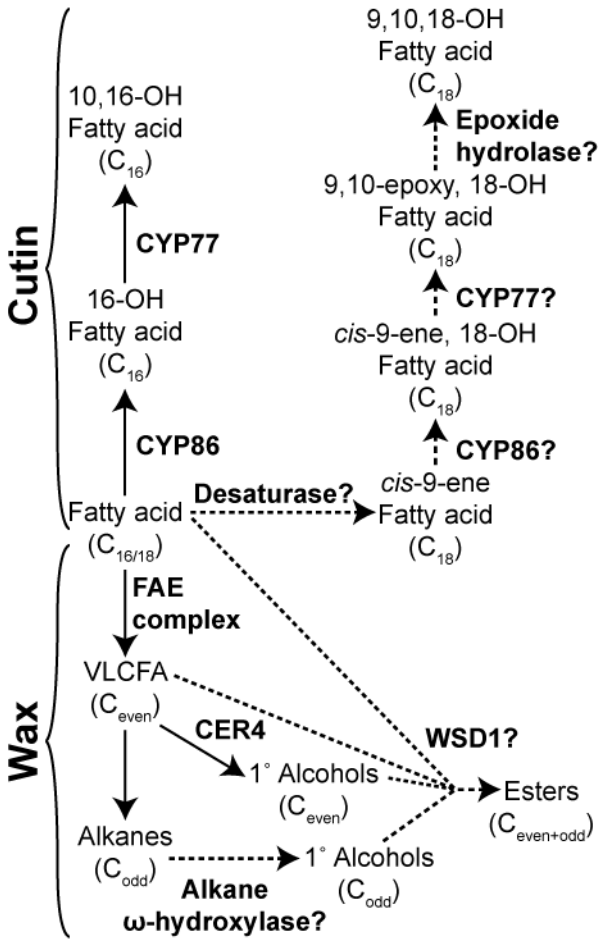


Figure 2.9. Biosynthesis of acyl lipids found in the cuticles of *Solanum spp.* Proposed synthesis pathways of lipid classes not found in *S. lycopersicum* are shown with dashed lines. For simplicity, coenzyme A is omitted from the diagram. FAE, fatty acid elongase; VLCFA, very long chain fatty acid.

Wax

The most striking observation regarding the cuticular wax of the species examined was the extremely high abundance of wax in *S. neorickii*, which was seven-fold higher than in *S. lycopersicum* (Figure 2.6a). It is tempting to draw the intuitive correlation between this dramatically increased wax accumulation and adaptation to an arid environment. However, *S. neorickii* is endemic to cooler and moister environments than the other species studied, with the exception of the closely related *S. chmielewskii* (Figure 2.1c). On the other hand, desiccating conditions may occur on smaller geographical scales and can depend on factors such as soil drainage, wind and altitude that are not accounted for when only considering average precipitation and temperature. Future studies looking at co-occurrence of wax accumulation and cuticular water permeability QTLs could provide new insight into the functional significance of increased wax accumulation. There is emerging evidence of genetic programs for enhanced wax production during drought stress (Seo *et al.*, 2011; Kosma *et al.*, 2009) and constitutive activation of these pathways could be an adaptive strategy for tolerating persistent water stress.

We observed a correlation between decreasing prevalence of triterpenoids and phylogenetic distance from *S. lycopersicum*, with *S. pennellii* exhibiting no detectable triterpenoids (Figure 2.6). Previous studies of tomato mutants have indicated that triterpenoids do not contribute to the water barrier properties of a cuticle (Vogg *et al.*, 2004; Leide *et al.*, 2007), suggesting that aliphatic, rather than triterpenoid, compounds would be favored in species adapted to arid environments. With the exception of *S. pimpinellifolium*, which is endemic to one of the warmer and drier

environments, our results might suggest a role for aliphatic compounds in dry adaptation for some species (Figures 2.1 and 2.6), but this has yet to be tested experimentally.

Identification of an unusual wax ester profile and potentially novel biosynthetic pathway

The final wax-related trend was the appearance of wax esters in three of the species, *S. chmielewskii, S. neorickii* and *S. habrochaites*. These are typical components of cuticular wax from many species, but they are not observed in wild type *S. lycopersicum* (Bauer *et al.*, 2004a), although the occurrence of a range of wax esters accompanied by depletion of alkanes was reported in the *positional sterile* (*ps*) tomato mutant (Leide *et al.*, 2011). The wax esters observed here were particularly remarkable for the abundance of odd chain lengths. Analysis of arabidopsis mutants suggests that the wax ester biosynthesis pathway depends on two key enzymes. The first, CER4, is a fatty-acid CoA reductase that produces primary alcohols by reduction of CoA esters of VLCFAs. The second is WSD1, a protein of the WS/DGAT family that catalyzes the synthesis of wax esters from fatty acid CoA esters and primary alcohols. Since fatty acid elongation occurs by addition of two-carbon units and synthesis of alkanes occurs by loss of one carbon, the typical chain length distribution of acids and primary alcohols is even, while alkanes are predominantly odd in chain length (Figure 2.9; Samuels *et al.*, 2009). Following this pattern, as they are typically synthesized from fatty acids and primary alcohols, esters are typically also even in chain length. The abundance of odd-numbered esters observed here, despite the typical

predominance of odd alkanes (Figure 2.6b,c), suggests an alternative biosynthetic route leading to odd-numbered esters. Analysis of the methanolyzed esters showed that the component responsible for the predominance of odd chain lengths was the alcohol moiety (Figure 2.7). Based on the existing model of wax ester biosynthesis, a relatively simple pathway that can be envisioned to explain the chain length distribution that we observed is depicted in Figure 2.9. First, synthesis of the even-chain length esters occurs through VLCFA reduction, catalyzed by an ortholog of CER4, followed by acyl transfer catalyzed by an ortholog of WSD1. This series of reactions is analogous to the known pathway in arabidopsis (Kunst and Samuels, 2009). The accumulation of esters in the *ps* mutant of *S. lycopersicum* (Leide *et al.*, 2011), and the occurrence of small amounts of primary alcohols in most of the species we examined (Figure 2.6a), suggests that these activities are present and that enhanced accumulation of their products may simply be a matter of variable expression of these enzymes.

The pathway leading to the odd chain length esters would be the same as that outlined above, except that the odd-chain primary alcohols would be synthesized by an alkane ω-hydroxylase. While hydroxylation of alkanes to yield secondary alcohols with an odd number of carbons is known to be catalyzed by CYP96A15 in arabidopsis (Greer *et al.*, 2007), an enzyme catalyzing ω-hydroxylation of alkanes is not yet known. The *Solanum* species provide a potentially excellent system to identify such an activity.

Genetic mapping of cuticle traits and future prospects

As a demonstration of the feasibility of mapping the genes underlying cuticular diversity in *Solanum* Sect. *Lycopersicon*, we selected a qualitative trait that clearly distinguishes *S. habrochaites* cuticles from those of *S. lycopersicum*. Although generally having a lower abundance of triterpenoids, *S. habrochaites* completely lacked all triterpenoids other than β-amyrin (Figure 2.6a,d). In the ILs that we analyzed, only LA3917 has the phenotype of the wild parent (Figure 2.8a). Since the introgressed genomic segments of this IL include the region on chromosome 12 encoding *SlTTS1* and *SlTTS2*, the two triterpene synthases responsible for wax triterpenoid biosynthesis in *S. lycopersicum*, these are promising candidates for genes underlying this aspect of cuticle diversity (Figure 2.8b). A simple loss of function in *TTS2*, leaving only *TTS1* functioning, would result in accumulation of solely β-amyrin (Figure 2.8c). Alternatively, since both *SlTTS1* and *SlTTS2* are very closely related, *SlTTS2* may be undergoing neofunctionalization in terms of product specificity following a tandem duplication, and *S. habrochaites TTS1* and *TTS2* may represent the ancestral state of a β-amyrin specific triterpene synthases. This second hypothesis is supported by the trend towards increasing complexity and abundance of triterpenoid isomers with species that are more closely related to *S. lycopersicum* (Figure 2.6d). Resolution of this question would require biochemical characterization of *S. habrochaites TTS1* and *TTS2* and will be the target of future studies.

As previously mentioned, aside from evolutionary adaptation, environmental growth conditions can have a substantial effect on cuticle properties. Although all plants grown for the majority of the experiments in this study were grown

simultaneously in the same conditions, we noticed a substantial difference in both the amount of cutin and the thickness of the cuticle of *S. lycopersicum* cv. M82, compared with our previous studies (Isaacson *et al.*, 2009), despite consistency in wax amount (Figures 2.2h, 2.5 and 2.6a). Similarly, although cuticular thickness is difficult to determine using SEM, the cuticles of plants grown in greenhouses in Israel for these experiments appear to be more substantial (Figure 2.4). We note that in both of these cases, these plants were grown in greenhouses with less sophisticated temperature control and may have been more stressed than the plants grown for light microscopy and chemical analysis in the present study. It is also worth noting that the substantial change in cuticle thickness and amount is similar to that observed with water-stressed arabidopsis plants (Kosma *et al.*, 2009). In *S. lycopersicum*, water stress has been shown to induce leaf wax biosynthesis and decrease cuticular transpiration, although cutin and fruit were not analyzed (Xu *et al.*, 1995). Taken together, it seems likely that there is an environmental component to induction of cuticle of biosynthesis in *Solanum* spp. that remains to be investigated. These results also suggest that future cuticle studies of wild species-derived QTLs must be carefully designed to control for environmental influence.

An additional consideration is the diversity that may exist between accessions of the same species. There are more than 1,200 catalogued accessions of *Solanum* Sect. *Lycopersicon* (Moyle, 2008, http://trgr.ucdavis.edu) and their distribution into various subspecies and species is still being defined (Peralta *et al.*, 2008). In this study we focused on a single accession of each species, but there is likely value in

investigating variation at the accession level, just as considerable cuticular diversity has been noted between cultivars of *S. lycopersicum* (Bauer *et al.*, 2004b).

Here we have shown that there is substantial diversity in the structure and chemical composition of the fruit cuticles of wild tomato species. Furthermore, despite the diversity, we have shown that the traits identified may be relatively simple in their genetic bases. *Solanum* Sect. *Lycopersicon* thus hits the 'sweet spot' of balancing recent evolutionary divergence with diversity that is required for future EEFG studies. Future genetic experiments aimed at correlating cuticle or fruit traits, such as water loss and pathogen resistance, with structural and chemical characteristics, will provide a means of untangling the complex interaction between cuticle structure, function and evolution.

Experimental procedures

Plant material

Seeds for all wild species were obtained from the Tomato Genetics Resource Center (UC Davis, http://tgrc.ucdavis.edu). For *S. lycopersicum,* the M82 cultivar was used (LA3475). The wild species accessions used were: *S. pimpinellifolium,* LA1589; *S. cheesmaniae,* LA0166; *S. chmielewskii,* LA1028; *S. neorickii,* LA2133; *S. habrochaites,* LA0407 and *S. pennellii*, LA0716. All species used for light microscopy, confocal microscopy and chemical analysis were grown in the same greenhouse in Ithaca, NY under standard conditions. For SEM experiments, plants were grown in the greenhouse in Bet Dagan, Israel. The ILs were grown in the summer of 2009 in the field in Freeville, NY. Since the wild parent of the *S.*

habrochaites ILs is the accession LA1777 (Monforte and Tanksley, 2000), this accession was also grown in the greenhouse for chemical analysis, which revealed that that there were no variations in cuticular triterpenoids between accessions (data not shown).

Microscopy

For light microscopy, tomato fruit pericarp tissue was fixed, embedded, cryosectioned, and post-fixed as described in (Buda *et al.*, 2009). Cryosections (4 µm) were melted on to VistaVision Histobond slides (VWR, www.vwr.com), dried and stained with Oil Red O (saturated in 60% isopropanol). Stain preparation and schedule was as described in Fukumoto and Fujimoto (2002): sections were rinsed with 50% isopropanol, followed by water, and mounted in water. Images were obtained using differential interference contrast (DIC) optics on an AxioImager A1 microscope equipped with an EC-Plan NeoFluar 40x/0.75 objective and an AxioCam Mrc color video camera (Zeiss, www.zeiss.com). For cuticle thickness measurements, sections were stained with Sudan IV (Buda *et al.*, 2009) and thickness was determined for each biological replicate, taking the average of 12 measurements (Figure 2.2). The average and standard error of the five biological replicates (fruits) is reported. For confocal microscopy and tomography, cryosections (30 µm) of fruit pericarp tissues were obtained and post-fixed as above, stained for 1 hour in Auramine O (0.1% w/v in 0.05 M Tris/HCl, pH 6.8) and mounted in water. Confocal microscopy and z-stack collection was performed as outlined in (Buda *et al.*, 2009). Z-stacks were pre-processed using LAS-AF v 1.8.2 (Leica, www.leica-microsystems.com) and Image J

(http://rsbweb.nih.gov) software. Cuticle image stacks were then assembled into 3D volume renderings using OsiriX software (Rosset *et al.*, 2004). SEM microscopy was performed as described in Hovav *et al.*, (2007).

Cuticle isolation

Mature fruits were harvested and three orthogonal diameters of each fruit were determined using calipers. The average of the three diameters was used to calculate the surface area, assuming a perfect sphere of this average diameter. For each biological replicate, 5-10 fruits were combined and manually dissected to remove the seeds and most of the pericarp. Cuticles were then isolated by the enzymatic method as previously described (Isaacson *et al.*, 2009).

Wax chemical analysis

Dried cuticles were spiked with tetracosane as an internal standard and extracted three times with a small volume of chloroform. The extracts were pooled and an aliquot was dried by heating under a gentle stream of nitrogen. The wax mixture was derivatized and subjected to GC analysis as previously described (Wang *et al.*, 2011) with the following exceptions: GC-FID quantitative analysis was performed on a model 6850 gas chromatograph (Agilent, www.agilent.com), while GC-MS qualitative analysis was as previously described. Since wax extracts from enzymatically isolated cuticles typically contain C16 and C18 fatty acids and other lipophilic contaminants absorbed from cellular debris during cuticle isolation

(Schönherr and Riederer, 1986), quantification only considered peaks eluting after the C27 *n*-alkane (~14 min).

To address the composition of wax esters observed in some species, the *S. neorickii* wax extract was fractionated by K6 silica gel TLC (Whatman, www.whatman.com) developed with chloroform. The ester band was recovered and eluted with chloroform. Methanolysis with sodium methoxide and subsequent derivatization with BSTFA/pyridine yielded a mixture of fatty alcohol TMS ethers and fatty acid methyl esters, which were identified by GC-MS and quantified by GC-FID as previously described.

Cutin chemical analysis

A protocol based on Bonaventure *et al.* (2004) was used for cutin analysis with some modifications. The dewaxed cuticular membranes were dried overnight and placed in a glass vial along with ω-pentadecalactone and methyl heptadecanoate as internal standards. A 10 mL reaction mixture, consisting of 6 mL methanol, 1.5 mL methyl acetate and 2.5 mL of 30% sodium methoxide in methanol, was added to each sample. The vials were capped and heated to 60˚C for 2 hours. After cooling to room temperature, 20 mL of diethyl ether and 2.5 mL of glacial acetic acid were added. To this, 5 mL of aqueous buffer (0.9% NaCl, 100 mM Tris pH 8.0) was added and the tubes were mixed thoroughly by vortexing. The phases were separated by centrifugation (2 min at 1500 x g) and an aliquot of the upper organic phase was removed. This was combined with an equal volume of 0.9% NaCl, vortexed and centrifuged again. An aliquot of the organic upper phase was taken and transferred to a

conical reaction vial. An equal volume of 2,2-dimethoxypropane was added to dry the sample and the vial was capped and incubated at 50˚C for 15 minutes. The solvent was evaporated by heating under a gentle stream of nitrogen before derivatization of the sample with 10 µL BSTFA and 10 µL pyridine. The reaction was heated for 10 minutes at 90˚C and dried again under nitrogen. The sample, resuspended in chloroform, was then subjected to GC-FID analysis as described for wax analysis. Compounds were identified by running the same samples on GC-MS and comparison to reference spectra from Holloway (1982).

Acknowledgments

We thank Jonathan Fuller and Samuel Mullin for technical assistance and Dr. J. Giovannoni for providing fruit from the *S. habrochaites* ILs. We thank Dr. S. Knapp for helpful comments and discussion. Confocal imaging was performed at the Plant Cell Imaging Center at the Boyce Thompson Institute which is supported by grants from the National Science Foundation (DBI-0618969) and the TRIAD foundation. This work was supported by grants from the National Science Foundation (Plant Genome Program; DBI-0606595), the United States-Israel Binational Agricultural Research and Development Fund (IS-4234-09), U.S.-Israel Binational Science Foundation (2005-168), a CUAES-Hatch grant (NYC-184462), the Canada Research Chairs Program and a Special Research Opportunities Grant from the Natural Sciences and Engineering Council of Canada (327937-2005). T.H.Y. was partly supported by a National Institutes of Health chemistry/biology interface training grant (T32 GM008500).

References

Baker, E.A., Bukovac, M.J. and Hunt, G.M. (1982) Composition of tomato fruit cuticle as related to fruit growth and development. In *The Plant Cuticle* (Cutler, D.F., Alvin, K.L. and Price, C.E., eds). London: Academic Press, pp. 33-44.

Barthlott, W. and Neinhuis, C. (1997) Purity of the sacred lotus, or escape from contamination in biological surfaces. *Planta* **202**, 1–8.

Bauer, S., Schulte, E. and Thier, H. (2004a) Composition of the surface wax from tomatoes: I. Identification of the components by GC/MS. *Eur. Food Res. Technol.* **219**, 223–228.

Bauer, S., Schulte, E. and Thier, H. (2004b) Composition of the surface wax from tomatoes: II. Quantification of the components at the ripe red stage and during ripening. *Eur. Food Res. Technol.* **219**, 487–491.

Bonaventure, G., Beisson, F., Ohlrogge, J. and Pollard, M. (2004) Analysis of the aliphatic monomer composition of polyesters associated with Arabidopsis epidermis: occurrence of octadeca-*cis*-6, *cis*-9-diene-1,18-dioate as the major component. *Plant J.* **40**, 920–930.

Buda, G.J., Isaacson, T., Matas, A.J., Paolillo, D.J. and Rose, J.K. (2009) Three-dimensional imaging of plant cuticle architecture using confocal scanning laser microscopy. *Plant J.* **60**, 378–385.

Buschhaus, C. and Jetter, R. (2011) Composition differences between epicuticular and intracuticular wax substructures: how do plants seal their epidermal surfaces? *J. Exp. Bot.* **62**, 841–853.

Croteau, R. and Kolattukudy, P.E. (1975a) Biosynthesis of hydroxyfatty acid polymers. Enzymatic epoxidation of 18-hydroxyoleic acid to 18-hydroxy-*cis*-9,10-epoxystearic acid by a particulate preparation from spinach (*Spinacia oleracea*). *Arch. Biochem. Biophys.* **170**, 61–72.

Croteau, R. and Kolattukudy, P.E. (1975b) Biosynthesis of hydroxyfatty acid polymers. Enzymatic hydration of 18-hydroxy-*cis*-9,10-epoxystearic acid to *threo*-9,10,18-trihydroxystearic acid by a particulate preparation from apple (*Malus pumila*). *Arch. Biochem. Biophys.* **170**, 73–81.

Dillon, M.O. (1989) Origins and diversity of the lomas formations in the Atacama and Peruvian Deserts of western South America. *Am. J. Bot.* **76** (Abstract), 212.

Edwards, D. (1993) Cells and tissues in the vegetative sporophytes of early land plants. *New Phytol.* **125**, 225–247.

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**, 1147–1162.

Fernandez, V., Khayet, M., Montero-Prado, P. *et al.* (2011) New insights into the properties of pubescent surfaces: peach fruit as a model. *Plant Physiol.* **156**, 2098–2108.

Franke, R., Briesen, I., Wojciechowski, T., Faust, A., Yephremov, A., Nawrath, C. and Schreiber, L. (2005) Apoplastic polyesters in *Arabidopsis* surface tissues–a typical suberin and a particular cutin. *Phytochem.* **66**, 2643–2658.

Fukumoto, S. and Fujimoto, T. (2002) Deformation of lipid droplets in fixed samples. *Histochem. Cell Biol.* **118**, 423–428.

Graca, J., Schreiber, L., Rodrigues, J. and Pereira, H. (2002) Glycerol and glyceryl esters of omega-hydroxyacids in cutins. *Phytochem.* **61**, 205–215.

Greer, S., Wen, M., Bird, D., Wu, X., Samuels, L., Kunst, L. and Jetter, R. (2007) The cytochrome P450 enzyme CYP96A15 is the midchain alkane hydroxylase responsible for formation of secondary alcohols and ketones in stem cuticular wax of Arabidopsis. *Plant Physiol.* **145**, 653–667.

Holloway, P.J. (1982) The chemical constiution of plant cutins. In *The Plant Cuticle* (Cutler, D.F., Alvin, K.L. and Price, C.E., eds). London: Academic Press, pp. 45-85.

Hovav, R., Chehanovsky, N., Moy, M., Jetter, R. and Schaffer, A.A. (2007) The identification of a gene (*Cwp1*), silenced during *Solanum* evolution, which causes cuticle microfissuring and dehydration when expressed in tomato fruit. *Plant J.* **52**, 627–639.

Isaacson, T., Kosma, D.K., Matas, A.J. *et al.* (2009) Cutin deficiency in the tomato fruit cuticle consistently affects resistance to microbial infection and biomechanical properties, but not transpirational water loss. *Plant J.* **60**, 363–377.

Javelle, M., Vernoud, V., Rogowsky, P.M. and Ingram, G.C. (2011) Epidermis: the formation and functions of a fundamental plant tissue. *New Phytol.* **189**, 17–39.

Jeffree, C.E. (2006) The fine structure of the plant cuticle. In *Biology of the Plant Cuticle* (Riederer, M. and Müller, C., eds). Oxford, UK: Blackwell, pp. 11-125.

Jenks, M.A. and Ashworth, E.N. (1999) Plant epicuticular waxes: function, production and genetics. *Hortic. Rev.* **23**, 1–68.

Jetter, R., Kunst, L. and Samuels, A.L. (2006) Composition of plant cuticular waxes. In *Biology of the Plant Cuticle* (Riederer, M. and Müller, C., eds). Oxford, UK: Blackwell, pp. 145-181.

Kolattukudy, P.E. (2001) Polyesters in higher plants. *Adv. Biochem. Eng. Biotechnol.* **71**, 1–49.

Kolattukudy, P.E., Walton, T.J. and Kushwaha, R.P. (1973) Biosynthesis of the C18 family of cutin acids: omega-hydroxyoleic acid, omega-hydroxy-9,10-epoxystearic acid, 9,10,18-trihydroxystearic acid, and their delta12-unsaturated analogs. *Biochemistry* **12**, 4488–4498.

Kosma, D.K., Bourdenx, B., Bernard, A., Parsons, E.P., Lu, S., Joubes, J. and Jenks, M.A. (2009) The impact of water deficiency on leaf cuticle lipids of Arabidopsis. *Plant Physiol.* **151**, 1918–1929.

Kunst, L. and Samuels, L. (2009) Plant cuticles shine: advances in wax biosynthesis and export. *Curr. Opin. Plant Biol.* **12**, 721–727.

Leide, J., Hildebrandt, U., Reussing, K., Riederer, M. and Vogg, G. (2007) The developmental pattern of tomato fruit wax accumulation and its impact on cuticular transpiration barrier properties: Effects of a deficiency in a beta-ketoacyl-coenzyme A synthase (LeCER6). *Plant Physiol.* **144**, 1667–1679.

Leide, J., Hildebrandt, U., Vogg, G. and Riederer, M. (2011) The *positional sterile* (*ps*) mutation affects cuticular transpiration and wax biosynthesis of tomato fruits. *J Plant Physiol.* **168**, 871–877.

Li-Beisson, Y., Pollard, M., Sauveplane, V., Pinot, F., Ohlrogge, J. and Beisson, F. (2009) Nanoridges that characterize the surface morphology of flowers require the synthesis of cutin polyester. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22008–22013.

Mitchell-Olds, T., Feder, M. and Wray, G. (2008) Evolutionary and ecological functional genomics. *Heredity* **100**, 101–102.

Monforte, A.J. and Tanksley, S.D. (2000) Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L. esculentum* genetic background: a tool for gene mapping and gene discovery. *Genome* **43**, 803–813.

Moyle, L.C. (2008) Ecological and evolutionary genomics in the wild tomatoes (*Solanum* sect. *Lycopersicon*). *Evolution* **62**, 2995–3013.

Mueller, L.A., Solow, T.H., Taylor, N. *et al.* (2005) The SOL Genomics Network: a comparative resource for Solanaceae biology and beyond. *Plant Physiol.* **138**, 1310–1317.

Nakazato, T., Warren, D.L. and Moyle, L.C. (2010) Ecological and geographic modes of species divergence in wild tomatoes. *Am. J. Bot.* **97**, 680–693.

Niklas, K.J. (1980) Paleobiochemical techniques and their applications to paleobotany. In *Progress in Phytochemistry* Vol. 6 (Reinhold, L., Harborne, J.B. and Swain, T., eds). Oxford, UK: Pergamon Press, pp. 143-181.

Peralta, I.E., Spooner, D.M. and Knapp, S. (2008) Taxonomy of wild tomatoes and their relatives (*Solanum* sect. *Lycopersicoides*, sect. *Juglandifolia*, sect. *Lycopersicon*; Solanaceae). *Syst. Bot. Monog.* **84**, 1-186.

Pfündel, E.E., Agati, G. and Cerovic, Z.G. (2006) Optical properties of plant surfaces. In *Biology of the Plant Cuticle* (Riederer, M. and Müller, C., eds). Oxford, UK: Blackwell, pp. 216-249.

Pinot, F. and Beisson, F. (2011) Cytochrome P450 metabolizing fatty acids in plants: characterization and physiological roles. *FEBS J.* **278**, 195–205.

Pollard, M., Beisson, F., Li, Y. and Ohlrogge, J.B. (2008) Building lipid barriers: biosynthesis of cutin and suberin. *Trends Plant Sci.* **13**, 236–246.

Prudent, M., Causse, M., Genard, M., Tripodi, P., Grandillo, S. and Bertin, N. (2009) Genetic and physiological analysis of tomato fruit weight and composition: influence of carbon availability on QTL detection. *J. Exp. Bot.* **60**, 923–937.

Reina-Pinto, J.J. and Yephremov, A. (2009) Surface lipids and plant defenses. *Plant Physiol. Biochem.* **47**, 540–549.

Riederer, M. and Schreiber, L. (2001) Protecting against water loss: analysis of the barrier properties of plant cuticles. *J. Exp. Bot.* **52**, 2023–2032.

Rosset, A., Spadola, L. and Ratib, O. (2004) OsiriX: an open-source software for navigating in multidimensional DICOM images. *J. Digit. Imaging* **17**, 205–216.

Samuels, A.L., Kunst, L. and Jetter, R. (2008) Sealing plant surfaces: cuticular wax formation by epidermal cells. *Annu. Rev. Plant Biol.* **59**, 683-707.

Sauveplane, V., Kandel, S., Kastner, P.E., Ehlting, J., Compagnon, V., Werck-Reichhart, D. and Pinot, F. (2009) Arabidopsis thaliana CYP77A4 is the first cytochrome P450 able to catalyze the epoxidation of free fatty acids in plants. *FEBS J.* **276**, 719–735.

Schönherr, J. and Riederer, M. (1986) Plant cuticles sorb lipophilic compounds during enzymatic isolation. *Plant Cell Environ.* 9, 459–466.

Seo, P.J., Lee, S.B., Suh, M.C., Park, M.J., Go, Y.S. and Park, C.M. (2011) The MYB96 transcription factor regulates cuticular wax biosynthesis under drought conditions in Arabidopsis. *Plant Cell* **23**, 1138–1152.

Vogg, G., Fischer, S., Leide, J., Emmanuel, E., Jetter, R., Levy, A.A., Riederer, M. (2004) Tomato fruit cuticular waxes and their effects on transpiration barrier properties: functional characterization of a mutant deficient in a very-long-chain fatty acid β-ketoacyl-CoA synthase. *J. Exp. Bot.* **55**, 1401–1410.

Walton, T.J. (1990) Waxes, cutin and suberin. In *Methods in plant biochemistry: lipids, membranes and aspects of photobiology.* (Harwood, J.L. and Bowyer, J.R., eds). London: Academic Press, pp. 105-158.

Wang, Z., Guhling, O., Yao, R., Li, F., Yeats, T.H., Rose, J.K. and Jetter, R. (2011) Two oxidosqualene cyclases responsible for biosynthesis of tomato fruit cuticular triterpenoids. *Plant Physiol.* **155**, 540–552.

Xu, H.L., Gauthier, L. and Gosselin, A. (1995) Stomatal and cuticular transpiration of greenhouse tomato plants in response to high solution electrical-conductivity and low soil-water content. *J. Am. Soc. Hortic. Sci.* **120**, 417–422.

CHAPTER 3

Mining the surface proteome of tomato (*Solanum lycopersicum*) fruit for proteins associated with cuticle biogenesis*

Trevor H. Yeats[1], Kevin J. Howe[2], Antonio J. Matas[1], Gregory J. Buda[1], Theodore W. Thannhauser[2] and Jocelyn K. C. Rose[1]

[1] Department of Plant Biology, Cornell University, Ithaca, NY 14853, USA.

[2] USDA-ARS, Robert W. Holley Center for Agriculture and Health, Ithaca, NY 14853, USA

* A version of this chapter has been published as:
Yeats, T.H., Howe, K.J., Matas, A.J., Buda, G.J., Thannhauser, T.W. and Rose, J.K. (2010) Mining the surface proteome of tomato (*Solanum lycopersicum*) fruit for proteins associated with cuticle biogenesis. *J. Exp. Bot.* **61**, 3759–3771. (Published by the Society of Experimental Biology and Oxford University Press)

Abstract

The aerial organs of plants are covered by the cuticle; a polyester matrix of cutin and organic-solvent soluble waxes that is contiguous with the polysaccharide cell wall of the epidermis. The cuticle is an important surface barrier between a plant and its environment, providing protection against desiccation, disease and pests. However, many aspects of the mechanisms of cuticle biosynthesis, assembly and restructuring are entirely unknown. To identify candidate proteins with a role in cuticle biogenesis, a surface protein extract was obtained from tomato (*Solanum lycopersicum*) fruits by dipping in an organic solvent and the constituent proteins were identified by several complementary fractionation strategies and two mass spectrometry techniques. Of the approximately 200 proteins that were identified, a subset is potentially involved in the transport, deposition, or modification of the cuticle, such as those with predicted lipid-associated protein domains. These include several lipid-transfer proteins, GDSL-motif lipase/hydrolase family proteins and an MD-2 related lipid recognition domain-containing protein. The epidermal-specific transcript accumulation of several of these candidates was confirmed by laser-capture microdissection and qRT-PCR, together with their expression during various stages of fruit development. This indicated a complex pattern of cuticle deposition and models for cuticle biogenesis and restructuring are discussed.

Introduction

The plant cuticle is a hydrophobic membrane that covers the aerial organs of land plants and provides protection against desiccation, pathogens, UV radiation and herbivory (Riederer, 2006). It is continuous with the outer-periclinal polysaccharide cell wall of the epidermis and consists of organic soluble waxes embedded in, and layered on, a non-soluble polyester matrix of ω-substituted fatty acids. The waxes include both aliphatic compounds, derived from very long chain fatty acids, and secondary metabolites, such as triterpenoids and flavonoids (Jetter *et al.*, 2006). In the majority of species analyzed to date, cutin is composed primarily of polymerized mid-chain substituted ω-hydroxy fatty acids, although arabidopsis (*Arabidopsis thaliana*) is a notable exception in that α,ω- dicarboxylic fatty acids predominate in stems and leaves (Bonaventure *et al.,* 2004; Franke *et al.,* 2005) and ω-hydroxy fatty acids only contribute significantly to the cutin of flowers (Beisson *et al.* 2007; Li-Beisson *et al.* 2009; Panikashvili *et al.*, 2009). In addition, the presence of glycerol in the cutin polymer is now well established (Graca *et al.*, 2002).

Both cutin monomers and waxes are produced within the epidermal cells and a clear picture of the molecular biology of their synthesis is emerging. This has been largely a result of the characterization of arabidopsis mutants (Pollard *et al.*, 2008; Samuels *et al.*, 2008), a species whose cutin is likely rich in glycerol (Pollard *et al.*, 2008). For example, the glycerol-3-phosphate acyltransferases GPAT4, GPAT6 and GPAT8 have been shown to be required for cutin synthesis (Li *et al.*, 2007; Li-Beisson *et al.,* 2009), and it was recently reported that an acyltransferase of the BAHD family, DCR, is required for cutin synthesis in arabidopsis floral organs (Panikashvili *et al.*,

2009). Both classes of enzymes appear to be intracellular: GPAT8 was localized to the endoplasmic reticulum (ER) (Gidda *et al.,* 2009) and DCR was shown to be in the cytoplasm (Panikashvili *et al.*, 2009). However, the subsequent extracellular aspects of cuticle biogenesis, including trafficking of the constituents and their assembly into a mature cuticle, as well as restructuring of cuticle architecture during growth and development, are far less well understood.

Current models hypothesize the involvement of several classes of extracellular proteins and enzymes, although few examples have yet been identified. Following biosynthesis in the endoplasmic reticulum (ER), wax and cutin monomers, or oligomers, are exported across the plasma membrane to the apoplast, in a process dependent on ABC transporters such as the arabidopsis proteins CER5 (Pighin *et al.*, 2004) and WBC11 (Bird *et al.*, 2007). Recently, a GPI-anchored lipid-transfer protein, LTPG, was shown to be required for wax secretion, possibly by acting as a membrane-anchored lipid binding protein that receives waxes as they are extruded by ABC transporters (Debono *et al.*, 2009). Trafficking of hydrophobic lipids across the polar environment of the polysaccharide cell wall is then often attributed to soluble extracellular lipid-transfer proteins (LTPs). However, their ability to bind wax or cutin monomers has not been confirmed and no cuticle mutant has been attributed to a lesion in a gene encoding a soluble LTP (Yeats and Rose, 2008).

Polymerization of the cutin polymer during development and organ expansion may also involve extracellular proteins. The protein BODYGUARD (BDG) is secreted by epidermal cells and is required for normal cuticle development in arabidopsis, although the *bdg* mutant paradoxically accumulates a larger amount of cutin

(Kurdyukov *et al.*, 2006a). While no biochemical activity for BDG has been identified, the protein is a member of the alpha/beta-hydrolase superfamily, leading the authors to suggest that it is a putative cutin synthase. A similar function has been proposed for AgaSGNH, a GDSL-motif lipase/hydrolase family protein from *Agave americana*, which was reported to have protein localization and gene expression patterns that correlated with cutin biosynthesis, although it was not associated with a genetic phenotype or biochemical activity (Reina *et al.*, 2007).

Thus, remarkably little is yet known about key mechanisms of cuticle biogenesis, and experimental strategies to identify new proteins that associate with cutin and waxes could provide a valuable means to identify new candidates. Cuticular waxes are easily extracted free of cellular lipid contamination by brief immersion of plant organs in organic solvents such as chloroform (Jetter *et al.*, 2006), while a small additional fraction of the recovered material is comprised of proteins (Martin and Juniper, 1970). Edman degradation peptide sequencing has previously been used to identify three proteins in plant cuticular waxes: an LTP from *Brassica oleracea* (Pyee *et al.*, 1994), and an endo-β-1,3-glucanase and a chitinase (glycosyl hydrolase family 17 and 18, respectively; www.cazy.org) from the wax of *Copernicia cerifera* (Cruz *et al.*, 2002). However, we hypothesized that generating a more comprehensive inventory of proteins that are associated with the outermost surface tissues of plant organs, using a range of complementary protein fractionation strategies coupled with modern sensitive mass spectrometry-based methods, would help identify new candidate proteins with a potential role in cuticle biosynthesis. To this end, we targeted the surface proteome of developing tomato (*Solanum lycopersicum*) fruit as a

model system. Although arabidopsis research has greatly accelerated the discovery of new cuticle-related genes, its cuticle poses some experimental limitations since it is relatively thin, fragile and difficult to isolate in substantial quantities. Conversely, tomato fruit cuticles are astomatous and large amounts of intact cuticular material can be isolated for chemical and biomechanical analyses. For example, the fruit accumulate on the order of 1 mg cm^{-2} cutin (Baker *et al.*, 1982), compared to the stem of Arabidopsis, which has 0.5-10 μg cm^{-2} (Franke *et al.*, 2005; Suh *et al.*, 2005). Thus, the typical 6 week period of tomato fruit development represents a remarkably rapid and extensive phase of cuticle biosynthesis, in a genetically tractable species for which there are now also many genomic resources (Mueller *et al.*, 2005; www.solgenomics.net).

We describe the proteomic analysis of tomato fruit cuticle extracts and the identification of several secreted proteins with lipid-related domains. The expression patterns of the genes encoding these proteins are further analyzed as to the specificity of their expression in the epidermis and during the time course of fruit development. Finally, based on these expression patterns and current models of cuticle biosynthesis, potential roles for these candidates in extracellular cutin and wax deposition and metabolism are discussed.

Materials and methods

Plant Materials

Solanum lycopersicum (cv. M82) plants were grown in the field (Freeville NY, summer 2007 and 2008) and 500 immature green fruits were harvested for protein extraction. To avoid bruising and damage during handling, fruits were harvested from all stages of expansion after the fruits had lost their visible trichomes and became glossy in appearance, at approximately 15-40 days post-anthesis (DPA). Prior to protein extraction, fruits were washed with deionized water and left to dry overnight. By first rinsing the fruits, we believe that our analysis excluded phylloplane proteins that are secreted to the outer surface of the cuticle by mechanisms discussed by Shepherd and Wagner (2007). Fruits used for confocal microscopy, laser-capture microdissection and developmental gene expression time course experiments were harvested from plants grown in the greenhouse (Ithaca, NY). To define the developmental stage of fruits during expansion, flowers were tagged at anthesis. The ripening stages were determined visually by color change according to standard conventions (Gonzalez-Bosch *et al.*, 1996). For RNA isolation, pericarp tissue from 3-10 fruits at each developmental stage was manually dissected, flash frozen, ground in liquid nitrogen and stored at -80˚C.

Microscopy

Confocal microscopy was performed as previously described (Buda *et al.*, 2009). To illustrate the different pericarp cell types harvested by laser-capture

microdissection, 10 μm paraffin sections of immature green fruits were prepared and stained with Toluidine blue O according to standard protocols (Ruzin, 1999).

Wax Extraction and Protein Isolation

Wax extraction and purification of polar components from the wax was conducted essentially as previously described (Pyee *et al.*, 1994). Fruits were dipped, without submerging the calyx scar, for 10 s in approximately 500 mL chloroform/methanol (2:1) that was gently stirred by a magnetic stir bar. For each set of extractions, 2-3 500 mL aliquots of fresh solvent were used and the extracts were pooled. The extract was then evaporated to dryness by rotary evaporation at 50˚C with reduced pressure. The residue was resuspended in 80 mL chloroform and 40 mL distilled water and transferred to a separatory funnel. The upper aqueous phase was recovered and lyophilized and the residue resuspended in 500 μL buffer (0.7 M sucrose, 0.1 M KCl, 0.5 M Tris-HCl pH 7.5, 50 mM ethylenediaminetetraacetic acid [EDTA], 2% β-mercaptoethanol and 1 mM phenylmethylsulphonyl fluoride). The protein component was then extracted into phenol and precipitated with 0.1 M ammonium acetate in methanol (Isaacson *et al.*, 2006). Calculation of approximate protein yield by densitometry of the gel-separated samples (see below) indicated that each extraction yielded approximately 8 μg of protein. Thus, assuming an average fruit surface area of 50 cm^2, the yield of protein was on the order of 0.3 ng cm^{-2} of surface. For comparison, the wax coverage of immature green tomato fruit is on the order of 5 μg cm^{-2} (15 000 fold greater).

Fractionation and Proteomic Analysis of Protein Extracts

Three independent extractions were analyzed using three different pre-fractionation schemes:

(1) Isolation of individual bands from 1D polyacrylamide gels. The pelleted protein extract was resuspended in 30 µL 1 × LDS sample buffer (Invitrogen, Carlsbad, CA) and separated on a 10% polyacrylamide gel (Novex 10% Bis-Tris Gel, 1.0 mm; Invitrogen) using MOPS running buffer, according to the manufacturer's instructions. The gel was fixed in 40% methanol/10% acetic acid and stained overnight with SYPRO Ruby (Invitrogen) according to the manufacturer's instructions. Gels were visualized with UV illumination and individual bands were excised (see Figure 3.2a) and frozen at -80˚C.

(2) Isolation of broad slabs from 1D polyacrylamide gels. Proteins were separated as above, except MES running buffer (Invitrogen) was used according the manufacturer's instructions. Slabs were excised (see Figure 3.2b), cut into small pieces and frozen at -80˚C.

(3) Gel-free in-solution trypsin digest. Precipitated proteins were resuspended in 100 µL of 50 mM ammonium bicarbonate, 6 M guanidinium chloride. To this, 5 µL of dithiothreitol (DTT) stock solution (200 mM DTT in 50 mM ammonium bicarbonate) was added and the mixture boiled for 10 min. Proteins were alkylated by addition of 4 µL of 1M iodoacetamide in 50 mM ammonium bicarbonate, followed by a 1 h room temperature incubation in the dark. To this, 40 µL of DTT stock was added and incubation was continued for an additional hour. The sample was then diluted by addition of 846 µL of 50 mM ammonium bicarbonate and digested by the addition of

5 μL of 200 ng μL^{-1} solution of sequencing grade trypsin (Promega, Madison, WI). The reaction was incubated overnight at 37˚C and then terminated by the addition of concentrated acetic acid to lower the pH below 6.0.

Analysis of the gel free extract was conducted by online liquid chromatography electrospray ionization tandem mass spectrometry (LC ESI-MS/MS), essentially as described by Yang *et al.* (2007). The sample was prefractionated by strong cation-exchange chromatography, eluting bound peptides in five fractions with a step gradient of 25 mM, 50 mM, 100 mM, 200 mM and 500 mM KCl. Each fraction was then analyzed by LC ESI-MS/MS as previously described. For the two gel-fractionated samples, in-gel trypsin digestion was performed as previously described (Shevchenko *et al.*, 1996), with modifications as described by Yang *et al.* (2007) and tryptic peptides were recovered with C18 ZipTips (Millipore, Bedford, MA), according to the manufacturer's directions. Peptides from each fraction were separated and analyzed by offline LC-MALDI TOF/TOF (liquid chromatography matrix assisted laser desportion ionization time of flight tandem mass spectrometry) analysis (Yang *et al.*, 2007).

Peak lists from the mass spectrometers were searched against the longest 6-frame translation of the Sol Genomics Network (SGN) Lycopersicum Combined unigene build from May 2009 (www.solgenomics.net) using MASCOT (Perkins *et al.*, 1999). For all experiments, the database was searched allowing for one missed cleavage, cysteine carboxyamidomethylation and variable methionine oxidation, requiring peptide scores corresponding to ≥95% confidence. For MALDI TOF/TOF experiments, a peptide mass tolerance of 10 ppm and fragment tolerance of 0.025 Da

was used. For ESI MS/MS experiments these tolerances were set to 1.5 and 0.6 Da, respectively. To limit the number of false positive results, the results were filtered by requiring that each identified protein be represented by at least two unique peptides in the same or multiple analyses.

Laser-capture microdissection, RNA amplification and cDNA synthesis

Tissue fixation and microdissection were performed based on the protocol of Nakazono *et al.* (2003). Pericarp tissue from 10 DPA immature green tomato fruits was manually dissected into 2 mm cubes using a razor and fixed by vacuum infiltration with 75% ethanol, 25% acetic acid. The ethanol/acetic acid was replaced with a fresh aliquot and the sample was left overnight at 4˚C. The fixative was decanted and replaced twice with a solution of 10% (w/v) sucrose in 100 mM phosphate buffered saline (PBS). Upon penetration of the solution into the tissue, as indicated by the tissue sinking, the solution was replaced twice more with a solution of 20% (w/v) sucrose in 100 mM PBS. The tissue was then embedded in TissueTek OCT medium (Sakura Finetek USA, Torrance, CA), frozen in a beaker submerged in a liquid nitrogen bath and the resulting cryoblocks stored at -80˚C until sectioning.

A Microm HM550 cryostat (ThermoFisher Scientific, Waltham, MA) was used to prepare 10 µm and 16 µm pericarp sections and the CryoJane tape-transfer system (Instrumedics, St. Louis, MO) was used to transfer sections to 0.5 × adhesive-coated slides, where they were adhered by UV-crosslinking. Slides were stored at -80˚C until later use. Immediately prior to laser-capture microdissection, slides were thawed and dehydrated as follows (all solvents at -20˚C): 1 min, 50% ethanol; 30 s, 95% ethanol;

1 min, 100% ethanol; 2 min, xylene; 2 min, fresh xylene. After air drying, cells were harvested into PALM adhesive cap tubes (Carl Zeiss, Oberkochen, Germany) using a PALM MicroBeam System (Carl Zeiss). Epidermal cells were captured from the 10 µm sections, while the larger, more vacuole-rich collenchyma cells were captured from the 16 µm sections. Total RNA was isolated from the harvested cells using an RNeasy Micro Kit (Qiagen, Valencia CA) and the mRNA amplified using the TargetAmp 2-Round aRNA Amplification Kit 2.0 (Epicentre Biotechnologies, Madison, WI), according to the manufacturers' instructions. 1.5 µg of amplified RNA was used for cDNA synthesis using SuperScript III reverse transcriptase and random hexamer primers (Invitrogen), according to the manufacturer's instructions.

RNA isolation and cDNA synthesis for developmental time course

RNA was isolated from frozen tissue (Schneiderbauer *et al.*, 1991) and 1.5 µg of total, DNAse treated RNA was used for cDNA synthesis using SuperScript II reverse transcriptase and oligo-dT primers (Invitrogen), according to the manufacturer's instructions.

Quantitative PCR

Quantitative PCR experiments were performed using an iQ5 system (BioRad, Hercules, CA). The cDNA samples were diluted 5 fold with water and 0.5 or 1 µL was used as a template for each 25 µL quantitative PCR reaction, prepared using HotStart-IT SYBR Green qPCR Master Mix (Affymetrix, Santa Clara, CA). For each gene, qPCR reactions were performed in technical triplicates. The sequences of

oligonucleotide primers are given in Supplementary Table S1 (Appendix 1). Specificity of the products was determined by gel electrophoresis, product sequencing and high resolution melt curve analysis. For tissue specificity, quantification was performed using REST 2008 software (Pfaffl *et al.*, 2002) with *RPL2* serving as a constitutive control, assuming PCR efficiency of 1.0 for all genes. For time course experiments, expression ratios for each gene and time point were calculated relative to *RPL2* expression. For each gene, expression was linearly normalized with a value of 0.0 assigned to the stage with lowest expression and 1.0 to the stage showing the highest expression.

Bioinformatics and Software

The area-proportional Venn diagram was constructed using BioVenn (Hulsen *et al.,* 2008; http://www.cmbi.ru.nl/cdd/biovenn/). Normalized gene-expression profile data was converted into a heat map using Cluster 3.0 (bonsai.ims.u-tokyo.ac.jp/~mdehoon/software/cluster/software.htm) and Java TreeView (Saldanha, 2004). Alignment of protein sequences was performed with Clustal W (Thompson *et al.*, 1994) and a neighbor-joining tree was constructed using MEGA4 (Tamura *et al.*, 2007). The alignment parameters and the settings for the phylogenetic reconstruction were the defaults of the MEGA4 package.

Results

Protein isolation and identification of candidate genes

As illustrated in Figure 3.1a-c, the fluorescently stained cuticle (Figure 3.1a) covers the surface of the tomato fruit but is separated from the epidermal cells by a subcuticular polysaccharide cell wall (Figure 1b). Previous studies have indicated that wax, rather than cutin, is the major barrier to the diffusion of polar molecules, including water (Leide *et al.*, 2007; Isaacson *et al.*, 2009), and presumably proteins, across the cuticle. We reasoned that, despite the relatively low abundance of wax in the tomato fruit cuticle, which is on the order of 5 μg cm^{-2} compared to 1 mg cm^{-2} for the cutin polymer (Baker *et al.*, 1982), a brief immersion of the fruits in an organic solvent would allow the isolation of proteins directly associated with cuticular wax, as well as those localized within the subcuticular epidermal cell wall and possibly epidermal intracellular proteins, depending on the degree to which the cells were compromised. We therefore used a standard protocol to remove waxes by immersion of intact plant organs in an organic solvent to obtain extracts for profiling of the fruit surface proteome, as was previously attempted on a smaller scale with *B. oleracea* leaves (Pyee *et al.*, 1994).

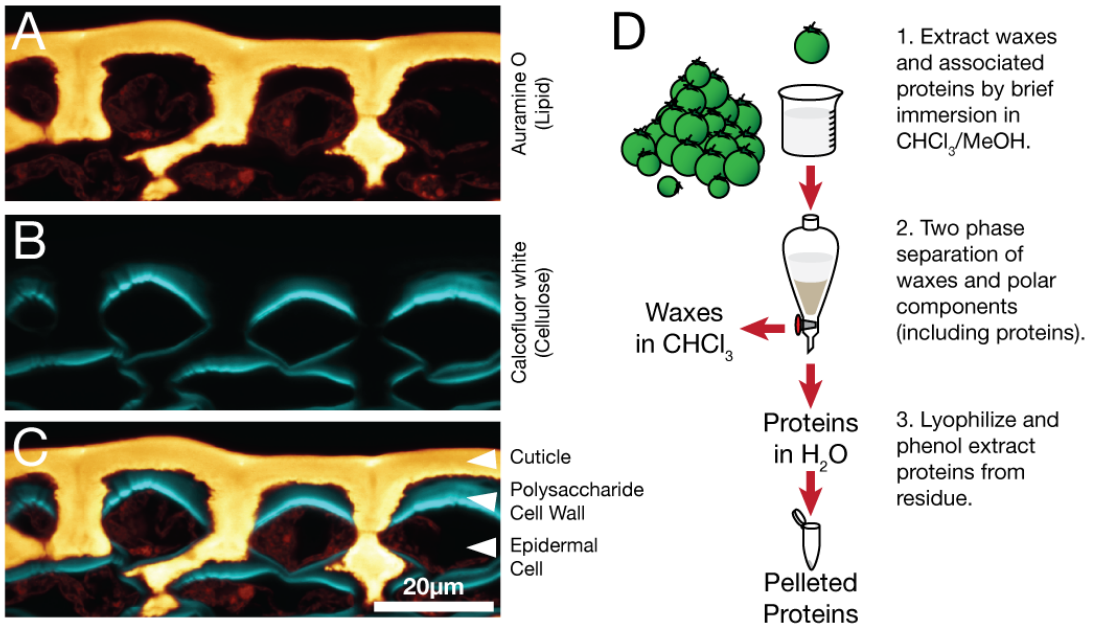


Figure 3.1. Epidermis structure and experimental design. Confocal microscopy of cryosectioned tomato breaker stage fruit epidermis, co-stained with the fluorescent lipid stain Auramine O (a) and the cellulose stain Calcofluor white M2R (b). The merged image (c) illustrates the cuticle and epidermal cell wall in the context of the epidermal cell layer. (d) Schematic representation of the extraction protocol used to isolate proteins from the cuticle and epidermal cell wall.

In order to more specifically target proteins that might be associated with cuticle biosynthesis, young, rapidly expanding tomato fruits were used in this study, since this represents the phase of most rapid cuticle deposition (Baker *et al.*, 1982; Mintz-Oron *et al.*, 2008). After extraction of cuticular waxes and other co-extracted components, the wax was separated from the more polar proteins by partitioning of polar constituents into an aqueous phase and wax into chloroform. The aqueous phase was then lyophilized and proteins were further purified from the residue by phenol

extraction and precipitation (Figure 3.1d).

In the initial analysis, the protein extract was separated by denaturing polyacrylamide gel electrophoresis (PAGE) and the 16 most distinct bands (Figure 3.2a) were excised and subjected to in-gel tryptic digestion, followed by offline LC-MALDI TOF/TOF analysis. This use of reverse phase liquid chromatography to separate tryptic peptides and robotic mixing of chromatographic fractions with a MALDI matrix (Bodnar *et al.*, 2003) combines the capacity for analyzing complex mixtures offered by on-line LC-ESI MS/MS analysis with the increased precision and reduced sensitivity to ion-suppression that is offered by MALDI TOF/TOF analysis (Yang *et al.*, 2007). Using this approach, a total of 44 different proteins were identified from the 16 bands following MASCOT searching of the mass spectra against a database of translated tomato unigene sequences (Supplementary Table S2, Appendix 1). Since an initial analysis using the spectra obtained from each band separately revealed some redundancy in the proteins identified in each band, as well as the presence of many proteins in each band (data not shown), the spectra from all bands were combined for this search.

Using a second experimental strategy and a new protein isolate, proteins were prefractionated by PAGE, but rather than cutting distinct bands, 10 contiguous gel slabs were excised and subjected to in-gel trypsin digestions (Figure 3.2b). We note that the banding pattern did not closely resemble that seen in the first analysis (Figure 3.2a). This likely reflects the fact that a different buffer system was used (MES), which favors the resolution of smaller proteins at the expense of larger proteins, or that the proteins may be subjected to varying degrees of post-extraction proteolysis. When

spectra from these 10 slabs were combined and a MASCOT search of the tomato predicted protein database was performed, a total of 25 proteins were identified (Supplementary Table S3, Appendix 1).

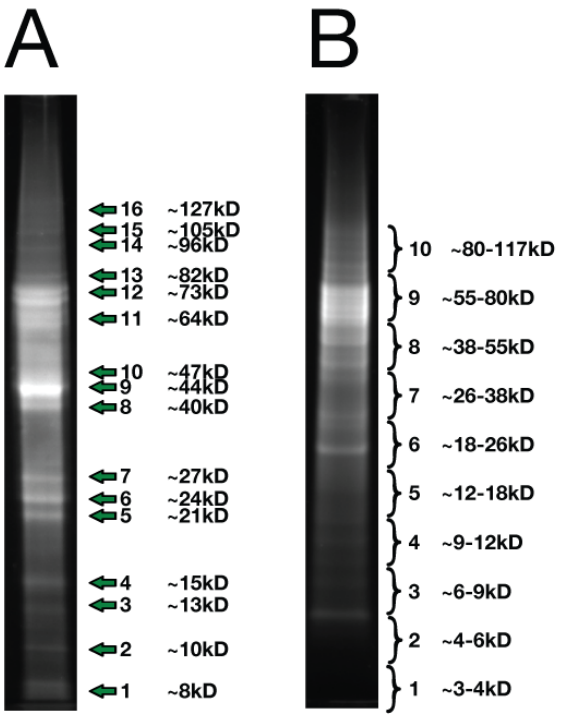


Figure 3.2. Denaturing polyacrylamide gel electrophoresis of protein extracts. Proteins were separated and distinct bands (a), or broad slabs covering the indicated ranges (b), were isolated.

A third protein extract was prepared as before, but rather than fractionating the sample by denaturing PAGE, the entire protein extract was subjected to in-solution tryptic digestion. The resulting solution of tryptic peptides was then pre-fractionated by step elution of a strong cation exchange solid phase extraction cartridge and each fraction was subjected to LC-ESI MS/MS. A database search using MASCOT with

the spectra from this analysis identified 192 unique proteins (Supplementary Table S4, Appendix 1). In addition to identifying more proteins, in the cases where a protein was identified by both a gel-based and gel-free approach, the latter strategy generally resulted in greater percentage protein coverage and total ion scores.

In summary, proteins corresponding to 202 distinct tomato unigenes were identified. The three analyses showed a substantial amount of overlap, as shown by the relatively high degree of redundancy between the sets of proteins identified in each analysis (Figure 3.3). Notably, only 5% of the proteins were identified only by the gel-based analysis and not by the gel-free approach. Given that so little is known about extracellular cuticle assembly and restructuring, we were particularly interested in the subset of proteins that are potentially secreted to the cell wall. Of the 202 proteins identified, 78 (39%) had secretory signal peptides (SPs) as predicted by SignalP 3.0 (Bendtsen *et al.*, 2004), and these were sorted into 40 putative functional families based on BLAST annotations (Table 1). Several of the putative secreted protein families had lipid-related domains, or similarity to proteins that have previously been implicated in cuticle biology. For example, we identified five LTPs and an MD-2 related lipid recognition domain-containing (ML) protein that is predicted to bind lipids. Also of interest were two GDSL-motif lipase/hydrolase family proteins. In addition to the proteins with putative roles in lipid metabolism, many defense-related proteins were also identified, including several PR-1 proteins, protease inhibitors, chitinases and endo-β-1,3-glucanases. A large number of proteins belonging to the category of cell wall modifying and structural proteins such as expansin, xyloglucan endotransglucosylase-hydrolase and extensin were also identified.

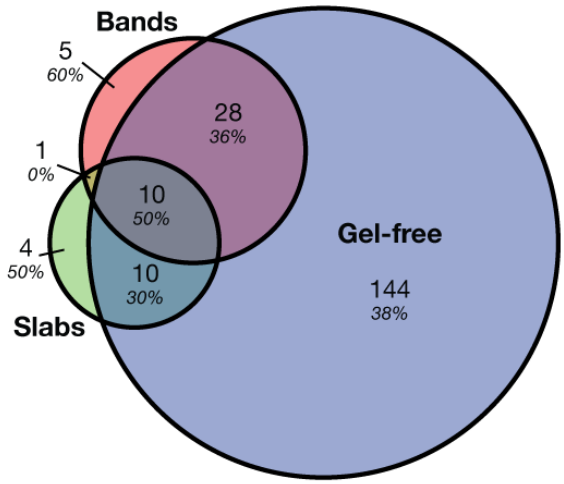


Figure 3.3. Venn diagram of proteins found in the three proteomic analyses and signal peptide prediction. The total number in each unique or overlapping set is shown with the percentage of each set with a predicted signal peptide (SignalP 3.0) indicated in italics.

Table 3.1. Proteins identified by MASCOT with predicted signal peptides.

Annotation/Gene Family[1]	SGN Unigene	Identified in Analysis[2]	Best Hit: Total Ion Score	Best Hit: Percent Coverage
Lipid and putative cuticle-related				
GDSL-motif lipase/hydrolase family protein	SGN-U583101	**A**	37	7.5
	SGN-U579520	AB**C**	173	22
inducible plastid-lipid associated protein	SGN-U577010	**C**	76	15.9
lipid-transfer protein (LTP)	SGN-U577838	**C**	43	11.7
	SGN-U579033	**C**	149	46.7
	SGN-U579687	**C**	252	55.7
	SGN-U580659	**C**	69	43.5
	SGN-U581465[3]	**C**	171	33.1
MD-2 related lipid recognition domain-containing (ML) protein	SGN-U577903	AB**C**	93	8.6
Defense-related				
allergen V5/Tpx-1-related family protein	SGN-U578890	**C**	105	13.8
Bet v I allergen family protein	SGN-U577856	A**C**	67	13.6
chitinase (GH family 18 and 19)[4]	SGN-U580366	B**C**	245	23.3
	SGN-U579068	**C**	244	27.3
	SGN-U579551	**C**	219	19.8
	SGN-U579696	**C**	72	8.7
	SGN-U581507	**C**	91	14.4
chitin-binding lectin	SGN-U562887	**C**	46	7.1
defensin	SGN-U577872	B**C**	341	47.4
	SGN-U591780	**C**	62	17.5
endo beta-1,3 glucanase (GH family 17)	SGN-U590837	**C**	102	7.8
hevein-like protein	SGN-U567805	**C**	103	12.3
	SGN-U579235	**C**	863	68.2
osmotin-like protein	SGN-U574403	A**C**	473	40.7
	SGN-U579414	**C**	505	31
	SGN-U581103	**C**	574	30.6
peroxidase	SGN-U581155	A**C**	588	21.9
	SGN-U583085	B**C**	550	23.5
	SGN-U564185	**C**	49	7.4
	SGN-U566251	**C**	200	25.1
	SGN-U571844	**C**	102	8.5
	SGN-U575184	**C**	243	36.8
	SGN-U578562	**C**	149	6.8
	SGN-U580369	**C**	199	17.3
	SGN-U580709	**C**	211	34.3
	SGN-U583086	**C**	461	23.8
peroxiredoxin	SGN-U579538	**C**	74	7.4
polygalacturonase inhibitor protein	SGN-U579059	A**C**	44	11.6
PR-1	SGN-U578279	**C**	180	48.7
	SGN-U579345	**C**	93	12.2

Table 3.1 (*Continued*)

Annotation/Gene Family[1]	SGN Unigene	Identified in Analysis[2]	Best Hit: Total Ion Score	Best Hit: Percent Coverage
	SGN-U579426	**C**	276	34.8
	SGN-U579545	**C**	771	52.8
	SGN-U579883	**C**	160	25
protease	SGN-U578421	A**C**	71	6.2
	SGN-U582837	A**C**	134	8.3
	SGN-U578351	**C**	76	4.7
	SGN-U578475	**C**	102	7.4
	SGN-U579972	**C**	159	6.1
protease inhibitor protein	SGN-U573941	AB**C**	194	18.3
	SGN-U574346	A**C**	83	7.3
	SGN-U577283	**C**	194	14.8
	SGN-U578389	**C**	163	20
	SGN-U578863	**C**	62	9.6
	SGN-U585465	**C**	134	15.5
snakin-like protein	SGN-U578258	**C**	168	9
Carbohydrate cell wall metabolism-related				
alpha-galactosidase (GH family 27)	SGN-U571081	**C**	102	5.9
beta glucosidase (GH family 1)	SGN-U580766	**A**	49	3.9
expansin	SGN-U577727	**C**	124	11.6
Ole e 1 allergen/extensin like	SGN-U563658	**C**	85	13.8
Other				
ADP/ATP translocator-like	SGN-U577960	**C**	91	5.1
ascorbate peroxidase	SGN-U578449	**C**	98	14
enolase	SGN-U579393	**C**	581	28.9
formate dehydrogenase	SGN-U579280	**C**	65	7.3
fructokinase	SGN-U586194	A**C**	250	9.2
fructose-bisphosphate aldolase	SGN-U578572	A**C**	209	17.6
glyceraldehyde 3-phosphate dehydrogenase	SGN-U580213	AB**C**	438	38.2
glycine-rich RNA binding protein	SGN-U578513	**B**	37	19.4
histone H2B	SGN-U579310	**C**	57	19.4
leucine-rich repeat transmembrane protein kinase	SGN-U579197	**C**	122	25.2
malate dehydrogenase	SGN-U565569	**C**	399	25.8
protein disulfide isomerase-like (PDIL) protein	SGN-U575297	**C**	61	6.5
	SGN-U577569	**C**	108	8.8
ribulose bisphosphate carboxylase large chain	SGN-U565452	AB**C**	50	4.8
SOUL heme-binding protein	SGN-U584870	**A**	48	9.2
strictosidine synthase family protein	SGN-U583542	A**C**	175	12.4
transketolase	SGN-U577918	**C**	109	4.2
unknown	SGN-U593950	**B**	23	11.3
	SGN-U565851	**C**	41	10.7
	SGN-U566943	**C**	52	18.2

(Previous Pages) **Table 3.1. Proteins identified by MASCOT with predicted signal peptides.** 1. Gene family groupings and annotation based on BLAST search of NCBI non-redundant database. 2. Analyses [(A) is gel-band based analysis, (B) is gel-slab based analysis, (C) is gel-free analysis] from which members of the protein family were identified. The analysis that yielded the highest protein total ion score is shown in bold. 3. The longest-six frame translation of SGN-U581465, corresponding to CAJ19705, has an incorrect start codon that was manually adjusted before SignalP analysis. 4. Glycosyl hydrolase (GH) families, www.cazy.org. The SGN annotation refers to the unigene identifier in the Sol Genomics Network database (www.solgenomics.net).

Several compelling candidates with homology to previously reported cuticle-related proteins were found in the set of proteins that were identified by only a single peptide and thus did not meet our stringent filtering criteria (Supplementary Table S1 and S3, Appendix 1). While we chose not to include these in the list of proteins that were confidently identified, their homology to previously identified cuticle-related proteins warranted further investigation. Thus, a glucose-methanol-choline (GMC) oxidoreductase family protein (SGN-U570812) with high similarity to the Arabidopsis protein HOTHEAD (HTH) (57% amino acid identity), which is involved in cuticle biosynthesis (Krolikowski *et al.*, 2003; Kurdyukov *et al.*, 2006b), and three additional GDSL-motif lipase/hydrolase family proteins (SGN-U577181, SGN-U583107 and SGN-U585129) were included in the expression and phylogenetic studies below.

Gene expression analysis of identified proteins

Since the initial aim of this study was to identify surface proteins with a possible role in cuticle formation, the further characterization of candidates with a previously reported cuticle association, or those with lipid-related domains, was of particular interest. Previous studies have shown many genes encoding cuticle biosynthetic enzymes are specifically expressed in the epidermis (Suh *et al.*, 2005; Mintz-Oron *et al.*, 2008) and so the cell type-specific expression of several candidate genes was investigated. Epidermis and collenchyma cells from immature green fruits were harvested using laser-capture microdissection (Figure 3.4a), RNA was isolated, amplified and qRT-PCR was performed. The expression of the epidermis-specific (Mintz-Oron *et al.*, 2008) cuticle biosynthesis gene *LeCer6* was used as a positive control and we also monitored the expression of transcripts encoding four defense-related proteins; a class that represented a substantial portion of the identified proteins. One of these, xyloglucan-specific endoglucanase inhibitor protein (XEGIP), was only identified by a single peptide (Supplementary Table S2, Appendix 1), but its well characterized expression and biological activity warranted its inclusion as a positive control for defense-related transcripts (Qin *et al.*, 2003). Of the ten genes selected for further characterization, five showed much greater expression in the epidermis relative to the collenchyma (90 to 1,700 fold), three showed more modest epidermal enrichment (4 to 8 fold) and two showed low expression ratios (2 fold and 0.03 fold), suggesting that their transcripts were not epidermis specific (Figure 3.4b). The positive control *LeCer6* and the six cuticle-related candidate genes all showed epidermal enrichment of greater than 4 fold while the four defense-related transcripts showed

mixed epidermal specificity: the *XEGIP* and PR-1 transcripts were more highly expressed in the epidermis while the defensin and chitinase both showed low expression ratios, indicating weak epidermal specificity and collenchyma specific expression, respectively.

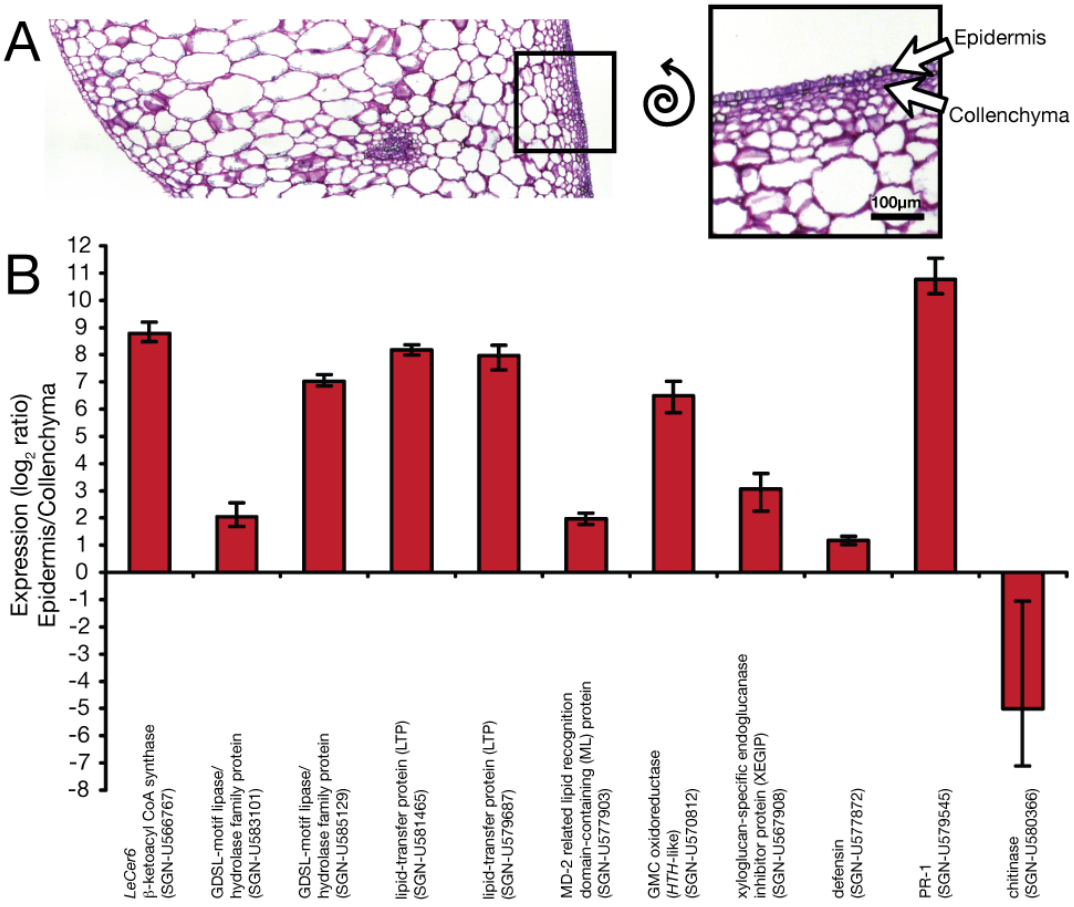


Figure 3.4. Tissue-specific expression of selected genes by qRT-PCR of RNA from microdissected cells. Epidermal cells and collenchyma cells were harvested from immature green tomato fruits by laser-capture microdissection as illustrated (A) and extracted, amplified RNA was used for qRT-PCR expression analysis of selected genes (B). The error bars are the standard error as determined by REST 2008 using three technical replicates.

Since deposition of wax and cutin follows a specific temporal pattern during fruit development, typified by maximal accumulation during fruit growth followed by a second phase of cuticle deposition during ripening (Baker *et al.*, 1982; Bauer *et al.*, 2004), we further characterized the expression of the eight epidermis upregulated genes during fruit growth and ripening using qRT-PCR (Figure 3.5). *LeCer6* expression was again used as a positive control, as it encodes a part of the fatty acid elongation complex required for aliphatic wax biosynthesis (Vogg *et al.*, 2004) and so its expression would be expected to correlate with wax deposition. The expression pattern of *LeCer6* was most similar to that of the GDSL-motif lipase/hydrolase family gene SGN-U585129 and GMC oxidoreductase, as all three were maximally expressed during the most rapid phase of fruit expansion, peaking at 15-20 DPA (Figure 3.5). The two defense-related transcripts, *XEGIP* and the PR-1 SGN-U579545, as well as the LTP SGN-U579687, showed related expression patterns with broad peaks of expression spanning the late phases of fruit growth and early ripening. We noted that the expression pattern of *XEGIP* corresponded well with a previously reported Northern-blot analysis of its expression (Qin *et al.*, 2003). Finally, the gene encoding the ML protein, the LTP SGN-U581465 and the GDSL-motif lipase/hydrolase family protein SGN-U583101 all showed similar expression patterns, with high levels of transcript in very young fruit and a substantial reduction by 15 DPA.

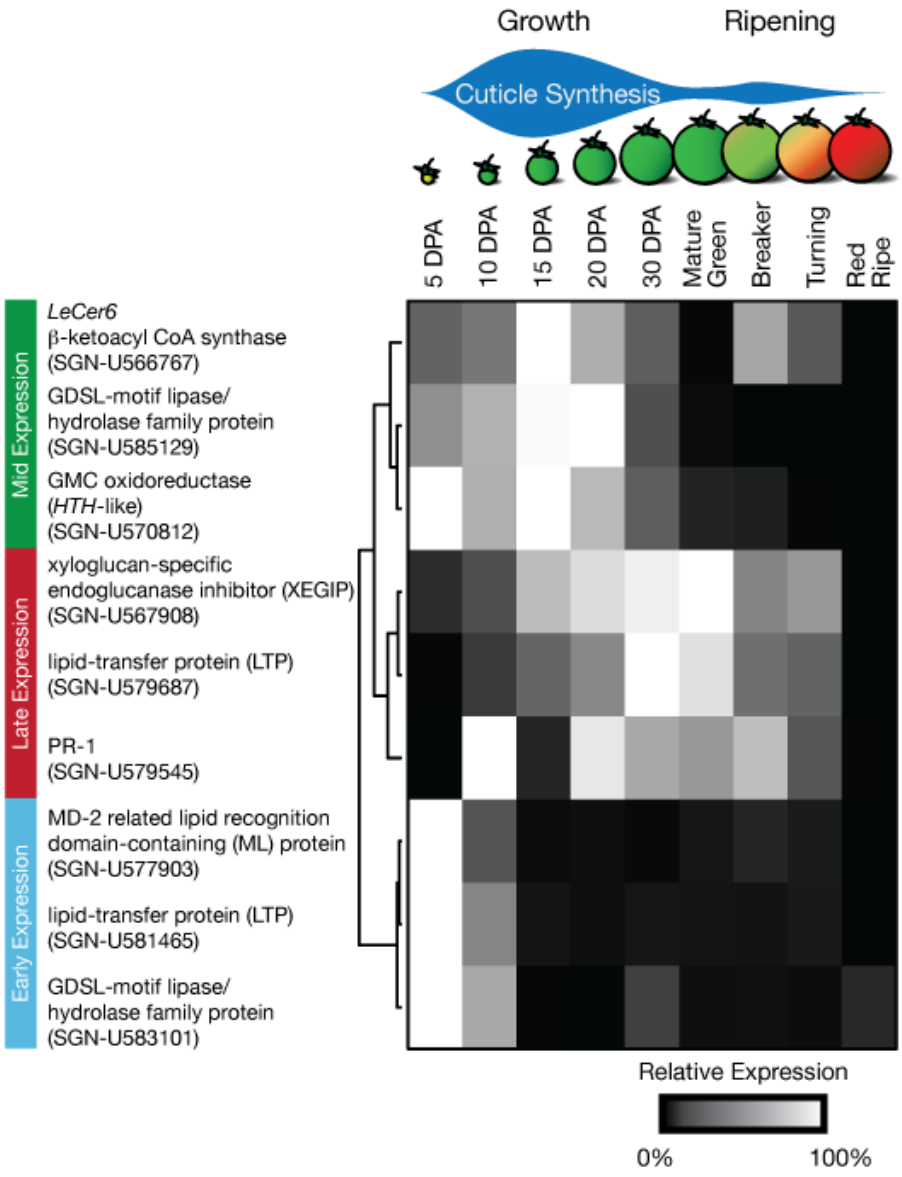


Figure 3.5. Time course expression of selected genes during fruit growth and ripening. Gene expression was determined by qRT-PCR relative to the constitutive control *RPL2* and normalized as described in the materials and methods section. The two phases of cuticle deposition are indicated above the fruit development stages considered.

Phylogenetic analysis of the GDSL-motif lipase/hydrolase family proteins

Two of the confidently identified proteins and three proteins identified by a single peptide belong to the GDSL-motif lipase/hydrolase family, which is widely distributed in both eukaryotes and prokaryotes (Akoh *et al.*, 2004). Plant GDSL-motif lipase/hydrolases comprise large gene families; for example, there are 113 predicted members in arabidopsis, although few have a known function. Several lines of circumstantial evidence suggest a role for these enzymes in cutin metabolism. First, biochemically characterized isozymes have been shown to have acyl hydrolase activity and the presence of a SP suggests that many are secreted (Akoh *et al.*, 2004). Second, microarray analysis of arabidopsis stem peels revealed that a subset of 18 members of the gene family is preferentially expressed in this cuticle-synthesizing tissue (Suh *et al.*, 2005). Furthermore, one of these, At2g04570, is highly induced by expression of the cuticle-associated transcription factor WIN1/SHN1 (Kannangara *et al.*, 2007). Microarrays with RNA from isolated tomato peel also identified three tomato GDSL-motif lipase/hydrolase family proteins that are preferentially expressed in the epidermis (Mintz-Oron *et al.*, 2008). Finally, the transcripts corresponding to a GDSL-motif lipase/hydrolase family protein, AgaSGNH, from *Agave americana* were shown to be highly abundant in the epidermis during leaf elongation, when cutin is being rapidly synthesized (Reina *et al.*, 2007).

Phylogenetic analysis of the five GDSL-motif lipase/hydrolase family proteins described in this present study, as well as the 18 epidermis-specific arabidopsis sequences, AgaSGNH and the three tomato sequences previously identified by Mintz-Oron *et al.* (2008) indicated that the candidate cuticle-related GDSL-motif

lipase/hydrolase family proteins can be grouped into four clades (Figure 3.6). The sequences identified in this study align within Clades I, III and IV. Co-expression analysis of the Arabidopsis members of Clades I and III using CressExpress (Srinivasasainagendra *et al.*, 2008; www.cressexpress.org) showed high levels of co-expression with nine cutin-biosynthesis related genes (*ATT1, LACS2, LCR, GPAT4, GPAT8, GPAT6, CYP86A4, CYP86A7* and *CYP77A6*), which were used as bait (Supplementary Table S5, Appendix 1). While co-expression is less pronounced in Clade IV, its smaller size and higher degree of conservation make it an attractive source of candidate cuticle-related GDSL-motif lipase/hydrolase family proteins, particularly in light of the expression patterns of *AgaSGNH* and SGN-U585129 that coincide with cutin deposition.

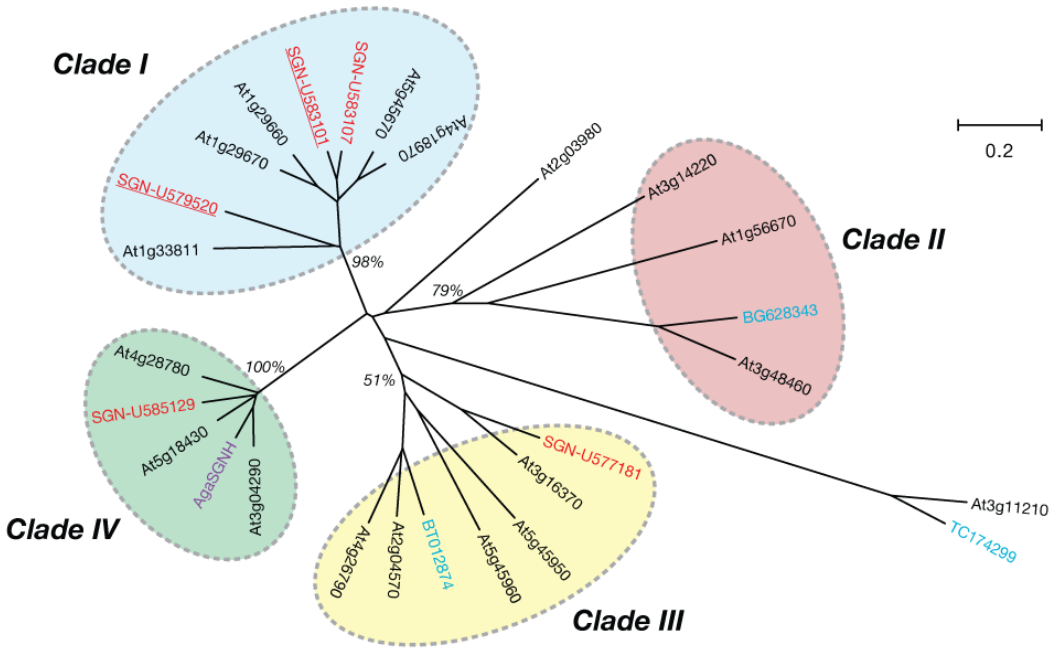


Figure 3.6. Phylogenetic analysis of GDSL-motif lipase/hydrolase family proteins. The *Arabidopsis* genes (AGI numbers, www.arabidopsis.org, in black) are those showing greater than two-fold enrichment in epidermal peels relative to whole stem tissue (Suh *et al.*, 2005). The blue TIGR plant transcript assembly numbers (plantta.jcvi.org) and NCBI EST accessions were identified as tomato transcripts enriched in the peel relative to tomato flesh (Mintz-Oron *et al.*, 2008). AgaSGNH (purple) was identified as an epidermal-specific transcript in *Agave americana* (Reina *et al.*, 2007). SGN unigenes (www.solgenomics.net) described in the current study are shown in red and are underlined if they were identified with confidence (spectra matching two or more peptides, see text). Bootstrap support of the four arbitrarily numbered clades is indicated in italic (500 replications). The bootstrap support of Clade III increases to 92% if the outlying At3g11210 and TC174299 sequences are discarded. Branch lengths are proportional to distance as indicated by the scale legend.

Discussion

In this study, the use of modern mass spectrometry-based proteomic techniques and a diverse set of protein fractionation strategies resulted in a large set of proteins putatively associated with the cuticle of the developing tomato fruit. During the first step of surface protein extraction, we were careful to minimize the time the fruits were submerged in the solvent, in order to reduce cell lysis and increase the proportion of secreted proteins. Bioinformatic analysis suggested that 39% of the cognate genes are predicted to encode N-terminal SPs that would direct their secretion and this represents a substantial enrichment. For comparison, when the arabidopsis predicted proteome (TAIR release 8, www.arabidopsis.org) is subjected to the same analysis, only 19% of proteins are predicted to have a SP (data not shown). Moreover, it is likely that the N-termini containing the SPs of some of the identified proteins are absent from the sequence databases, since a full genome sequence is not yet available for tomato. However, the presence of known intracellular proteins can be taken to indicate the lysis of some epidermal cells. In addition, we emphasize that a subset of extracellular proteins will not be extracted or successfully fractionated with the protocols used here due to the recalcitrant nature of the cell wall proteome (reviewed in Lee et al., 2004; Isaacson and Rose, 2006). Moreover, computational tools for predicting SPs are imperfect, and so the presence, or absence, of a predicted SP is not a *de facto* indication of protein extracellular or intracellular localization, respectively. Nonetheless, the enrichment we observed suggest that the protein extracts will provide a valuable starting point for researchers interested in cuticle assembly and restructuring.

The three fractionation strategies employed were generally complementary and helped to confirm findings in other analyses, as indicated by the significant overlap between the sets of proteins found in each analysis (Figure 3.3). However, the gel-free approach has clear advantages in terms of the number of proteins that were identified and the higher identification confidence scores, as indicated by MASCOT total ion score. Conversely, the gel-slab based analysis (Figure 3.2b) yielded the fewest identified proteins (Figure 3.3). Since the initial goal was to identify candidate proteins that might be involved in cuticle metabolism, several proteins attracted our attention because they had lipid-associated domains, or shared sequence similarity with proteins that are known, or proposed to have, roles in cuticle biogenesis.

Putative lipid-binding proteins

Of the five LTPs that were identified, four belong to the family 1 of LTPs, and one to family 2 (SGN-U577838) (Yeats and Rose, 2008). The cDNA le16, corresponding to SGN-U579033, was previously identified as being up-regulated by drought and ABA (Plant *et al.*, 1991), and the same gene, as well as the gene corresponding to SGN-U581465, was later shown to encode the tomato Lyc e 3 allergen (Le *et al.*, 2006). In a microarray analysis of tomato peel transcripts, SGN-U579687 was seen to be more highly expressed in the exocarp than the inner pericarp (Mintz-Oron *et al.*, 2008), a result that supports our finding that this transcript is more highly expressed in the epidermis than underlying collenchyma.

Aside from LTPs, SGN-U577903, which encodes an ML protein, and has predicted extracellular localization and lipid-binding activity, is also a candidate for contributing to cuticle biogenesis. The ML domain is shared by proteins from diverse

eukaryotic species and takes its name from MD-2, a soluble extracellular protein in humans that binds lipopolysaccharide (LPS) in the first step of a signaling cascade that triggers the innate immune response (Jerala, 2007). Other members of this family include the human cholesterol binding-protein NPC2 (Friedland *et al.*, 2003) and the dust mite allergen Der f 2, which was recently also shown to bind LPS (Ichikawa *et al.*, 2009). Structurally, the domain is composed of two β-sheets that enclose a deep lipid-binding pocket (Ohto *et al.*, 2007), although no ligand is known or function proposed for the protein family in plants. The transcript abundance of the ML protein was approximately four fold greater in the epidermis than in the underlying collenchyma cells (Figure 3.4b) and its expression, like that of the LTP SGN-U581465, was highest at the earliest stage of fruit development before rapidly declining (Figure 3.5). This precedes the extensive cutin and wax deposition that occurs during the phase of greatest fruit expansion, from 10 DPA to 30 DPA. However, this does not necessarily lead us to reject the LTP or ML proteins as candidates for wax or cutin transporters: both proteins have extremely stable folds that may result in the protein remaining functional far longer than steady state mRNA levels are maintained. While no structural or biochemical characterization of any plant-derived ML proteins has been reported, we suggest that the large lipid binding cavity of this domain may accommodate the pentacyclic triterpenoids that are abundant in tomato cuticular wax. In contrast, the family 1 LTPs that have been previously proposed as lipid-binding proteins that transport wax across the cell wall are unable to bind planar sterols that are structurally analogous to triterpenoids (Cheng *et al.*, 2004).

Putative HTH ortholog

The GMC oxidoreductase gene that was tentatively identified and shown to have epidermis specific expression, SGN-U570812, has 57% amino-acid identity with the arabidopsis gene *HTH*. The *hth* mutant has a fused floral organ phenotype that is attributed to a defective cuticle (Krolikowski *et al.*, 2003), and a biochemical activity for HTH has been proposed based on *hth* cutin polymer composition (Kurdyukov *et al.*, 2006b). Mutant plants accumulate increased amounts of ω-hydroxy fatty acids and lower levels of α,ω-dicarboxylic fatty acids that predominate in arabidopsis cutin. Thus, the authors propose that HTH oxidizes ω-hydroxy fatty acids to ω-oxo fatty acids prior to formation of α,ω- dicarboxylic fatty acids.

In tomato, α,ω- dicarboxylic fatty acids comprise only ~1% by weight of cutin monomer composition (Leide *et al.*, 2007; Isaacson *et al.*, 2009). Nevertheless, this monomer may play an important structural role in determining either the degree of cross-linking between cutin chains and the polysaccharide cell wall, or other cutin chains. Expression analysis of the *HTH*-like SGN-U570812 indicated that it is highly expressed (approximately 90 fold) in the epidermis relative to the collenchyma (Figure 3.4) and that its expression during fruit development coincides with the rapid expansion and cuticle deposition that occurs 10-20 DPA (Figure 3.5).

GDSL-motif lipase/hydrolase family proteins

The genes encoding two confidently identified and three tentatively identified GDSL-motif lipase/hydrolase family proteins were of particular interest and qRT-PCR

characterization of the expression patterns of the five genes was attempted. Despite repeated attempts, this was only successful for SGN-U583101 and SGN-U585129. Both were shown to be more highly expressed in the epidermis than in the underlying collenchyma, although the ratio for SGN-U585129 was much greater (Figure 3.4b). The time course of their expression during fruit development was also distinct as SGN-U583101 was highly expressed only very early in fruit development, while SGN-U585129 was expressed throughout fruit expansion (Figure 3.5).

GDSL-motif lipase/hydrolase family proteins have previously been proposed as cutin-synthases (Reina *et al.*, 2007), or enzymes involved in modification or recycling of the cutin polymer (Pollard *et al.*, 2008). Thus, in the absence of genetic or biochemical evidence for their activity, three specific biochemical activities that may be required for cutin metabolism can be imagined. As cutin synthases, they may incorporate either cutin monomers or oligomers. Second, controlled hydrolysis of cutin during organ expansion may be required, so they may act as cutin hydrolases. A third hypothetical enzyme activity is that of a cutin transacylase, wherein the cutin polymer could be loosened by simultaneous cleavage and religation (transacylation) of ester bonds, allowing for organ expansion during growth. This activity, combined with synthesis of cutin oligomers by intracellular enzymes such as GPATs and BAHD family acyltransferases, could be sufficient for cutin polymer synthesis, allowing the oligomers to be "stitched" into the growing cutin polymer matrix only by exchange of existing ester bonds.

In conclusion, using a proteomic approach, a diverse collection of proteins with putative roles in lipid metabolism was identified. Several of these have gene

expression patterns that correlate with cuticle biosynthesis; that is they are specifically expressed in the epidermis and their expression coincides or precedes the deposition of wax and cutin. Our results further suggest that there are discrete phases of cuticular lipid metabolism and/or trafficking, which are associated with different gene classes, and even more interestingly, distinct members of the same gene family (e.g. GDSL-motif lipase/hydrolase family proteins and LTPs). However these remain candidates for cuticle biogenesis and reverse-genetic experiments are currently underway for functional confirmation. Analysis of the phenotypes of these plants coupled with *in vitro* demonstration of their proposed biochemical activities will advance the goal of better understanding cuticle biosynthesis.

Supplementary Material

The following items are included in Appendix 1:

Supplementary Table S1. PCR primers use for gene expression analysis.

Supplementary Table S2. Complete MASCOT results for gel-band analysis.

Supplementary Table S3. Complete MASCOT results for gel-slab analysis.

Supplementary Table S4. Complete MASCOT results for gel-free analysis.

Supplementary Table S5. Co-expression results of arabidopsis GDSL-motif lipase/hydrolase family proteins and known cutin-biosynthesis genes.

Acknowledgements

The authors gratefully acknowledge Dr. Mike Scanlon for generously providing access to the PALM MicroBeam System. This work was supported by NSF Plant Genome Research Program grant (DBI-0606595) and T.H.Y. was supported in part by a NIH chemistry/biology interface training grant (grant number T32 GM008500). LC ESI MS/MS analysis was performed by Dr. Sheng Zhang and the Cornell Proteomics and Mass Spectrometry Core Facility.

References

Akoh, C.C., Lee, G.C., Liaw, Y.C., Huang, T.H. and Shaw, J.F. (2004) GDSL family of serine esterases/lipases. *Prog. Lipid Res.* **43**, 534–552.

Baker, E.A., Bukovac, M.J. and Hunt, G.M. (1982) Composition of tomato fruit cuticle as related to fruit growth and development. In *The Plant Cuticle* (Cutler, D.F., Alvin, K.L. and Price, C.E., eds). London: Academic Press, pp. 33-44.

Bauer, S., Schulte, E. and Thier, H. (2004) Composition of the surface wax from tomatoes: II. Quantification of the components at the ripe red stage and during ripening. *Eur. Food Res. Technol.* **219**, 487–491.

Beisson, F., Li, Y., Bonaventure, G., Pollard, M. and Ohlrogge, J.B. (2007) The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of Arabidopsis. *Plant Cell* **19**, 351–368.

Bendtsen, J.D., Nielsen, H., von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340**, 783–795.

Bird, D., Beisson, F., Brigham, A., Shin, J., Greer, S., Jetter, R., Kunst, L., Wu, X., Yephremov, A. and Samuels, L. (2007) Characterization of Arabidopsis ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion. *Plant J.* **52**, 485–498.

Bodnar, W.M., Blackburn, R.K., Krise, J.M. and Moseley, M.A. (2003) Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. *J. Am. Soc. Mass Spectrom.* **14**, 971–979.

Bonaventure, G., Beisson, F., Ohlrogge, J. and Pollard, M. (2004) Analysis of the aliphatic monomer composition of polyesters associated with Arabidopsis epidermis: occurrence of octadeca-cis-6, cis-9-diene-1,18-dioate as the major component. *Plant J.* **40**, 920–930.

Buda, G.J., Isaacson, T., Matas, A.J., Paolillo, D.J. and Rose, J.K. (2009) Three-dimensional imaging of plant cuticle architecture using confocal scanning laser microscopy. *Plant J.* **60**, 378–385.

Cheng, C.S., Samuel, D., Liu, Y.J., Shyu, J.C., Lai, S.M., Lin, K.F. and Lyu, P.C. (2004) Binding mechanism of nonspecific lipid transfer proteins and their role in plant defense. *Biochemistry* **43**, 13628–13636.

Cruz, M.A.L., Gomes, V.M., Fernandes, K.V.S., Machado, O.L.T. and Xavier, J. (2002) Identification and partial characterization of a chitinase and a beta-1,3-glucanase from Copernicia cerifera wax. *Plant Physiol. Biochem.* **40**, 11–16.

Debono, A., Yeats, T.H., Rose, J.K., Bird, D., Jetter, R., Kunst, L. and Samuels, L. (2009) Arabidopsis LTPG is a glycosylphosphatidylinositol-anchored lipid

transfer protein required for export of lipids to the plant surface. *Plant Cell* **21**, 1230–1238.

Franke, R., Briesen, I., Wojciechowski, T., Faust, A., Yephremov, A., Nawrath, C. and Schreiber, L. (2005) Apoplastic polyesters in Arabidopsis surface tissues–a typical suberin and a particular cutin. *Phytochemistry* **66**, 2643–2658.

Friedland, N., Liou, H.L., Lobel, P. and Stock, A.M. (2003) Structure of a cholesterol-binding protein deficient in Niemann-Pick type C2 disease. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 2512–2517.

Gidda, S.K., Shockey, J.M., Rothstein, S.J., Dyer, J.M. and Mullen, R.T. (2009) Arabidopsis thaliana GPAT8 and GPAT9 are localized to the ER and possess distinct ER retrieval signals: functional divergence of the dilysine ER retrieval motif in plant cells. *Plant Physiol. Biochem.* **47**, 867–879.

Gonzalez-Bosch, C., Brummell, D.A. and Bennett, A.B. (1996) Differential Expression of Two Endo-1,4-[beta]-Glucanase Genes in Pericarp and Locules of Wild-Type and Mutant Tomato Fruit. *Plant Physiol*. **111**, 1313–1319.

Graca, J., Schreiber, L., Rodrigues, J. and Pereira, H. (2002) Glycerol and glyceryl esters of omega-hydroxyacids in cutins. *Phytochemistry* **61**, 205–215.

Hulsen, T., de Vlieg, J. and Alkema, W. (2008) BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **9**, 488.

Ichikawa, S., Takai, T., Yashiki, T., Takahashi, S., Okumura, K., Ogawa, H., Kohda, D. and Hatanaka, H. (2009) Lipopolysaccharide binding of the mite allergen Der f 2. *Genes Cells* **14**, 1055–1065.

Isaacson, T., Damasceno, C.M., Saravanan, R.S., He, Y., Catala, C., Saladie, M. and Rose, J.K. (2006) Sample extraction techniques for enhanced proteomic analysis of plant tissues. *Nat. Protoc.* **1**, 769–774.

Isaacson, T., Kosma, D.K., Matas, A.J. *et al.* (2009) Cutin deficiency in the tomato fruit cuticle consistently affects resistance to microbial infection and biomechanical properties, but not transpirational water loss. *Plant J.* **60**, 363–377.

Isaacson, T. and Rose, J.K.C. (2006) The plant cell wall proteome, or secretome. In *Plant Proteomics* (Finnie, C., ed). Oxford: Blackwell, pp. 175-209.

Jerala, R. (2007) Structural biology of the LPS recognition. *Int. J. Med. Microbiol.* **297**, 353–363.

Jetter, R., Kunst, L. and Samuels, A.L. (2006) Composition of plant cuticular waxes. In *Biology of the Plant Cuticle* (Riederer, M. and Müller, C., eds). Oxford, UK: Blackwell, pp. 145-181.

Kannangara, R., Branigan, C., Liu, Y., Penfield, T., Rao, V., Mouille, G., Hofte, H., Pauly, M., Riechmann, J.L. and Broun, P. (2007) The Transcription Factor WIN1/SHN1 Regulates Cutin Biosynthesis in Arabidopsis thaliana. *Plant Cell* **19**, 1278–1294.

Krolikowski, K.A., Victor, J.L., Wagler, T.N., Lolle, S.J. and Pruitt, R.E. (2003) Isolation and characterization of the Arabidopsis organ fusion gene HOTHEAD. *Plant J.* **35**, 501–511.

Kurdyukov, S., Faust, A., Nawrath, C. *et al.* (2006a) The epidermis-specific extracellular BODYGUARD controls cuticle development and morphogenesis in Arabidopsis. *Plant Cell* **18**, 321–339.

Kurdyukov, S., Faust, A., Trenkamp, S., Bar, S., Franke, R., Efremova, N., Tietjen, K., Schreiber, L., Saedler, H. and Yephremov, A. (2006b) Genetic and biochemical evidence for involvement of HOTHEAD in the biosynthesis of long-chain alpha-,omega-dicarboxylic fatty acids and formation of extracellular matrix. *Planta* **224**, 315–329.

Le, L.Q., Lorenz, Y., Scheurer, S., Fötisch, K., Enrique, E., Bartra, J., Biemelt, S., Vieths, S. and Sonnewald, U. (2006) Design of tomato fruits with reduced allergenicity by dsRNAi-mediated inhibition of ns-LTP (Lyc e 3) expression. *Plant Biotechnology Journal* **4**, 231–242.

Lee, S.J., Saravanan, R.S., Damasceno, C.M., Yamane, H., Kim, B.D. and Rose, J.K. (2004) Digging deeper into the plant cell wall proteome. *Plant Physiol. Biochem.* **42**, 979–988.

Leide, J., Hildebrandt, U., Reussing, K., Riederer, M. and Vogg, G. (2007) The Developmental Pattern of Tomato Fruit Wax Accumulation and Its Impact on Cuticular Transpiration Barrier Properties: Effects of a Deficiency in a beta-Ketoacyl-Coenzyme A Synthase (LeCER6). *Plant Physiol.* **144**, 1667–1679.

Li-Beisson, Y., Pollard, M., Sauveplane, V., Pinot, F., Ohlrogge, J. and Beisson, F. (2009) Nanoridges that characterize the surface morphology of flowers require the synthesis of cutin polyester. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22008–22013.

Li, Y., Beisson, F., Koo, A.J., Molina, I., Pollard, M. and Ohlrogge, J. (2007) Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18339–18344.

Martin, J.T. and Juniper, B.E. (1970). *The Cuticles of Plants*. London: Edward Arnold.

Mintz-Oron, S., Mandel, T., Rogachev, I. *et al.* (2008) Gene expression and metabolism in tomato fruit surface tissues. *Plant Physiol.* **147**, 823–851.

Mueller, L.A., Solow, T.H., Taylor, N. *et al.* (2005) The SOL Genomics Network: a comparative resource for Solanaceae biology and beyond. *Plant Physiol.* **138**, 1310–1317.

Nakazono, M., Qiu, F., Borsuk, L.A. and Schnable, P.S. (2003) Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: identification of genes expressed differentially in epidermal cells or vascular tissues of maize. *Plant Cell* **15**, 583–596.

Ohto, U., Fukase, K., Miyake, K. and Satow, Y. (2007) Crystal structures of human MD-2 and its complex with antiendotoxic lipid IVa. *Science* **316**, 1632–1634.

Panikashvili, D., Shi, J.X., Schreiber, L. and Aharoni, A. (2009) The Arabidopsis DCR encoding a soluble BAHD acyltransferase is required for cutin polyester formation and seed hydration properties. *Plant Physiol.* **151**, 1773–1789.

Perkins, D.N., Pappin, D.J., Creasy, D.M. and Cottrell, J.S. (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567.

Pfaffl, M.W., Horgan, G.W. and Dempfle, L. (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res.* **30**, e36.

Pighin, J.A., Zheng, H., Balakshin, L.J., Goodman, I.P., Western, T.L., Jetter, R., Kunst, L. and Samuels, A.L. (2004) Plant cuticular lipid export requires an ABC transporter. *Science* **306**, 702–704.

Plant, A.L., Cohen, A., Moses, M.S. and Bray, E.A. (1991) Nucleotide sequence and spatial expression pattern of a drought- and abscisic acid-induced gene of tomato. *Plant Physiol.* **97**, 900–906.

Pollard, M., Beisson, F., Li, Y. and Ohlrogge, J.B. (2008) Building lipid barriers: biosynthesis of cutin and suberin. *Trends Plant. Sci.* **13**, 236–246.

Pyee, J., Yu, H. and Kolattukudy, P.E. (1994) Identification of a lipid transfer protein as the major protein in the surface wax of broccoli (Brassica oleracea) leaves. *Arch. Biochem. Biophys.* **311**, 460–468.

Qin, Q., Bergmann, C.W., Rose, J.K., Saladie, M., Kolli, V.S., Albersheim, P., Darvill, A.G. and York, W.S. (2003) Characterization of a tomato protein that inhibits a xyloglucan-specific endoglucanase. *Plant J.* **34**, 327–338.

Reina, J.J., Guerrero, C. and Heredia, A. (2007) Isolation, characterization, and localization of AgaSGNH cDNA: a new SGNH-motif plant hydrolase specific to Agave americana L. leaf epidermis. *J Exp. Bot.* **58**, 2717–2731.

Riederer, M. (2006) Introduction: biology of the plant cuticle. In *Biology of the Plant Cuticle* (Riederer, M. and Müller, C., eds). Oxford, UK: Blackwell, pp. 1-10.

Ruzin, S.E. (1999). *Plant Microtechnique and Microscopy*. Oxford: Oxford University Press.

Saldanha, A.J. (2004) Java Treeview--extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248.

Samuels, L., Kunst, L. and Jetter, R. (2008) Sealing plant surfaces: cuticular wax formation by epidermal cells. *Annu. Rev. Plant Biol.* **59**, 683–707.

Schneiderbauer, A., Sandermann, H.J. and Ernst, D. (1991) Isolation of functional RNA from plant tissues rich in phenolic compounds. *Anal. Biochem.* **197**, 91–95.

Shepherd, R.W. and Wagner, G.J. (2007) Phylloplane proteins: emerging defenses at the aerial frontline? *Trends Plant Sci.* **12**, 51–56.

Shevchenko, A., Wilm, M., Vorm, O. and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**, 850–858.

Srinivasasainagendra, V., Page, G.P., Mehta, T., Coulibaly, I. and Loraine, A.E. (2008) CressExpress: a tool for large-scale mining of expression data from Arabidopsis. *Plant Physiol.* **147**, 1004–1016.

Suh, M.C., Samuels, A.L., Jetter, R., Kunst, L., Pollard, M., Ohlrogge, J. and Beisson, F. (2005) Cuticular lipid composition, surface structure, and gene expression in Arabidopsis stem epidermis. *Plant Physiol.* **139**, 1649–1665.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596–1599.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.

Vogg, G., Fischer, S., Leide, J., Emmanuel, E., Jetter, R., Levy, A.A. and Riederer, M. (2004) Tomato fruit cuticular waxes and their effects on transpiration barrier properties: functional characterization of a mutant deficient in a very-long-chain fatty acid beta-ketoacyl-CoA synthase. *J Exp. Bot.* **55**, 1401–1410.

Yang, Y., Zhang, S., Howe, K., Wilson, D.B., Moser, F., Irwin, D. and Thannhauser, T.W. (2007) A comparison of nLC-ESI-MS/MS and nLC-MALDI-MS/MS for GeLC-based protein identification and iTRAQ-based shotgun quantitative proteomics. *J Biomol. Tech.* **18**, 226–237.

Yeats, T.H. and Rose, J.K. (2008) The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). *Protein Sci.* **17**, 191–198.

CHAPTER 4

Enzymatic synthesis of the plant biopolyester cutin from hydroxyacylglycerol precursors

Laetitia B. B. Martin[1]*, Trevor H. Yeats[1]*, Tal Isaacson[1], Yonghua He[1], Lingxia Zhao[2], Antonio J. Matas[1], Gregory J. Buda[1], David S. Domozych[3], Jocelyn K. C. Rose[1]

[1] Department of Plant Biology, Cornell University, Ithaca, NY 14853, USA.

[2] Plant Biotechnology Research Center, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai 200240, China

[3] Department of Biology and Skidmore Microscopy Imaging Center, Skidmore College, Saratoga Springs, NY 12866, USA.

* These authors contributed equally to this work. L.B.B.M. performed gene and protein expression experiments and cutin analysis. T.H.Y. purified recombinant protein, performed chemical analysis of surface soluble lipids and acyltransferase assays. T.I. identified and performed initial characterization and rough mapping of the mutant. Y.E. and L.Z. performed fine genetic mapping experiments and constructed the transgenic complementation vector. A.J.M performed the tissue specific gene expression experiments. G.J.B. conducted light microscopy experiments. L.B.B.M. and D.S.D. performed immunolocalization experiments. L.B.B.M, T.Y. and J.K.C.R. designed the study, analyzed the data and wrote the paper.

Summary

A waxy cuticle covers the aerial epidermis of all plants, providing essential protection from desiccation (Nawrath, 2006). Indeed, fossil evidence suggests that evolution of this structure was necessary for terrestrial colonization by plants ~400 million years ago (Edwards, 1993). Additionally, as the primary interface between plants and their environment, it plays key roles in defense against pests and pathogens, as well as establishing organ boundaries during development (Nawrath, 2006). The cuticle consists of an insoluble polyester of hydroxy fatty acids, or dicarboxylic acids, known as cutin, which is covered and infiltrated with a variety of waxes. While the generic composition of the cutin polymer is known, the mechanism and site of cutin polymerization have remained long standing questions (Heredia *et al.*, 2009; Pollard *et al.*, 2008). Here we identify an enzyme that catalyzes cutin polymerization, which is absent in the *cutin deficient 1* (*cd1*) mutant of tomato (*Solanum lycopersicum*). While cutin is typically highly abundant in tomato fruit cuticles, it is almost entirely absent from *cd1* fruits. We show, by map based cloning, that this mutant lacks a protein (CD1) of the GDSL-motif lipase/hydrolase family. The *cd1* fruits accumulate free 2-mono(10,16-dihydroxyhexadecanoyl)glycerol (2-MHG), a putative monomeric cutin precursor. CD1 protein shows acyltransferase activity, with a preference for 2-monoacylglycerol as an acyl donor, and localizes in developing cuticles. Our results suggest that CD1 catalyzes extracellular cutin polymerization at the site of cuticle formation via successive transesterification of 2-MHG.

Main Text

We previously identified several tomato mutants with dramatic deficiencies in cutin (Isaacson *et al.*, 2009). One of these, *cutin deficient 1* (*cd1*), has approximately 5-10% levels of fruit cutin compared with the wild type (M82) genotype, an extremely thin cuticle and increased sensitivity to water loss and pathogen susceptibility (Figure 4.1a, b; Isaacson *et al.*, 2009). Fine mapping of the *cd1* mutation revealed it to lie within a five-exon gene (*CD1*) (Figure 4.1c, d) that is predicted to encode a member of the GDSL motif lipase/hydrolase (GDSL) family of proteins (Figure 4.2). GDSLs collectively exhibit diverse functions and substrate specificities (Akoh *et al.*, 2004) and a broad taxonomic distribution, including prokaryotes and eukaryotes. In plants they are present as large gene families (Volokita *et al.*, 2011) and, based on their expression patterns, it has been speculated that GDSLs may play a role in cuticle biosynthesis (Irshad *et al.*, 2008; Matas *et al.*, 2010; Mintz-Oron *et al.*, 2008; Reina *et al.*, 2007; Yeats *et al.*, 2010).

The *cd1* mutant has a point mutation introducing a stop codon upstream of the fifth conserved domain, which contains two of the three predicted catalytic amino acid residues (Figure 4.1d and 4.2). In the mutant, *cd1* transcript levels are reduced (Figure 4.3a) but the CD1 protein was not detected by Western blot analysis (Figure 4.3c, d), indicating that it is a null mutant. Complementation of the *cd1* mutant with the wild type gene driven by the constitutive Cauliflower Mosaic Virus 35S promoter rescued the phenotype (Figure 4.4a, b), confirming that the mutation in *CD1* is responsible for the cutin deficiency.

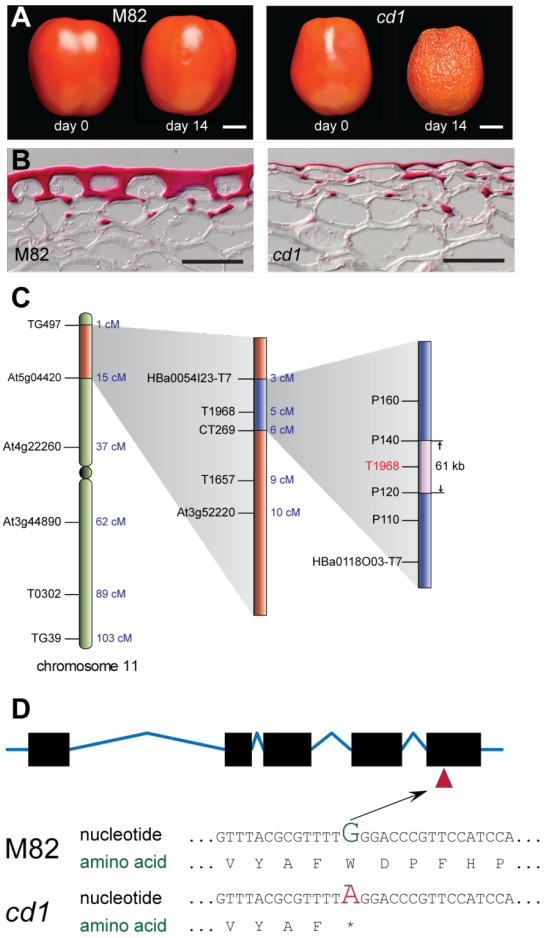


Figure 4.1. Fine mapping of the *CD1* gene. (a) Fruits of the M82 wild type tomato cultivar and the *cd1* mutant on the day of harvest at the fully ripe stage and 14 days after storage at room temperature. Scale bars = 1 cm. (b) Light microscopy showing the cuticle of M82 and *cd1* ripe fruit stained with Sudan Red 7b. Scale bars = 50 µm. (c) Schematic diagram showing the mapping of *CD1* to chromosome 11 between the markers P120 and P140. Marker T1968 cosegregates with *CD1*. (d) *CD1* gene splicing model with exons represented by black boxes and introns by blue lines. In *cd1*, the substitution G to A introduces a stop codon (TAG) in the last exon leading to the truncation of the last 41 amino acids including 2 residues of the catalytic triad.

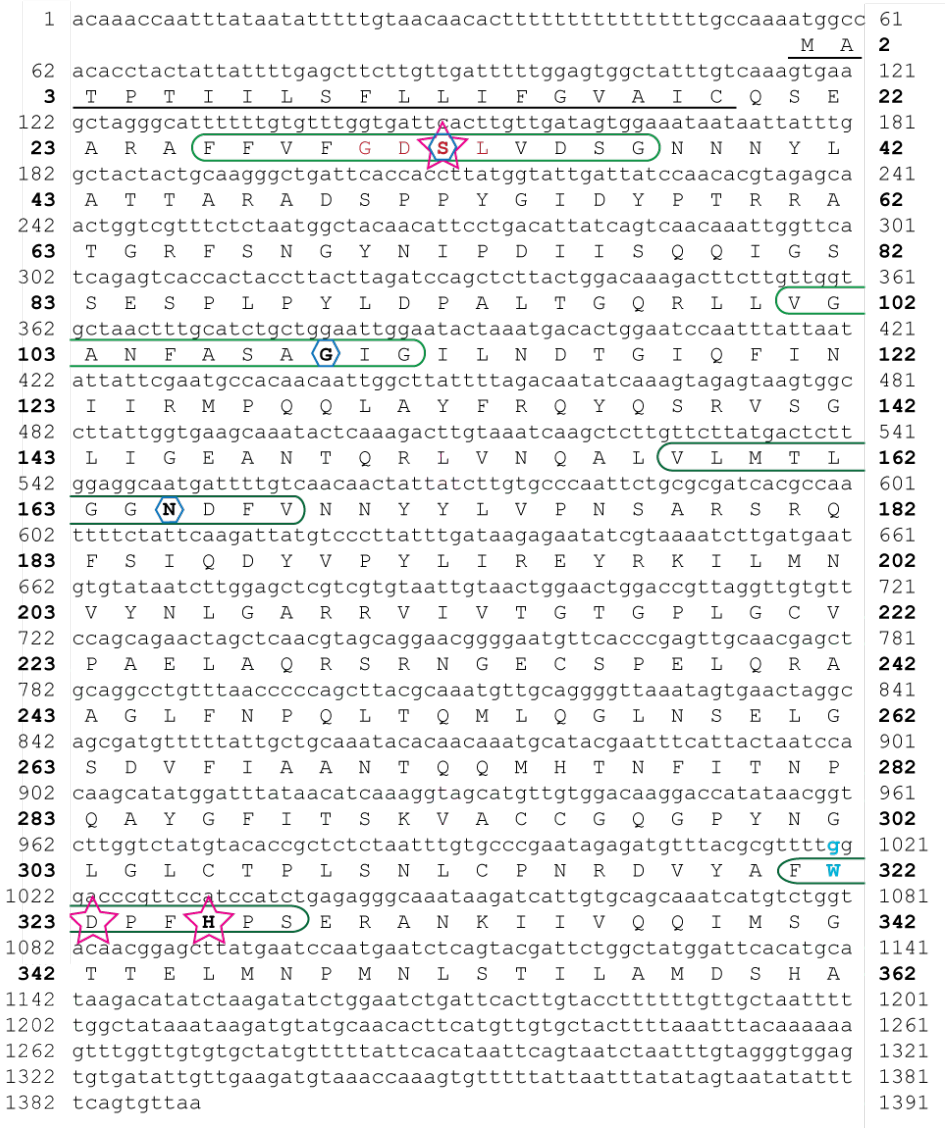


Figure 4.2. CD1 nucleotide and deduced amino acid sequences. The four conserved domains of GDSL lipases are circled in green. The catalytic triad residues are indicated by a magenta star and the oxyanion hole residues by a blue hexagon. The conserved residues giving its name to GDSLs are in red while the four residues strictly conserved are in bold font. The predicted secretory signal peptide (first 19 residues) is underlined. The mutated nucleotide in *cd1* and its corresponding amino acid residue are shown in light blue and bold font.

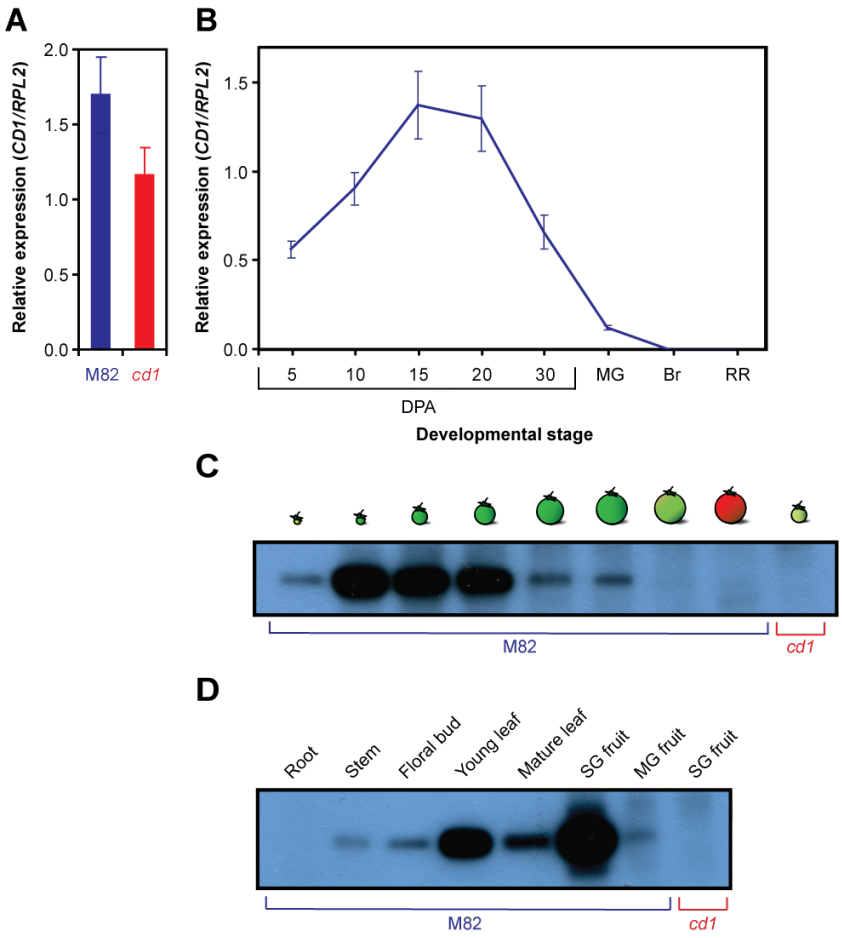


Figure 4.3. CD1 expression through fruit ontogeny and in different organs. (a) CD1 relative transcript levels in *cd1* and M82 15 days post anthesis (DPA) fruit pericarp obtained by qRT-PCR. (b) CD1 transcript levels in M82 fruit pericarp during fruit ontogeny, determined by qRT-PCR. (c) Immunoblot analysis of CD1 protein abundance levels in M82 fruit pericarp during fruit ontogeny. (d) Immunoblot analysis of CD1 protein expression in various tomato plant organs. SG, small green (15 DPA); MG, mature green; Br, breaker; RR, red ripe stages. RPL2: Ribosomal protein L2 (constitutive control). Error bars are SE for n = 3 with 3 technical replicates for (a) and n = 2 with 3 technical replicates for (b).

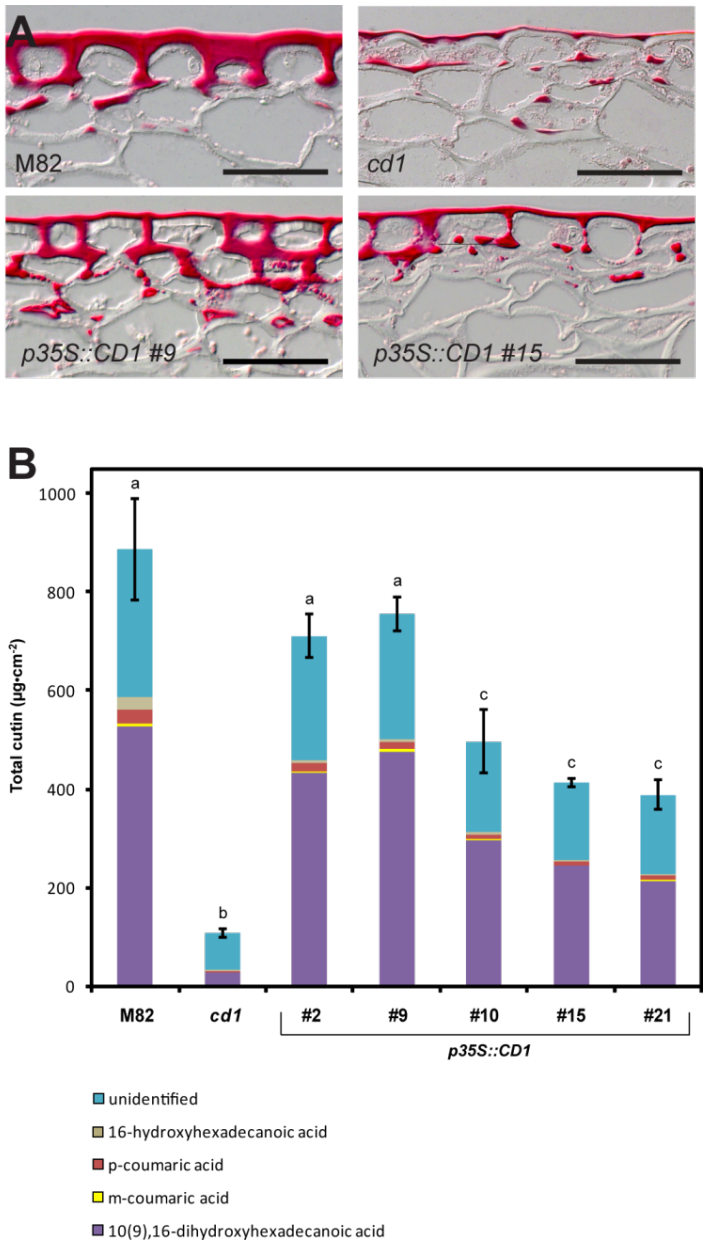


Figure 4.4. Complementation of the *cd1* mutant. (a) Light micrographs of mature green (MG) fruit pericarp sections showing fruit cuticles of M82, *cd1* and 2 independent transgenic complementation lines expressing the wild type *CD1* sequence driven by the 35S constitutive promoter, stained with Sudan Red 7b. Scale bars = 50 µm. (b) Cutin analysis of enzymatically isolated MG fruit cuticles from M82, *cd1* and 5 independent transgenic complementation lines driven by the 35S

Figure 4.4. (*Continued*)

constitutive promoter. 9,16-dihydroxyhexadecanoic acid and 10,16-dihydroxyhexadecanoic were not separated chromatographically, so they are reported together. Typically, the 10-isomer predominates by a ratio of ~ 10:1 (Baker *et al.*, 1982). Data represent the mean of three replicates. Error bars = SE of the total cutin load. *t*-tests were done on the complementation lines versus M82 and versus *cd1* at $\alpha = 0.05$. All complementation lines are statistically different from *cd1* and lines #2 and #9 are not statistically different from M82.

An analysis of the spatial distribution of CD1 proteins or transcripts showed that expression is highest in expanding organs, which require rapid cuticle synthesis to accommodate growth (Baker *et al.*, 1982), but is undetectable in roots, which have no cuticle (Figure 4.3b-d). Moreover, we used laser-capture microdissection of various pericarp tissues from young fruit to show that *CD1* transcript levels are highest in the epidermis (Figure 4.5), the tissue responsible for cuticle synthesis. Thus, *CD1* expression parallels spatial and temporal patterns of cuticle deposition at several levels. Immunolocalization of CD1 in M82 fruits indicated that the protein is almost exclusively localized in the cuticle (Figure 4.6, Figure 4.7). More specifically, labeling density followed the contour of the cuticle over both the periclinal and anticlinal cell walls (Figure 4.6b). This localization pattern suggests a role for CD1 very late in the cutin biosynthetic pathway, leading us to investigate whether CD1 is directly involved in cutin polymerization.

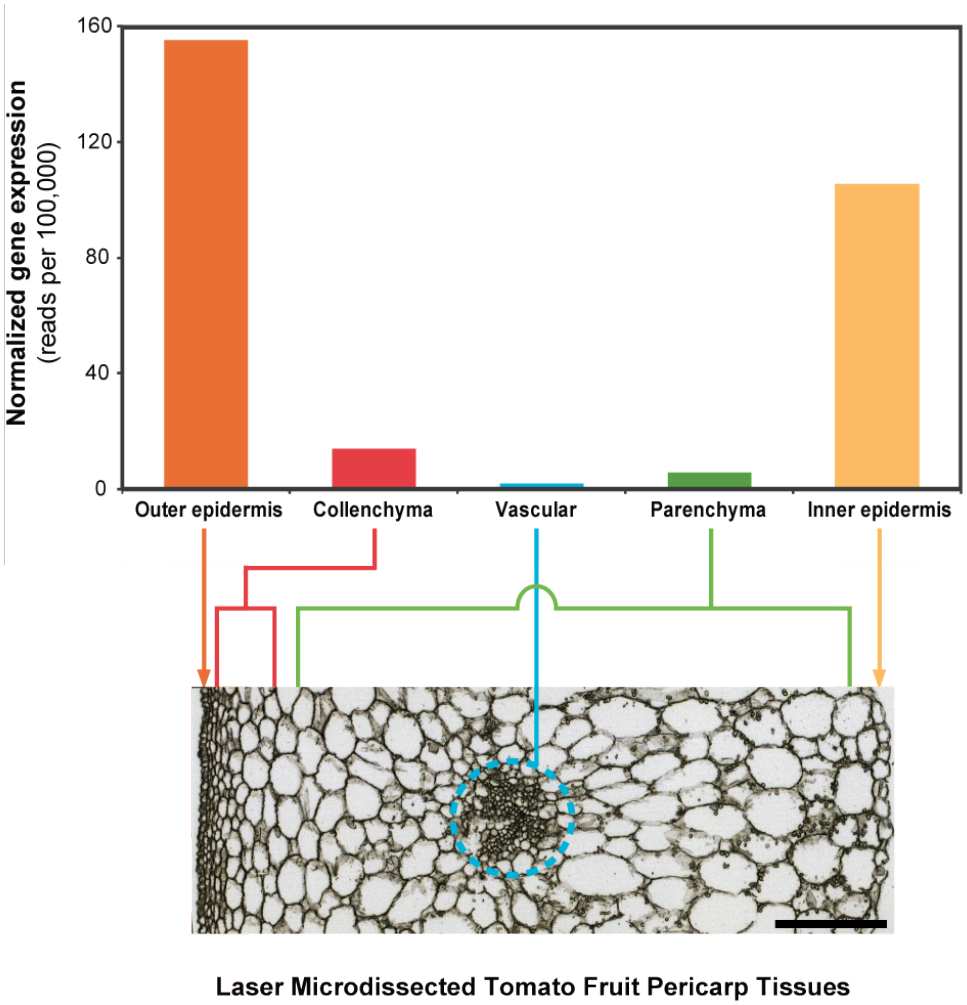


Figure 4.5. Expression of *CD1* transcripts in the tomato fruit pericarp tissues. Tissues are represented in a light micrograph of a 10 DPA fruit pericarp section, based on data derived from 454 sequencing of laser capture derived cDNA libraries. Scale bar = 100 μm.

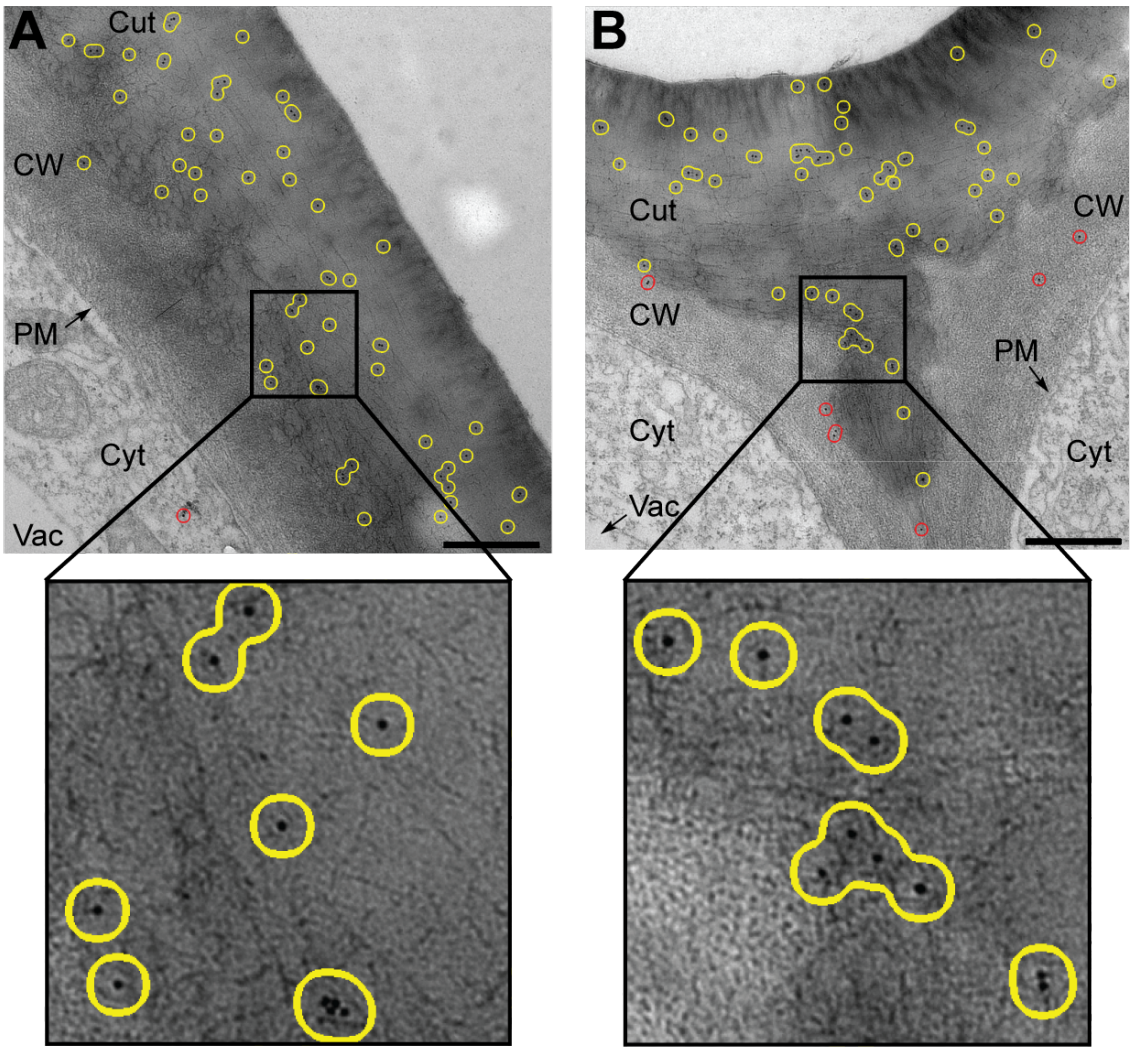


Figure 4.6. Immunolocalization of CD1 in the cuticle of 15 days post-anthesis (DPA) M82 fruits. (a) TEM micrograph showing the cuticle over the periclinal wall of an epidermal cell. (b) The cuticle between two adjacent epidermal cells (anticlinal peg). Yellow circles were drawn around gold particles localizing in the cuticle, red circles around the gold particles localizing elsewhere. Scale bars = 500 nm. Cut, cuticle; CW, cell wall; Cyt, cytoplasm; PM, plasma membrane; Vac, vacuole. Magnified areas of each image are shown in the lower panels.

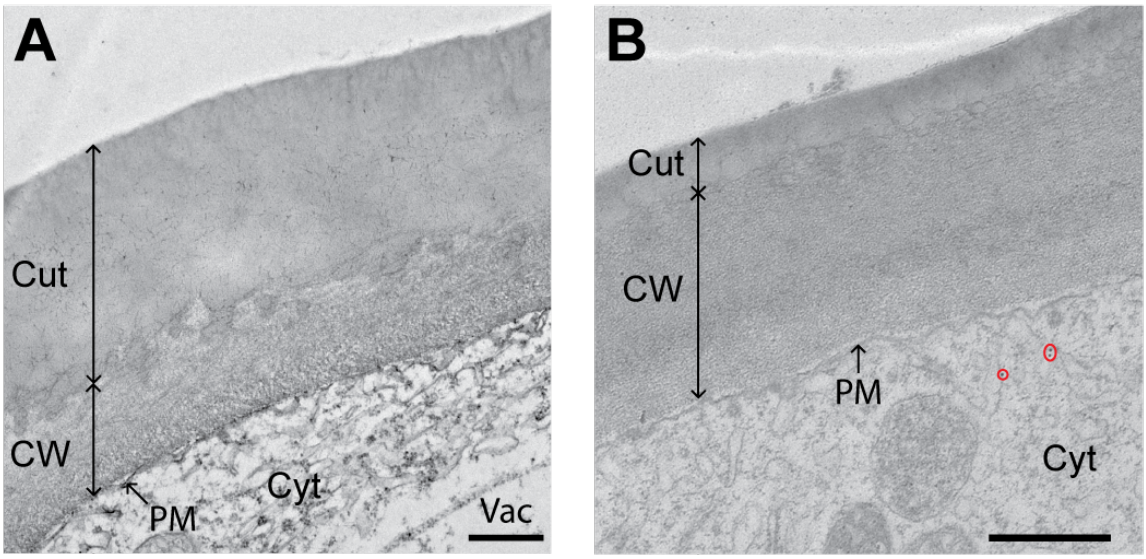


Figure 4.7. Immunolocalization negative controls corresponding to Figure 4.6. (a) Pre-immune serum tested on M82 15 days post anthesis (DPA) fruits. (b) Immunogold labeling of *cd1* 15 DPA fruits. Red circles were drawn around the gold particles. Scale bars = 500 nm. Cut, cuticle; CW: cell wall; Cyt, cytoplasm, PM, plasma membrane; Vac, vacuole.

Analyses of cuticle mutants have shown that several enzymes are required for formation of the cutin polymer, including glycerol phosphate acyltransferases (GPATs) (Pollard *et al.*, 2008). Recently, biochemical characterization of GPAT4 and GPAT6 showed them to possess both glycerol-3-phosphate acyltransferase activity specific to the *sn-2* position, and phosphatase activity (Yang *et al.*, 2010). This may indicate a structural role for 2-monoacylglyceryl esters (2-MAGs) in the cutin polymer, as they were identified in small quantities in the products of partially depolymerized cutin (Graca *et al.*, 2002). Alternatively, the 2-MAG products of GPAT4 and GPAT6 may act primarily as acyl donors for the polymerization reaction.

If this is the case, and CD1 is indeed a cutin polymerase, we would expect that 2-MAGs would accumulate as soluble lipids in the surface tissues of the *cd1* mutant fruit, but not in the M82 wild type genotype.

To test this, we looked for the accumulation of 2-MAGs in *cd1* fruit soluble surface lipids. This fraction, containing the cuticular waxes, can readily be extracted from plants by brief immersion of intact organs in organic solvents (Jetter *et al.*, 2006). In tomato fruits this consists primarily of a mixture of high melting-point alkanes and triterpenoids, while the cutin, a polyester of principally 10,16-dihydroxyhexadecanoic acid, is insoluble under these conditions. Although soluble 2-MAGs can be found in the waxes associated with suberin, they are not observed in cuticular waxes (Beisson *et al.*, 2007). We identified the 2-MAG species 2-mono(10,16-dihydroxyhexadecanoyl)glycerol (2-MHG) by GC-MS analysis of the soluble surface lipids from *cd1* fruits at the rapidly expanding stage, when CD1 is most highly expressed (Figure 4.3c), but not in equivalent extracts from M82 fruit (Figure 4.8). The coincident single ion chromatograms of diagnostic fragments clearly showed the specific accumulation of 2-MHG in the mutant (Figure 4.8a), despite partial coelution with two neighboring peaks. An additional, later-eluting trace peak of these ions likely corresponds to the thermodynamically favored 1-mono(10,16-dihydroxyhexadecanoyl)glycerol (1-MHG) isomer. The identity of the larger of the two peaks as representing the 2-isomer is confirmed by its earlier elution and the absence of the M-103 = 547 ion produced by α-cleavage between the 2- and 3- carbon in 1-MHG (Figure 4.8b; Graca *et al.*, 2002). Additional GC-MS analysis of the

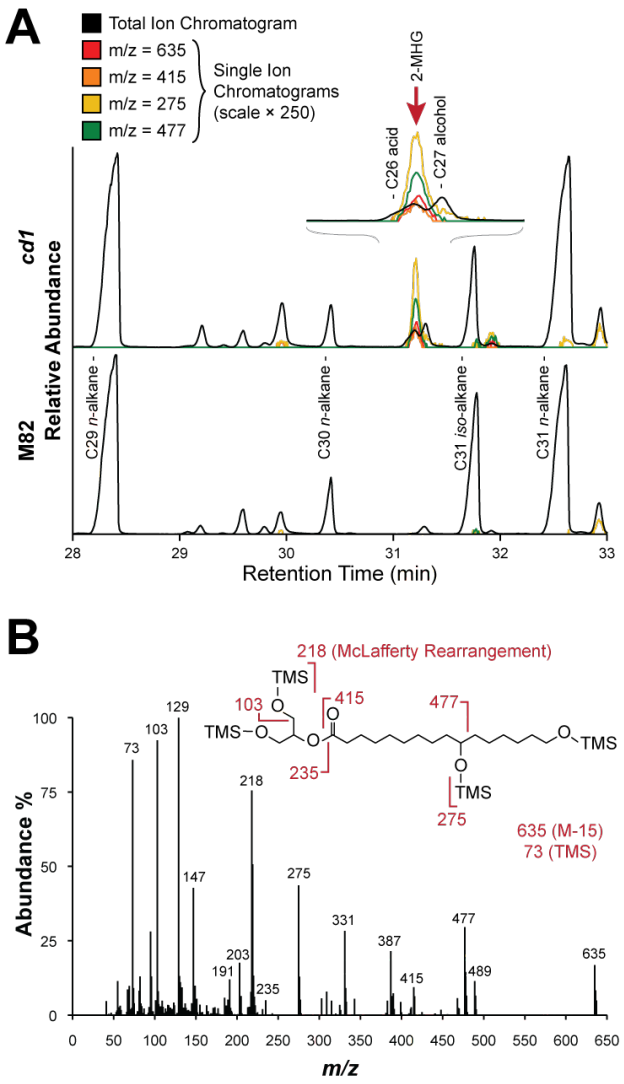


Figure 4.8. The identification of 2-MHG in the soluble surface lipids of *cd1*. (a) GC-MS chromatograms of the *cd1* mutant and M82 wild type TMS-derivatized extracts. The total ion chromatograms and several single ion chromatograms corresponding to characteristic fragments of 2-MHG are shown. Inset is an enlargement of the region surrounding the 2-MHG peak. For reference, several of the wax compounds common to both mutant and wild type are labeled. (b) Mass spectrum from the 2-MHG peak found in the *cd1* extract, with interpretation of the spectrum inset. For more detailed discussion of the spectrum of 2-MHG, see Graca *et al*. (2002).

acetylated *cd1* soluble surface lipids yielded a mass spectrum that is consistent with the acetyl derivative of 2-MHG (Figure 4.9).

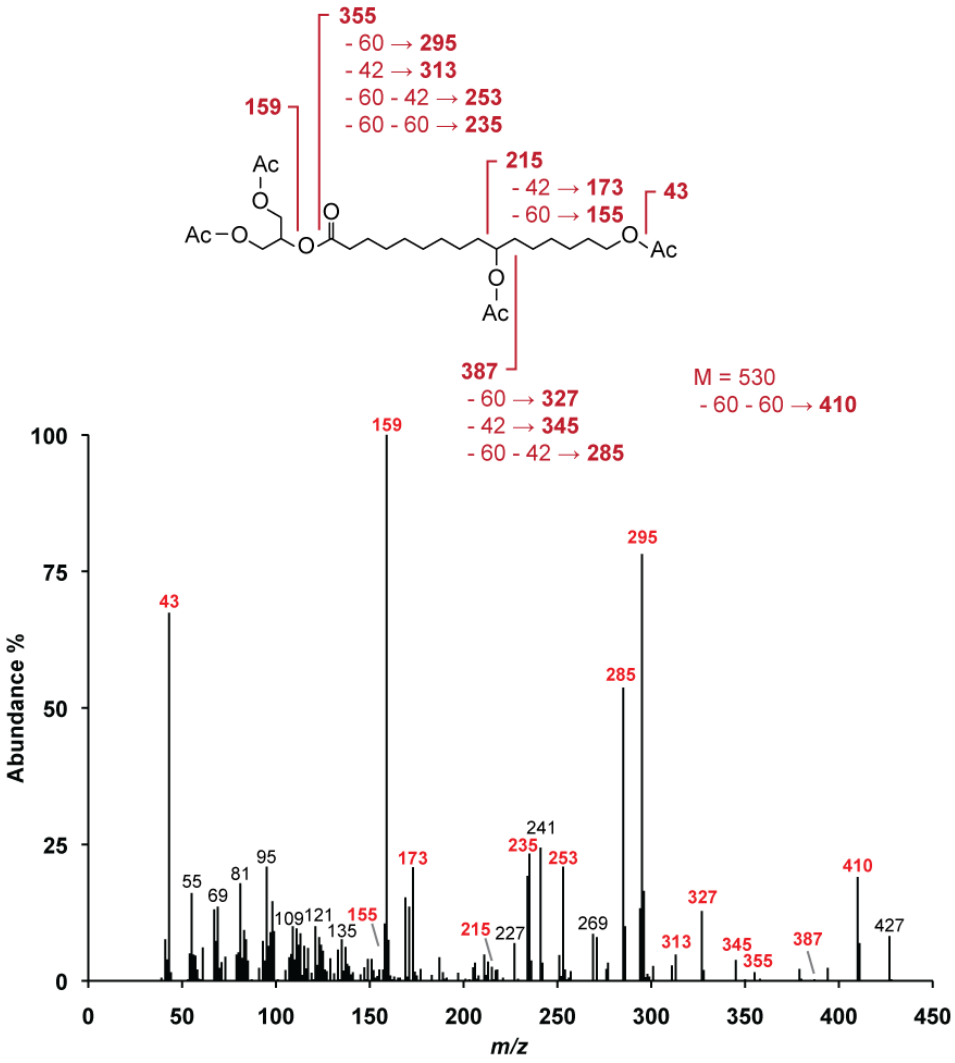


Figure 4.9. Mass spectrum of the acetyl derivative of 2-MHG. The fragmentation pattern is consistent with α-cleavages followed by loss of acetic acid (60 Da) and/or ketene (42 Da) that is typically observed for acetyl derivatives of hydroxy fatty acid esters (Nicolaides *et al.*, 1983). Observed ions are highlighted in bold red.

We propose a model for cutin polymerization wherein CD1 transfers the hydroxyacyl group of 2-MHG to either another molecule of 2-MHG, or to the growing cutin polymer itself (Figure 4.10a). Experiments involving partial depolymerization of tomato cutin have identified oligomers primarily consisting of directly coupled 10,16-dihydroxyhexadecanoic acid monomers (Graca and Lamosa, 2010; Osman *et al.*, 1999). This result, together with the observation that glycerol is quantitatively a minor component of tomato cutin (Graca *et al.*, 2002), indicates that the principal linkage in tomato cutin is between the carboxylic acid and hydroxyl groups of 10,16-dihydroxyhexadecanoic acid. The detection of small amounts of 2-MHG in the cutin polymer (Graca *et al.*, 2002) may therefore reflect the presence of 2-MHG 'primers' remaining in the polymer. The presence of polymerized 1-MHG could be a consequence of spontaneous acyl migration accelerated by the alkaline conditions used for *in vitro* depolymerization. To test our hypothesis that CD1 acts as an acyltransferase, we purified recombinant CD1 protein following expression in *Nicotiana benthamiana*. 2-MHG is not available commercially and its multiple hydroxyl groups present a synthetic challenge, so we used a simplified acyltransferase assay with decanol as an acyl acceptor and 2-monohexadecanoylglycerol as an acyl donor. Incubation of an emulsion of these two compounds with CD1 led to formation of the decyl hexadecanoate ester (Figure 4.10b, c). In contrast, when 1-monohexadecanoyl-*rac*-glycerol or hexadecanoic acid were used as acyl donors, the production of decyl hexadecanoate was greatly reduced (Figure 4.10c). Given appropriate conditions, both esterification and transesterification reactions can be

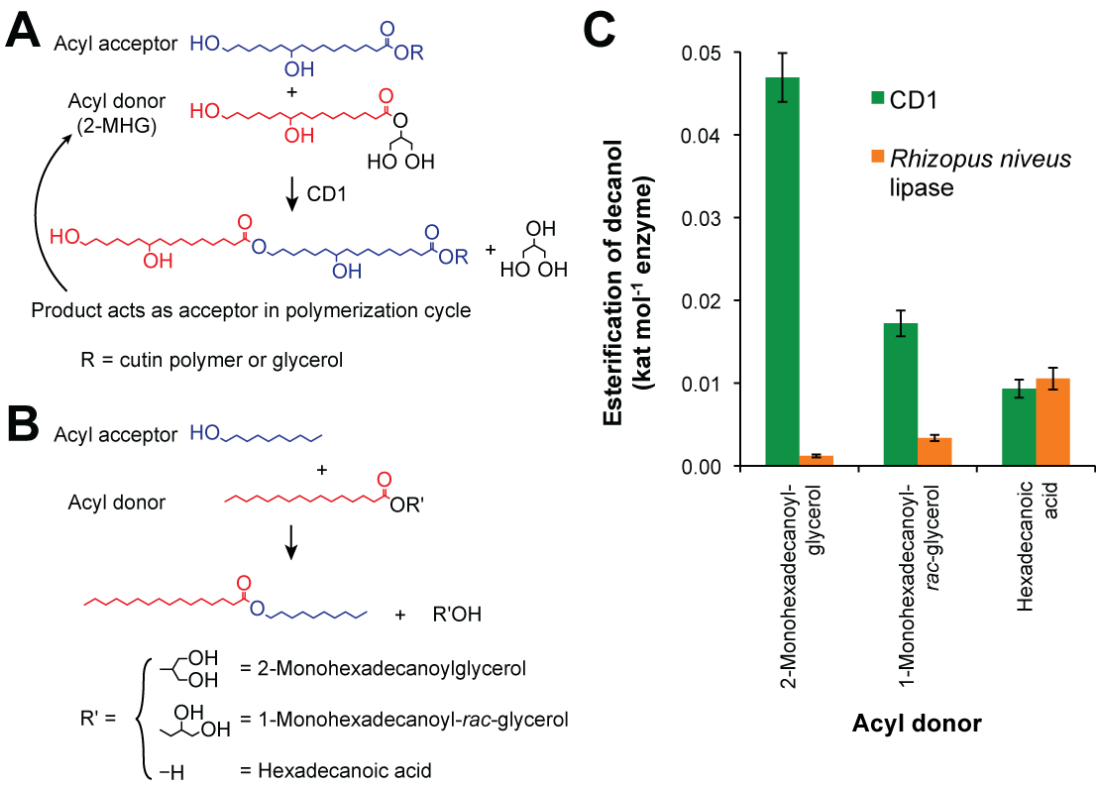


Figure 4.10. Acyltransferase activity of CD1. (a) Proposed model for CD1 catalyzed cutin biosynthesis by transfer of the hydroxyacyl group from 2-MHG to the growing polymer. (b) Acyltransferase assay and structure of the substrates tested as acyl donors for the esterification of decanol. (c) Acyltransferase activity of CD1 using three different acyl donors compared to that of a *Rhizopus niveus sn-1,3* specific lipase. Error bars are SE for $n = 3$. For each enzyme, all substrates showed statistically significant differences in activity (*t*-test, $\alpha = 0.05$).

catalyzed by a variety of lipases *in vitro* (Yahya *et al.*, 1998). Therefore, to address the substrate specificity of CD1, we compared its activity with a lipase from *Rhizopus niveus* that has a preference for the *sn-1* and *sn-3* positions of acyl glycerols (Haas and Joerger, 1995). The *R. niveus* lipase catalyzed acyl transfer to a greater extent with 1-monohexadecanoyl-*rac*-glycerol than with 2-monohexadecanoylglycerol and the greatest acyltransferase activity was observed using hexadecanoic acid as an acyl donor (Figure 4.10c). Thus, the transesterification reaction catalyzed by the *R. niveus* lipase likely occurs first by hydrolysis, as has been described for a closely related *R. oryzae* lipase (Haas and Joerger, 1995). In contrast, CD1 acyltransferase activity was much lower with hexadecanoic acid as an acyl donor, suggesting that the CD1 catalyzed transesterification reaction occurs without initial hydrolysis and proceeds directly via an acyl-enzyme intermediate. Two additional substrates, methyl hexadecanoate and trihexadecanoylglycerol, were also tested in this assay, but the decyl hexadecanoate product was not produced by either enzyme. Taken together, these results suggest that CD1 has a narrow substrate specificity with a preference for 2-MAGs.

In vivo ester synthesis from transesterification of acyl glycerol by a lipase-like enzyme is not without precedent. For example, in animals, the extracellular acylation of cholesterol by transesterification of lecithin is catalyzed by lecithin cholesterol acyltransferase (LCAT). In the absence of cholesterol as an acyl acceptor, LCAT has acyl esterase activity (Jonas, 2000). The unique feature necessary for this transesterification reaction is the action of the enzyme at the lipid-aqueous interface of high density lipoproteins. Here, cholesterol concentrations are high enough to favor

the resolution of the acyl-enzyme intermediate by transesterification rather than hydrolysis. We propose that CD1 acts through a similar mechanism at the interface between the aqueous environment of the plant cell wall and the lipid phase of the cuticle. The potential aqueous solubility of 2-MHG and the insolubility of the cutin polyester may thermodynamically favor polyester synthesis.

In vitro cutin polymerization by crude plant enzyme preparations was first reported more than thirty years ago (Croteau and Kolattukudy, 1974). Recent molecular genetic characterization of this process has identified several intracellular acyltransferases that are involved in biosynthesis of presumed cutin precursors (Panikashvili *et al.*, 2009; Pollard *et al.*, 2008). However, the molecular basis of cutin polymerization following secretion of the precursors into the cell wall has remained a mystery. Here we show that CD1, an extracellular enzyme that localizes in the developing cuticle, has acyltransferase activity with a 2-MAG acyl donor. Moreover, we identify accumulation of 2-MHG, the corresponding 2-MAG of the major cutin monomer of tomato, in the *cd1* mutant. Based on these complementary lines of evidence, we propose that CD1 is the principal catalyst of cutin polymerization and that the polymerization process is extracellular, at the site of cuticle deposition. Furthermore, a survey of the DNA sequence databases revealed *CD1* homologs in a taxonomically broad range of plant species (Figure 4.11). This suggests an evolutionarily conserved and ubiquitous mechanism of cutin biosynthesis in land plants.

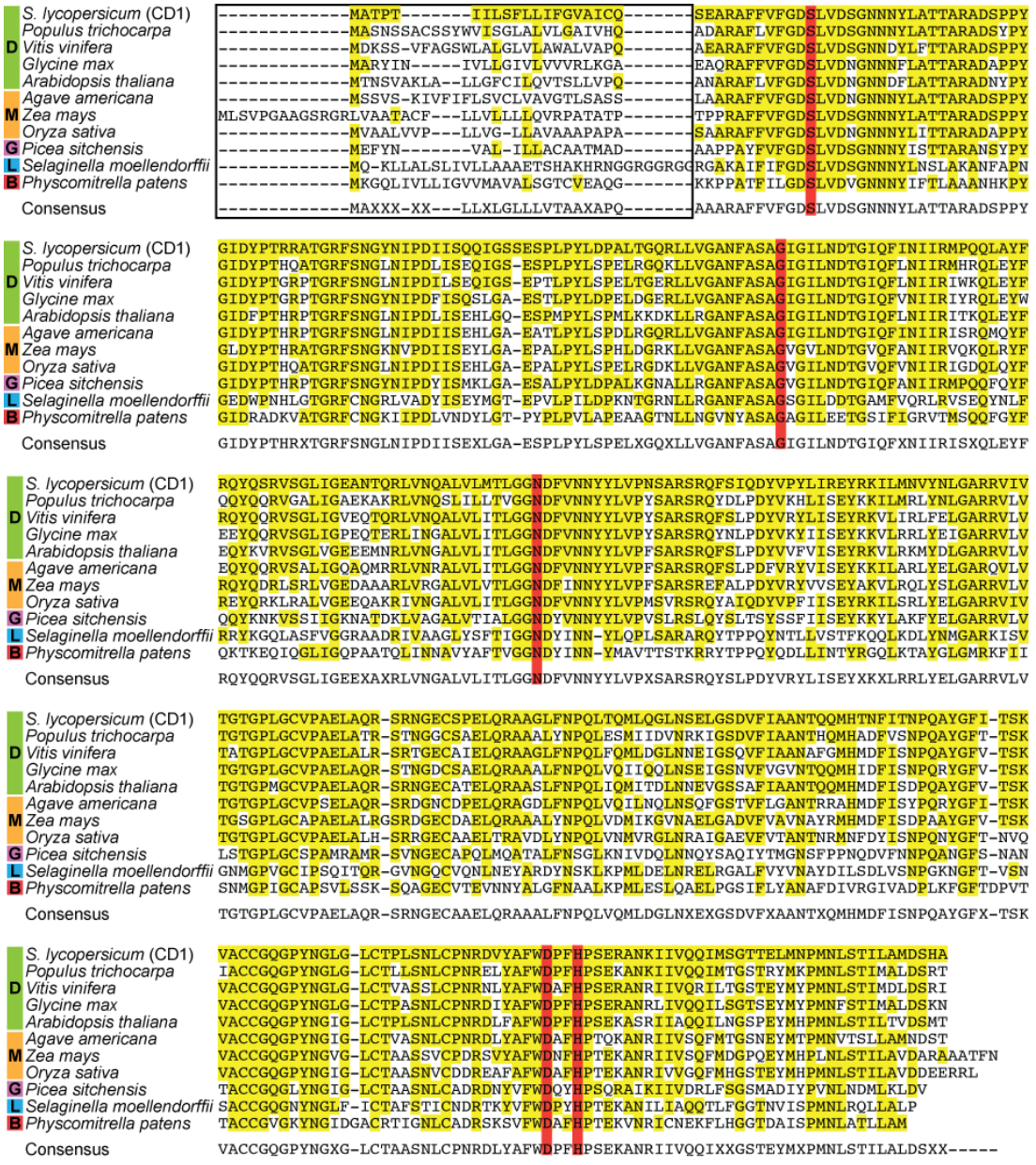


Figure 4.11. Multiple sequence alignment of CD1 and homologs from representative land plant species. Residues with identity to the CD1 sequence are shaded yellow, the conserved active site residues are shaded red and the signal peptide is boxed. D, dicotyledons; M, monocotyledons; G, gymnosperms; L, lycopods; B, bryophytes.

Figure 4.11 *(Continued)*

Sequences correspond to GenBank accessions XP_002319125, XP_002268296, ACU23610, NP_198322, AAS75127, NP_001146251, EAY86703, ABK25359, XP_002975076 and XP_001764401, respectively.

Materials and methods

Fine mapping of the CD1 gene

For fine genetic mapping, 880 F2 plants of *cd1* (in *S. lycopersicum* cv. M82 background) × *S. pimpinellifolium* were screened for recombinants using markers TG497 and At5g04420 (position 11.001 and 11.015 respectively, according to EXPEN2000 map, http://solgenomics.net). A total of 105 recombinant plants were generated and their fruit cuticles were phenotyped (Isaacson *et al.*, 2009) and the fruits genotyped using newly designed CAPS (or dCAPS) markers. The location of *CD1* was narrowed down first to the 3.22 Mbp tomato WGS scaffold SL1.03sc01386, (corresponding to SL2.40sc03748 in version 2.3 of the ITAG tomato genome annotation, http://solgenomics.net) and subsequently to a 61 kb region between markers P120 and P140 (Figure 4.1c and Table 4.1). Six putative genes were found in this 61 kb region and sequencing of one of these (subsequently termed *CD1*) revealed a polymorphism (G→A) between M82 and *cd1* in the last exon that showed perfect cosegregation with Marker T1968 (position 11.005), a previously developed marker corresponding to an EST encoded by SL2.40sc03748. The chromosomal position of *CD1* is: SL2.40ch11:1004362…1007293. Genbank accession number for CD1: JF968592.

Table 4.1. Primer sequences, PCR product length and phenotyping methods used for mapping markers P110, P120, P140 and P160.

Primers	Primer sequence	PCR product length (bp)	Polymorphism (position on scaffold SL1.03sc01386, *S. lycopersicum/ S. pimpinellifolium*) or Enzyme
P110F	CGCCCATCATATGCCAACCTCCT	1009	2154299 GA/AT
P110R	CCCTTTTTGGTCTGCGGAAG		
P120F	CACTATGCAGTGGGCGTTTG	1025	2183872 C/T
P120R	ACGTTTTCAGCCACCTAAACC		
P140F	GGAGAGCGAGATTCGCTAGG	980	2244420 T/A
P140R	GCAATTTAAGAGTGCAGGCGTAC		
P160F	TCTGTTTCGGCAGGCAACTG	1014	*Fok*I
P160R	TCAAACCCGCGACCGGTGTGA		

General molecular biology techniques

RNA extraction, cDNA synthesis and qRT-PCR analysis were performed as described in ref. 8. For transgenic complementation, the coding sequence of *CD1* was introduced into pCAMBIA1305.1 using standard PCR-based cloning protocols. The resulting *CaMV35S::CD1* plasmid was introduced via *Agrobacterium tumefaciens*-mediated transformation into calli generated from *cd1* mutant seeds at the Boyce Thompson Institute for Plant Research, Ithaca, NY.

Tissue specific CD1 gene expression

Pericarp tissues (epidermis, collenchyma, parenchyma, vascular tissues and inner epidermis) were laser microdissected from 10 DPA (days post anthesis) tomato fruits

following the protocol described in Nakazono *et al.* (2003). Total RNA was isolated from each tissue sample and mRNA amplified as described in Matas *et al.* (2010). A double strand cDNA library from each sample of amplified RNA was sequenced with a GS-FLX (454, Roche; Genomics core at Memorial Sloan-Kettering Cancer Center Laboratory Facility) following manufacturer's protocols. The raw 454 sequences were base called using the Pyrobayes base caller (Quinlan *et al.*, 2008) then processed to remove low quality regions and adaptor sequences using the programs LUCY (Quinlan *et al.*, 2008) and SeqClean (http://compbio.dfci.harvard.edu/tgi/software). The resulting high quality sequences were then screened against the NCBI UniVec database, *E. coli* genome sequences, and tomato ribosomal RNA, to remove sequence contamination. Sequences shorter than 30bp were discarded. Tomato cDNA sequences were assembled into unigenes using the iAssembler program (Guo *et al.*, 2010). Following cDNA sequence assembly, digital expression information of each unigene was derived following normalization to the total number of sequenced transcripts per sample.

Light Microscopy

Fixation and embedding was performed as in Buda *et al.* (2009). Four micron cryosections of each sample were obtained using a Microm HM550 cryostat (ThermoFisher Scientific). Sections were melted onto room temperature VistaVision Histobond (VWR) slides and air dried. The slides were then heated on a hot plate at 200°C for 2 min immediately prior to staining. Preparation of Sudan Red 7b followed the protocol outlined in Brundrett *et al.* (1991). The staining solution was applied

directly to the slides in a humidity chamber and left for 1 h. The slides were washed and mounted with 75% glycerol. The stained slides were viewed on an AxioImager A1 microscope (Zeiss,) using Zeiss EC-Plan NeoFluar 40x/0.75 dry and 100x/1.3 oil immersion objectives, a Zeiss AxioCam MRc color video camera and Zeiss Axio Vs40 4.6.3.0 software. The images were taken using differential interference contrast optics (DIC). Images were processed using Photoshop CS4 software (Adobe) to adjust levels and color balance.

Recombinant protein production in E. coli

The *CD1* coding region lacking the 57 nucleotides encoding the predicted native signal peptide and the stop codon was amplified by PCR and cloned into the pET-26b(+) vector (Novagen/Merck). The resulting *pet26b(+)::CD1* plasmid, encoding the *pelB* leader sequence followed by the *CD1* mature sequence and a C-terminal hexahistidine tag was transformed into BL21(DE3) *E. coli*. Protein expression was induced and inclusion bodies were harvested according to the pET system manual (Novagen/Merck). Protein was solubilized and purified using a 1 mL HisTrap FF column with denaturing conditions according to the product manual (GE Healthcare Life Sciences).

Antibody-based techniques

A polyclonal antibody to recombinant CD1 was produced in rabbits (Pacific Immunology Corporation). To generate a high titer antiserum, CD1 specific antibodies were purified by absorption to purified CD1 immobilized on a PVDF membrane

(Ritter, 1991). For Western blot analysis, protein was extracted by boiling frozen ground tissue in 3X Laemmli buffer (6% SDS, 30% glycerol, 300 mM DTT, 187 mM Tris, pH 6.8) for 5 min. Protein concentration was determined by Bradford assay (Bio-Rad) and 25 µg were analyzed by Western blot using an HRP conjugated secondary antibody and Pierce ECL Western Blotting Substrate (Thermo Scientific). For immunolocalization experiments, tissue sections on Formvar-coated nickel grids were prepared as previously described (Domozych *et al.*, 2009) with minor modifications: Cacodylate buffers were 0.1 M, initial fixation was with 1% glutaraldehyde at room temperature. The grids were incubated in 5% H_2O_2 for 5 min, washed with deionized water (dH_2O), incubated in 0.1 M NH_4Cl for 20 min, washed with dH_2O and then blocked for 30 min with 1% fat free milk in phosphate buffered saline containing 0.1% Tween-20 (PBS-T). The grids were washed with dH_2O and incubated in primary antibody diluted 1/25 in PBS-T at 4^{o}C overnight. The grids were washed, blocked again and incubated in the anti-rabbit antibody conjugated to 15 nm gold particles (diluted 1/100 in PBS-T) for 2 h at 37^{o}C. The grids were washed extensively with dH_2O, stained in conventional uranyl acetate/lead citrate, washed again with dH_2O and dried. The sections were viewed with a Zeiss Libra 120 transmission electron microscope (TEM).

Chemical analysis

All cuticle isolations were performed as described in ref. 4. For cutin monomer analysis, 0.32 cm^2 disks of dewaxed cuticles were depolymerized using a base-catalyzed transmethylation protocol (Li-Beisson *et al.*, 2010). The subsequent dry

extracts were derivatized by heating with 10 µL N,O-bis(trimethylsilyl)trifluoroacetamide (BSTFA) and 10 µL pyridine for 10 min at 90˚C. They were then evaporated to dryness by heating under a gentle stream of nitrogen, resuspended in 100 µL of chloroform and analyzed by GC protocol 1 (Table 4.2). Compounds were identified based on retention behavior compared to authentic standards and by separate analyses employing the GC-MS instrument described in protocol 2 (Table 4.2). For chemical analysis of soluble surface lipids, M82 and *cd1* fruits were harvested at the immature green developmental stage. Soluble surface lipids were extracted by immersing 6 fruits from each genotype twice for 30 s in two beakers containing 2:1 mixture of chloroform and methanol. The extracts were combined and dried over anhydrous sodium sulfate. The solutions were then filtered and the solvent was evaporated by rotary evaporation. Trimethylsilyl (TMS) derivatization of the surface extracts was performed as described above and the samples were analyzed by GC-MS according to protocol 2 (Table 4.2). Acetyl derivatives of the extract were prepared by heating to 50˚C with a 1:1 mixture of pyridine and acetic anhydride for 3 h and analyzed by the same GC-MS method. For the spectra shown, background from neighboring peaks was subtracted.

Table 4.2. Gas chromatography parameters.

	Protocol 1	Protocol 2	Protocol 3
Application	Cutin analysis	Identification of 2-MHG	Acyltransferase assay
Instrument	Agilent 6850	Agilent 6890	Agilent 6890
Inlet	Cool-on-column	Splitless, 320˚C	Splitless, 320˚C
Carrier gas	He, 1.4 mL.min^{-1}	He, 1 mL.min^{-1}	He, 1 mL.min^{-1}
Column	DB-1 (30 m x 320 µm x 0.1 µm)	DB-5MS (30 m x 250 µm x 0.25 µm)	HP-5 (30 m x 250 µm x 0.25 µm)
Oven program	2 min 50˚C ramp 40˚C.min^{-1} to 120˚C hold 2 min ramp 10˚C.min^{-1} to 320˚C hold 15 min	2 min 50˚C ramp 40˚C.min^{-1} to 200˚C hold 2 min ramp 3˚C.min^{-1} to 320˚C hold 12 min	2 min 140˚C ramp 20˚C.min^{-1} to 320˚C hold 2 min
Detector	FID	Mass spectrometer (JEOL GC Mate II, Electron Impact mode with accelerating voltage of 2500, scanning from 35-750 at a rate of 0.3 s.scan^{-1} with a 0.1 s interscan delay)	FID

Expression of functional CD1 in Nicotiana benthamiana

Despite trying a variety of bacterial expression conditions and vectors, no soluble recombinant CD1 protein was produced. We therefore used a transient expression system with agroinfiltrated *Nicotiana benthamiana* leaves. The coding sequence of CD1 lacking the stop codon was inserted into pEAQ-DEST3 (Sainsbury *et al.*, 2009) using a combination of PCR and Gateway cloning. The resulting construct encoded CD1 followed by an engineered TEV protease cleavage site, several amino

acids from the vector and finally a hexahistidine tag. Infiltration of *N. benthamiana* with *A. tumefaciens* was as described by Sainsbury *et al.* (2009) except that strain GV2260 was used and plants were harvested 5 days after infiltration. Tissue was flash frozen, ground in liquid nitrogen and stored at -80˚C until use.

For protein purification, 50 g of frozen tissue was added to 150 mL of chilled buffer A (50 mM Na_2HPO_4, 500 mM NaCl, 10 mM β-mercaptoethanol, pH 7.0) and 10 g of polyvinylpolypyrrolidone. Tissue was homogenized using a PowerGen 125 homogenizer (Fisher Scientific), the extract was clarified and imidazole was added to a 5 mM final concentration. The extract was rocked on ice with 200 µL (settled volume) of equilibrated HisPur Ni-NTA agarose (Pierce). After 2 h, the resin was collected by centrifugation and transferred to a 5 mL column. The beads were then washed with 5 mL of buffer with 50 mM imidazole and without β-mercaptoethanol. Protein was then eluted by 250 µL of buffer with 300 mM imidazole. The elution step was repeated 3 more times and the eluates were pooled and applied to an equilibrated Superdex 75 HR 10/30 column (GE Healthcare Life Sciences). Fractions containing the 41 kD CD1 protein were pooled and concentrated by ultrafiltration. Protein concentration was determined by absorbance at 280 nm using the theoretical extinction coefficient base on the sequence of the protein (Gill and von Hippel, 1989). The correct identity of the protein was confirmed by trypsin digest and MALDI TOF/TOF (Applied Biosystems, 4700 Analyzer; http://cores.lifesciences.cornell.edu/brcinfo/?f=2) mass spectrometry (MASCOT total protein score 843; Matrix Science) searching predicted proteins from the tomato genome (http://solgenomics.net).

Acyltransferase assay

2-monohexadecanoylglycerol and 1-monohexadecanoyl-*rac*-glycerol were obtained from Santa Cruz Biotechnology and hexadecanoic acid, methyl hexadecanoate and trihexadecanoylglycerol from Sigma Aldrich. The *Rhizopus niveus* lipase was purified from crude enzyme (Sigma Aldrich) by size exclusion chromatography. Substrate emulsions were prepared by dissolving 1 mol equivalent of acyl donor in 33 mol equivalent of decanol. Ten microliters of the solution, containing 1.5 μmol acyl donor and 50 μmol decanol, was added to 100 μL of a 5% aqueous gum arabic solution. The mixture was vortexed for 30 s and then sonicated for 5 min at room temperature. For each 10 μL reaction, 2.5 μL of the substrate emulsion was further diluted and buffered to a final concentration of 4 mM acyl donor, 130 mM decanol, 50 mM Na_2HPO_4 (pH 6.0), 150 mM NaCl, 1.25% (w/v) gum arabic. Reactions were initiated by addition of 12.3 pmol of either purified CD1 or the purified *Rhizopus niveus* lipase and incubated at 37˚C for 3 h with 300 rpm shaking. The reaction was stopped by immediately freezing at -80˚C. For analysis of the reaction products, 500 μL of 1M acetic acid was added to the thawed sample. Lipids were extracted with 500 μL of chloroform containing 0.02 μg μL^{-1} methyl heptadecanoate as an internal standard. The extracts were dried over anhydrous sodium sulfate and then derivatized as previously described. Products were analyzed by GC using protocol 3 (Table 4.2). An authentic standard of decyl hexadecanoate was synthesized by solvent free acid catalyzed esterification of hexadecanoic acid in decanol. The product accumulation was quantified relative to the internal standard

assuming a mass response factor of 1 and dividing by the molar mass of decyl hexadecanoate. No acyltransferase activity was detected when CD1 was first denatured by boiling for 10 min.

Acknowledgements

We thank George Lomonossoff (John Innes Centre) and Plant Bioscience Limited for the pEAQ vector, Mark Toso for help with the TEM and Mike Pollard (Michigan State University) and Karl Niklas (Cornell University) for critical discussion. This work was supported by grants from the National Science Foundation (Plant Genome Program; DBI-0606595), the United States-Israel Binational Agricultural Research and Development Fund (IS-4234-09) and the USDA Cooperative State Research, Education and Extension Service (grant number #2006-35304-17323).

References

Akoh, C.C., Lee, G.C., Liaw, Y.C., Huang, T.H. and Shaw, J.F. (2004) GDSL family of serine esterases/lipases. *Prog. Lipid Res.* **43**, 534–552.

Baker, E.A., Bukovac, M.J. and Hunt, G.M. (1982) Composition of tomato fruit cuticle as related to fruit growth and development. In *The Plant Cuticle* (Cutler, D.F., Alvin, K.L. and Price, C.E., eds). London: Academic Press, pp. 33-44.

Beisson, F., Li, Y., Bonaventure, G., Pollard, M. and Ohlrogge, J.B. (2007) The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of Arabidopsis. *Plant Cell* **19**, 351–368.

Brundrett, M.C., Kendrick, B. and Peterson, C.A. (1991) Efficient lipid staining in plant material with sudan red 7B or fluorol yellow 088 in polyethylene glycol-glycerol. *Biotech. Histochem.* **66**, 111–116.

Buda, G.J., Isaacson, T., Matas, A.J., Paolillo, D.J. and Rose, J.K. (2009) Three-dimensional imaging of plant cuticle architecture using confocal scanning laser microscopy. *Plant J.* **60**, 378–385.

Croteau, R. and Kolattukudy, P.E. (1974) Biosynthesis of hydroxyfatty acid polymers. Enzymatic synthesis of cutin from monomer acids by cell-free preparations from the epidermis of *Vicia faba* leaves. *Biochemistry* **13**, 3193–3202.

Domozych, D.S., Sorensen, I. and Willats, W.G. (2009) The distribution of cell wall polymers during antheridium development and spermatogenesis in the Charophycean green alga, *Chara corallina*. *Ann. Bot.* **104**, 1045–1056.

Edwards, D. (1993) Cells and tissues in the vegetative sporophytes of early land plants. *New Phytologist* **125**, 225–247.

Gill, S.C. and von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319–326.

Graca, J. and Lamosa, P. (2010) Linear and branched poly(omega-hydroxyacid) esters in plant cutins. *J. Agric. Food Chem.* **58**, 9666–9674.

Graca, J., Schreiber, L., Rodrigues, J. and Pereira, H. (2002) Glycerol and glyceryl esters of omega-hydroxyacids in cutins. *Phytochemistry* **61**, 205–215.

Guo, S., Zheng, Y., Joung, J.G., Liu, S., Zhang, Z., Crasta, O.R., Sobral, B.W., Xu, Y., Huang, S. and Fei, Z. (2010) Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types. *BMC Genomics* **11**, 384.

Haas, M.J. and Joerger, R.D. (1995) Lipases of the Genera *Rhizopus* and *Rhizomucor*: Versalitle Catalysts in Nature and the Laboratory. In *Food Biotechnology: Microorganisms* (Hui, Y.H. and Khachatourians, G.G., eds). New York: VCH Publishers, pp. 549-588.

Heredia, A., Heredia-Guerrero, J.A., Dominguez, E. and Benitez, J.J. (2009) Cutin synthesis: A slippery paradigm. *Biointerphases* **4**, P1–3.

Irshad, M., Canut, H., Borderies, G., Pont-Lezica, R. and Jamet, E. (2008) A new picture of cell wall protein dynamics in elongating cells of *Arabidopsis thaliana*: confirmed actors and newcomers. *BMC Plant Biol.* **8**, 94.

Isaacson, T., Kosma, D.K., Matas, A.J. *et al.* (2009) Cutin deficiency in the tomato fruit cuticle consistently affects resistance to microbial infection and biomechanical properties, but not transpirational water loss. *Plant J.* **60**, 363–377.

Jetter, R., Kunst, L. and Samuels, A.L. (2006) Composition of plant cuticular waxes. In *Biology of the Plant Cuticle* (Riederer, M. and Müller, C., eds). Oxford, UK: Blackwell, pp. 145-181.

Jonas, A. (2000) Lecithin cholesterol acyltransferase. *Biochim. Biophys. Acta* **1529**, 245–256.

Li-Beisson, Y., Shorrosh, B., Beisson, F. *et al.* (2010) Acyl-Lipid Metabolism. *The Arabidopsis Book* **8**, e0133.

Matas, A.J., Agusti, J., Tadeo, F.R., Talon, M. and Rose, J.K. (2010) Tissue-specific transcriptome profiling of the citrus fruit epidermis and subepidermis using laser capture microdissection. *J. Exp. Bot.* **61**, 3321–3330.

Mintz-Oron, S., Mandel, T., Rogachev, I. *et al.* (2008) Gene expression and metabolism in tomato fruit surface tissues. *Plant Physiol.* **147**, 823–851.

Nakazono, M., Qiu, F., Borsuk, L.A. and Schnable, P.S. (2003) Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: identification of genes expressed differentially in epidermal cells or vascular tissues of maize. *Plant Cell* **15**, 583–596.

Nawrath, C. (2006) Unraveling the complex network of cuticular structure and function. *Curr. Opin. Plant. Biol.* **9**, 281–287.

Nicolaides, N., Soukup, V.G. and Ruth, E.C. (1983) Mass spectrometric fragmentation patterns of the acetoxy and trimethylsilyl derivatives of all the positional isomers of the methyl hydroxypalmitates. *Biol. Mass Spectrom.* **10**, 441–449.

Osman, S.F., Irwin, P., Fett, W.F., O'Connor, J.V. and Parris, N. (1999) Preparation, isolation, and characterization of cutin monomers and oligomers from tomato peels. *J. Agric. Food. Chem.* **47**, 799–802.

Panikashvili, D., Shi, J.X., Schreiber, L. and Aharoni, A. (2009) The Arabidopsis *DCR* encoding a soluble BAHD acyltransferase is required for cutin polyester formation and seed hydration properties. *Plant Physiol.* **151**, 1773–1789.

Pollard, M., Beisson, F., Li, Y. and Ohlrogge, J.B. (2008) Building lipid barriers: biosynthesis of cutin and suberin. *Trends Plant Sci.* **13**, 236–246.

Quinlan, A.R., Stewart, D.A., Stromberg, M.P. and Marth, G.T. (2008) Pyrobayes: an improved base caller for SNP discovery in pyrosequences. *Nat. Methods* **5**, 179–181.

Reina, J.J., Guerrero, C. and Heredia, A. (2007) Isolation, characterization, and localization of *AgaSGNH* cDNA: a new SGNH-motif plant hydrolase specific to Agave americana L. leaf epidermis. *J. Exp. Bot.* **58**, 2717–2731.

Ritter, K. (1991) Affinity purification of antibodies from sera using polyvinylidenedifluoride (PVDF) membranes as coupling matrices for antigens presented by autoantibodies to triosephosphate isomerase. *J. Immunol. Methods* **137**, 209–215.

Sainsbury, F., Thuenemann, E.C. and Lomonossoff, G.P. (2009) pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol. J.* **7**, 682–693.

Volokita, M., Rosilio-Brami, T., Rivkin, N. and Zik, M. (2011) Combining comparative sequence and genomic data to ascertain phylogenetic relationships and explore the evolution of the large GDSL-lipase family in land plants. *Mol. Biol. Evol.* **28**, 551–565.

Yahya, A.R.M., Anderson, W.A. and Moo-Young, M. (1998) Ester synthesis in lipase-catalyzed reactions. *Enzyme Microb. Tech.* **23**, 438–450.

Yang, W., Pollard, M., Li-Beisson, Y., Beisson, F., Feig, M. and Ohlrogge, J. (2010) A distinct type of glycerol-3-phosphate acyltransferase with *sn-2* preference and phosphatase activity producing 2-monoacylglycerol. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 12040–12045.

Yeats, T.H., Howe, K.J., Matas, A.J., Buda, G.J., Thannhauser, T.W. and Rose, J.K. (2010) Mining the surface proteome of tomato (*Solanum lycopersicum*) fruit for proteins associated with cuticle biogenesis. *J. Exp. Bot.* **61**, 3759–3771.

CHAPTER 5

Conclusions and future directions

These studies have demonstrated the utility of studying cuticle biology using tomato fruit as a model system. However, an important consideration is how relevant conclusions drawn from the study of tomato fruit are to cuticles in general, such as leaf cuticles, and to cuticles in other species. As described in the introduction, arabidopsis has an unusual leaf and stem cuticle with a unique cutin composition. While the chemical composition of tomato fruit is fairly typical, it is about 100 times more substantial than typical leaf cuticles. More specifically, this represents an extreme accumulation of cutin beyond what is apparently necessary for 'normal' cuticle function. In contrast, tomato fruit wax is relatively typical in its abundance. Does this discrepancy limit tomato fruit as a model system, and why are fruit cuticles so substantial, particularly in the amount of cutin?

It is well established that wax rather than cutin is the essential determinant of cuticular water permeability (Burghardt and Riederer, 2006), so the requirement for large amounts of cutin in the tomato fruit is not likely an adaptation to specifically resist fruit desiccation. The relatively modest increase in cuticular water permeability seen in the severely cutin deficient *cd* mutants suggests that this is also the case in tomato fruit (Isaacson *et al.*, 2009). Rather, it has been suggested that the thick, cutin-rich cuticle of the tomato and many other fleshy fruits provides an important barrier against microbial penetration (Saladié *et al.*, 2007). Thus, the likely role of the cuticle of the mature fleshy fruit is to provide a resilient surface structure that preserves the

quality of nutrient-rich fleshy fruits so that they can attract seed dispersing animals before undergoing microbial spoilage.

In many cases, the structure and the composition of the cuticle are organ specific (Walton, 1990), suggesting the presence of organ-specialized cuticle biosynthesis pathways. Thus, while mutation of components of the vegetative cuticle biosynthetic pathways may be lethal, mutants with fruit-specific phenotypes are likely to be viable, as was the case with the *cd* mutants (Isaacson *et al.*, 2009). Studies of gene expression in tomato peels have also indicated that many putative orthologs of the cuticle biosynthetic pathways known from arabidopsis are expressed in tomato fruit (Mintz-Oron *et al.*, 2008), suggesting that the dramatic accumulation of cutin in tomato fruit is the result of existing pathways being upregulated rather than being newly evolved. In summary, although somewhat specialized in its function, the tomato fruit cuticle is a valid model for the study of general cuticle biosynthetic pathways. However, the specific functional adaptation of cuticles of fleshy fruits must be considered when interpreting phenotypes in a physiological context.

In Chapter 2, the survey of wild tomato species confirmed that there were significant differences in the structure and chemical composition of the cuticles of these species. As discussed in that chapter, an important future prospect will be to explore the functional significance of these differences by directly measuring functional parameters of the cuticle, particularly cuticular water permeability. A well designed QTL-mapping study combining chemical and structural phenotyping with measurement of cuticular water permeability in an introgression line population will help to dissect the complex interaction between cuticle structure, chemistry and the

core function of the cuticle as a transpirational barrier.

A more challenging project will be to dissect the genetics that underlie the greatly enhanced wax accumulation in *S. neorickii* because there are no existing introgression line populations from this species. An interesting consideration is that *S. neorickii* is resistant to *Botrytis cinerea* and putative QTLs associated with this trait have been identified (Finkers *et al.*, 2008). Given the previously demonstrated dependence of microbial resistance on the properties of the cuticle (L'haridon *et al.*, 2011; Isaacson *et al.*, 2009; Bourdenx *et al.*, 2011), it will be useful to see how a dramatic increase in wax coverage affects fruit resistance to microbial infection. Additionally, if cuticular water permeability is shown to be decreased by enhanced accumulation of cuticular wax, there may be significant improvements to postharvest shelf life and texture, as has been previously discussed with the DFD cultivar (Saladie *et al.*, 2007). Introgression of cuticle related traits from *S. neorickii* could thus form the basis of a plant breeding program aimed at improving tomato texture and shelf life.

The proteomic approach undertaken in Chapter 3 identified several candidate genes/proteins that may be involved in cuticle biology. RNAi lines aimed at knocking down several of these genes have been created and T1 seeds have been harvested and stored. The genes targeted include two GDSLs, represented by the unigenes SGN-U579520 and SGN-U583101, which are in a distinct clade from *CD1* (SGN-U585129). Since the GDSL family is particularly large, it is possible that any phenotype in these knockdown lines may be subtle because of genetic redundancy. However, since these represent a distinct clade of GDSLs, they may have functions related to different aspects of cutin metabolism, perhaps acting to modify the cutin

polymer during fruit growth. The other genes that I targeted are candidates potentially involved in extracellular lipid trafficking. These include two LTPs (SGN-U581465 and SGN-U579033) and an MD-2-related lipid recognition domain-containing (ML) protein (SGN-U577903).

The focus of Chapter 4 was characterization of the CD1 protein, an extracellular acyltransferase that is required for cutin synthesis. The biochemical characterization of CD1, together with the accumulation of 2-mono(10,16-dihydroxyhexadecanoylglycerol) (2-MHG) in the *cd1* mutant strongly suggests that this is the native substrate of CD1. A priority for future experiments aimed at confirming the role of CD1 in cutin polymerization is the purification or synthesis of this compound in sufficient quantities for *in vitro* biochemical assays to test the polymerase activity of CD1. The structural characterization of cutin oligomers formed in such assays should reveal additional aspects of substrate and product specificity. More specifically, studies of tomato cutin have indicated that essentially all ω-hydroxyl groups of 10,16-dihydroxyhexadecanoic acid are esterified, while roughly half of the midchain hydroxyls are esterified (Deas and Holloway, 1976). The question of whether the CD1-acyl intermediate has a preference for esterification of primary or secondary hydroxyls has not yet been addressed. If CD1 catalyzes the same proportion of linkages *in vitro* as that observed *in vivo*, this will support the hypothesis that CD1 is the primary catalyst of cutin polymerization, rather than an enzyme whose activity is prerequisite for additional activities.

Although cutin deficiency in *cd1* fruit is severe, about 5% of the total cutin remains, suggesting additional mechanisms of cutin formation. These could be

explained by the action of enzymes similar to CD1, since *CD1* is one of five closely related tomato genes. This question can be addressed by expression profiling of the *CD1* gene family and the generation of RNAi knockdown plants specifically targeting the other members of this family. It will also be interesting to see whether members of the CD1 family, or other less closely related GDSL-motif lipase/hydrolase family proteins, are involved in the synthesis of the suberin polyester. Suberin shares several structural similarities with cutin and there appears to be some overlap between their respective biosynthetic pathways (Pollard *et al.*, 2008).

Experimental plant biology is entering a golden age for non-arabidopsis model organisms. While plant biology will always rely on the foundation built on arabidopsis research, with the increasing ease and quality of genome sequencing, essentially any plant can be a genetic model organism. Useful and easy to characterize phenotypes are again becoming the key characteristics of a 'model' organism, just as they were for classical studies in biochemistry and physiology. These studies and the interpretation of their results highlight this shifting paradigm. The genetic resources of tomato and known biosynthesis pathways from arabidopsis allowed for a survey of the wild relatives of tomato to be interpreted in a genetic context that can be tested, establishing a promising model system for studying the evolution and genetics of cuticular diversity. The most obvious feature of the tomato cuticle, the fact that it is extremely thick and cutin rich, provided an ideal system for identifying a critical gene of cutin biosynthesis, *CD1*, both through proteomics and forward genetics. Moreover, it is unlikely that the putative substrate of this enzyme, 2-MHG, would have been detected in the corresponding arabidopsis mutant, where it would be expected to be several

orders of magnitude lower in abundance. In conclusion, arabidopsis has established a framework for understanding plant cuticle biosynthesis, but its greatest utility in the future will be in consolidating and interpreting results from tomato and other species.

References

Bourdenx, B., Bernard, A., Domergue, F. *et al.* (2011) Overexpression of Arabidopsis *ECERIFERUM1* promotes wax very-long-chain alkane biosynthesis and influences plant response to biotic and abiotic stresses. *Plant Physiol.* **156**, 29–45.

Burghardt, M. and Riederer, M. (2006) Cuticular transpiration. In *Biology of the Plant Cuticle* (Riederer, M. and Müller, C., eds). Oxford, UK: Blackwell, pp. 292-311.

Deas, A.H.B. and Holloway, P.J. (1976) The intermolecular structure of some plant cutins. In *Lipids and Lipid Polymers in Higher Plants* (Tevini, M. and Lichtenthaler, H.K., eds). Berlin: Springer-Verlag, pp. 293-299.

Finkers, R., Bai, Y.L., van, d.B., P, van, B., R, Meijer-Dekens, F., ten, H., A, van, K., J, Lindhout, P. and van, H., AW (2008) Quantitative resistance to *Botrytis cinerea* from *Solanum neorickii*. *Euphytica* **159**, 83–92.

Isaacson, T., Kosma, D.K., Matas, A.J. *et al.* (2009) Cutin deficiency in the tomato fruit cuticle consistently affects resistance to microbial infection and biomechanical properties, but not transpirational water loss. *Plant J.* **60**, 363–377.

L'haridon, F., Besson-Bard, A., Binda, M. *et al.* (2011) A permeable cuticle is

associated with the release of reactive oxygen species and induction of innate immunity. *PLoS Pathog.* **7**, e1002148.

Mintz-Oron, S., Mandel, T., Rogachev, I. *et al.* (2008) Gene expression and metabolism in tomato fruit surface tissues. *Plant Physiol.* **147**, 823–851.

Pollard, M., Beisson, F., Li, Y. and Ohlrogge, J.B. (2008) Building lipid barriers: biosynthesis of cutin and suberin. *Trends Plant Sci.* **13**, 236–246.

Saladie, M., Matas, A.J., Isaacson, T. *et al.* (2007) A reevaluation of the key factors that influence tomato fruit softening and integrity. *Plant Physiol.* **144**, 1012–1028.

Walton, T.J. (1990) Waxes, cutin and suberin. In *Methods in plant biochemistry: lipids, membranes and aspects of photobiology.* (Harwood, J.L. and Bowyer, J.R., eds). London: Academic Press, pp. 105-158.

APPENDIX

Supplementary material for Chapter 3

Supplementary Table S1. PCR primers used for gene expression analysis

Target	Forward Primer Sequence	Reverse Primer Sequence
SGN-U581377 (*RPL2*)	CAGCGGATGTCGTGCTATGAT	GGGATGCTCCACTGGATTCA
SGN-U566767 (*LeCer6*)	CCAGTGTTCATCCCAGAGATTGTC	GTCTGAGAGCTCACACACGTT
SGN-U583101	CTGCAGCTGCTGGAATTAGAGATG	ACCACTTGTTGCACTGTGTTCCTG
SGN-U585129	GTAGCATGTTGTGGACAAGGACCA	TTTGCCCTCTCAGATGGATGGAAC
SGN-U581465	CCCTCTACTGACTGCTCTAAAGTTC	TCGAAACAAGACTCGAGTACATACG
SGN-U579687	GACTGCTCGAAGGTTCAGTAAGGT	ACAACCAACCACCAAAGTTCATTAC
SGN-U577903	GCTTCAGCATCAGTTTGTTCGACG	TTGTCACAAGGTTCTTGGACTGCC
SGN-U570812	GGAGGGTGTCAAGTTGGCAATGTT	CTTCCAAGCATCATGACAGTGGCT
SGN-U567908	TGCAGCATCAAGATTGGGATTCAC	CAACTTAGGTCCACACTACCTAAGCA
SGN-U577872	GTCATGGCTACTGGACCAATGAGA	ATCCACGACAATCACCACCGGAAA
SGN-U579545	ACTCAAGTAGTCTGGCGCAACTCA	AGTAAGGACGTTGTCCGATCCAGT
SGN-U580366	AATCACCAAAGCCTTCTTGCCACG	ACGACCACATTCTAGGCCACCATT

Supplementary Table S2. Complete MASCOT results for gel-band analysis

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
1	SGN-U583542	38058	116	3	6.9	34.83	GVGPYTGVGDGR	
1						41.4	VLIALSGDTTGR	
1						74.62	VLIALSGDTTGR	
2	SGN-U577720	85298	114	4	7.6	26.34	IPSTEEIADR	
2						47.59	GVTAFGFDLVR	
2						70.36	YGAGIGPGVVYDIHSPR	
2						26.79	ALSGAKDEAFFSANAAAQASR	
3	SGN-U577900	67191	105	9	7.1	21.01	YQLATSR	
3						44.89	YQLATSR	
3						30.42	YQLATSR	
3						23.64	IPPMFDR	
3						23.66	VFPLAKLDR	
3						19.85	FDVFLNVDK	
3						24.06	EGSSLYDEKR	
3						19.12	EGSSLYDEKR	
3						47.07	EGSSLYDEKR	
4	SGN-U585664	63110	90	4	8.3	46.24	IGLFGGAGVGK	
4						37.28	VVDLLAPYQR	
4						44.6	AHGGFSVFAGVGER	
4						19.52	FTQANSEVSALLGR	
5	SGN-U572041	105097	79	7	1.8	29.71	STFVYVPR	
5						28.06	STFVYVPR	
5						36.31	STFVYVPR	
5						32.15	STFVYVPR	
5						23.11	STFVYVPR	
5						19.02	STFVYVPR	
5						20.68	SSLFNIVPR	
6	SGN-U577665	24219	65	1	7.3	64.51	NAVFGDSSALAPGGVR	
7	SGN-U585234	33590	62	2	8	46.24	IGLFGGAGVGK	
7						36.16	AHGGVSVFGGVGER	
8	SGN-U579520	41889	60	6	7.7	27.4	GVNYASGAAGIR	
8						30.29	GVNYASGAAGIR	
8						27.81	ANYPPYGIDFPDGPTGR	
8						25.69	ANYPPYGIDFPDGPTGR	
8						21.97	ANYPPYGIDFPDGPTGR	
8						24.86	ANYPPYGIDFPDGPTGR	
9	SGN-U581155	39394	59	1	3.3	58.99	AVVDSAIDAETR	
10	SGN-U573941	26066	57	4	3.5	32.51	YFVLPSLR	
10						23.81	YFVLPSLR	
10						24.03	YFVLPSLR	
10						28.01	YFVLPSLR	
11	SGN-U582837	48616	57	3	6.1	34.79	SFVPIASGR	
11						33.24	SPNFSGTLR	
11						33.23	ILFDVPNSR	
12	SGN-U580132	69361	55	1	1.6	54.64	LHDDLVAGFR	
13	SGN-U586603	17334	54	2	5.2	43.27	VVSVSIPR	
13						28.95	VVSVSIPR	
14	SGN-U580766	59101	49	4	3.9	19.02	FEGLDAFR	
14						24.83	FEGLDAFR	
14						39.69	SNGDIALDFYHR	
14						23.12	SNGDIALDFYHR	
15	SGN-U584870	27327	48	2	9.2	26.67	WGDTYVAVR	
15						38.51	QFSGFIADDDLPR	
16	SGN-U578349	26593	47	1	3.8	46.68	GLVGEIISR	
17	SGN-U579420	42380	43	2	4.7	35.94	GQNPVFPR	
17						27.51	GTFFGNYKPR	
18	SGN-U586194	36110	41	2	8.3	33.83	APGGAPANVAIAVTR	
18						26.19	EAGALLSYDPNLR	
19	SGN-U577711	42200	40	3	9	37.6	AGFAGDDAPR	
19						19.84	GEYDESGPSIVHR	
19						19.98	IWHHTFYNELR	
20	SGN-U564978	22435	40	1	5.2	39.73	NCNPSIQYGR	
21	SGN-U583862	9444	38	1	11.8	38.4	LTHGAPGDEIR	
22	SGN-U579059	38956	38	1	2.6	38.19	TSQVIDLSR	
23	SGN-U563303	26302	38	1	4.8	37.59	VLVPYDFLAGR	
24	SGN-U584963	57821	37	2	4.5	35.29	LIESAAPGIISR	
24						19.18	VVNALAKPIDGR	
25	SGN-U583101	41833	37	2	7.5	29.28	ISFSGQVNNYR	
25						24.7	ANYLPYGIDFSGGPTGR	
26	SGN-U570812	63434	36	1	1.9	35.66	IGGTIFDAAGR	
27	SGN-U581597	34199	35	2	3.7	23.2	AFQDLDGPEYR	
27						30.07	AFQDLDGPEYR	
28	SGN-U578421	52835	35	1	3.2	35.13	NGGIDTEEDYPYKER	
29	SGN-U577856	20019	34	2	12.4	29.7	INFVEGGPIK	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
29						25	YSLIEGDVLGDK	
30	SGN-U565452	23005	33	1	3.9	32.93	DDFVEQDR	
31	SGN-U577505	39365	32	1	4.5	32.44	NIQNAISGAGLGNQIK	
32	SGN-U579794	30156	32	1	3.6	31.58	GQIYDVSQSR	
33	SGN-U569290	29567	31	2	8.4	28.32	LLANILYSYR	
33						21.61	GPGLYYVDSEGGR	
34	SGN-U574403	28843	31	2	8.9	28.3	GGFTLHSLTHR	
34						21.71	SFPAPNAHWSGR	
35	SGN-U562693	22357	30	1	6.4	30.24	VGYGGIPGPSGGR	
36	SGN-U572824	40253	30	1	2.4	29.89	ESALAQIIR	
37	SGN-U564876	24122	30	1	4.2	29.73	DTDILAAFR	
38	SGN-U577339	16858	30	1	6.1	29.71	NITVNEAQSR	
39	SGN-U578441	17786	30	2	14.1	25.63	QMNFVEGGPIK	Oxidation (M)
39						25	YSLIEGDVLGDK	
40	SGN-U569603	22988	29	1	3.9	29.41	NVALYQFR	
41	SGN-U576441	34153	29	1	3.9	29.13	AKPYPGHPTIVR	
42	SGN-U565096	27786	29	1	4	28.9	HITIFSPEGR	
43	SGN-U566611	28579	28	1	4.7	28.32	QDKSTSSTPKPR	
44	SGN-U579307	24065	28	1	3.7	28.01	LLDVYESR	
45	SGN-U577370	28337	28	1	5.1	27.67	VNQAIYLLTTGAR	
46	SGN-U578139	32942	28	2	2.6	19.02	LLILTDPR	
46						24.12	LLILTDPR	
47	SGN-U580213	38900	28	2	7.9	24.37	IGINGFGR	
47						23.19	GILGYTEDDVVSTDFVGDSR	
48	SGN-U581867	38591	27	3	3.3	22.06	LTMSESLGGKR	Oxidation (M)
48						21.74	LTMSESLGGKR	Oxidation (M)
48						21.61	LTMSESLGGKR	Oxidation (M)
49	SGN-U571799	29487	27	1	3.1	27.37	EVHNYLTR	
50	SGN-U573751	49012	27	1	2.2	27.29	AYLFPEGAAR	
51	SGN-U577839	20500	27	1	6.1	26.69	VCGHYTQVVWR	
52	SGN-U579787	17946	27	1	6.1	26.56	EVEYPGQVLR	
53	SGN-U578572	42060	26	1	3.1	26.05	VAPEVIAEYTVR	
54	SGN-U577585	54410	26	1	2.2	25.9	KFVSQGAYDTR	
55	SGN-U565260	27056	26	1	7.9	25.88	QAMTASSLALSGDTIAQLR	
56	SGN-U578288	13947	26	1	8.9	25.88	HLIEAADPQQK	
57	SGN-U579788	38480	26	2	4.5	24.37	IGINGFGR	
57						21.37	AVTVFGFR	
58	SGN-U582452	15293	25	1	11.2	25.39	IDYAPGGINPPHTHPR	
59	SGN-U578074	33067	25	1	4.1	25.35	DPDGYLFEIIQR	
60	SGN-U577903	19892	25	2	13.7	22.75	TTTFSITAETGR	
60						22.22	VSQIDITPYPIK	
61	SGN-U577181	26987	25	1	3.3	24.57	ANYPPYGR	
62	SGN-U590345	15845	24	1	6.5	24.5	GNVDIFTGR	
63	SGN-U569783	42863	24	1	2.3	24.22	VQTQPGFAR	
64	SGN-U577789	57697	24	1	2.3	24.19	FLVSDSFPGNER	
65	SGN-U583388	56028	24	1	2.3	23.9	EFPADVILGDDR	
66	SGN-U572904	13067	24	1	7.4	23.88	DEILEVAGK	
67	SGN-U579972	57924	23	1	1.7	23.45	YHTVFDYEK	
68	SGN-U603965	8564	23	1	14.7	23.42	RPRPRGCLQPR	
69	SGN-U579474	30447	23	1	3.5	23.27	YTTLKPLGDR	
70	SGN-U584746	27122	23	1	3.4	22.92	RQLSWNTR	
71	SGN-U578814	28460	23	1	4	22.87	LGIHEDSQNR	
72	SGN-U563994	47662	23	1	2.1	22.67	IEKNSIGGR	
73	SGN-U563063	24057	22	1	7.1	22.23	GGSPYGAGTFAGDGSR	
74	SGN-U583107	45188	22	1	2.8	22.08	ITFGGQVNNYR	
75	SGN-U569028	28670	22	1	4.6	22.05	TFIAIKPDGVQR	
76	SGN-U579512	66146	22	1	1.4	21.93	LISEFENAR	
77	SGN-U577277	72048	22	1	2.5	21.61	ATAGDTHLGGEDFDNR	
78	SGN-U564000	51428	21	1	2.2	21.36	LAVNLIPFPR	
79	SGN-U568214	44592	21	1	1.9	21.3	VLAFEAGR	
80	SGN-U594469	21118	20	1	7.9	19.94	EVIAVNQDPLGVQGR	
81	SGN-U574346	26287	20	1	3.9	19.9	LLHCPPYVK	
82	SGN-U571012	13062	20	1	5.6	19.58	KYYDTR	
83	SGN-U580857	29209	19	1	4.5	19.31	WSPSVADSAAGR	
84	SGN-U580653	36739	19	1	2.4	19.02	YLFEDGSR	

Supplementary Table S3. Complete MASCOT results for gel-slab analysis

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
1	SGN-U579520	41889	102	4	12.4	19.24	TLYHYGAR	
1						44.89	DESGIHLGDR	
1						59.35	GVNYASGAAGIR	
1						31.3	ANYPPYGIDFPDGPTGR	
2	SGN-U581590	17859	78	2	13.1	46.84	EHGAPEDEVR	
2						47.57	QIPLTGPQSIIGR	
3	SGN-U577711	42200	64	3	6.9	36.22	AGFAGDDAPR	
3						27.92	AGFAGDDAPR	
3						37.63	NYELPDGQVITIGAER	
4	SGN-U577903	19892	60	2	13.7	49.94	TTTFSITAETGR	
4						29.93	VSQIDITPYPIK	
5	SGN-U578193	49384	53	1	4	52.91	AAVPSGASTGIYEALELR	
6	SGN-U575256	65125	49	2	4.3	47.21	VGSDLSAFYR	
6						20.83	VITSSTEAQAYTPGR	
7	SGN-U565346	39755	45	1	3.3	44.65	LIFQYASFNNSR	
8	SGN-U580213	38900	44	2	9.6	38.3	VPTADVSVVDLTVR	
8						24.95	GILGYTEDDVVSTDFVGDSR	
9	SGN-U579942	15928	43	2	8.6	33.45	DCANDAFDQGGK	
9						22.47	DCANDAFDQGGK	
10	SGN-U578146	25969	43	1	4.3	42.85	YVPVLFNAVR	
11	SGN-U578441	17786	43	1	9.2	42.84	SIEIVEGDGGAGSIK	
12	SGN-U569028	28670	42	1	4.6	41.54	TFIAIKPDGVQR	
13	SGN-U578601	26237	40	1	7.8	39.88	NNCPYTVWAASTPIGGGR	
14	SGN-U570307	24859	39	3	3.2	28.23	RMVSLPR	
14						28.12	RMVSLPR	
14						26.64	RMVSLPR	Oxidation (M)
15	SGN-U579788	38480	39	2	6.2	23.64	AVTVFGFR	
15						33.49	VPTVDVSVVDLTVR	
16	SGN-U577720	85298	39	2	3.3	27.03	GVTAFGFDLVR	
16						30.04	AGITVIQIDEAALR	
17	SGN-U578513	15973	37	2	19.4	20.88	EGGGGGYGGGGYGGGR	
17						30.74	DAIEGMNGQDLDGR	Oxidation (M)
18	SGN-U585664	63110	36	1	1.7	35.6	VVDLLAPYQR	
19	SGN-U568301	40198	34	1	3.1	34.22	IPLNLVYDDIR	
20	SGN-U575645	47394	34	2	2.7	30.07	HAQFSSFVALR	
20						25.27	HAQFSSFVALR	
21	SGN-U580366	38598	33	1	4	32.56	NNFYSYNAFINAAR	
22	SGN-U573941	26066	32	1	3.5	32.25	YFVLPSLR	
23	SGN-U577277	72048	32	2	4.1	29.2	VEIIANDQGNR	
23						20.85	ATAGDTHLGGEDFDNR	
24	SGN-U578246	20806	31	2	4.3	26.55	IIGFDNVR	
24						24.01	IIGFDNVR	
25	SGN-U578219	52689	31	1	3.2	30.65	AVFVDLEPTVIDEVR	
26	SGN-U563073	61448	30	1	2.5	29.87	IDAYQDTLYAHSQR	
27	SGN-U577872	9137	27	1	21.8	26.99	NCASVCETEGFSGGDCR	
28	SGN-U586603	17334	27	1	5.2	26.99	VVSVSIPR	
29	SGN-U593591	15483	27	1	9.9	26.64	GTSNHVLDEICVSR	
30	SGN-U603754	9805	26	1	18.8	25.59	RALVGCVGGLGVLPMR	Oxidation (M)
31	SGN-U578748	23932	26	1	4.2	25.58	GNLDIFSGK	
32	SGN-U577182	36918	25	1	2.4	24.6	IGLDGFGR	
33	SGN-U563994	47662	24	1	2.1	24.44	IEKNSIGGR	
34	SGN-U593950	13395	23	2	11.3	22.12	KEIPENANQSTTR	
34						21.59	KERPENANQSTTR	
35	SGN-U598086	25029	23	1	4.7	22.54	NLINMLCMKR	Oxidation (M)
36	SGN-U583862	9444	22	1	11.8	22.39	LTHGAPGDEIR	
37	SGN-U599822	8544	22	1	11.3	22.24	LYFVNESR	
38	SGN-U585943	66609	22	1	1.4	22.05	VNDLESLR	
39	SGN-U565330	33342	22	1	4.7	21.78	VSWEKIGEDYTAQR	
40	SGN-U578851	28703	22	1	6.3	21.71	DDQTSTFTCPAGTNYR	
41	SGN-U581281	27932	21	1	3.9	21.45	EYYYLSVLTR	
42	SGN-U572904	13067	21	1	7.4	21.39	DEILEVAGK	
43	SGN-U568794	24215	21	1	7.3	20.9	DEEEEEEEETAAEKR	
44	SGN-U565452	23005	21	1	5.3	20.88	DITLGFVDLLR	
45	SGN-U570961	76624	21	1	2.2	20.62	DAVVTVPAYFNDAQR	
46	SGN-U577165	12587	21	1	7.6	20.6	AFQSAYYNR	
47	SGN-U583085	39609	20	1	4.7	20.47	DSVVLTGGPNYDVPLGR	
48	SGN-U577913	9746	20	1	18.1	20.41	IASQSLSLYMCAEQR	
49	SGN-U564000	51428	20	1	2.2	20.05	LAVNLIPFPR	
50	SGN-U594445	16608	20	2	5.5	19.79	DEVLSLGK	
50						22.27	DEVLSLGK	
51	SGN-U581867	38591	20	1	3.3	19.71	LTMSESLGGKR	Oxidation (M)
52	SGN-U599587	17507	19	1	5.1	19.02	RNLPESPR	

Supplementary Table S4. Complete MASCOT results for gel-free analysis

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
1	SGN-U579235	17017	863	22	68.2	34.71	IVDQCR	
1						24.77	LDTNGLGYQR	
1						93.91	VTNTGTGTQETVR	
1						93.48	VTNTGTGTQETVR	
1						91.64	VTNTGTGTQETVR	
1						93.52	VTNTGTGTQETVR	
1						82.46	NGGLDLDVNVFNR	
1						82.13	CLRVTNTGTGTQETVR	
1						50.97	CLRVTNTGTGTQETVR	
1						73.57	ATYHLYNPQNINWDLR	
1						37.1	ATYHLYNPQNINWDLR	
1						100.34	YGWTAFCGPAGPTGQASCGR	
1						71.64	YGWTAFCGPAGPTGQASCGR	
1						106.23	YGWTAFCGPAGPTGQASCGR	
1						40.28	VTNTGTGTQETVRIVDQCR	
1						25.37	VTNTGTGTQETVRIVDQCR	
1						24.55	VTNTGTGTQETVRIVDQCR	
1						26.84	TASVYCATWDADKPLEWR	
1						87.66	RYGWTAFCGPAGPTGQASCGR	
1						66.96	RYGWTAFCGPAGPTGQASCGR	
1						20.91	TASVYCATWDADKPLEWRR	
1						56.52	TASVYCATWDADKPLEWRR	
2	SGN-U579545	18249	771	26	52.8	42.69	GGGDFTGR	
2						31.77	AQNYANSR	
2						46.18	AQNYANSR	
2						47.17	AQNYANSR	
2						40.99	AQNYANSR	
2						45.84	AQNYANSR	
2						45.58	AQNYANSR	
2						39.61	HYTQVVWR	
2						40.07	CRHYTQVVWR	
2						36.81	CRHYTQVVWR	
2						97.92	AGDCNLIHSGAGENLAK	
2						43.95	AGDCNLIHSGAGENLAK	
2						52.92	AGDCNLIHSGAGENLAK	
2						100.36	AQVGVGPMSWDANLASR	
2						101.07	AQVGVGPMSWDANLASR	
2						60.98	AQVGVGPMSWDANLASR	
2						117.24	AQVGVGPMSWDANLASR	Oxidation (M)
2						72.51	AQVGVGPMSWDANLASR	Oxidation (M)
2						94.63	AGDCNLIHSGAGENLAKGGGDF TGR	
2						61.6	AGDCNLIHSGAGENLAKGGGDF TGR	
2						38.53	AGDCNLIHSGAGENLAKGGGDF TGR	
2						44.09	AGDCNLIHSGAGENLAKGGGDF TGR	
2						36.03	AGDCNLIHSGAGENLAKGGGDF TGR	
2						48.17	AAVQLWVSERPSYNYATNQCV GGK	
2						38.6	AAVQLWVSERPSYNYATNQCV GGKK	
2						28.96	AAVQLWVSERPSYNYATNQCV GGKK	
3	SGN-U585664	63110	632	20	28.2	50.39	IGLFGGAGVGK	
3						47.75	VVDLLAPYQR	
3						77.77	VVDLLAPYQR	
3						28.48	TIAMDGTEGLVR	
3						63.48	TIAMDGTEGLVR	Oxidation (M)
3						81.45	TIAMDGTEGLVR	Oxidation (M)
3						34.81	VCQVIGAVVDVR	
3						64.24	VLNTGSPITVPVGR	
3						67.25	VLNTGSPITVPVGR	
3						53.37	TVLIMELINNVAK	
3						50.47	TVLIMELINNVAK	Oxidation (M)
3						29.1	TIAMDGTEGLVRGQR	
3						77.54	LVLEVAQHLGENMVR	
3						54.6	LVLEVAQHLGENMVR	
3						54.87	MLSPHILGEDHYNTAR	Oxidation (M)
3						50.59	QISELGIYPAVDPLDSTSR	
3						54.54	QISELGIYPAVDPLDSTSR	
3						48.97	QISELGIYPAVDPLDSTSR	
3						66.07	EAPAFVEQATEQQILVTGIK	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
3						57.97	IPSAVGYQPTLATDLGGLQER	
4	SGN-U578441	17786	613	28	66.9	31.49	LLSHDVK	
4						22.38	IHVIDDK	
4						63.85	QMNFVEGGPIK	
4						46.05	QMNFVEGGPIK	Oxidation (M)
4						74.57	FEAAGDGGCVCK	
4						83.23	FEAAGDGGCVCK	
4						37.58	GLVLDFDSLVPK	
4						25.48	GLVLDFDSLVPK	
4						35.23	GLVLDFDSLVPK	
4						34.43	GLVLDFDSLVPK	
4						25.05	GLVLDFDSLVPK	
4						35.01	GLVLDFDSLVPK	
4						40.87	GLVLDFDSLVPK	
4						37.64	GLVLDFDSLVPK	
4						20.3	GLVLDFDSLVPK	
4						44.99	IHVIDDKNLVTK	
4						39.47	IHVIDDKNLVTK	
4						73.19	SIEIVEGDGGAGSIK	
4						54.1	SIEIVEGDGGAGSIK	
4						56.53	GDHVVSEEEHNVGK	
4						34.22	QMNFVEGGPIKYLK	
4						37.21	QMNFVEGGPIKYLK	
4						38.85	QMNFVEGGPIKYLK	
4						36	QMNFVEGGPIKYLK	Oxidation (M)
4						38.72	GDHVVSEEEHNVGKGK	
4						89.58	YSLIEGDVLGDKLESIAYDVK	
4						68.44	YSLIEGDVLGDKLESIAYDVK	
4						69.29	SIEIVEGDGGAGSIKQMNFVEGG PIK	Oxidation (M)
5	SGN-U581155	39394	588	21	21.9	47.17	MGASLIR	Oxidation (M)
5						29.24	MGASLIR	Oxidation (M)
5						20.38	GYEVIAQAK	
5						32.96	GYEVIAQAK	
5						28.1	GYEVIAQAK	
5						41.18	GYEVIAQAK	
5						28.9	LGGQTYSVALGR	
5						80.44	LGGQTYSVALGR	
5						65.36	LGGQTYSVALGR	
5						28.81	LGGQTYSVALGR	
5						81.78	LGGQTYSVALGR	
5						72	AVVDSAIDAETR	
5						58.76	AVVDSAIDAETR	
5						76.38	AVVDSAIDAETR	
5						68.68	EMVALAGAHTVGFAR	
5						76.52	EMVALAGAHTVGFAR	Oxidation (M)
5						47.28	MGDLPPSAGAQLEIR	Oxidation (M)
5						82.18	DSVAKLGGQTYSVALGR	
5						70.5	DSVAKLGGQTYSVALGR	
5						41.72	MGDLPPSAGAQLEIRDVCSR	
5						44.4	MGDLPPSAGAQLEIRDVCSR	Oxidation (M)
6	SGN-U579393	49872	581	16	28.9	22.02	TCNALLLK	
6						28.12	ARQIFDSR	
6						33.95	AGWGVMTSHR	
6						61.01	LAKYNQLLR	
6						53.14	LAKYNQLLR	
6						39.93	KAGWGVMTSHR	Oxidation (M)
6						88.36	VNQIGSVTESIEAVK	
6						101.6	VNQIGSVTESIEAVK	
6						109.36	AAVPSGASTGIYEALELR	
6						97.7	AAVPSGASTGIYEALELR	
6						88.34	AAVPSGASTGIYEALELR	
6						21.32	VNQIGSVTESIEAVKMSK	Oxidation (M)
6						49.4	IEEELGSEAVYAGASFRKPVEPY	
6						37.35	IEEELGSEAVYAGASFRKPVEPY	
6						55.07	AAVPSGASTGIYEALELRDGGS DYLGK	
6						74.76	YGQDATNVGDEGGFAPNIQENK EGLELLK	
7	SGN-U581103	28230	574	21	30.6	37.9	TNCNFDGAGR	
7						48.21	TNCNFDGAGR	
7						62.14	TNCNFDGAGR	
7						42.68	TNCNFDGAGR	
7						43.64	TNCNFDGAGR	
7						43.38	GQTWVINAPR	
7						52.7	GQTWVINAPRGTK	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
7						76.65	LDRGQTWVINAPR	
7						28.79	LDRGQTWVINAPR	
7						31.61	LDRGQTWVINAPR	
7						36.5	IWGRTNCNFDGAGR	
7						61.99	NNCPYTVWAASTPIGGGR	
7						63.04	NNCPYTVWAASTPIGGGR	
7						84.19	NNCPYTVWAASTPIGGGR	
7						60.53	NNCPYTVWAASTPIGGGR	
7						21.18	NNCPYTVWAASTPIGGGRR	
7						34.26	NNCPYTVWAASTPIGGGRR	
7						41.7	NNCPYTVWAASTPIGGGRR	
7						80.13	CPDAYSYPQDDPTSTFTCPSGSTNYR	
7						56.21	CPDAYSYPQDDPTSTFTCPSGSTNYR	
7						72.03	QRCPDAYSYPQDDPTSTFTCPSGSTNYR	
8	SGN-U583085	39609	550	11	23.5	58.57	QIKDDVGQAAGLLR	
8						83.08	QGLFTSDQDLYTDR	
8						88.29	DSVVLTGGPNYDVPLGR	
8						34.96	QGLFTSDQDLYTDRR	
8						30.69	QGLFTSDQDLYTDRR	
8						33.16	MGQMNVLTGGQGEIRNR	Oxidation (M)
8						59.79	DSVVLTGGPNYDVPLGRK	
8						91.87	DSVVLTGGPNYDVPLGRK	
8						94.53	IQDECGQVVSCSDIVAIAAR	
8						126.83	IQDECGQVVSCSDIVAIAAR	
8						105.99	IQDECGQVVSCSDIVAIAAR	
9	SGN-U578851	28703	547	13	43.5	65.61	TGCNFDASGK	
9						70.9	TGCNFDASGK	
9						74.63	TGCNFDASGKGK	
9						55.6	TGCNFDASGKGK	
9						73.96	GQTWTINAPPGTK	
9						40.8	CQTGDCNGLLVCK	
9						45.37	LDRGQTWTINAPPGTK	
9						27.14	NNCPYTVWAAGVPAGGGK	
9						102.92	NNCPYTVWAAGVPAGGGK	
9						100.49	NNCPYTVWAAGVPAGGGK	
9						81.77	NNCPYTVWAAGVPAGGGKR	
9						62.86	AEINQQCPNELKAPGGCNNPCTVFK	
9						55.02	CPDAYSYPKDDQTSTFTCPAGTNYR	
10	SGN-U578193	49384	511	13	24	45.84	LTAEIGQK	
10						22.02	TCNALLLK	
10						28.12	ARQIFDSR	
10						61.01	LAKYNQLLR	
10						53.14	LAKYNQLLR	
10						40.56	QAGWGVMTSHR	Oxidation (M)
10						88.36	VNQIGSVTESIEAVK	
10						101.6	VNQIGSVTESIEAVK	
10						109.36	AAVPSGASTGIYEALELR	
10						97.7	AAVPSGASTGIYEALELR	
10						88.34	AAVPSGASTGIYEALELR	
10						21.32	VNQIGSVTESIEAVKMSK	Oxidation (M)
10						74.76	YGQDATNVGDEGGFAPNIQENKEGLELLK	
11	SGN-U579414	28517	505	17	31	86.57	TGCNFNAAGR	
11						73.47	TGCNFNAAGR	
11						89.49	TGCNFNAAGR	
11						43.38	GQTWVINAPR	
11						52.7	GQTWVINAPRGTK	
11						57.57	LNRGQTWVINAPR	
11						25.42	LNRGQTWVINAPR	
11						25.02	LNRGQTWVINAPR	
11						61.99	NNCPYTVWAASTPIGGGR	
11						63.04	NNCPYTVWAASTPIGGGR	
11						84.19	NNCPYTVWAASTPIGGGR	
11						60.53	NNCPYTVWAASTPIGGGR	
11						21.18	NNCPYTVWAASTPIGGGRR	
11						34.26	NNCPYTVWAASTPIGGGRR	
11						41.7	NNCPYTVWAASTPIGGGRR	
11						34.74	ALKVPGGCNNPCTTFGGQQYCCTQGPCGPTELSK	
11						58.7	ALKVPGGCNNPCTTFGGQQYCCTQGPCGPTELSK	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
12	SGN-U574403	28843	473	11	40.7	51.81	GGFTLHSLTHR	
12						50.49	FYCATGDCGGR	
12						43.18	SFPAPNAHWSGR	
12						105.58	ANLLESCPAVLQFR	
12						85.16	ANLLESCPAVLQFR	
12						96.57	ANLLESCPAVLQFR	
12						60.87	ANLLESCPAVLQFR	
12						51.33	SACEAFKSDEFCCR	
12						51.7	SACEAFKSDEFCCR	
12						53.84	NHYNSPQTCKPSSYSQFFK	
12						48.81	HACPATFTYAHDSPSLMHECSSPR	
13	SGN-U583086	40092	461	15	23.8	36.21	IIEDLR	
13						20.07	IIEDLRR	
13						20.71	IIEDLRR	
13						65.35	QDIGQAAGLLR	
13						78.47	VVSCADITAIAAR	
13						48.27	VVSCADITAIAAR	
13						91.16	MGQLNVLTGTQGEIR	Oxidation (M)
13						83.08	QGLFTSDQDLYTDR	
13						34.96	QGLFTSDQDLYTDRR	
13						30.69	QGLFTSDQDLYTDRR	
13						88.44	DSVFFSGGPDYDLPLGR	
13						86.05	DSVFFSGGPDYDLPLGR	
13						53.34	LQNVFRQDIGQAAGLLR	
13						44.1	LQNVFRQDIGQAAGLLR	
13						30.27	DSVFFSGGPDYDLPLGRR	
14	SGN-U580213	38900	438	17	38.2	49.03	SSIFDAK	
14						26.42	IGRLVAR	
14						29.98	IGRLVAR	
14						28.35	TLLFGEK	
14						60.29	IGINGFGR	
14						31.16	VLPQLNGK	
14						62.84	AGIALSKNFVK	
14						23.65	VIDLICHMAKA	
14						30.96	VIDLICHMAKA	Oxidation (M)
14						52.98	AASFNIIPSSTGAAK	
14						69.66	VVSWYDNEWGYSSR	
14						53.75	AASFNIIPSSTGAAKAVGK	
14						32.5	AASFNIIPSSTGAAKAVGK	
14						69.06	FGIVEGLMTTVHAMTATQK	
14						84.16	GILGYTEDDVVSTDFVGDSR	
14						113.4	GILGYTEDDVVSTDFVGDSR	
14						46.21	VINDRFGIVEGLMTTVHAMTATQK	
15	SGN-U565569	38570	399	12	25.8	39.45	SLCTAIAK	
15						29.03	SLCTAIAK	
15						65.65	TQDGGTEVVEAK	
15						47.09	TQDGGTEVVEAK	
15						22.06	DDLFNINAGIVK	
15						69.48	ANLSDEEIVALTK	
15						54.2	VAVLGAAGGIGQPLSLLMK	Oxidation (M)
15						62.09	VAVLGAAGGIGQPLSLLMK	Oxidation (M)
15						58.98	VAVLGAAGGIGQPLSLLMK	Oxidation (M)
15						56.51	KVAVLGAAGGIGQPLSLLMK	Oxidation (M)
15						111.83	LNPLVSSLSLYDIAGTPGVAADVSHINTR	
15						64.17	LNPLVSSLSLYDIAGTPGVAADVSHINTR	
16	SGN-U577365	19542	372	9	34.7	83.45	TPGPGAQSALR	
16						68.8	TPGPGAQSALR	
16						34.19	CKELGINALHIK	
16						78.51	IEDVTPIPTDSTR	
16						82.25	IEDVTPIPTDSTR	
16						46.07	IEDVTPIPTDSTRR	
16						40.95	IGRIEDVTPIPTDSTR	
16						71.42	ADRDESSPYAAMLAAQDVSQR	
16						63.09	ADRDESSPYAAMLAAQDVSQR	Oxidation (M)
17	SGN-U579572	19091	363	24	60.9	20.92	QAMELFK	Oxidation (M)
17						25.41	SITEYHTK	
17						38.44	NIEAEGDGSIK	
17						48.64	NIEAEGDGSIK	
17						55.15	MNFVEGSPIK	Oxidation (M)
17						57.07	NIEAEGDGSIKK	
17						26.16	ALVVDSDNLIPK	
17						37.16	ALVVDSDNLIPK	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
17						30.05	ALVVDSDNLIPK	
17						39.16	ALVVDSDNLIPK	
17						30.54	ALVVDSDNLIPK	
17						26.52	ALVVDSDNLIPK	
17						32.88	ALVVDSDNLIPK	
17						23.03	ALVVDSDNLIPK	
17						44.64	ALVVDSDNLIPK	
17						43.17	FEAHGNGGCVCK	
17						47.17	IHVVDDKNLVTK	
17						56.44	IHVVDDKNLVTK	
17						37.51	MNFVEGSPIKYLK	
17						39.55	MNFVEGSPIKYLK	
17						35.65	MNFVEGSPIKYLK	
17						39.07	MNFVEGSPIKYLK	Oxidation (M)
17						22.79	ALVVDSDNLIPKLMPQVK	Oxidation (M)
17						65.55	YSMIEGDVLGDKLESISYDLK	Oxidation (M)
18	SGN-U577872	9137	341	11	47.4	22.83	CFCTRPC	
18						44.45	FKGPCVSEK	
18						24.37	FKGPCVSEK	
18						38.02	RCFCTRPC	
18						22.67	RCFCTRPC	
18						94.2	NCASVCETEGFSGGDCR	
18						94.57	NCASVCETEGFSGGDCR	
18						94.49	NCASVCETEGFSGGDCR	
18						52.41	NCASVCETEGFSGGDCR	
18						69.28	NCASVCETEGFSGGDCR	
18						34.11	NCASVCETEGFSGGDCRGFR	
19	SGN-U585234	33590	326	10	38.9	50.39	IGLFGGAGVGK	
19						58.97	AVAMSATEGLTR	
19						50.55	AVAMSATEGLTR	Oxidation (M)
19						56.52	AVAMSATEGLTR	Oxidation (M)
19						29.23	TVLIMELINNIAK	
19						53.68	TVLIMELINNIAK	Oxidation (M)
19						38.29	VALVYGKMNEPPGAR	Oxidation (M)
19						81.65	GMAVIDTGAPISVPVGGATLGR	Oxidation (M)
19						52.03	DSVGQPINVACEVQQLLGNNRV R	
19						93.25	IFNVLGEPVDNLGPVDTSTTSPI HR	
20	SGN-U577742	56554	323	7	7	46.7	ATDVMIAGK	
20						50.25	ATDVMIAGK	Oxidation (M)
20						71.05	LVGVSEETTTGVK	
20						87.48	LVGVSEETTTGVK	
20						57.7	AEFGPSQPFKGAK	
20						94.15	LVGVSEETTTGVKR	
20						66.12	LVGVSEETTTGVKR	
21	SGN-U578520	49770	314	14	20.7	52.92	QTVAVGVVK	
21						63.01	QTVAVGVVK	
21						56.48	IGGIGTVPVGR	
21						46.9	IGGIGTVPVGR	
21						36.58	IGGIGTVPVGR	
21						54.69	DMRQTVAVGVVK	Oxidation (M)
21						27.06	GYVASNSKDDPAK	
21						53.85	YDEIVKEVSSYLK	
21						41.43	MIPTKPMVVETFAEYPPLGR	
21						49.01	MIPTKPMVVETFAEYPPLGR	Oxidation (M)
21						57.34	MIPTKPMVVETFAEYPPLGR	2 Oxidation (M)
21						38.09	MIPTKPMVVETFAEYPPLGR	2 Oxidation (M)
21						45.66	VETGVIKPGMVVTFGPTGLTTE VK	
21						44.9	VETGVIKPGMVVTFGPTGLTTE VK	Oxidation (M)
22	SGN-U579667	12759	296	7	30.1	64.23	AYMEAEFQR	Oxidation (M)
22						31.03	AYMEAEFQR	Oxidation (M)
22						89.79	VSADVVQMLLR	
22						85.22	VSADVVQMLLR	Oxidation (M)
22						35.76	LEQTSGDSGANVK	
22						106.32	KLEQTSGDSGANVK	
22						45.79	KLEQTSGDSGANVK	
23	SGN-U577277	72048	295	7	6.3	60.46	NALENYAYNMR	Oxidation (M)
23						68.13	TTPSYVGFTDSER	
23						74.38	IINEPTAAAIAYGLDK	
23						70.41	IINEPTAAAIAYGLDK	
23						37.46	IINEPTAAAIAYGLDKK	
23						59.74	IINEPTAAAIAYGLDKK	
23						64.85	IINEPTAAAIAYGLDKK	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
24	SGN-U578839	59028	281	7	11.8	42.27	LTVKPTM	Oxidation (M)
24						69.72	IFAYADTQR	
24						51.83	IFAYADTQR	
24						49.15	GPVLLEDYYLIEK	
24						58.29	GPVLLEDYYLIEK	
24						95	IGPNYMQLPVNAPK	Oxidation (M)
24						67.69	EGNFDLVGNNVPVFFNR	
25	SGN-U577869	57572	279	9	16.7	25.02	VVSVGDGIAR	
25						38.03	AVDSLVPIGR	
25						69.08	TGSIVDVPAGK	
25						69.23	TGSIVDVPAGK	
25						64.79	AAELTSLLESR	
25						73.36	AAELTSLLESR	
25						58.3	VVDGLGVPIDGR	
25						54.27	AILNSVKPELLQSFLEK	
25						51.72	LTEVLKQPQYAPLPIEK	
26	SGN-U570963	61158	277	7	8.5	39.88	NQIDEIVLVGGSTR	
26						64.37	QATKDAGVIAGLNVAR	
26						74.38	IINEPTAAAIAYGLDK	
26						70.41	IINEPTAAAIAYGLDK	
26						37.46	IINEPTAAAIAYGLDKK	
26						59.74	IINEPTAAAIAYGLDKK	
26						64.85	IINEPTAAAIAYGLDKK	
27	SGN-U579426	18416	276	8	34.8	43.56	GSGDFTGR	
27						33.24	VCGHYTQVVWR	
27						51.61	VCGHYTQVVWR	
27						95.16	GQVGVGPMSWDDALATK	
27						73.16	GQVGVGPMSWDDALATK	Oxidation (M)
27						86.3	GQVGVGPMSWDDALATK	Oxidation (M)
27						53.42	RGDCNLIHSGPGENLAK	
27						45.16	GQVGVGPMSWDDALATKAQR	
28	SGN-U584410	27132	269	10	20.8	28.78	KLEEAVR	
28						29.1	KLEEAVR	
28						34.22	LVPVGYGIK	
28						28.85	WYQAVSAK	
28						44.25	LVPVGYGIKK	
28						66.82	FGSQAAPAGAAPAK	
28						59.35	FGSQAAPAGAAPAK	
28						73.19	SVQMDGLLWGASK	
28						73.95	SVQMDGLLWGASK	Oxidation (M)
28						69.77	SVQMDGLLWGASK	Oxidation (M)
29	SGN-U578978	38632	263	11	27.5	49.03	SSIFDAK	
29						26.42	IGRLVAR	
29						29.98	IGRLVAR	
29						60.29	IGINGFGR	
29						62.84	AGIALSKNFVK	
29						20.46	VVDLIKHMASVQ	
29						52.98	AASFNIIPSSTGAAK	
29						74.64	LVSWYDNEMGYSTR	Oxidation (M)
29						53.75	AASFNIIPSSTGAAKAVGK	
29						32.5	AASFNIIPSSTGAAKAVGK	
29						73.45	FGIVEGLMTTVHSITATQK	
30	SGN-U579687	12205	252	10	55.7	29.84	TACTCLK	
30						43.61	TACTCLK	
30						43.75	TACTCLK	
30						52.05	GPLGGCCR	
30						58.12	GVKGLLGAAK	
30						36.03	SAANAIKGLNLGK	
30						61.03	SAANAIKGLNLGK	
30						58.54	ISPFTDCSKVQ	
30						45.49	AAGIPSACGVSIPYK	
30						46.24	AAGIPSACGVSIPYKISPFTDCSK	
31	SGN-U586194	36110	250	3	9.2	88.07	IVDDQTILEDEAR	
31						117.73	TVDTTGAGDSFVGALLTK	
31						99.14	TVDTTGAGDSFVGALLTK	
32	SGN-U577711	42200	249	12	30.1	59.76	AGFAGDDAPR	
32						49.45	AGFAGDDAPR	
32						38.46	HTGVMVGMGQK	
32						31.05	HTGVMVGMGQK	Oxidation (M)
32						27.48	EITALAPSSMK	Oxidation (M)
32						37.15	AVFPSIVGRPR	
32						35.49	AVFPSIVGRPR	
32						34.44	VAPEEHPVLLTEAPLNPK	
32						35.45	AGFAGDDAPRAVFPSIVGRPR	
32						75.49	SSSSIEKNYELPDGQVITIGAER	
32						48.81	TTGIVLDSGDGVSHTVPIYEGYA	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
							LPHAILR	
32						32.55	TTGIVLDSGDGVSHTVPIYEGYA LPHAILR	
33	SGN-U580366	38598	245	6	23.3	31.15	WQPSGADR	
33						67.09	SFRGFGTTGDNTAR	
33						83.66	NNFYSYNAFINAAR	
33						67.05	AIGADLLNNPDLVATDPVISFK	
33						87.16	AIGADLLNNPDLVATDPVISFK	
33						27.9	SAIWFWMTPQSPKPSCHDVITG R	Oxidation (M)
34	SGN-U579068	12579	244	12	27.3	25.36	CPNKLCCSK	
34						28.92	CPNKLCCSK	
34						29.94	CPNKLCCSK	
34						47.68	CPNKLCCSK	
34						28.97	CPNKLCCSK	
34						30.63	FGWCGTSCDYCGSGCQSNCR	
34						76.92	FGWCGTSCDYCGSGCQSNCR	
34						21.69	FGWCGTSCDYCGSGCQSNCR	
34						40.3	FGWCGTSCDYCGSGCQSNCR	
34						33.29	FGWCGTSCDYCGSGCQSNCRK	
34						56.23	FGWCGTSCDYCGSGCQSNCRK	
34						78.47	FGWCGTSCDYCGSGCQSNCRK	
35	SGN-U575184	36154	243	13	36.8	47.17	MGASLLR	Oxidation (M)
35						29.24	MGASLLR	Oxidation (M)
35						53.86	NMDASLAR	Oxidation (M)
35						37.15	AQPNFNSAR	
35						30.36	GFEVIDNIK	
35						24.44	SCPKLYQTVK	
35						20.5	SCPKLYQTVK	
35						73.24	LISSFTAVGLSTK	
35						86.33	SVVNSAIQKETR	
35						56.62	SVVNSAIQKETR	
35						23.87	DSVVILGGPNWDVK	
35						63.01	SYSNNPSSFISDFVTAMIK	
35						25.66	GLLHSDQQLFNGGSVDSIVK	
36	SGN-U577630	18469	223	6	30.1	100.76	VVMELFADTTPK	
36						73.25	VVMELFADTTPK	Oxidation (M)
36						67.01	VVMELFADTTPK	Oxidation (M)
36						44.9	CSKPVVIADCGQL	
36						44.88	CSKPVVIADCGQL	
36						41.92	VIPGFMCQGGDFTAGNGTGGES IYGAK	Oxidation (M)
37	SGN-U579551	35549	219	5	19.8	68.02	YCGILGVSPGENLDCGNQR	
37						66.83	AIGVDLLNNPDLVATDPVISFK	
37						86.69	AIGVDLLNNPDLVATDPVISFK	
37						58.31	AIGVDLLNNPDLVATDPVISFK	
37						27.9	SAIWFWMTPQSPKPSCHDVITG R	Oxidation (M)
38	SGN-U591875	16675	212	3	17	83.09	SAGAGDDSQKAGEESGSTTELF ASAK	
38						89.8	SAGAGDDSQKAGEESGSTTELF ASAK	
38						82.3	SAGAGDDSQKAGEESGSTTELF ASAK	
39	SGN-U580709	36174	211	13	34.3	47.17	MGASLLR	Oxidation (M)
39						29.24	MGASLLR	Oxidation (M)
39						27.63	AAPNVNSVR	
39						30.36	GFEVIDNIK	
39						21.52	SCPKLYQTIK	
39						46.04	LISSFSAVGLSTK	
39						73.61	STVQSAINKETR	
39						60.28	STVQSAINKETR	
39						20.71	DSVVILGGPNWNVK	
39						34.19	DSVVILGGPNWNVK	
39						39.11	AAPNVNSVRGFEVIDNIK	
39						47.7	GLLHSDQQLFNGGSADSIVR	
39						76.01	SYINNPSSFNSDFVTAMIK	
40	SGN-U578572	42060	209	12	17.6	27.89	VLAACYK	
40						27.49	VLAACYK	
40						22.91	YYEAGAR	
40						38.91	CADVTER	
40						44.6	ALQQSTLK	
40						54.68	ALQQSTLK	
40						32.28	CQKYYEAGAR	
40						53.94	VAPEVIAEYTVR	
40						73.38	VAPEVIAEYTVR	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
40						49.56	IGANEPSQLAINDNANGLAR	
40						46.21	IGANEPSQLAINDNANGLAR	
40						35.03	AVLKIGANEPSQLAINDNANGLAR	
41	SGN-U590345	15845	205	4	33.8	90.27	ISVTLGDASGK	
41						51.3	TGSIIKGGTDSK	
41						44.49	GNVDIFTGRGK	
41						93.98	WGLMGPNYDYYER	Oxidation (M)
42	SGN-U564978	22435	204	9	19.6	37.37	ICTNCCAGSK	
42						45.42	ICTNCCAGSK	
42						52.52	NCNPSIQYGR	
42						48.18	NCNPSIQYGR	
42						67.1	IAYSICPGNKR	
42						41.22	IAYSICPGNKR	
42						46.47	NCNPSIQYGRCPK	
42						26.04	NCNPSIQYGRCPK	
42						40.43	SEGRICTNCCAGSK	
43	SGN-U579480	29420	202	6	13.8	43.27	LAEQAER	
43						44.39	LAEQAER	
43						46.15	NLLSVAYK	
43						66.09	NLLSVAYKNVIGAR	
43						80.9	VVAALNGEELTVEER	
43						68.83	VVAALNGEELTVEER	
44	SGN-U580253	15749	201	8	26.2	29.84	TACTCLK	
44						43.61	TACTCLK	
44						43.75	TACTCLK	
44						44.52	GPLGGCCGGVK	
44						47.58	GPLGGCCGGVK	
44						48.33	SAANAIKGIDLNK	
44						70.5	GPLGGCCGGVKNLLGSAK	
44						50.09	GPLGGCCGGVKNLLGSAK	
45	SGN-U566251	35801	200	10	25.1	47.29	SDFAAAMVK	Oxidation (M)
45						40.89	SDFAAAMVK	Oxidation (M)
45						21.36	NAMTQAVNR	Oxidation (M)
45						57.6	NAMTQAVNR	Oxidation (M)
45						22.76	NAMTQAVNREAR	
45						27.94	NAMTQAVNREAR	
45						25.47	NAMTQAVNREAR	Oxidation (M)
45						65.87	DMTALSGSHTIGQAR	Oxidation (M)
45						27.17	EGTVLLGGPSWAVPLGRR	
45						88.19	TASQSAANTQIPAPSSSLATLISMFSAK	Oxidation (M)
46	SGN-U580369	39192	199	8	17.3	47.17	MGASLIR	Oxidation (M)
46						29.24	MGASLIR	Oxidation (M)
46						22.64	GFEVIAQAK	
46						89.74	LGGQTYTVALGR	
46						68.68	EMVALAGAHTVGFAR	
46						76.52	EMVALAGAHTVGFAR	Oxidation (M)
46						34.82	MGNLPPSAGAQLEIRDVCSR	
46						32.1	MGNLPPSAGAQLEIRDVCSR	Oxidation (M)
47	SGN-U572912	34831	199	6	15.1	35.72	QVAAQEAER	
47						41.47	QVAAQEAER	
47						70.64	TLGENYNER	
47						22.98	VLPSIIHETLK	
47						82.46	SAQLIGQAIANNPAFITLR	
47						77.84	SAQLIGQAIANNPAFITLR	
48	SGN-U577283	25637	194	4	14.8	73.85	VGVINQDGR	
48						74.07	VGVINQDGR	
48						53.87	SSAPCLDGVFR	
48						78.07	LALVNENPLGVYFK	
49	SGN-U573941	26066	194	6	18.3	29.31	GTLGALNR	
49						88.82	LAASDNELPFSVYFK	
49						85.57	LAASDNELPFSVYFK	
49						47.58	RLAASDNELPFSVYFK	
49						34.07	RLAASDNELPFSVYFK	
49						26.67	YFVLPSLRGSGGGLVLSR	
50	SGN-U581590	17859	186	4	18.9	41.54	AVVVHADPDDLGK	
50						78.94	QIPLTGPQSIIGR	
50						82.42	QIPLTGPQSIIGR	
50						46.22	AVVVHADPDDLGKGGHELSK	
51	SGN-U583361	27249	186	4	14.6	23.69	VLLTLEEK	
51						40.2	IFPTFVSFLK	
51						96.07	AAVGAPDVLGDCPFSQR	
51						95.8	AAVGAPDVLGDCPFSQR	
52	SGN-U578748	23932	181	3	18.4	78.36	GNLDIFSGKGPCVNGPICK	
52						60.77	GNLDIFSGKGPCVNGPICK	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
52						89.84	IKNIEAWGGLMGPGYNYFER	
53	SGN-U578279	12929	180	7	48.7	70.13	LFVNILGDVVQIPR	
53						62.55	LFVNILGDVVQIPR	
53						25.87	LFVNILGDVVQIPR	
53						42.16	QFWPELIGVPAQYAK	
53						42.62	QFWPELIGVPAQYAK	
53						67.44	LFVNILGDVVQIPRVT	
53						28.36	GIIEKENPSIANIPILLNGSPVTK	
54	SGN-U580220	33812	176	5	12.2	43.27	LAEQAER	
54						44.39	LAEQAER	
54						46.15	NLLSVAYK	
54						66.09	NLLSVAYKNVIGAR	
54						103.92	VSTSLGSEELTVEER	
55	SGN-U578082	42419	175	10	38.3	33.89	YLMQNGAR	Oxidation (M)
55						42.26	YSLKPLVPR	
55						32.48	LSELLGIEVK	
55						41.18	LSELLGIEVK	
55						33.98	VILASHLGRPKGVTPK	
55						35.64	LASLADLYVNDAFGTAHR	
55						59.87	VDLNVPLDDNFKITDDTR	
55						42.34	LAELSGKGVTTIIGGGDSVAAVEK	
55						27.38	IVPASEIPDGWMGLDIGPDAIK	Oxidation (M)
55						44.64	MSHISTGGGASLELLEGKQLPGVLALDDA	
56	SGN-U583542	38058	175	5	12.4	55.84	GVGPYTGVGDGR	
56						53.85	VLIIPTIISTGQR	
56						68.01	LQLPLQAFGPDSSAFDTK	
56						44.96	LQLPLQAFGPDSSAFDTK	
56						58.05	LQLPLQAFGPDSSAFDTKGVGPYTGVGDGR	
57	SGN-U577810	16798	175	6	19.4	54.38	FTGGHCSKLQR	
57						60.98	FTGGHCSKLQR	
57						28.42	APSQTFPGLCFMDSSCR	Oxidation (M)
57						60.52	APSQTFPGLCFMDSSCR	Oxidation (M)
57						42.51	APSQTFPGLCFMDSSCR	Oxidation (M)
57						39.05	APSQTFPGLCFMDSSCR	Oxidation (M)
58	SGN-U580093	25101	175	5	20.3	71.4	IAGIASAIR	
58						28.33	LGDKPLFVLVS	
58						50.02	AFKDTIDLFVER	
58						70.01	GFIFGPPIALAIGAK	
58						64.57	GFIFGPPIALAIGAK	
59	SGN-U579520	41889	173	7	22	36.94	FIEPFATVK	
59						67.49	GQCLKGGGACSDR	
59						37.9	GQCLKGGGACSDR	
59						35.48	GQCLKGGGACSDR	
59						36.48	ANYPPYGIDFPDGPTGR	
59						58.04	GVNYASGAAGIRDESGIHLGDR	
59						44.7	ASHYFWDGFHPTEIPNKVTAIR	
60	SGN-U581465	17428	171	6	33.1	39.11	QAACTCLK	
60						55.31	ISPSTDCSKVQ	
60						53.69	SAASSFTGLNLGK	
60						67.61	SAASSFTGLNLGK	
60						48.62	TTVDRQAACTCLK	
60						49.89	AAALPNTCSVNIPYKISPSTDCSK	
61	SGN-U578258	13704	168	3	9	80.12	IDCGGACAAR	
61						83.59	IDCGGACAAR	
61						71.22	KIDCGGACAAR	
62	SGN-U577774	23393	167	7	28.8	32.34	SAALINQK	
62						23.78	SAALINQK	
62						31.49	QIEVEGPRGK	
62						51.88	FLDGIYVSEK	
62						52	FLDGIYVSEK	
62						99.4	KVDMLDGVTVVR	Oxidation (M)
62						45.24	MKTILSSETMDIPDGITIK	
63	SGN-U578389	23785	163	7	20	47.1	IQFVISNVEK	
63						47.09	YGQVGPMGTPVR	Oxidation (M)
63						59.93	YGQVGPMGTPVR	Oxidation (M)
63						35.23	LLGYELLTCDGALVGTMGQR	Oxidation (M)
63						36.92	LLGYELLTCDGALVGTMGQR	Oxidation (M)
63						42.49	LLGYELLTCDGALVGTMGQR	Oxidation (M)
63						42.89	LLGYELLTCDGALVGTMGQR	Oxidation (M)
64	SGN-U577975	19226	162	8	31.1	41.07	IGPLGLSPK	
64						40.17	IGPLGLSPK	
64						27.32	IGPLGLSPK	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
64						51.57	VSVVPSAAALVIK	
64						35.04	VSVVPSAAALVIK	
64						69.29	VTGGEVGAASSLAPK	
64						55.24	VTGGEVGAASSLAPK	
64						36.18	NIKHNGNISLDDVIEIAK	
65	SGN-U584963	57821	160	6	9.7	20.82	IEQYNREVK	
65						54.95	LIESAAPGIISR	
65						46.5	LIESAAPGIISRR	
65						42.81	IAQIPVSEAYLGR	
65						61.55	IAQIPVSEAYLGR	
65						69.27	IVNTGTVLQVGDGIAR	
66	SGN-U581254	14200	160	4	17.1	59.46	LTEGCSFR	
66						63.2	LVQSPNSFFMDVK	
66						62.75	LVQSPNSFFMDVK	Oxidation (M)
66						46.42	LVQSPNSFFMDVK	Oxidation (M)
67	SGN-U579883	24225	160	6	25	42.67	VCTQQCDPK	
67						35.59	VCVNCCTAKPGCK	
67						38.02	VCVNCCTAKPGCK	
67						54.96	VAYMTCPPESTKLTR	
67						58.21	VAYMTCPPESTKLTR	Oxidation (M)
67						54.48	TPICTTCCAGYKGCK	
68	SGN-U579972	57924	159	6	6.1	33.15	SVVSEYGK	
68						31.49	TILDLLESK	
68						33.3	TILDLLESK	
68						33.51	TILDLLESK	
68						53.48	AAPQQICSQIGLCSR	
68						102.07	AAPQQICSQIGLCSR	
69	SGN-U578527	17168	154	7	33.7	51.85	GSGGGGGGFGGGR	
69						26.55	IINDRETGR	
69						29.42	IINDRETGR	
69						41.71	NITVNEAQAR	
69						44.43	NITVNEAQAR	
69						53.55	GFGFVTFKDEK	
69						85.09	TLSDAFSTYGEVVDSK	
70	SGN-U578637	61601	154	4	7.4	58.88	CLGEAAAAAAK	
70						37.56	SLDILGLGTGPEIEK	
70						83.78	IALVGLGSPTSSTAAYR	
70						55.82	IALVGLGSPTSSTAAYR	
71	SGN-U577977	39762	152	6	17.2	54.21	YGGDEVDLR	
71						53.69	TIANEYQKR	
71						56.18	LLVPLVSSYR	
71						43.53	RGTEEDHLTR	
71						33.12	AIAKDTGGDYENMLVALLGQEE	Oxidation (M)
71						44.26	AIAKDTGGDYENMLVALLGQEE	Oxidation (M)
72	SGN-U591123	14984	151	6	38.5	39.11	QAACTCLK	
72						55.31	ISPSTDCSKVQ	
72						53.69	SAASSFTGLNLGK	
72						67.61	SAASSFTGLNLGK	
72						48.62	TTVDRQAACTCLK	
72						37.38	AAALPNTCSVNIPYKISPSTDCSEVQ	
73	SGN-U579400	34583	150	4	11.8	32.38	ESGSTMAVVAAQTK	
73						51.13	ESGSTMAVVAAQTK	Oxidation (M)
73						73.65	ESGSTMAVVAAQTK	Oxidation (M)
73						70.93	AILGESNEFVGDKVAYALSQGLK	
74	SGN-U580740	18428	149	11	20.1	22.38	IHVIDDK	
74						63.85	QMNFVEGGPIK	
74						46.05	QMNFVEGGPIK	Oxidation (M)
74						28.62	ALVLDFDNLVPK	
74						34.58	ALVLDFDNLVPK	
74						27.45	ALVLDFDNLVPK	
74						26.33	ALVLDFDNLVPK	
74						34.22	QMNFVEGGPIKYLK	
74						37.21	QMNFVEGGPIKYLK	
74						38.85	QMNFVEGGPIKYLK	
74						36	QMNFVEGGPIKYLK	Oxidation (M)
75	SGN-U578562	36686	149	4	6.8	41.03	CGTFQQR	
75						95.73	GVMQQAQSTDVR	
75						56.31	GVMQQAQSTDVRAGAK	
75						45.94	GVMQQAQSTDVRAGAK	Oxidation (M)
76	SGN-U579033	13037	149	8	46.7	29.84	TACTCLK	
76						43.61	TACTCLK	
76						43.75	TACTCLK	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
76						44.52	GPLGGCCGGVK	
76						47.58	GPLGGCCGGVK	
76						40.35	SAANSIKGIDTGK	
76						38.12	AAGLPGVCGVNIPYK	
76						29.99	AAGLPGVCGVNIPYKISPSTDCS TVQ	
77	SGN-U578074	33067	148	5	11.2	25.5	AIKFYEK	
77						30.36	QPGSIPGLNTK	
77						58.23	VVNLAIQELGGK	
77						75.63	VVNLAIQELGGK	
77						59.19	VVNLAIQELGGKITR	
78	SGN-U580382	16244	146	6	24.3	46.77	VLNITTR	
78						48.19	VCADLVR	
78						29.29	ITLSSKNVK	
78						54.59	VIDLFSSPDVVK	
78						49.59	VIDLFSSPDVVK	
78						62.78	VIDLFSSPDVVK	
79	SGN-U581290	32024	146	3	6.1	53.78	IVIGLYGDDVPQTAENFR	
79						78.27	IVIGLYGDDVPQTAENFR	
79						58.34	IVIGLYGDDVPQTAENFR	
80	SGN-U575407	28842	139	7	13.2	38.39	HIDALLR	
80						28.78	KLEEAVR	
80						29.1	KLEEAVR	
80						34.22	LVPVGYGIK	
80						44.25	LVPVGYGIKK	
80						68.06	SPSAEYVNASR	
80						62.42	SPSAEYVNASR	
81	SGN-U577370	28337	138	5	18.5	30.47	IGSAGVVRR	
81						46.96	LTNSLMMHGR	2 Oxidation (M)
81						50.85	VNQAIYLLTTGAR	
81						87.05	VNQAIYLLTTGAR	
81						36.51	TIAECLADELINAAK	
82	SGN-U578195	71080	135	4	6.4	37.08	IYPTKAVNGAAR	
82						20.18	QVDLQPGSIELLR	
82						81.94	LLVDHSIVESFAQGGR	
82						59.24	LLVDHSIVESFAQGGR	
83	SGN-U585465	25252	134	2	15.5	110.91	LVLNDETYAFGFSK	
83						48.03	FVITGATLGFPGPNNIKNWFK	
84	SGN-U582837	48616	134	5	8.3	73.32	ASCNLVPR	
84						24.47	SFVPIASGR	
84						49.41	ILFDVPNSR	
84						41.48	QILQTPTYIVK	
84						46.54	QILQTPTYIVK	
85	SGN-U577779	23021	131	4	17.9	34.08	TIDWDGMAK	Oxidation (M)
85						52.46	AFDDVNSQLQTK	
85						64.62	AFDDVNSQLQTK	
85						49.25	FSQEPEPINWEYYR	
86	SGN-U577558	12920	131	2	12.4	84.9	LIDNILGVVVQIPR	
86						72.67	LIDNILGVVVQIPR	
87	SGN-U585251	40306	128	3	6.6	55.28	LADCAIGFGK	
87						65.96	QTGAGASSSSTYAR	
87						53.71	QTGAGASSSSTYAR	
88	SGN-U578825	40198	126	3	8.9	66.31	EGISAEVINLR	
88						51.16	VLAPYSSEDAR	
88						60.69	LAVPQVEDIVR	
89	SGN-U579765	14171	124	2	23.3	77.7	QGPKPGEQAAGGSGSVPDAASS AQASSTSR	
89						68.8	QGPKPGEQAAGGSGSVPDAASS AQASSTSR	
90	SGN-U577727	30594	124	4	11.6	27.7	SVVNFKV	
90						23.39	GCGACYQVK	
90						50.23	SGKGCGACYQVK	
90						93.48	LDSGSSLVAPFSLK	
91	SGN-U579203	38246	123	7	19.5	26.33	NSVTLDR	
91						41.64	MLLALIGHGDA	
91						35.32	MLLALIGHGDA	Oxidation (M)
91						45.39	AIAGDTSGDYEK	
91						36.11	VVLLWTLSPAER	
91						64.06	ISDKAYNDEEIIR	
91						34.24	VVLLWTLSPAERDAYLVNEATK	
92	SGN-U577900	67191	123	5	8.7	25.49	YQLATSR	
92						25.1	AISFSITRPASSR	
92						57.92	IRPPAHAADEEYVAK	
92						32.76	TTQEKNEQEEILTFNK	
92						76.59	IRPPAHAADEEYVAKYQLATSR	
93	SGN-U579197	14879	122	6	25.2	43.8	GGGGGYGRGGGGGYGR	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
93						72.92	GGGGGYGRGGGGGYGR	
93						26.55	GGGGGYGRGGGGGYGR	
93						38.03	GGYCQYGCCGHGDNGCYR	
93						26.4	GGYCQYGCCGHGDNGCYR	
93						38.4	RGGYCQYGCCGHGDNGCYR	
94	SGN-U576976	26863	117	2	12.4	64.48	VTPNNVDIAR	
94						77.65	VSPTYHLYSPSEVEDVISRL	
95	SGN-U577726	16308	115	5	24	32.43	VMPAVIVR	Oxidation (M)
95						44.74	VMPAVIVR	Oxidation (M)
95						23.31	NLYIISVK	
95						41.23	MSLGLPVAATVNCADNTGAK	
95						63.41	MSLGLPVAATVNCADNTGAK	Oxidation (M)
96	SGN-U580023	39828	114	6	22.5	24.52	VPFLFTIK	
96						43.02	RLTYDEIQSK	
96						57.53	FCLEPTSFTVK	
96						46.26	KFCLEPTSFTVK	
96						31.2	GTGTANQCPTIEGGVGSFAFKP GK	
96						22.86	QLVASGKPESFSGEFLVPSYRGS SFLDPK	
97	SGN-U577906	43212	113	3	8.3	66.31	EGISAEVINLR	
97						51.16	VLAPYSSEDAR	
97						54.89	MAVPQIEDIVR	Oxidation (M)
98	SGN-U564876	24122	113	3	23.7	26.93	TFQGPPHGIQVER	
98						61.45	VTPQPGVPPEEAGAAVAAESST GTWTTVWTDGLTSLDR	
98						68.36	VTPQPGVPPEEAGAAVAAESST GTWTTVWTDGLTSLDR	
99	SGN-U565096	27786	110	2	11.2	69.86	AAGITSIGVR	
99						68.09	VLTTEEIDEHLTAISERD	
100	SGN-U577918	70304	109	2	4.2	56.56	NLSQQNLNALAK	
100						79.01	ALPTYTPESPADATR	
101	SGN-U577569	40355	108	3	8.8	69	GADYANNEIQR	
101						20.74	GADYANNEIQRLER	
101						67.92	SAEALAEYVNSEAGTNVK	
102	SGN-U580521	45292	108	4	15.1	34.48	ALQNTVLK	
102						49.53	AKANSLAQLGK	
102						66.4	GILAIDESNATAGKR	
102						44.96	YAAISQDNGLVPIVEPEILLDGD HPIER	
103	SGN-U579745	24614	107	4	13.4	41.48	LRAEYLR	
103						43.27	AWPYVQNDLR	
103						47.69	FYLQPLTPAEAAQR	
103						55.69	FYLQPLTPAEAAQR	
104	SGN-U578890	20796	105	3	13.8	57.47	AAQEFLDVHNK	
104						23.42	SIELGCAQATCSK	
104						70.85	AAQEFLDVHNKAR	
105	SGN-U579474	30447	104	3	14.4	49.52	TGAQVIYSK	
105						29.43	YTTLKPLGDRVLVK	
105						76.84	YAGSEFKGADGSDYITLR	
106	SGN-U577591	19049	104	2	5.7	61.51	LATSGANFAR	
106						70.85	LATSGANFAR	
107	SGN-U567805	25333	103	4	12.3	40.19	VTNTRTGAQTTVR	
107						42.25	VTNTRTGAQTTVR	
107						53.05	YGWTAFCGPVGPR	
107						51.06	SKYGWTAFCGPVGPR	
108	SGN-U571844	36555	102	2	8.5	47.5	CSSFSER	
108						86.22	LVEAECPGVVSCADIVALVAR	
109	SGN-U578475	25995	102	3	7.4	36.34	DQPLDVLFEEIK	
109						61.31	LALTKDQPLDVLFEEIK	
109						48.91	LALTKDQPLDVLFEEIK	
110	SGN-U590837	38320	102	2	7.8	38.35	TYNNNLIQHVK	
110						87.18	YIAVGNEVSPFNENSK	
111	SGN-U571081	46977	102	3	5.9	26.05	TFASWGVDYLK	
111						44.89	LGVYSDAGTQTCSK	
111						81.13	LGVYSDAGTQTCSK	
112	SGN-U580502	13628	101	4	33	55.56	FVCEGESDEPK	
112						44.17	FVCEGESDEPK	
112						47.56	FICEGESDPKRPNACTFNCDPNI AYSR	
112						31.29	FICEGESDPKRPNACTFNCDPNI AYSR	
113	SGN-U577720	85298	100	3	5.6	50.61	TIRTQLASAK	
113						20.8	ASAALQGSDHR	
113						78.27	KLNLPILPTTTIGSFPQTVELR	
114	SGN-U580420	17223	99	2	14.4	62.72	FNKVNYANVSTNNYALDEVEE	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
							VK	
114						55.44	FNKVNYANVSTNNYALDEVEE VK	
115	SGN-U578449	29144	98	3	14	56.51	TGGPFGTMR	Oxidation (M)
115						21.58	NCAPIMLR	
115						69.42	QMGLSDKDIVALSGAHTLGR	Oxidation (M)
116	SGN-U580606	17646	97	6	27.1	26.55	IINDRETGR	
116						29.42	IINDRETGR	
116						32.76	NITVNEAQSR	
116						38.49	NITVNEAQSR	
116						73.31	GGGGGGSDGNWRN	
116						40.17	EGGYGGGGGGYGGGDR	
117	SGN-U578197	44508	96	2	8.8	64.81	KSEYQEPSSEYGSGYGR	
117						60.66	KSEYEEPTPQYGSGYGR	
118	SGN-U577463	20846	96	3	21.7	44.32	ELGTVMR	Oxidation (M)
118						41.11	VFDKDQNGFISAAELR	
118						64.59	EADVDGDGQINYDEFVK	
119	SGN-U586248	44587	96	2	2.7	95.14	AAGEEAVTR	
119						28.46	AKAAGEEAVTR	
120	SGN-U579254	11499	96	2	14.2	47.15	AANVTVEPYWPLLFAK	
120						65.88	AANVTVEPYWPLLFAK	
121	SGN-U578802	43198	95	4	12.5	40.57	VAINGFGR	
121						75.82	AAALNIVPTSTGAAK	
121						23.49	VVDLADIVANQWK	
121						33.64	GTMTTTHSYTGDQR	Oxidation (M)
122	SGN-U577903	19892	93	2	8.6	53.14	VSQIDITPYPIKGGR	
122						60.13	VSQIDITPYPIKGGR	
123	SGN-U579345	21489	93	3	12.2	72.53	LAAFAQNYANQR	
123						33.24	VCGHYTQVVWR	
123						51.61	VCGHYTQVVWR	
124	SGN-U577960	45477	91	4	5.1	40.98	TAAAPIER	
124						48.2	TAAAPIER	
124						46.99	TAAAPIER	
124						43.62	AVAGAGVLAGYDK	
125	SGN-U581507	28426	91	2	14.4	39.28	ELFEQMLSFR	Oxidation (M)
125						76.27	AGQGIGVGQELVNNPDLVATDP IISFK	
126	SGN-U579857	33395	90	3	14.1	43.16	LLILTDPR	
126						68.5	FAQYTGAHAIAGR	
126						33.76	VIVAIENPQDIIVQSARPYGQR	
127	SGN-U562749	11134	90	2	11.3	48.74	IMTQPINLIFR	Oxidation (M)
127						64.02	IMTQPINLIFR	Oxidation (M)
128	SGN-U581320	16189	89	4	10.6	59.68	YGEVVEAR	
128						59.78	YGEVVEAR	
128						21.28	IIYDRESGR	
128						20.7	IIYDRESGR	
129	SGN-U580870	30802	89	2	4.5	64.03	AAAVSPADDELAK	
129						54.73	AAAVSPADDELAK	
130	SGN-U573751	49012	88	2	7.3	61.11	SVDEYDYLPYFYSR	
130						50.19	IVGAFLESGSPEENKAIAK	
131	SGN-U570208	8312	88	3	43.4	24.07	FGLALALKP	
131						80.83	SLFTVSGEVDTK	
131						32.53	ASALLQHEWRPK	
132	SGN-U563658	28246	85	3	13.8	37.05	LLSSGAYHK	
132						27.8	GLVYCKPCKFR	
132						72.5	GINTLNQAKPLQGAK	
133	SGN-U577200	12671	85	4	18.3	37.45	QYVNSPNAK	
133						37.42	QYVNSPNAK	
133						41.88	VAKTCGVSTPSC	
133						46.13	VAKTCGVSTPSC	
134	SGN-U579126	31819	85	2	6.5	84.64	ITIDPEDPAAVSEYAK	
134						28.72	EKITIDPEDPAAVSEYAK	
135	SGN-U577771	38634	84	4	13.7	29.83	SQASALEK	
135						51.79	MELVDAAFPLLK	Oxidation (M)
135						32.43	LDHNRALGQISER	
135						48.75	VLVVANPANTNALILK	
136	SGN-U574346	26287	83	2	7.3	70.35	LVTVDGEGFTPFVFIKA	
136						30.5	LVTVDGEGFTPFVFIKA	
137	SGN-U583987	27700	83	1	5.2	82.83	IGFIDFATTSPLR	
138	SGN-U581393	23283	82	3	13.6	40.71	IQQVSSALLK	
138						28.61	IFLTPISDVIR	
138						66.05	TGEIGDGKIFLTPISDVIR	
139	SGN-U562985	31996	82	1	7	81.91	LFLFPANTPATVGSGVSQSR	
140	SGN-U580783	46276	80	2	3.6	71.53	FVIGGPHGDAGLTGR	
140						37.03	FVIGGPHGDAGLTGR	
141	SGN-U578717	67463	80	1	3.3	79.61	SAMMTTADTLNLANSPILDER	2 Oxidation (M)

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
142	SGN-U579217	59135	80	5	11.3	29.02	SAIGEGMTR	Oxidation (M)
142						29.95	FVSQGAYDTR	
142						37.13	DISGSSINPSER	
142						34.7	QIYPPINVLPSLSR	
142						52.37	VTLFLNLANDPTIER	
143	SGN-U570662	18417	78	4	11.9	49.39	VLVVDGGGSLR	
143						47.82	VLVVDGGGSLR	
143						49.19	VFEDNVLVR	
143						30.27	VFEDNVLVR	
144	SGN-U569163	20772	78	1	4.9	78.14	SGEEVYICR	
145	SGN-U579080	19441	78	4	8.5	63.84	VCGNPHAIIR	
145						34.47	VCGNPHAIIRK	
145						30.31	TCRVCGNPHAIIR	
145						36.25	TCRVCGNPHAIIR	
146	SGN-U578351	42981	76	3	4.7	23.54	AGGLQLEK	
146						59.88	GVSCPLICFK	
146						44.18	GVSCPLICFK	
147	SGN-U576441	34153	76	1	3.9	75.82	VLLTLDGSNPVK	
148	SGN-U577010	22492	76	2	15.9	73.95	STFQAAGLPLNAK	
148						24.36	TTIMLADLGDFQKVNAIYAK	Oxidation (M)
149	SGN-U579307	24065	76	5	33.8	29.35	TIELSKQ	
149						25.63	VVATLKEK	
149						48.93	LLDVYESR	
149						36.71	AITQYIAHTYADKGNQLLPNDP K	
149						41.39	YLGGESFTLADLHHAPSLHYLS GSKVK	
150	SGN-U562704	16112	76	1	8.5	75.66	AGALGDSVTVTR	
151	SGN-U563622	17857	75	2	16.8	48.84	YGTLPQDEASETAR	
151						53.36	LIESLSTPSILSKR	
152	SGN-U563289	7015	75	2	37.5	55.7	ASGVSYSSVVK	
152						38.34	ATYQVAALPMNAR	Oxidation (M)
153	SGN-U579709	30520	74	4	23	43.67	SVDETLR	
153						50.88	AYNVLIPDQGIALR	
153						27.11	SGGLGDLNYPLISDVTK	
153						28.35	TLQALQYVQENPDEVCPAGWK PGEK	
154	SGN-U577216	13855	74	3	9.1	50.22	AGLQFPVGR	
154						41.86	SSKAGLQFPVGR	
154						41.2	SSKAGLQFPVGR	
155	SGN-U579538	31506	74	2	7.4	43.67	SVDETLR	
155						58.2	SYNVLIPDQGIALR	
156	SGN-U570979	16616	74	7	30.6	34.06	IYGGNQR	
156						21.54	IYGGNQR	
156						32.07	DLDQVAGR	
156						37.96	AYAAHLKR	
156						39.5	ITSNGQRDLDQVAGR	
156						33.01	ELAPYDPDWYYIR	
156						28.73	LKELAPYDPDWYYIR	
157	SGN-U575256	65125	73	2	3.6	36.2	NMVTAQGR	Oxidation (M)
157						63.58	TDPNQNTGIVIQK	
158	SGN-U579609	26931	72	1	5.1	72.46	LGNAHVTATCGAR	
159	SGN-U579696	42422	72	3	8.7	60.3	AAINTEAR	
159						39.91	TAYGIMAR	Oxidation (M)
159						20.73	SNAGSIDDGSMTYNQIK	Oxidation (M)
160	SGN-U577797	25107	72	1	4	72.21	VVDIVDTFR	
161	SGN-U581258	14456	72	2	8.5	60.97	FIAPILADIAK	
161						43.41	FIAPILADIAK	
162	SGN-U581255	42349	72	1	3.4	71.82	ALQESLASELASR	
163	SGN-U578421	52835	71	3	6.2	23.55	NGVCDQYRK	
163						28.9	NGVCDQYRK	
163						61.53	NVASSSGLCGLAIEPSYPVK	
164	SGN-U569271	67674	71	1	2	71.23	VSAANSRPPNPQ	
165	SGN-U578270	11089	71	2	11.7	42.11	STASISETPITK	
165						57.78	STASISETPITK	
166	SGN-U585601	27104	70	2	4.3	46.77	GLCTNCCAGK	
166						46.77	GLCTNCCAGK	
167	SGN-U572068	16588	69	1	6.6	69.2	TLGEWAGLCK	
168	SGN-U580659	12507	69	5	43.5	39.11	KAACTCLK	
168						53.51	SAAATITGINYR	
168						31.23	LLFPCLAYLR	
168						23.64	LLFPCLAYLR	
168						29.53	DKGGIGSCCSGVSSLANAAK	
169	SGN-U579054	32610	68	1	5.9	68.27	AASSEGVALIVPDTSPR	
170	SGN-U567576	8711	68	2	15.8	56.36	AVVYALSPFQQK	
170						32.82	AVVYALSPFQQK	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
171	SGN-U562785	26374	68	3	12.9	35.33	QAEQYYR	
171						47.27	AANGVVIATEK	
171						38.41	LYKEPIPVTQLVR	
172	SGN-U578780	14105	68	2	28.1	61.5	IASESDVSVHSTFASR	
172						30.66	EAAYQIINDELMLDGNPR	Oxidation (M)
173	SGN-U577856	20019	67	3	13.6	43.95	ALVIDGDNLIPK	
173						46.45	ALVIDGDNLIPK	
173						34.75	IHAIDDKNLVTK	
174	SGN-U592277	18697	67	1	7.1	66.94	YFQVCVSNPTAR	
175	SGN-U593370	20957	66	4	18.7	20.34	HVPGFIEK	
175						35.25	YALLVDDLEVK	
175						36.81	YALLVDDLEVK	
175						42.64	VIIFAVPGAFTPTCSMK	Oxidation (M)
176	SGN-U578385	17276	66	3	13.2	48.88	AGVVRQELAK	
176						34.04	ALVDAPDMVR	
176						40.42	ALVDAPDMVR	Oxidation (M)
177	SGN-U581550	27763	66	2	8.7	22.88	YIAGLQQK	
177						65.77	LTVEDPVTVEYITR	
178	SGN-U583104	73904	65	2	2.5	25.08	IAGLDVQR	
178						64.79	VIENSEGAR	
179	SGN-U579280	45500	65	2	7.3	45.08	TVGTVGAGR	
179						50.16	YMPNQAMTPHISGTTIDAQLR	2 Oxidation (M)
180	SGN-U577345	16404	65	1	7	64.68	VLITTDLLAR	
181	SGN-U567463	37224	64	1	3.8	63.86	VGGADDVYIGDIR	
182	SGN-U579540	23018	64	1	11.2	63.81	SHIANLAQVTSNALALINQYAA NH	
183	SGN-U577318	24561	63	1	7.9	63.2	ALVPTDLSIAVPQGTYAR	
184	SGN-U580828	26934	63	5	14.3	22.1	LLREPINF	
184						21.7	LLREPINF	
184						27.53	LLREPINF	
184						44.67	VKAMTDISSEAK	Oxidation (M)
184						38.09	FSQFMKDLSVPTETR	Oxidation (M)
185	SGN-U578250	20790	62	1	4.2	62.26	VVEVSTSK	
186	SGN-U591780	11460	62	4	17.5	39.81	CFCTKPC	
186						37.53	CFCTKPC	
186						35.58	CFCTKPC	
186						21.46	FSGGNCRGFR	
187	SGN-U572631	30361	62	1	3.7	61.55	VHITDAKEQR	
188	SGN-U578863	23175	62	2	9.6	53.69	ATGCSFICPR	
188						35.82	HLAVNRPVFK	
189	SGN-U577338	10823	61	3	13.5	31.58	TAGPPVVMNPISR	
189						44.04	TAGPPVVMNPISR	Oxidation (M)
189						33.09	TAGPPVVMNPISR	Oxidation (M)
190	SGN-U577638	10335	61	2	20.8	28.47	TACSCLK	
190						55.32	SAASIIKGIDMSK	Oxidation (M)
191	SGN-U580980	18120	61	1	9.5	60.86	IVQLNDAIDDISNQLR	
192	SGN-U575297	56351	61	2	6.5	51.22	AASELSSHDPPIVLAK	
192						36.1	SEPIPEVNDEPVKVVVR	
193	SGN-U583862	9444	61	1	30.1	60.55	ALVVHELEDDLGKGGHELSLTT GNAGGR	
194	SGN-U579937	20930	60	1	7.8	59.96	ASDVTGPNEAAVKGGSR	
195	SGN-U576271	32715	60	2	6.1	46.49	TPNVSYLPK	
195						41.58	QVAQQEAER	
196	SGN-U564094	27391	60	3	17.7	30.06	KDDTAAVLK	
196						33.78	TVVQAEKLDVMLQNR	Oxidation (M)
196						51.74	FAPPKVPGIDDVTGEPLIQR	
197	SGN-U576269	31601	59	2	6.2	44.81	TPNVAYLPK	
197						41.58	QVAQQEAER	
198	SGN-U578755	32509	59	1	3.9	59.02	IVAAALNPVDFK	
199	SGN-U569119	22019	59	2	4.9	34.8	NVAQADASVK	
199						45.72	NVAQADASVK	
200	SGN-U566921	61840	58	2	2.7	37.71	TIDEFPPIVFAGEAR	
200						52.55	TIDEFPPIVFAGEAR	
201	SGN-U578219	52689	58	1	3.2	58.14	AVFVDLEPTVIDEVR	
202	SGN-U578349	26593	58	1	3.8	57.65	GLVGEIISR	
203	SGN-U579891	11799	57	2	28.2	39.45	LIPTLNRVLIEK	
203						41.07	ITAPAKTSAGILLPENSSK	
204	SGN-U579310	17773	57	4	19.4	21.41	LVLPGELAK	
204						26.53	LVLPGELAK	
204						40.91	INKKPTITSR	
204						42.11	HAVSEGTKAVTK	
205	SGN-U568571	59533	56	1	1.9	56.44	LAANAFLAQR	
206	SGN-U563063	24057	56	1	4.9	56.32	AFLDATGGLWR	
207	SGN-U569146	65633	56	1	1.9	55.98	ALAENEGEGLAK	
208	SGN-U584066	19943	56	1	4.7	55.93	GVEDVILR	
209	SGN-U581281	27932	56	1	8.1	55.66	SITDYGSPEEFLSKVDYLLGK	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
210	SGN-U583757	26197	55	1	7.7	55.26	LSARFDSLDSEGGPGPER	
211	SGN-U581617	28960	55	1	6.3	54.57	GVDTLLGASPIQGAVVK	
212	SGN-U572927	12116	55	1	10.8	54.52	TVGNFNTLDYLR	
213	SGN-U580596	31504	54	1	5.9	54.46	VETTPFEGQKPGTSGLR	
214	SGN-U579846	16287	54	1	17	54.17	TNQALIIGIYDEPMTPGQCNMIVER	2 Oxidation (M)
215	SGN-U574380	33150	54	1	3.7	53.85	IITETNESWAK	
216	SGN-U574622	86505	54	1	2.1	53.82	ERPIFSLSDDHAMFKK	Oxidation (M)
217	SGN-U585756	36599	53	1	6.1	53.15	LVAVSNTVAFTNPNAPLYPR	
218	SGN-U578980	12846	53	2	13.3	44.14	LSWPELIGVPAQYAK	
218						37.38	LSWPELIGVPAQYAK	
219	SGN-U578546	34466	53	1	3	52.77	AMGAGVGAR	Oxidation (M)
220	SGN-U583670	37539	53	1	2.6	52.63	VAVTVDAPR	
221	SGN-U593766	26058	52	3	21	36.34	LPPPEPK	
221						31.8	VQVEEDNVLLISGER	
221						35.67	AMAATPADVKEYPNSYVFVVDMPGLK	2 Oxidation (M)
222	SGN-U583388	56028	52	2	3	50.11	NVGGNPNAVYEVK	
222						26.27	VVKNVGGNPNAVYEVK	
223	SGN-U566943	17485	52	3	18.2	21.64	LVQDIVLK	
223						52.55	ACGTCCVR	
223						25.13	CSLHSRPNVCFR	
224	SGN-U569989	41653	52	3	4.3	26.88	MKIFVK	
224						29.5	MKIFVK	Oxidation (M)
224						51.54	LIQEHQADFLR	
225	SGN-U579794	30156	51	1	3.6	51.45	GQIYDVSQSR	
226	SGN-U577566	22221	51	1	7.3	51.07	VLEQLSGQSPVFSK	
227	SGN-U578342	25100	51	2	10.3	26.23	IIGFDNVR	
227						50.72	QVQCISFIAYKPEGY	
228	SGN-U568406	12472	50	2	8	38.01	LQPTQQQTR	
228						40.49	LQPTQQQTR	
229	SGN-U565452	23005	50	2	4.8	26.62	VALEACVK	
229						50.42	VALEACVKAR	
230	SGN-U580587	29614	50	3	14.3	22.31	ASTVVSPK	
230						50.27	ALSVSPGNTVLYSK	
230						23.99	YAGNDFKGADGSEYITLR	
231	SGN-U576354	52562	50	1	3	50.15	VLSASPSAYSANPK	
232	SGN-U582384	28634	50	1	4.7	49.88	FTPVDANENIVR	
233	SGN-U580742	17877	50	3	18.6	36.6	PKMSLIPR	
233						35.54	FRLPENAK	
233						38.21	IDWKETPEAHVFK	
234	SGN-U579013	25020	49	1	5.6	49.42	SKVTIADSGELPL	
235	SGN-U568670	28752	49	1	3.5	49.07	TTIFSPEGR	
236	SGN-U568182	11768	49	1	13.5	49.04	CNTKPICLTLCLAK	
237	SGN-U578700	37530	49	1	2.7	49.01	FIINDSVLFK	
238	SGN-U597130	21796	49	2	9.3	20.34	AMVAHNNAR	Oxidation (M)
238						48.99	IPVFLDGGVR	
239	SGN-U564185	37355	49	3	7.4	47.17	MGASLIR	Oxidation (M)
239						29.24	MGASLIR	Oxidation (M)
239						30.46	DSVAMLGGIPYPVSLGRR	Oxidation (M)
240	SGN-U579112	50916	49	1	4.5	48.56	VSQMIEDHEPFEGQALLDAK	Oxidation (M)
241	SGN-U579420	42380	48	1	3.4	48.39	TLKGTFFGNYKPR	
242	SGN-U583105	73221	48	2	2.5	25.08	IAGLDVQR	
242						48.26	VIENAEGAR	
243	SGN-U579652	17111	48	1	5.5	48.21	AVDAVMGPR	Oxidation (M)
244	SGN-U576036	44278	48	2	2.7	33.25	VLVTPTSDLGK	
244						43.79	VLVTPTSDLGK	
245	SGN-U578288	13947	48	1	6.5	47.93	LIASVEMK	Oxidation (M)
246	SGN-U578721	11677	47	1	9.8	47.12	SQDGNEVFFR	
247	SGN-U577811	57887	47	1	3.1	46.8	AAEILSQNDPPVVLAK	
248	SGN-U579020	66201	47	1	2.5	46.75	GSGFVAFSTPEEASR	
249	SGN-U564605	16830	47	2	14.3	43.01	IILGGKER	
249						32.55	HFKQVVDIDDAAK	
250	SGN-U573980	31582	46	1	3.5	46.46	AIYSLGSIVR	
251	SGN-U577505	39365	46	1	4.5	46.42	NIQNAISGAGLGNQIK	
252	SGN-U577470	30033	46	1	3.7	45.91	TVAVAEIIKR	
253	SGN-U562887	30688	46	2	7.1	45.83	QCPAPYPEGR	
253						21.42	YCDPQHCQK	
254	SGN-U577182	36918	45	3	7.2	42.07	IGLDGFGR	
254						22.22	IGLDGFGR	
254						31.89	FANGGAYPPDLSLITK	
255	SGN-U577273	10473	45	1	11.5	45.44	ATLEVSCKPK	
256	SGN-U579155	14226	45	2	12.1	45.08	TVGTVGAGR	
256						21.55	KDYVDLK	
257	SGN-U585129	41802	45	1	2.7	44.89	NGECSPELQR	
258	SGN-U577521	16840	45	1	10.8	44.78	VGSIVSCNYNEAGQKK	

Protein Hit							Peptide Hit	
#	SGN Unigene	Mass	Total Ion Score	Peptide Matches	Protein Coverage	Peptide Score	Peptide Sequence	Variable Modification
259	SGN-U579059	38956	44	4	11.6	29.08	VLLQIKK	
259						26.97	CLCGSPLPK	
259						25.41	CLCGSPLPK	
259						37.15	DLGNPYHLASWDPNTDCCYWYVVK	
260	SGN-U582906	32304	44	1	2.7	44.18	QATYTTAR	
261	SGN-U578651	25278	44	1	6.8	43.79	LSGKDVVFEYPITEA	
262	SGN-U580444	23464	43	1	3.6	42.78	IATIKAVR	
263	SGN-U577838	12533	43	2	11.7	24.99	FGISPEAAMNLPK	Oxidation (M)
263						36.71	FGISPEAAMNLPK	Oxidation (M)
264	SGN-U573933	50141	42	1	5	42.2	LAKENAPAIIFIDEVDAIATAR	
265	SGN-U572197	17564	42	1	4.6	41.84	WLNISAK	
266	SGN-U578077	30974	41	1	5.8	41.28	IAIITGGASGIGAATTR	
267	SGN-U565851	18289	41	2	10.7	22	IDNPSSR	
267						41.17	TSCSSPKYTR	
268	SGN-U570385	9612	40	1	15.3	40.31	AALFDGIEEGGIR	
269	SGN-U571799	29487	40	1	3.8	40.3	SAVNKLQIAK	
270	SGN-U578607	39886	40	1	2.6	40	VIAQTDGTR	
271	SGN-U567329	41520	40	1	3.2	39.67	AATLSSPMLKGK	Oxidation (M)
272	SGN-U580879	20399	39	2	4.7	38.55	MAGDFLSKK	
272						27.99	MAGDFLSKK	Oxidation (M)
273	SGN-U579867	57244	39	1	2.3	38.51	SIPSIIDLGSLK	
274	SGN-U602617	8453	37	1	11.8	37.24	QRLSWSSCK	
275	SGN-U588010	3011	37	2	31.4	32.76	GGGGGYQGGDR	
275						30.97	GGGGGYQGGDR	
276	SGN-U564174	18530	37	2	10.4	35.31	YLEEEAR	
276						31.12	VGESVYNSGK	
277	SGN-U573714	51894	37	1	1.9	36.67	AAAQANLAK	
278	SGN-U577413	38097	36	1	2.9	36.35	AESIVQSTVR	
279	SGN-U585819	24772	36	2	8.1	36.2	NMVTAQGR	Oxidation (M)
279						20.7	VGADQSVINR	
280	SGN-U567538	41484	36	1	2.5	36.05	ELPKEFDAR	
281	SGN-U582752	11526	36	1	13.7	36.05	ILFSPSSSESAATR	
282	SGN-U566615	23246	36	1	3.4	35.88	IDDPASR	
283	SGN-U577721	16915	36	1	6.4	35.72	LRPTDCKPR	
284	SGN-U564846	30089	35	1	5.1	35.33	FRKPVIAGDTLVMR	
285	SGN-U573432	71083	35	1	2.2	35.27	TPSAFSLLDPSFTR	
286	SGN-U601140	5292	35	1	15.9	35.06	QINRALK	
287	SGN-U577815	19604	35	1	8.4	35.03	IFFIAWCPDTSKVR	
288	SGN-U580665	54341	35	2	5.1	34.98	IGTPAMTSR	Oxidation (M)
288						23.37	NAVFGDSSALAPGGVR	
289	SGN-U604783	25267	34	3	11.4	33.86	LAPCLAWR	
289						34.27	LAPCLAWR	
289						20.63	LLVLCLLDNMGVTGIAR	Oxidation (M)
290	SGN-U579097	30985	34	1	3	34	TQELVAAK	
291	SGN-U581940	41355	34	1	3.3	33.89	LLGNFLQEPVPR	
292	SGN-U565098	19834	34	1	3.6	33.88	GWETNR	
293	SGN-U579697	19798	34	1	5.6	33.74	ANQAANTISK	
294	SGN-U601373	29900	34	3	3	32.59	KDGVLSPR	
294						27.24	KDGVLSPR	
294						34.85	KDGVLSPR	

Supplementary Table S5. Co-expression results of arabidopsis GDSL-motif lipase/hydrolase family proteins and known cutin-biosynthesis genes

		R^2 by Bait Gene[1]									Average R^2	
Category	Co-Expressed Gene	*ATT1* (At4g00360)	*LACS2* (At1g49430)	*LCR* (At2g45970)	*GPAT4* (At1g01610)	*GPAT8* (At4g00400)	*GPAT6* (At2g38110)	*CYP86A4* (At1g01600)	*CYP86A7* (At1g63710)	*CYP77A6* (At3g10570)	By Co-Expressed Gene	By Category
Clade I	At1g33811	**0.2604**	0.0011	0.0871	0.0638	**0.3759**	0.0059	0.0003	0.0389	0.0907	0.1027	0.1585
	At1g29670	**0.4204**	0.0006	0.0544	0.1349	**0.5572**	0.0206	0.0009	0.0109	0.0391	0.1377	
	At1g29660	**0.5477**	0.0156	0.0658	0.1320	**0.4389**	0.0726	0.0103	0.0036	0.0006	0.1430	
	At5g45670	0.0131	**0.4079**	**0.4149**	0.1506	0.2061	0.1050	**0.3597**	**0.2395**	**0.2714**	**0.2409**	
	At4g18970	0.0876	0.1268	0.2364	0.1186	**0.4830**	0.0087	0.0809	0.1446	0.2296	0.1685	
Clade II	At2g03980	**0.2626**	0.0085	0.0108	0.0217	**0.2379**	0.0368	0.0302	0.0004	0.0187	0.0697	0.0853
	At3g14220	0.0226	0.0155	0.1795	0.0572	0.1686	0.0164	0.0377	0.0055	0.0184	0.0579	
	At1g56670	0.0011	0.0164	0.1373	0.0367	0.0289	0.0061	0.0604	0.0154	0.0079	0.0345	
	At3g48460	0.1594	0.1144	0.1600	0.1671	**0.3119**	0.0275	0.1078	0.1893	0.2352	0.1636	
Clade III	At3g11210	0.0208	0.2286	0.2021	0.0613	0.1054	0.0855	**0.2593**	0.1943	0.2160	0.1526	0.1927
	At3g16370	**0.3774**	0.0192	0.1793	0.2295	**0.6545**	0.0035	0.0133	0.0284	0.0657	0.1745	
	At5g45960	0.1760	**0.3108**	0.0553	0.0037	0.0339	**0.3838**	0.2314	**0.5233**	**0.6133**	**0.2590**	
	At5g45950	0.1429	0.0870	0.1992	0.1385	**0.5651**	0.0163	0.0418	0.1380	**0.2470**	0.1751	
	At2g04570	0.0412	0.1773	0.1015	0.1174	**0.3520**	0.0621	0.1164	**0.3140**	**0.3918**	0.1860	
	At4g26790	0.0136	**0.2427**	0.2363	0.0809	0.0470	0.1357	**0.2490**	**0.2389**	**0.2738**	0.1687	
Clade IV	At3g04290	0.0034	0.2126	**0.2980**	0.1256	**0.3561**	0.0589	0.1441	**0.2734**	**0.3639**	0.2040	0.1390
	At5g18430	0.0902	0.0065	0.1678	0.0235	0.2189	0.0001	0.0271	0.0542	0.1009	0.0766	
	At4g28780	0.0343	0.1075	0.1846	0.0770	**0.2811**	0.0313	0.0672	0.2042	**0.2404**	0.1364	
Bait	ATT1 (At4g00360)	**1.0000**	0.0456	0.0302	0.1919	**0.3094**	0.1525	0.0464	0.0838	0.0685	0.1160	**0.2374**
	LACS2 (At1g49430)	0.0456	**1.0000**	0.2101	0.2004	0.0623	**0.2743**	**0.3121**	**0.3201**	**0.2905**	0.2144	
	LCR (At2g45970)	0.0302	0.2101	**1.0000**	**0.4532**	**0.4246**	0.0776	**0.3584**	0.1543	0.1687	0.2346	
	GPAT4 (At1g01610)	0.1919	0.2004	**0.4532**	**1.0000**	**0.4219**	0.0472	0.1742	0.0834	0.0600	0.2040	
	GPAT8 (At4g00400)	**0.3094**	0.0623	**0.4246**	**0.4219**	**1.0000**	0.0001	0.0899	0.0740	0.0953	0.1847	
	GPAT6 (At2g38110)	0.1525	**0.2743**	0.0776	0.0472	0.0001	**1.0000**	**0.3990**	**0.4990**	**0.5144**	**0.2455**	
	CYP86A4 (At1g01600)	0.0464	**0.3121**	**0.3584**	0.1742	0.0899	**0.3990**	**1.0000**	**0.5866**	**0.4512**	**0.3022**	
	CYP86A7 (At1g63710)	0.0838	**0.3201**	0.1543	0.0834	0.0740	**0.4990**	**0.5866**	**1.0000**	**0.8168**	**0.3272**	
	CYP77A6 (At3g10570)	0.0685	**0.2905**	0.1687	0.0600	0.0953	**0.5144**	**0.4512**	**0.8168**	**1.0000**	**0.3082**	

1. Square of Pearson's correlation coefficient (R^2) for each gene's expression profile versus the nine bait genes tested (CressExpress 3.0, www.cressexpress.org). For comparison, the co-expression of the nine bait genes is indicated (gray cells) and R^2 values exceeding the average for the pairwise comparison of baits (0.2374) are shown in **bold**.

APPENDIX 2

Related publications

The following publications include significant collaborations and a literature review that are related to the work described in this thesis.

Debono, A., Yeats, T.H., Rose, J.K., Bird, D., Jetter, R., Kunst, L. and Samuels, L. (2009) Arabidopsis LTPG is a glycosylphosphatidylinositol-anchored lipid transfer protein required for export of lipids to the plant surface. *Plant Cell* **21**, 1230–1238.

Matas, A.J., Yeats, T.H., Buda, G.J., *et al.* (2011) Tissue and cell type specific transcriptome profiling of expanding tomato fruit provides insights into metabolic and regulatory specialization and cuticle formation. *Plant Cell* (Submitted)

Yeats, T.H. and Rose, J.K. (2008) The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). *Protein Sci.* **17**, 191–198.