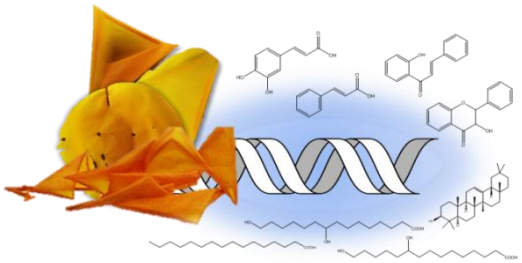
UNIVERSITAT POLITÈCNICA DE VALÈNCIA

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

IBMCP
Instituto de Biología Molecular y Celular de Plantas

TOMATO FLESHY FRUIT QUALITY IMPROVEMENT:

CHARACTERIZATION OF GENES AND GENOMIC REGIONS ASSOCIATED TO SPECIALIZED METABOLISM IN TOMATO FLESHY FRUIT



Reported by **JOSEFINA PATRICIA FERNÁNDEZ MORENO**

to obtain the degree of **Doctor in Biotechnology**

Supervisor: **ANTONIO GRANELL RICHART**

Co-Supervisor: **DIEGO ORZÁEZ CALATAYUD**

July, 2015

Front composed image: Tomato fruit isolated cuticle membranes (original), DNA double helix, phenylpropanoid, pentacyclic triterpenoid and very-long-chain fatty acid molecules (ChemBioDraw Ultra 13.0.0.3015).
Author: Josefina Patricia Fernández Moreno

Part of this work has been published in:

Methods in Molecular Biology. Springer Life Editorial. Volume 975, 2013, Chapter 14 (pp183-196).

Fernández-Moreno J.-P., Orzáez, D. and Granell, A. (2013). Virus-Induced Gene Silencing: a tool to study fruit development in *Solanum lycopersicum*. Methods in Molecular Biology. Springer Life Editorial. Volume 975, 2013, Chapter 14 (pp183-196).

Part of this work is submitted to:

***Metabolomics* journal**.

Fernández-Moreno, J.-P., Malitsky, S., Lashbrooke, J., Racovita R., Jetter, R., Orzaez, D., Aharoni, A. and Granell, A. "*An efficient method for the cuticular wax composition analysis of different plant species and organs*".

***Plant Physiology* journal**

Fernández-Moreno, J.-P., Tzfadia, O., Forment, J., Rogachev, I., Meir, S., Orzáez, D., Aharoni, A. and Granell, A. "*Characterization of a new 'pink fruit' tomato mutant results in the identification of a null allele of the SlMYB12 transcription factor*".

El trabajo aquí expuesto ha sido realizado por **Josefina Patricia Fernández Moreno** para la obtención del grado de Doctor en Biotecnología, en el *Instituto de Biología Molecular y Celular de Plantas*, centro mixto UPV-CSIC, bajo la tutela del director de tesis Prof. **Antonio Granell Richart**, y el co-director Dr. **Diego Orzáez Calatayud**.

The document here exposed have been done by **Josefina Patricia Fernández Moreno** to obtain the degree of Doctor in Biotechnology, at the *Instituto de Biología Molecular y Celular de Plantas*, a UPV-CSIC joint institute, under the supervision of Prof. **Antonio Granell Richart** and the co-supervision of Dr. **Diego Orzáez Calatayud**.

Prof. Antonio Granell Richart

(firmar arriba / to sign up)

Dr. Diego Orzáez Calatayud

(firmar arriba / to sign up)

In all your activities, apply the maxims

Train yourself impartially in any field. It is crucial that you do it every time and fully of enthusiastic eagerness

Atiśa Dīpa Kara Śrījñāna

ACKNOWLEDGMENTS

This work was possible thanks to the supervision of Prof. *Antonio Granell*, the co-supervision of Dr. *Diego Orzáez* and the financial supporting of the following instruments from the corresponding funding organisms: FPU fellowship (2007-2012) and short term grants (2010 and 2011) from the *Spanish Ministry of Education, Culture and Sport*; EUSOL project from EU; BIO2008-034034 grant from the *Spanish Ministry of Science and Technology*; and STSM grant from COST ACTION FA1106 'Fleshy Fruit' (2012-2013).

I am really grateful to Prof *Asaph Aharoni* for offering me the opportunity to develop new skills and gather new knowledge in his laboratory; to the A. Aharoni's Lab people for making enjoyable, refreshing and intriguing the art of the science, and especially to *Sergey Malitsky*, *Ilana Rogachev*, *Sagit Meir* and *Justin Lashbrooke* whose were collaborating with me producing both the general wax method and the characterization of new '*pink fruit*' mutants manuscripts; to Prof *Jian Xin Shi* for his inestimable help in explaining the details of the acyl-lipid extraction processes; to *Antonio Monforte*, for his incredible help in the QTL analyses; to both *Oren Tzfadia* and *Javier Forment* bioinformaticians for their essential job in this thesis; to Prof. *Reinhard Jetter* and *Radu Racovita* for their incredible help identifying cuticular waxes from tomato, apple and hybrid aspen; to *Asunción Fernández*, '*Asun*', for all the advices about how to be self-sufficient doing science during my first years of the Ph.D.; to *Silvia Presa* for be my hands

when I could not be in the lab; to *Ewa J. Mellerowicz* and *Ajaya Biswal* for their collaboration in the cuticular wax analysis on hybrid aspen leaves; to *Dana Elyahud* because her important help in growing the IL population during 2010 and 2011 summer seasons; to *Eyal Shimoni* for his great help in teaching me sample preparation and imaging for TEM electron microscopy; to *Calanit Raanan* for teaching me in preparing the paraffin blocks for histology; to *Dani Zamir* and *Itai Ofner* for providing the seeds collection of the *Solanum pennellii* introgression line population; to Prof *Dinesh Kumar* who kindly provided the TRV-based silencing vectors pTRV1 and pTRV2; to Prof *Cathie Martin* and *Eugenio Butelli* for providing Del/Ros1 transgenic tomato lines; and to *Clara Pons* for her help in the identification of Solyc IDs from *Affymetrix* gene IDs.

I cannot finish these acknowledgment section without expressing my gratitude to all those, (friends, labmates and dharma-friends), that supported me in one way or another during this long process leading to my PhD. Specially I would like to thanks to my Guru *Gueshe Lobsang Lamsang* who teach me about patient and enthusiastic endeavor in the study and in the practice.

Finally, the last words are for deeply thank the invaluable help and support that my family, and particularly my dear parents, provide me among the full process, because, without their care this way definitely would have been harder if possible.

Sincerely, to all of them: ***thank you***!

Josefina Patricia Fernández Moreno
(Pathy)

CONTENTS

ABBREVIATIONS

4CL 4-Coumarate-CoA ligase;

ADT Arogenate dehydratase;

ANS Anthocyanidin synthase;

C3H *p*-Coumaroyl 3-hydroxylase;

C4H Cinnamate 4-hydroxylase;

CCR Cinnamoyl-CoA reductase;

CHI Chalcone isomerase;

CHS Chalcone synthase;

CM1 Corismate mutase;

COMT Caffeic acid O-methyltransferase;

CW Cuticular waxes;

DAHP Phospho-2-dehydro-3-deoxyheptonate-aldolase synthetase;

DFR Dihydroflavonol reductase;

EPSP 3-phosphoshikimate-1-carboxivinil-transferase synthetase;

F3'H Flavonone 3'-hydroxylase;

F3H Flavonone 3-hydroxylase;

FLS Flavonol synthase;

GOI Gene of interest;

GT Glycosyl transferase;

HCA Hierarchical clustering analysis;

HCT Cinnamoyl-CoA transferase;

HQT Hydroxycinnamoyl-CoA quinate transferase;

IFS Isoflavonol synthase;

PAL Phenylalanine amonio lyase;

PDT Prephrenate dehydratase;

PP Phenylpropanoid pathway;

PPA-AT Prephenic acid amino transferase;

PPY-AT Phenylpyruvate aminotransferase;

QTL Quantitative-trait-loci;

RT Rhamnosyl transferase;

TF Transcription factor;

VIGS Virus-induced gene silencing;

VLCFA Very-long-chain fatty acid;

y-lines Defines a set of lines with pink fruit phenotype described by Ballester et al. (2010) and Lin et al., (2014).

y-mutant Defines the mutant described by Adato et al. (2009).

FIGURES CONTENT

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

TABLES CONTENT

CHAPTER II

CHAPTER IV

ADDITIONAL FIGURES CONTENT

CHAPTER I

CHAPTER II

CHAPTER IV

ADDITIONAL TABLES CONTENT

CHAPTER I

CHAPTER II

CHAPTER IV

Table IV.S1 UPLC-QTOF-MS-detected metabolites differentially accumulated in the skin of breaker fruits for both mutant *pf*^{1} and *pf*^{2}.

Table IV.S2 Differentially expressed genes in breaker fruit skin from the *pf* mutants and shared with the *y* mutant.

Table IV.S3 Common genes differentially expressed in the *pf* mutants.

Table IV.S4 Transcription factor-binding elements by an MEME-TOMTOM analysis.

Table IV.S5 Down-regulated genes in the *pf* mutants that contained transcription factor DNA-binding domains.

ABSTRACT

Until recently, the genetic improvement of tomato (*Solanum lycopersicum*) was focused in agronomic traits, such as yield and biotic or abiotic stresses; therefore the interest in tomato fruit quality is relatively new. The tomato fruit surface can be considered both an agronomic trait as well as a quality trait, because it has an effect on consumer impression in terms of color and glossiness but also it underlies the resistance/sensitivity to cracking or water loss with consequences on fruit manipulation (e.g. transport and processing). The cuticle is deposited over the cell wall surrounding the epidermal cells and it is the first barrier in the plant-environment interface. The cuticle composition includes two main groups of metabolites: cuticular waxes and cutin. Other metabolites can be founded into the cuticle matrix, as triterpenoids and flavonoids. Those minor cuticular components are involved in the correct functionality of the cuticle. Understanding cuticle biosynthesis and genetic regulation requires the development of fast and simple analytical methodologies to study those specialized metabolites using large populations (e.g. mutant collections or introgression lines), together with the identification of genes and genomic regions responsible of their production. This thesis aims to contribute to our understanding of the molecular programs underlying tomato fruit quality by providing: (**i**) a general protocol to profile cuticular waxes in different species, including tomato; (**ii**) a QTL map for cuticular composition (i.e. cuticular waxes and cutin monomers) using the *Solanum pennellii* introgression line population; (**iii**) a detailed protocol of the reverse genetic tool so-called *Fruit-VIGS* to assist in the study of gene function in tomato fruit; and (**iv**) a thorough characterization of the first null allele for the transcription factor SlMYB12 (i.e. *Slmyb12-pf*) in tomato fruit which provides new insights into the regulation of the flavonoid biosynthetic pathway in the fruit peel by high resolution mass spectrometry and RNA-Seq approaches.

RESUMEN

Hasta hace poco, la mejora genética del cultivo del tomate (*Solanum lycopersicum*) había estado centrada principalmente en caracteres agronómicos, como la productividad y la resistencia a estreses, tanto bióticos como abióticos. Así, el interés en la calidad del fruto de tomate es relativamente reciente. La superficie del fruto del tomate puede considerarse tanto un carácter agronómico como de calidad, pues influye en la primera impresión de los consumidores en términos de color y brillo, así como también en los procesos de resistencia o sensibilidad a la rotura ('*cracking*') o a la pérdida de agua. Estos factores determinan el aspecto del fruto y condicionan atributos relacionados con su manipulación (transporte y procesado). La cutícula se deposita sobre la pared celular de las células epidérmicas y es la primera barrera que interacciona con el ambiente. Está constituida por dos grandes tipos de metabolitos: las ceras cuticulares y la cutina. Otros metabolitos pueden aparecer embebidos en la matriz cuticular, como es el caso de los triterpenoides y los flavonoides. Estos metabolitos contribuyen a la correcta funcionalidad de la cutícula. La compresión de la biosíntesis y regulación génica de la cutícula requiere del desarrollo de metodologías de análisis sencillas y rápidas para el estudio de estos metabolitos especializados en grandes poblaciones (colecciones de mutantes o líneas de introgresión), así como para la identificación de genes y regiones génicas responsables de la producción y acumulación de dichos compuestos, pudiendo ser muy útiles para implementar programas de mejora de la calidad del tomate. El objetivo de esta tesis es contribuir a la comprensión sobre los programas moleculares subyacentes a la calidad del fruto de tomate, proporcionando: (**i**) un protocolo general de análisis del contenido de ceras cuticulares en diferentes especies, incluyendo el tomate; (**ii**) un mapa de QTL de la composición cuticular (incluyendo ceras y monómeros de cutina) obtenido con la población de líneas de introgresión de *Solanum pennellii*; (**iii**) un protocolo detallado de uso de la herramienta de genética reversa *Fruit-*

VIGS con el que realizar estudios de funciones génicas en fruto de tomate; y (**iv**) una minuciosa caracterización de un nuevo alelo nulo del factor de transcripción SlMYB12 (*Slmyb12-pf*) en fruto de tomate, proporcionando nueva información sobre la regulación de la ruta biosintética de los flavonoides en la piel del fruto, utilizando espectrometría de masas de alta resolución y de nuevas tecnologías de secuenciación.

RESUM

Fins fa poc de temps, la millora genètica de la tomata (*Solanum lycopersicum*) anava dirigida fonamentalment als caràcters de tipus agronòmic, com la productivitat i la tolerància a estressos biòtics o abiòtics, resultant que l'interés per la qualitat dels fruits és relativament nou. La superfície de la tomata pot ser considerada tant com un caràcter agronòmic com un de qualitat, ja que és l'aspecte de la superfície del fruit el que confereix al consumidor la primera impressió de color, brillantor, però és també la pell del fruit la responsable de la diferent susceptibilitat del fruit a desenvolupar clevills o que el fruit sofrisca més o menys pèrdues d'aigua, tot tenint importants conseqüències en la manipulació (i.e. transport i processament del fruit). La cutícula és dipositada per sobre de la paret cel·lular que envolta la capa de cèl·lules epidèrmiques i constitueix la primera barrera en la interfase planta-medi ambient. La composició de la cutícula presenta dos grups principals de metabòlits: les ceres i la cutina. També es poden trobar altres metabòlits els triterpenoids i el flavonoids. Aquests darrers components cuticulars menors són implicats en el correcte funcionament de la cutícula. Per tal de comprendre la biosíntesi i la regulació genètica de la cutícula cal desenvolupar tecnologies analítiques senzilles i rapides que permeten estudiar aquests metabòlits especialitzats en poblacions grans de plantes (i.e. Col·leccions de mutants o de línies d'introgressió), a més de la identificació de gens i regions genòmiques que són responsables de la seua producció. Aquesta tesi té com a objectiu contribuir a millorar la nostra comprensió dels programes moleculars que afecten determinats aspectes de la qualitat de la tomata mitjançant els següents objectius: (**i**) proporcionar un protocol general per obtenir perfils de ceres cuticulars en diferents espècies, inclosa la tomata; (**ii**) obtenir un mapa de QTL per a la composició cuticular (i.e. ceres cuticulars i monòmers de cutina) mitjançant la utilització de la població de línies d'introgressió de Solanum pennelli; (**iii**) descriure amb detall el protocol d'una eina de revers genètica denominada

Fruit-VIGS que resulta molt adequada per estudiar funció gènica a la tomata; y (**iv**) fer una caracterització exhaustiva del primer al·lel nul del factor de transcripció SlMYB12 (ie. *Slmyb12-pf*) en tomata la qual proporciona informació nova sobre la regulació de la ruta de biosíntesi de flavonoides en la pell de la tomata mitjançant espectrometria de masses d'alta resolució i RNAseq.

GENERAL INTRODUCTION

GENERAL INTRODUCTION

The tomato (*Solanum lycopersicum*) is one of the most consumed vegetables worldwide (USDA, 2013. http://www.ers.usda.gov/) and, therefore, an economically important crop. Tomato breeding has been focused on agronomic traits (Grandillo et al., 1999; Llácer et al., 2006) and most recently on quality traits, especially those related to organoleptic or nutritional aspects (Casañas and Costell, 2006; Camara, M., 2006; Martin et al., 2011). Tomato quality is a complex trait which involves a set of attributes related with chemical composition as e.g. flavor, aroma, color or texture (Fernie et al., 2006; Luo et al., 2008; Ballester et al., 2010; Adato et al., 2009; Saladié et al., 2007; Rambla et al., 2014). These attributes are influenced not only by varietal differences or by the plant environmental and nutritional regime, but also by the ripening stage and post-harvest storage conditions (Causse et al., 2002; Giovannoni, J.J., 2004; Kosma et al., 2010; Lara et al., 2014). Moreover, most of the quality traits show a continuous variation resulting from the joint action of many genes on a number of specific genome regions which could be identified /mapped by genetic markers (Quantitative-Trait-Loci, QTL) and which are strongly influenced by the environment (Eshed and Zamir, 1995; Koenig et al., 2013; Bolger et al., 2014). Thus, the study of tomato quality traits requires both genetic diversity and analytical techniques to reveal the different components associated to attributes of interest and to determine which are their molecular basis and genetic determinants in order to use them in plant breeding programs.

Tomato genetic diversity is small for cultivated varieties (less than 5% of the total). In contrast, wild tomato relatives have a large genetic variability that has been used mainly for improving tomato biotic and abiotic stress. The allelic variation found in wild tomato species has been successfully exploited by developing interspecific populations containing chromosomal introgressions from wild relatives into the cultivated genetic background (Paterson et al., 1988; Goldman et al., 1995; Eshed and Zamir,

1995; Tanksley et al., 1996; Bernacchier al., 1998; Chen et al., 1999; Alseekh et al., 2013). One of the most extensively used population is the *Solanum pennellii* introgression line population (Eshed and Zamir, 1995) which has been used to analyze valuables traits of interest, to identify associated QTL and to introduce such regions into elite tomato varieties by introgression breeding (Zamir, 2001). In the last decade a large effort has been put in complex quantitative traits such as those affecting fruit chemical composition, which are highly affected by a large number of QTL (Tadmor et al., 2002; Liu et al., 2003; Rousseaux et al., 2005; Bermudez et al., 2008; Chapman et al., 2012; Alseekh et al., 2013). Such studies have benefited from the development of analytical techniques to study chemical and metabolite compositions (Moço et al., 2006; Overy et al., 2005) and resulted in a better understanding of the metabolic pathways contributing to organoleptic and nutritional traits (Martin et al., 2011), how they are affected by the environment (Zoratti et al., 2014) and their role in plant-pathogen interactions (Schenke et al., 2014).

The study of the metabolism in tomato fruit has been important to understand maturation, ripening and postharvest behaviors (Giovannoni, J.J., 2001 and 2004; Lara et al., 2014). Tomato fruit ripening involves a large number of interrelated processes which include changes in primary metabolism (Causse et al., 2002; Overy et al., 2004) and in specialized metabolism such as volatiles, carotenoid and polyphenol compounds (Fulton et al., 2002; Liu et al., 2003; Rosseaux et al., 2005). Ripening also produces changes in texture, cell wall ultrastructure and in the fruit surface, including cuticular composition and cuticle biomechanics (Saladié et al., 2007; Chapman et al., 2012; Domínguez et al., 2012). Although the tomato fruit surface has been important for breeders during the last forty years (Kolattukudy, P.E., 1970), the interest on it has increased in the last few years (Yeats et al., 2012a; Petit et al., 2014).

The tomato fruit surface is composed by the epidermal cell layer, which is covered by the cuticle (Heredia, A., 2003; Riereder and Müller, 2006; Yeats and Rose, 2013). The cuticle is a layered and mostly lipophilic structure that represents the first barrier between the plant and the environment (Buschhaus and Jetter, 2011) and constitutes the interface between the plant and pathogen interaction (Martin and Rose, 2014; Schenke et al., 2014). Further, the cuticle is an essential cytoskeleton for organ development and integrity (Hen-Aviv et al., 2014). Actually, the cuticle morphology and nomenclature is still in dispute, however there is an agreement in the scientific community to define the cuticle as a layered structure (Heredia and Domínguez, 2009; Riereder and Müller, 2006). Thus, the structure of the cuticle consists of two layers, the *cuticle proper* which is the outermost layer, and the *cuticular layer* which is placed below (Heredia and Domínguez, 2009; Riereder and Müller, 2006). The cuticle proper is composed mostly by the cutin matrix, a tridimensional structure made of hydroxylated and epoxy-hydroxylated fatty acids polymerized covalently (Pollard et al., 2008; Beisson et al., 2012). Embedded into and deposited onto the matrix there are free cuticular waxes, a mixture of dozens of compounds, mostly very-long-chain fatty acids (i.e. fatty acids, alkanes, aldehydes, alcohols and esters) but also flavonoids, triterpenoids, sterols or glycoalkaloids between others (Samuels and Jetter, 2008; Buschhaus and Jetter, 2011; España et al., 2014). On the other hand, the cuticular layer, which is a cutinizated secondary cell wall, contains cutin polymers linked to polysaccharides (i.e. pectin, hemicellulose and cellulose) and proteins (Heredia and Domínguez, 2009; Lopez-Casado et al., 2007; Riederer and Müller, 2006). The cuticle proper thus could be isolated by degrading the polysaccharides of the cuticular layer using enzymatic cocktails (Heredia and Domínguez, 2009; Hovav et al., 2007).

In tomato, the cuticle of the fruit is thick, continue (without stomatas) and easily extractable enzymatically (Hovav et al., 2007). These

characteristics, together with the well-developed *omics* techniques have proposed the tomato as a model system for fruit cuticle studies (Yeats and Rose, 2013; Hen-Aviv et al., 2014; Martin and Rose, 2014). The chemical composition and biophysical properties of tomato fruit cuticle have been extensively studied, mainly for those aspects affecting fruits quality attributes such as water loss and permeability, susceptibility to infections and physiological disorders (e.g. *cracking*) (Hovav et al., 2007; Isaacson et al., 2009; Domínguez and Heredia, 2011; Domínguez, et al., 2012: Lashbrooke et al., 2015). In these studies, mutants showing affected fruit cuticle (Isaacson et al., 2009; Petit et al., 2014) or wild species (Yeats et al., 2013) have been analyzed.

Large interspecific populations have only recently started to be used in cuticle composition studies. Thus, a study in tomato leaves (Bolger et al., 2014) and another on tomato fruit addressing the triterpenoid composition in a *Solanum habrochaites* introgression line population (Yeats et al., 2013) have been only reported. The methodology to study tomato cuticle composition in mid/high-throughput experiments is not yet well established. Rapid methodologies for cuticle metabolite screening, the development of specific chemical databases or analytical software for mass spectrometry outputs are desirable goals to facilitate the study of such complex trait that is the tomato fruit cuticle composition. Furthermore, QTL analyses for cuticle composition in tomato fruit would allow us to investigate the molecular basis underlying cuticle biosynthesis, development and regulation processes, providing new opportunities to improve tomato fruit quality attributes.

Objectives

This thesis aims to increase our understanding of the tomato fruit by providing information on metabolite composition and gene function, characterization of genetic resources for metabolites and transcripts, and identifying genes and genomics regions associated to metabolite changes in the context of tomato fruit surface background. To reach these goals, the following specific objectives are proposed:

First, to develop a general method for the metabolomics analysis of cuticular wax composition in fruit, but also extending it to other plant organs.

Second, to define genomic regions associated to variations in cuticle composition by QTL analysis on the *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression line population and propose candidate genes.

Third, to optimize Fruit-VIGS tool to study gene functions in tomato fruit.

And ***fourth***, to conduct the metabolic and the transcriptomic characterizations of two new *pink fruit* mutants.

The results obtained to achieve these objectives are described and discussed in the next four chapters after a short introduction to the field for fruit quality. At the end of the manuscript we present the final remarks derived from the results, a list of goals, a statement on the future perspectives and the main conclusions arising from this work.

CHAPTER I

AN EFFICIENT METHOD FOR THE CUTICULAR WAX COMPOSITION ANALYSIS OF DIFFERENT PLANT SPECIES AND ORGANS

Josefina-Patricia Fernández-Moreno[1], Sergey Malitsky[2], Justin Lashbrooke[2], Radu Racovit[3], Reinhard Jetter[3], Diego Orzaez[1], Asaph Aharoni[2] and Antonio Granell[1]

(Adapted from the manuscript submitted to *Metabolomics*)

[1] Fruit Genomics and Biotechnology Laboratory. Instituto de Biología Molecular y Celular de Plantas (CSIC-UPV). Valencia, Spain.

[2] Department of Plant Sciences and the Environment. Weizmann Institute of Science. Rehovot, Israel.

[3] Department of Botany. Faculty of Science. The University of British Columbia. Vancouver, Canada.

SUMMARY

Most aerial plant organs are covered by a cuticle, which largely consists of cutin and wax. Cuticular waxes are mixtures of dozens of compounds, mostly very-long-chain aliphatics that are easily extracted by solvents. Over the last four decades, diverse cuticular wax analysis protocols have been developed, most of which are complex and time-consuming, and need to be adapted for each plant species or organ. Plant genomics and breeding programs often require mid- or high-throughput metabolic phenotyping approaches to screen large numbers of individuals and obtain relevant biological information. Therefore, a fast, simple and user-friendly methodology able to capture most wax complexity independently of the plant, cultivar and organ must be

developed. Here we present a simple method for screening relatively small wax amounts, sampled by short extraction with a versatile, uniform solvent. The method is validated herein for three different crop species: tomato (*Solanum lycopersicum*), apple (*Malus domestica*) and hybrid aspen (*Populus tremula* x *tremuloides*). Consistent results were obtained in tomato cv. M82 across three consecutive years (2010-2012), two organs (leaf and fruit), and also in two different tomato (M82 and MicroTom) and apple (Golden Delicious and Granny Smith) cultivars. Our results on tomato wax composition match those reported previously, while our apple and hybrid aspen analyses provide the first comprehensive cuticular wax profile of these species. Consequently, the protocols developed here allow standardized identification and quantification of most cuticular wax components in a range of species.

Key words: metabolic profiling, cuticular waxes, fruit surface, fleshy fruit.

I.1 INTRODUCTION

The cuticle is a flexible self-healing barrier which is covering almost all aerial surfaces of plants. It plays important roles in preventing water loss and desiccation, and in interactions with pathogen spores and herbivores (Belding et al., 1998; Leide et al., 2007; Riederer and Müller, 2006; Vogg et al., 2004). It is thus playing a key role in plant adaptation to abiotic and biotic stresses (Buschhaus and Jetter, 2011; Isaacson et al., 2009). Furthermore, the cuticle is crucial for correct growth and differentiation of developing surface tissues by establishing boundaries between nascent organs, and also appears to be associated with the process of fleshy fruit development and ripening (Hovav et al., 2007; Hen-Avivi et al., 2014; Kimbara et al., 2012).

The two major components that make up the cuticle are cutin and waxes (Riederer and Müller, 2006; Hen-Avivi et al., 2014; Yeats and Rose, 2013). Cutin is a three-dimensional polymeric structure which, by definition, cannot be extracted by solvents. It is a polyester formed mainly by hydroxylated and epoxy-hydroxylated C_{16} and C_{18} fatty acids and glycerol (Riederer and Müller, 2006; Yeats and Rose, 2013). Wax, in contrast, is a complex mixture made up of dozens of compounds derived from very-long-chain fatty acids (VLCFAs), which range from C_{20} to C_{40}, and are easily removed by lipophilic solvent extraction (Buschhaus and Jetter, 2011; Riederer and Müller, 2006; Samuels et al., 2008). Presence of different functional groups in these VLCFA derivatives generates different wax classes, including alkanes, alcohols, aldehydes, ketones and alkyl esters (Buschhaus and Jetter, 2011; Riederer and Müller, 2006; Yeats and Rose, 2013). Finally, secondary metabolites with ring structures (e.g., pentacyclic triterpenoids or flavonoids) may also form part of the cuticular wax mixture in some species (Bauer et al., 2004a; Buschhaus and Jetter, 2011). While cuticles present only a little variation in qualitative chemical composition between diverse species, their wax mixtures may differ drastically in quantitative terms. The different proportions of those basic components and

their absolute amounts are thought to cause the specific cuticle properties of each species, organ or developmental stage (Buschhaus and Jetter, 2011).

Cuticular properties are associated with major horticultural traits, and deficient cuticle performance often leads to serious economic losses (Albert et al., 2013; Belding et al., 2000; Cameron et al., 2002; Hovav et al., 2007; Lara et al., 2014). Understanding cuticle characteristics, chemical composition, biosynthesis, regulation, development and interaction with the environment will therefore provide valuable information about plant development and adaptation, and can also help to reveal important targets for breeding. Thus, the analysis of cuticle composition is important for example in the fleshy fruit crops marketed as fresh commodities (e.g., apples, tomatoes, grapes or berries). In such fruit, water loss, surface cracking and pathogen infection have a negative impact on the quality of the final product (e.g., appearance, texture or post-harvest shelf life; Belding et al., 1998; Hovav et al., 2007; Isaacson et al., 2009; Lara et al., 2014). Thus, studying the cuticle in commercial fruit is particularly interesting since the processes that affect it also have an impact on the storability and marketability of commodities (Domínguez et al., 2011; Lara et al., 2014). Tomato (*Solanum lycopersicum*) and apple (*Malus domestica*) are two such important fresh commodity crops (Belding et al., 1998; Hovav et al., 2007; Isaacson et al., 2009; Veraverbeke et al., 2001).

The tomato fruit is an excellent model for cuticle studies, as it has an astomatous, thick and continuous cuticle that is easy to isolate. Moreover, the availability of the tomato genome sequence provides an excellent opportunity to study the genetic basis that controls different cuticle processes. In apple fruits, long storage periods in controlled-atmosphere chambers and the high accumulation of bioactive compounds such as triterpenoid acids represent quality components associated with the cuticle, which are of much interest in breeding programs (Caligiani et al., 2013; Szakiel et al., 2012; Veraverbeke et al., 2001). Furthermore, the russetting (i.e., micro-crackings in the cuticle)

is an important process affecting post-harvest shelf life and producing a negative impact for consumers (Lashbrooke et al., 2015). However, the interest in studying the cuticle in crop plants is not restricted to fruits. Studying cuticular wax composition is also relevant in photosynthetic organs for woody crops such as *Populus* and *Salix* species, in which the cuticle has essential physiological functions and affects final dry matter production (Cameron et al., 2002; Jones 1992).

The methodologies used to extract and analyze cuticular wax composition in crop plants vary vastly (Buschhaus and Jetter 2011; Riederer and Müller 2006). So far, the use of wax analyses in mid-throughput approaches such as for breeding population screening has been limited, as they involve costly, time-consuming protocols and a complex experimental set-up. For example, different protocols have been used to screen cuticular waxes in different organs of a single plant species (Belding et al., 1998; Dobson et al., 2012; Hovav et al., 2007). Similarly, no consensus methodology exists to study the cuticular wax composition of the same organ across different species (Adato et al., 2009; Bauer et al., 2004a and 2004b; Isaacson et al., 2009; Leide et al., 2007 and 2011; Yeats et al., 2012). Crop breeding programs could benefit from the development of a mid/high-throughput method, such as those required for metabolic profiling of a medium or large numbers of individuals (Eshed and Zamir, 1995; Liu et al., 2005; Watanabe et al., 2007). Therefore, optimizing and testing a simple methodology to be used across species that would be able to reliably capture wax complexity would be valuable.

The aim of this study was to develop a rapid method for profiling cuticular waxes that could be used for different crop surfaces independently of the species, cultivar or organ studied. We, herein, demonstrate the applicability of this procedure to determine the cuticular wax composition of diverse plant material, including three different crop species (tomato – *Solanum lycopersicum*–, apple – *Malus domestica*– and hybrid aspen –

Populus tremula x tremuloides–), four different cultivars (the tomato cvs. M82 and MicroTom, and the apple cvs. Golden Delicious and Granny Smith), and two different aerial organs (fleshy fruits in tomato and apple, and leaves in tomato and hybrid aspen).

I.2 MATERIALS AND METHODS

Plant Material

Cuticular wax analyses were carried out using the following organs and species, which represent a range of cuticle thicknesses and compositions: tomato fruit (*Solanum lycopersicum* cultivars M82 and MicroTom) and tomato leaf (*S. lycopersicum* cv. M82); apple fruit (*Malus domestica* cultivars Golden Delicious and Granny Smith); and hybrid aspen leaf (*Populus tremula* x *tremuloides*). Three biological replicates were analyzed for tomato and apple samples, and 10 for hybrid aspen. Each one represented a mix of tissue samples that originated from 5-10 different fruits or leaves. Tomato fruits were harvested in the mature red stage, and tomato leaves from 1-month-old plants grown in a greenhouse in Rehovot, Israel. Apple fruits were picked in the mature stage from an orchard in Trentino, Italy. Hybrid aspen leaves were harvested from 4-month-old plants and were grown in a greenhouse in Sweden.

General Considerations

Unless otherwise specified, all chemicals were purchased from Sigma Aldrich®: *n*-alkanes (C_8-C_{41}), fatty acids (C_{29}-C_{31}), α- and β-amyrins, BSTFA (bis-trimethylsilyltrifluoroacetamide), pyridine, pectinase and cellulase enzymes (*Aspergillus niger*), sodium azide, sodium tetraborate decahydrate, except for sodium acetate (Merck®), glacial acetic acid (Furtarom®) and chloroform (ACS, ISO, Reag. Ph Eur chloroform, EMSURE®). Alkanes and fatty acids were used as internal standards or as a standard for identification at individual concentrations of 0.2 mg/ml. The analysis of fruit cuticular waxes was performed using enzymatically isolated cuticular membranes (Riederer & Müller, 2006). Leaf cuticular waxes were extracted from both the abaxial and adaxial sides of the same leaf together.

Cuticular Wax Extraction

Fruit cuticle extraction

The tomato and apple fruits were washed with tap water and dried gently before further handling. Tissue discs containing both exocarp and mesocarp were dissected with the help of a cork borer (1 cm diameter for apple and tomato cv. M82, and 0.78 cm diameter for tomato cv. MT) and a sharp knife. The cuticular membrane was isolated from fruit discs by digestion with an enzymatic cocktail in buffered solution (EB) during a 3-day process (Hovav et al., 2007) in tomato, and a 4-day process in apple. The EB solution contained 1% pectinase and 0.5% cellulase in 0.1 M sodium acetate buffer, pH 3.8, which included 1 mM sodium azide to prevent microbial growth. Incubation of fruit skin in EB was performed at 37ºC for 24 h under gentle shaking. Subsequently, the enzymatic solution was replaced and the reaction was allowed to proceed for another 24-hour period under the same conditions in tomato, and for an extra 24-hour period in apple to ensure the hydrolysis of the apple mesocarp. The remaining mesocarp from the tissue discs was cleaned gently after each digestion. The isolated cuticular membrane discs were folded in Petri dishes to prevent them from rolling in before being treated with sodium tetraborate decahydrate solution (10 mM TTBS, pH 9-9.2) for 1 minute, followed by two washing steps in distilled water, 1 minute each. Then, isolated cuticular membranes were laid flat on a glass vial and dried at 60ºC for 48 h. Finally, they were weighed and stored in a non-oxidative atmosphere by flushing with N_2 gas.

Tissue sampling in leaves

Fully expanded tomato and hybrid aspen leaves were harvested, washed with tap water and dried gently. The leaf discs were cut with a cork borer (1 cm diameter) from fully expanded leaves (avoiding the mid-veins) and directly immersed in chloroform.

Cuticular wax extraction

Cuticular wax extraction was performed over a total surface area of 18.84 cm^2 for tomato fruits and leaves, and over a total surface area of 12.56 cm^2 for apple fruits and hybrid aspen leaves. The tissue discs were dipped into chloroform at room temperature and shaken gently for 30 sec. in the case of aspen samples and for 30 sec. twice in the case of tomato and apple samples. Immediately after extraction, *n*-tetracosane (0.2 mg/ml) was added as the internal standard (ISTD). A total of 5 μg of ISTD was added to tomato fruit and leaf extracts, and of 2.5 μg to apple fruit and hybrid aspen extracts. The final volume was reduced to 1 ml using a stream of N_2, and the solution stored at -20ºC until further analyses.

Derivatization and GC-MS Analysis

Before Gas Chromatography – Mass Spectrometry (GC-MS) analysis, the hydroxyl groups in the cuticular wax compounds were transformed into the corresponding trimethylsilyl derivatives. The volume of wax extracts was first reduced to 100 μl under a stream of N_2. Then, 40 μl of BSTFA and pyridine (1:1 v/v) were added and the solutions mixed by vortexing for 1 minute. Derivatization was allowed to proceed at 70ºC for 40 minutes. After having cooled to room temperature, the whole reaction mixture was transferred to a vial and an aliquot injected into the GC-MS.

The qualitative and quantitative composition of the cuticular waxes was analyzed with a GC (Trace GC, Thermo Scientific), connected to an electron impact MS detector (DSQ2; Thermo Scientific) set at 70 eV and m/z 40-850 Da (± 0.5 Da). MS tuning was performed by using FC-43 perfluorotributylamine (PFTBA, Thermo Scientific) in positive mode after each sequence run. GC column quality was monitored using a standard acid mix (C_{29}, C_{30} and C_{31}) with a final concentration of 0.2 mg/ml for each standard acid, which was run using the standard GC-MS program. The ratio

between the areas of the first standard acid (C_{29}) and the third standard acid (C_{31}) was kept at less than 3, ideally between 1.4 and 1.7. Additionally, a standard mix (alkanes C_{20} to C_{40}) 1:1 (v:v) of *n*-alkane mix and chloroform was used to test the reproducibility of compound retention and detection by comparing peak times and areas of the 20 *n*-alkanes across three injections. Solutions containing the internal standard *n*-tetracosane (alkane C_{24}), the standard acid mix and the standard *n*-alkane mix were injected at the beginning of each sequence of samples and after every five samples.

Samples were injected using a solvent split mode (50 ml/min of split flow) in a PTV injector, and compounds were separated in a Zebron DB-1 column (Phenomenex ZB-1MS, 30 m length, 0.25 mm I.D. and 0.50 μm film thickness). GC was carried out in an oven programmed for 0 min at 65ºC, 20ºC min^{-1} to 220ºC, 3 min at 220ºC, 3ºC min^{-1} to 300ºC, 0 min at 300ºC, 5ºC min^{-1} to 330ºC and 35 min at 330ºC using a helium flow of 1.2 ml/min. However, these conditions had to be modified for hybrid aspen leaf wax analysis in order to capture its wax complexity: 0.5 min at 100ºC, 30ºC min^{-1} to 210ºC, 0 min at 210ºC, 5ºC min^{-1} to 330ºC and 35 min at 330ºC with the helium flow at 1.2 ml/min.

Different wax constituents were identified in the total ion chromatogram using their Kovat's indices and by comparing their mass spectra with those of authentic standards as well as data from the literature. Only compounds eluting after minute 15 were analyzed, to exclude any contaminating membrane lipids (Schönherr and Riederer, 1986; Yeats *et al.*, 2012). Online Resource 1 gives signature fragments for each compound class and example spectra for the different wax classes identified in this study. Detailed information about the specific fragments (i.e., $[M]^+$, $[M-15]^+$, $[M-18]^+$ or $[M-29]^+$) for each wax compound is available in Online Resources 2 to 6. Overall, a set of three to seven ions characteristic for each compound, including typical wax class fragments (e.g., 117, 129, 132, 145 m/z for fatty

acids) and specific wax compound fragments (e.g., $[M]^+$ and $[M-15]^+$ or $[M-18]^+$) was used.

The ICIS algorithm (Xcalibur software, Thermo Fisher Scientific) was used for the detection and peak integration using the following parameters: an area noise factor of 5, a peak noise factor of 10 and a minimum peak height (s/n) of 3. The quantification of each wax compound was performed in the same way for all the wax classes, and quantities calculated based on the ratio between each peak area and the area of the ISTD. In all the six wax mixtures analyzed, coverages were $\geq$0.001 μg/cm^2. Peak integration was confirmed manually. Finally, each ratio was normalized using the Excel software (Microsoft) by the total amount of ISTD (5 μg in tomato and apple mixtures, and 2.5 μg in hybrid aspen mixtures) and by the total area of extraction (18.84 cm^2 in tomato fruits and leaves, and 12.56 cm^2 in apple fruits and hybrid aspen leaves) to produce the final dataset (μg/cm^2).

I.3 RESULTS AND DISCUSSION

An efficient standard method for cuticular wax analysis in tomato fruit

It is often important to study the variability in cuticular wax composition in both breeding and mutagenized populations, to investigate the genetic and molecular basis of cuticle-related processes in different crops such as tomato. However, the study of wax composition in tomato fruit in relatively large populations has not been achieved yet, in part because this requires the use of time-consuming methodologies. In general, the analysis of wax composition is divided to four main steps: (i) sample harvesting; (ii) wax extraction; (iii) GC-MS separation and detection; and (iv) cuticular wax compound identification and quantification. Parts of the last two steps are normally performed automatically, while steps one and two are done manually and are often time-consuming, tedious and cumbersome, representing two important limiting factors in the study of medium/ large breeding populations. Then, to optimize initial processing steps and to generate a method able to capture most wax variability we modified the procedure described by Hovav et al. (2007). The modified method thus produced, proved to be efficient for cuticular wax analysis in tomato fruit for low- and mid-throughput experiments using large collection of plants. We also observed that the modified method worked to profile wax composition in other organs and crops. The modifications performed were the following: *first*, we used a smaller cuticular surface (~19 cm^2 instead of 25 cm^2) from only three biological replicates instead of 12 (Hovav et al., 2007), allowing us to reduce the enzymatic buffer volume required during cuticular membrane isolation in tomato fruit (12.5 ml instead of 20 ml). Then, we performed the wax extraction by using $CHCl_3$ at room temperature in glass vials for twice 30 sec instead of several minutes with hot $CHCl_3$ in a Soxhlet apparatus (Hovav et al., 2007). Finally, we adapted the GC injection method

to a split method in a PTV injector instead of on-column injections (Hovav et al., 2007) and the gas chromatographic conditions to 0 min at 65ºC, 20ºC min^{-1} to 220ºC, 3 min at 220ºC, 3ºC min^{-1} to 300ºC, 0 min at 300ºC, 5ºC min^{-1} to 330ºC and 35 min at 330ºC in order to resolve closely eluting peaks.

The adapted method was tested on red fruits of the tomato cv. M82, as this is one of the tomato varieties used as genetic background in breeding populations (Eshed and Zamir, 1995). The adapted GC conditions provided well separated individual peaks, facilitating wax compound identifications (**Figure I.1A**). We were able to identify seven classes of wax compounds based on their typical fragmentation patterns recorded with high accuracy (±0.05 Da): fatty acids, *n*-aldehydes, *n*-alkanes, branched alkanes, *n*-alcohols, flavonoids and pentacyclic triterpenols (**Figure S1**). Further, we quantified 42 cuticular wax components (**Table S1**), including those present in lower amounts, due to the high sensitivity achieved by our methodology (≥ 0.001 ± 0.0001 μg/cm^2). Thus, the wax composition obtained for tomato cv. M82 fruits included: odd-numbered *n*-alkanes C_{25}-C_{33}, even-numbered *n*-alkanes C_{28}-C_{34}, pentacyclic triterpenols (i.e., α-, β-and δ-amyrin) and flavonoids (i.e., naringenin) as major wax components quantified (white bars in **Figure. I.2A**); and *iso*-alkanes C_{29}-C_{32}, *anteiso*-alkane C_{32}, even-numbered fatty acids C_{20}-C_{34} and even-numbered *n*-alcohols C_{26}-C_{34}, except C_{30} (white bars in **Figure I.2A**) as minor wax components. Those waxes contributing to lesser extent to the total wax coverage were also quantified (**Table I.S1**): odd-numbered fatty acids C_{21}-C_{25}, even-numbered *n*-aldehydes C_{24} and C_{26}, odd-numbered *n*-alkane C_{23} and odd-numbered *n*-alcohols C_{23}-C_{29}.

Similar results were obtained for tomato cv. M82 red fruits in a second year (**Figure I.1B** and see black bars in **Figure I.2A**) and also in a third year (**Figure I.1C** and see hatched bars in **Figure I.2A**). Overall, the adapted method enabled the identification of 38 wax compounds in the tomato samples from three consecutive years, showing small fluctuations in the identification only in the case of trace compounds accumulating to

amounts close to 0.001 ± 0.0001 μg/cm^2. Therefore, this method was found to be efficient, robust and reproducible for cuticular wax screening in tomato fruit.

Next, we investigated the performance of the standard method in fruits from a different tomato cultivar. We selected the cultivar MicroTom (MT), as this is one of the most used genetic backgrounds in tomato mutant collections such as the ethylmethanesulfonate (EMS) mutant collection (Watanabe et al., 2007) or the γ irradiation-induced mutant lines (Matsukura et al., 2007). In this case, similar GC resolution was achieved as for cv. M82 (**Figure I.1C**). We were also able to identify six of the seven wax classes (*n*-aldehydes were not detected). Furthermore, we were able to identify in cv. MT 36 of the 42 wax compounds previously identified in cv. M82, with high sensitivity (≥ 0.001 ± 0.0001 μg/cm^2) (**Table I.S2**). We also observed that in cv. MT the odd-numbered *n*-alkanes (but not the even-numbered *n*-alkanes), the pentacyclic triterpenols and the flavonoids were the major cuticular wax constituents, with higher coverages than in cv. M82 (**Figure I.2A-B**). Thus, the standard method revealed qualitatively similar wax compositions in cvs. MT and M82 (with only six wax compounds missing: C_{24} and C_{26} *n*-aldehydes, C_{29} *iso*-alkane and C_{20}, C_{22} and C_{25} *n*-alcohols), and quantitative differences in specific compounds between the cultivars. Therefore, our standard method should be also efficient in the screening of cuticular wax composition in different tomato cultivars.

The tomato fruit wax composition observed with our standard method was consistent with previous reports (i.e. cv. M82 fruit cuticles Isaacson et al., 2009; Yeats et al., 2012). The extraction method used previously involved extraction for 24 h in a Soxhlet apparatus with a chloroform:methanol solution (Isaacson et al., 2009; Yeats et al., 2012). Herein, our standard method required only 1 min of extraction with chloroform at room temperature to obtain almost the same wax composition as those previous reports.

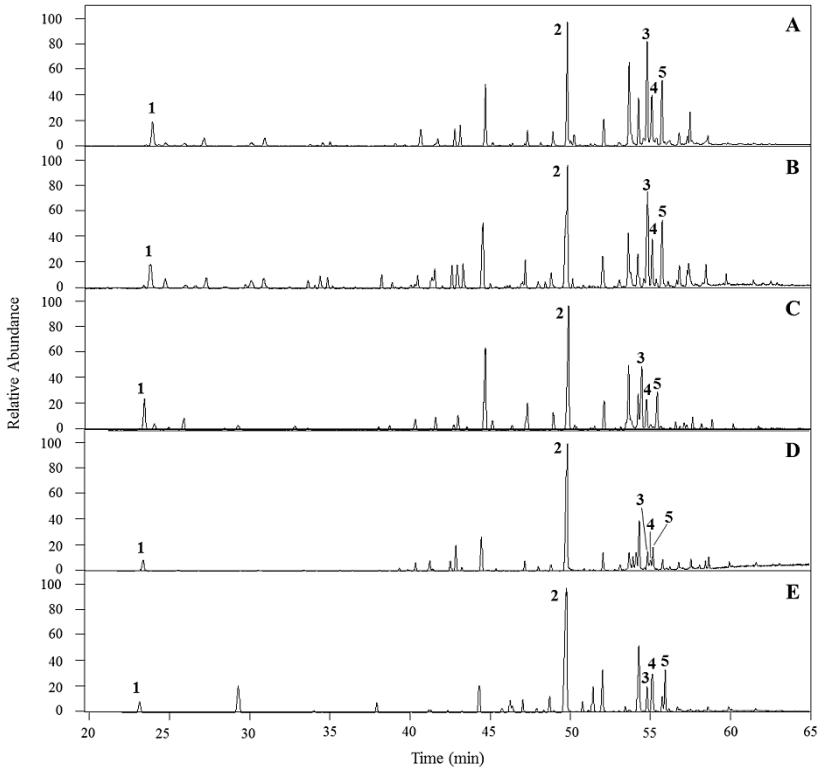


Figure I.1 Gas chromatograms of tomato (Solanum lycopersicum) waxes. *Tomato fruit cv. M82 during year 2010 (**A**), 2011 (**B**) and 2012 (**C**), tomato fruit cv. MT (**D**) and tomato leaf cv. M82 (**E**) total ion count (TIC) chromatograms are presented. Chromatogram parts from the C_{24} n-alkane (1) internal standard (~ min 24) to the last peak found (~ min 62) are shown. Plots are normalized to the largest peak (C_{31} n-alkane) in each chromatogram. The y-axis represents the relative abundance of the different compounds and the x-axis the time of the GC sequence in minutes. The GC conditions allowed good peak resolution, even between peaks with vastly differing relative abundances (e.g., 53-57 min). The major cuticular wax compounds in the five tomato genotypes fruit samples were (1) C_{31} n-alkane (2), δ-amyrin (3), β-amyrin (4) and α-amyrin (5).*

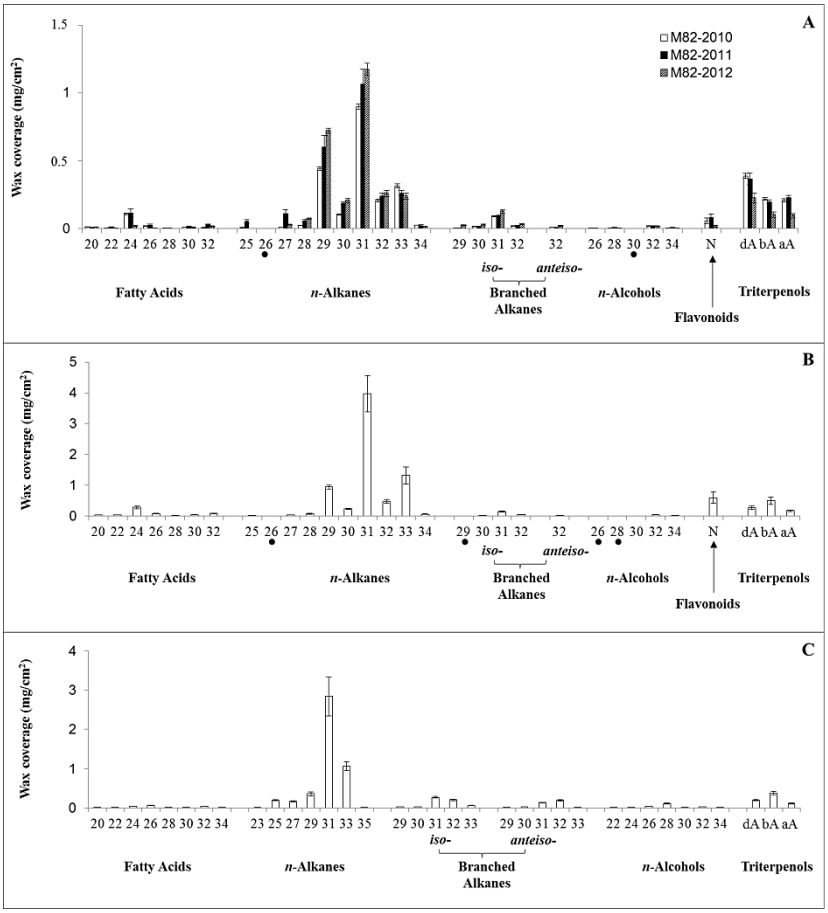


Figure I.2 Cuticular wax composition of tomato. *The cuticular wax composition in tomato fruit was similar among the three years in cv. M82 (**A**), between the two cultivars M82 and MT (**B**), and also between tomato fruit and tomato leaf in cv. M82 (**C**). In the three sets of data, n-alkanes and amyrin-type triterpenols were the major compounds detected and quantified by our standard method. In tomato fruits (**A**, **B**) flavonoids, fatty acids, branched alkanes and n-alcohols were also detected and quantified by our standard method. Those even-numbered alkane and alcohol homologs marked by a dot (●) were found in tomato leaves but not in tomato fruit cuticles. Error bar, standard error (n = 3 in **A**, **B** and **C**).*

The main differences between our results and those described by Isaacson et al. (2009) and Yeats et al. (2012) were in the identification of a number of pentacyclic triterpenols (i.e., multiflorenol, taraxerol, traraxasterol and ψ-taraxasterol), and two other minor wax classes (i.e., alkenols and alkenes) which were reported mainly by Isaacson et al. (2009) but only a couple of pentacyclic triterpenols by Yeats et al. (2012). Although we failed to identify such compounds in our tomato fruit wax mixtures, we were able to identify four *iso*-alkanes, an *anteiso*-alkane and three odd-chain fatty acids which were not reported previously (Isaacson et al., 2009; Yeats et al., 2012).

For tomato cv. MT, a similar extraction method (dipping in chloroform at room temperature for 1 or 2 min) had revealed 11 wax classes, five more than our standard method (i.e., sterols, alkenes, alkenols, alkadienols and dihydroxyalkanoics), and 77 wax compounds (Leide et al., 2007). However, the total surface extracted in this latter report was the full fruit, in contrast to the ~19 cm^2 from the equatorial region of the fruit used here. On the other hand, Adato et al. (2009) also performed cuticular wax analyses of tomato cv. MT exocarp discs from the equatorial region of the fruit, but using long extraction periods (dipping in chloroform at room temperature twice for 1 or 2 h). The wax composition reported by Adato et al. (2009) included seven wax classes (sterols were identified but not n-aldehydes) and only 16 wax compounds (Figure S7 in Adato et al., 2009), seven of them not identified in our study (stigmasterol, β-sitosterol, taraxerol, multiflorenol, ψ-taraxasterol and taraxasterol; Adato et al., 2009). The extracted area and period of extraction seem to affect the final results on wax composition in tomato. Thus, our standard protocol may be considered a good compromise able to provide a richer wax composition than that of Adato et al. (2009) but not as complete as that of Leide et al. (2007).

Despite these differences, our standard method is capable of capturing most cuticular wax complexity in tomato fruit cuticles from different cultivars and could prove useful to identify mutants or lines that

contain variable amounts of *n*-alkanes, triterpenols, branched alkanes, fatty acids, *n*-alcohols or total cuticular wax load, and it is thus of special interest as a screening tool for large collections of individuals in medium-throughput approaches. While our GC runs were quite long (75 min), this does not involve personnel time and could be a compromise as it ensures detection and quantification of the fatty acids and alcohols accumulating in lower amounts. This standard method has been used in our lab to process ~200 samples in four days from the cuticular membrane isolation to the wax extraction processes (unpublished data). In case of a special interest in any of the minor components of the selected individuals or lines can be studied using extended protocols.

The same method can be efficiently applied for wax composition analysis in tomato leaves

To test the applicability of our method on different plant organs, we studied the cuticular wax composition of tomato cv. M82 leaves. Thus, we obtained leaf discs (~19 cm^2) and followed the same procedure for wax extraction and GC-MS analysis as that used for tomato fruit. As for tomato fruits, we also achieved good peak resolution for the cuticular wax from tomato leaves (**Figure I.1E**). We found six of the seven wax classes identified in cv. M82 fruits but no flavonoids (**Figure I.2C**). Furthermore, we were able to identify 53 cuticular wax compounds in tomato cv. M82 leaves with the same sensitivity as in fruits (**Figure I.2C** and **Table I.S3**): odd-numbered (C_{23}-C_{35}) *n*-alkanes, pentacyclic triterpenols (i.e., amyrins) and branched alkanes (C_{29}-C_{33} for both *iso*- and *anteiso*-alkanes) dominated the wax composition, while even-numbered fatty acids (C_{20}-C_{34}) and *n*-alcohols (C_{22}-C_{34}) accumulated to lower amounts. Odd-numbered fatty acids (C_{21}-C_{33}), even-numbered *n*-alkanes (C_{26}-C_{34}) and odd-numbered *n*-alcohols (C_{25}-C_{33}) were also present but contributed to a lesser extent to the total wax coverage.

Only a few reports have described the cuticular wax composition of tomato leaves (Smith et al. 1996; Vogg et al. 2004) but only one of them was performed using cv. M82 (Wang et al., 2011). In those previous reports, the major cuticular wax classes (alkanes, branched alkanes and pentacyclic triterpenols) were described in leaves from the same cultivar (cv. M82, Wang et al., 2011) or different tomato cultivars (cv. MT, Vogg et al., 2004) or species (*Solanum hirsutum*, Smith et al., 1996). Moreover, *n*-alcohols were also reported only in leaves from cv. M82 (Wang et al., 2011). However, minor wax components such as fatty acids were not reported in the literature. Thus, our standard method was also suitable for analyzing cuticular wax composition in tomato leaves, as it identified all major wax components described in the literature. Similarly to the mid-throughput cuticular wax screening performed on tomato fruit, we also have used this standard method in our lab to screen cuticular waxes in tomato leaves derived from a large population. In fact, we were able to obtain leaf discs and extract the cuticular waxes of ~200 samples in a few hours (unpublished data).

The standard method is suitable to explore the wax composition in other crops

In order to test the versatility of the standard method, other crops such as apple fruits (*Malus domestica*) and hybrid aspen leaves (*Populus tremula* x *tremuloides*) were analyzed, two important species in the food and wood industries, respectively. *Firstly*, we used apple fruits from two commercial cultivars: cv. Golden Delicious (GD) and cv. Granny Smith (GS). Similar to tomato fruit, we obtained exocarp discs from the equatorial part of the apple fruit. However, we had to extend the cuticular membrane isolation process by one day to ensure the complete digestion of the mesocarpic tissue. On the other hand, roughly half the surface area (12.56 cm^2) was sufficient for sensitive detection and accurate quantification of all compounds of interest. We then performed the wax analysis on the exocarp discs in a similar way as for tomato fruit (identical GC-MS conditions). We

obtained good peak resolution during the GC sequence (**Figure I.3A-B**), and were able to identify seven wax classes: fatty acids, *n*-aldehydes, *n*-alkanes, *n*-alcohols, secondary alcohols, alkyl esters and pentacyclic triterpenoid acids (**Figure I.S1**). A total of 33 wax compounds were identified and quantified (**Table I.S5**): odd-numbered *n*-alkanes (C_{25}-C_{31}), pentacyclic triterpenoic acids (i.e., ursolic acid and oleanolic acid) and the secondary alcohol C_{29} were dominating the cuticular waxes (**Figure I.4A**), while even-numbered fatty acids (C_{20}-C_{30}), odd- and even-numbered *n*-alcohols (C_{25}-C_{29} and C_{22}-C_{26}, respectively) and *n*-aldehydes (C_{24}-C_{30}) appeared to be minor components (**Figure I.4A**). Odd-numbered fatty acids (C_{23}-C_{29}), even-numbered *n*-alkanes (C_{28} and C_{30}) and the alkyl ester C_{40} were also identified and quantified, but they contributed to a lesser extent to the total wax coverage (**Table I.S4**).

In the sparse literature on apple wax composition there is a single report on cuticular wax composition from an apple cultivar related to those used in our study, i.e., Golden Delicious (Belding et al., 1998). Belding et al. (1998) used whole fruit immersions in chloroform for three times 30 sec each, and found three *n*-alcohols (C_{24}, C_{26} and C_{28}), the major *n*-alkane (C_{29}) and the major triterpenoid acid (ursolic acid). In other apple cultivars, Veraverbeke et al. (2001) identified a few fatty acids, *n*-alkanes and esters, as well as a single *n*-aldehyde and *n*-alcohol, by dipping the entire fruits into hexane for 45 sec (Veraverbeke et al., 2001). Interestingly, our standard method proved to be highly effective when applied to apple fruit, allowing the identification of 33 cuticular wax compounds, 20 more than previously described in the literature for any apple cultivar (Belding et al., 1998; Veraverbeke et al., 2001). Thus, our standard method is able to capture a large portion of the wax complexity present in apple fruit surfaces, representing an improvement over other reports in the literature, suggesting that it could be used for medium-throughput screening of apple germplasm collections.

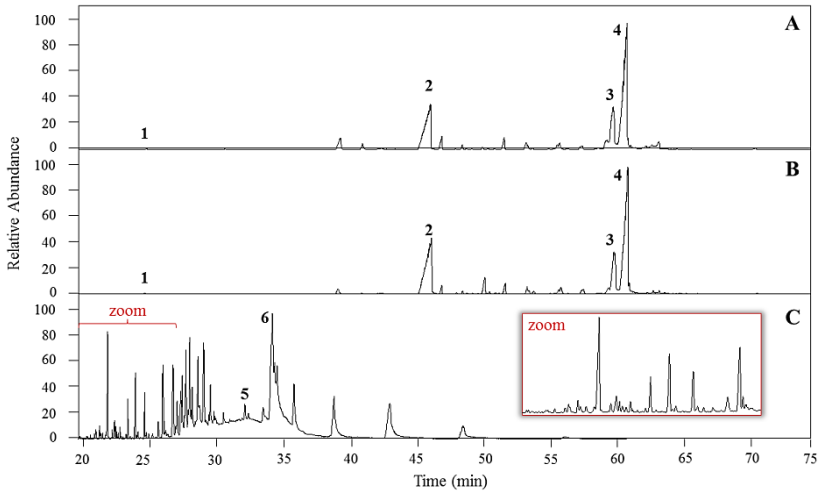


Figure I.3 Gas chromatograms of apple (Malus domestica) and hybrid aspen (Populus tremula x tremuloides) waxes. *Fruit of apple cultivars Golden Delicious (**A**) and Granny Smith (**B**), as well as hybrid aspen leaf (**C**) total ion count (TIC) chromatograms are represented. Chromatogram parts from the C_{24} n-alkane (1) internal standard (~ min 24) to the end of the run sequence (min 75) are shown for an apple sample, and from ~min 20 to the end of the run (min 60) for a hybrid aspen sample. In the hybrid aspen sample, the internal standard was C_{36} n-alkane (5) (~min 27). Plots are normalized to the largest peak (i.e., ursolic acid in apple (4); and a salicylate derivative in hybrid aspen (6)) in each chromatogram. The y-axis represents the relative abundance of the different compounds and the x-axis the time of the GC sequence in minutes. The GC conditions allowed good peak resolution for both apple and hybrid aspen waxes, even between peaks eluting closely (see zoom-in of hybrid aspen chromatogram, (**C**)).The major cuticular wax compounds in the two apple samples are also shown, C_{29} n-alkane (2), oleanolic acid (3) and ursolic acid (4).*

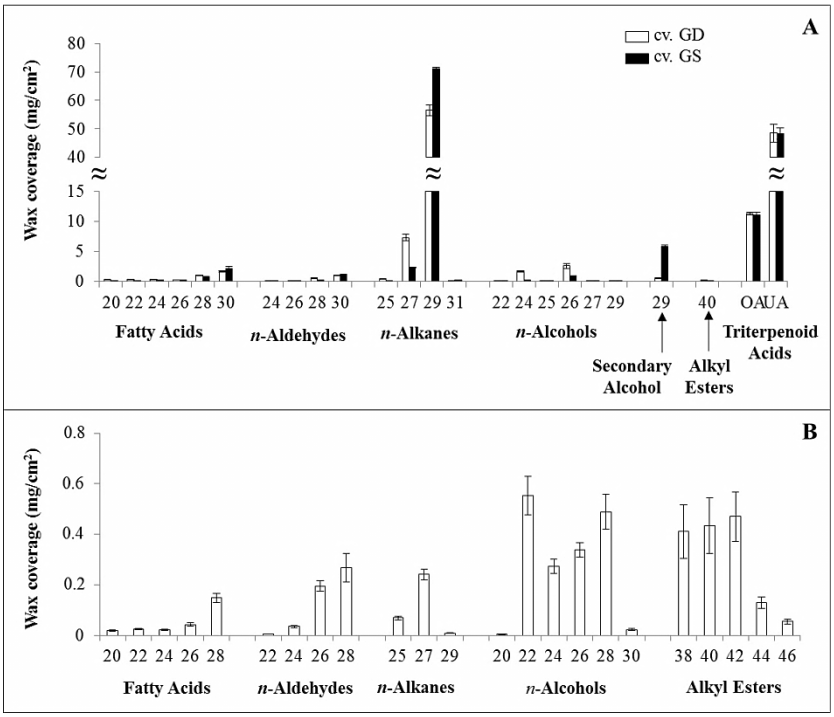


Figure I.4 Cuticular wax composition of apple fruit (Malus domestica) cultivars Golden Delicious (GD) and Granny Smith (GS), and of hybrid aspen leaf (Populus tremula x tremuloides). *In apple fruits (**A**) n-alkanes and triterpenoid acids were the major cuticular wax compounds quantified by our method, while fatty acids, n-aldehydes and n-alcohols were found to be minor compounds. The secondary alcohol n-nonacosan-10-ol was also present at high concentration in apple fruits. In hybrid aspen leaves (**B**) n-alcohols and alkyl esters were major cuticular wax constituents, together with n-aldehydes and n-alkanes. Fatty acids were also detected in hybrid aspen leaf wax by our standard method. Error bar, standard error (n = 3 in apple; n = 10 in hybrid aspen).*

A different crop plant explored by our standard method was hybrid aspen (*P. tremula* x *tremuloides*). We obtained leaf discs similarly to those from tomato leaves and dipped them into chloroform for extraction. For this species, we had to reduce the time of extraction to only 30 sec and to modify

the GC separation parameters to ensure detection of compounds with high molecular weight and low polarity eluting at the end of the run (i.e., alkyl esters). Thus, the GC oven was programmed for 0.5 min at 100ºC, 30ºC min^{-1} to 210ºC, 0 min at 210ºC, 5ºC min^{-1} to 330ºC and 35 min at 330ºC with a helium flow of 1.2 ml/min. The modified method enabled good peak resolution (**Figure I.3C**) even for very closely eluting peaks, and to identify five wax classes in hybrid aspen leaf wax: fatty acids, *n*-aldehydes, *n*-alkanes, *n*-alcohols and alkyl esters (**Table I.S5**). We were also able to detect a set of phenolic compounds, probably salicylate derivatives, based on their typical fragmentation pattern, but could not determine their specific structures. Thus, we were able to identify and quantify 28 wax compounds with similar high resolution as for the other crop samples tested (**Table I.S5**): odd-numbered *n*-alkanes (C_{25}-C_{29}), even-numbered fatty acids (C_{20}-C_{28}), *n*-aldehydes (C_{22}-C_{28}), *n*-alcohols (C_{20}-C_{30}) and alkyl-esters (C_{38}-C_{46}) (**Figure I.3B**). Odd-numbered fatty acids (C_{21}-C_{25}) and *n*-alcohols (C_{21}-C_{25}) were also present, but they contributed to a lesser extent to the total wax coverage (**Table I.S5**). The wax mixture was dominated by the unidentified phenolics (~ 32% of the total wax), followed by *n*-alcohols (~23%) and alkyl esters (~20%).

There is a single previous report on cuticular wax composition for two different *Populus* species hybrids (Cameron et al. 2002). Dipping leaf discs (12 mm^2) in chloroform for 30 sec revealed similar fatty acids, *n*-alcohols and *n*-alkanes composition but not *n*-aldehydes. However, the fatty acids and *n*-alcohols reported by Cameron et al. (2002) included not only the respective free fatty acids and *n*-alcohols, but also those derived from the hydrolysis of alkyl esters before the GC analysis (Cameron et al., 2002). Here, we were able to identify free fatty acids and *n*-alcohols separately from those forming the alkyl esters. Thus, our standard method is also suitable to capture the wax complexity in hybrid aspen leaves.

One method, several crop surfaces

In summary, we established the suitability of our brief method in tomato among batches form different years (**Figure I.2A**), cultivars (**Figure I.2B**) and organs (**Figure I.2C**). We were also able to explore using the same method the cuticular wax composition in apple fruits and hybrid aspen leaves (**Figure I.4**) and provide the first comprehensive cuticular wax profile for these two species. Moreover, the standard method was effective in capturing a broad range of wax variability, spanning 10 classes of wax compounds identified in the three crops of interest (**Figure I.5**), and total wax coverages ranging from ~4 μg/cm^2 in tomato fruits to ~150 μg/cm^2 in apple fruits (**Figure I.6**). Thus, a detailed comparative study of the cuticular wax composition for the six surfaces revealed differences in wax class distribution in different crops and organs. Fatty acids, *n*-alcohols and *n*-alkanes were present in all six surfaces, but in different proportions depending on the crop/organ (**Figure I.5**). The remaining compound classes identified were crop- or organ-specific (**Figure I.5**). For example, branched alkanes and pentacyclic triterpenols (i.e., amyrins) were present in tomato fruits and leaves, while flavonoids (i.e., naringenin) were present only in tomato fruits and pentacyclic triterpenoic acids (i.e., ursolic and oleanolic acids) only in apple fruits.

Our comparative study also revealed differences in the predominating wax constituents. Thus, *n*-alkanes were the major compound class in tomato fruits and leaves (i.e., alkane C_{31}) as well as in apple fruits (i.e., alkane C_{29}), followed by the pentacyclic triterpenoids (δ-amyrin in tomato fruits from cv. M82, β-amyrin in tomato leaves from cv. M82 and tomato fruits from cv. MT) and ursolic acid in fruits from both apple cultivars Golden Delicious and Granny Smith (**Figure I.2** and **Figure I.4A**). However, in hybrid aspen leaves the major wax compounds were phenolics (~32%), followed by *n*-alcohols (~23%) and alkyl esters (~20%). C_{22} alcohol and C_{42}

alkyl ester were the major compounds in hybrid aspen leaf wax (**Figure I.4B**).

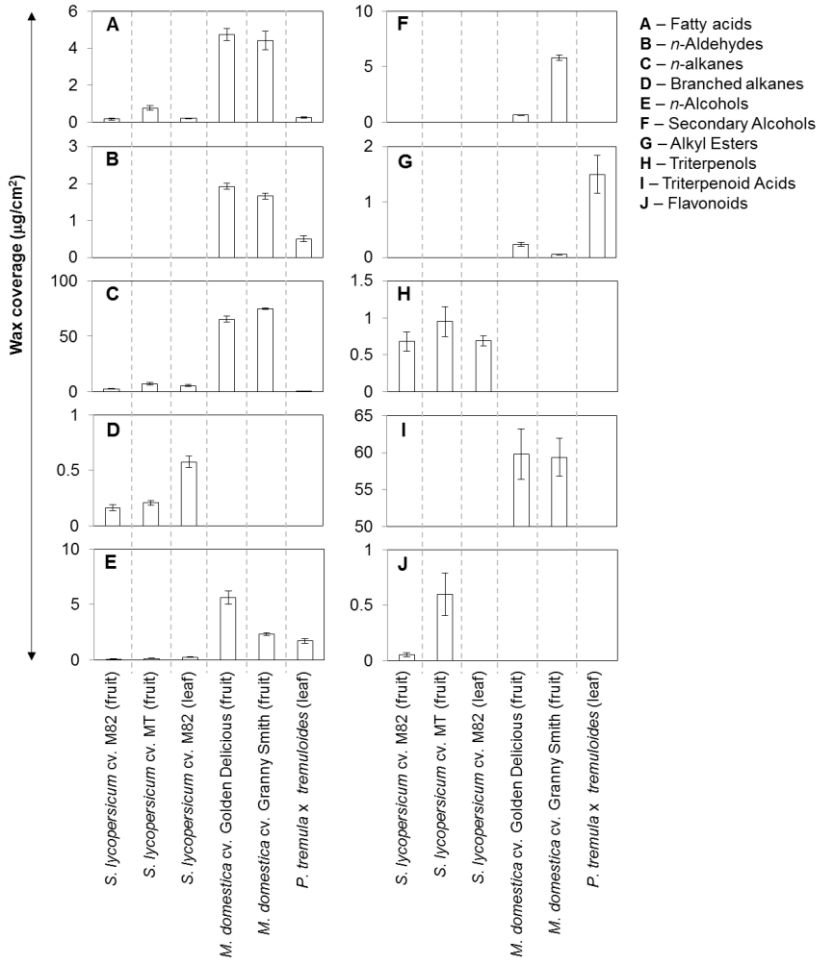


Figure I.5 Cuticular wax compound classes identified from the six crop surfaces. *The standard method was effective capturing the wax compositional complexity among the different samples studied. The fatty acids, n-alkanes and n-alcohols were found on all six surfaces. The other wax compound classes were mostly specific (e.g., triterpenols in tomato (**H**)) or even organ-specific (e.g., flavonoids in tomato fruits (**J**)). Values for tomato fruit cv. M82 represent a three-year average. Error bar, standard error (n = 3 for tomato and apple; n = 10 for hybrid aspen).*

Finally, the standard method represents a compromise between both rapid sampling and extraction periods (manual process) and slow GC sequences (automatic process), to ensure a good peak resolution that enable the identification and quantification of cuticular wax constituents present in low amounts. It allowed us to identify most of the wax compounds detected in the six mixtures (**Figure I.5**), and from the total wax mixture 85% of the peaks detected were identified for tomato fruit cv. M82, 91% for tomato fruit cv. MT, 94% for tomato leaf cv. M82, 99% for apple fruit cv. Golden Delicious, 98% for apple fruit cv. Granny Smith and 54% for hybrid aspen leaf. Additional files (**Figure I.S1** and **Tables I.S1** to **I.S5**) provide detailed information about the mass spectra for those identified compounds.

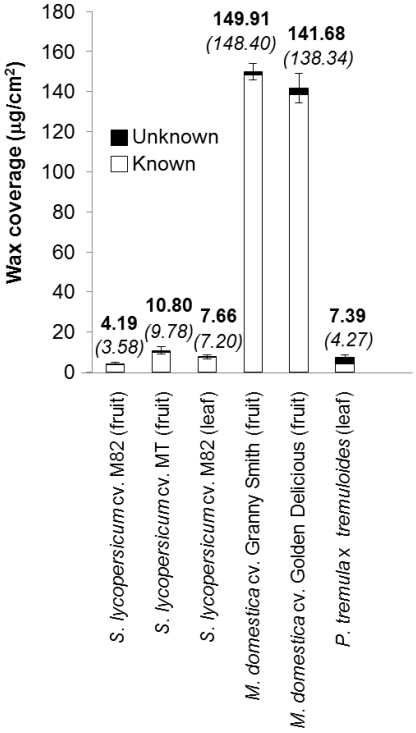


Figure I.6 Total wax coverages on six crop surfaces. *Over each bar, the wax coverage is given in bold face and, within it, the amount of identified compounds in italics. The value of tomato fruit cv. M82 represents a three-year average. Error bar, standard error (n = 3 for tomato and apple; n= 10 for hybrid aspen).*

I.4 CONCLUDING REMARKS

A standard method for cuticular wax analysis has been adapted. This method is suitable for different crop plant surfaces (i.e., organs, cultivars and species) and is effective in capturing most of the wax compositional complexity. The smaller surface area needed for analysis and the shorter extraction periods using a uniform and versatile solvent makes the standard method a cheap, fast, useful and desirable tool for medium-throughput approaches. Therefore, we propose this standard method as an effective tool for screening large populations or collections of any surface samples of crop plants.

I.5 ANNEXES I

Additional Figures

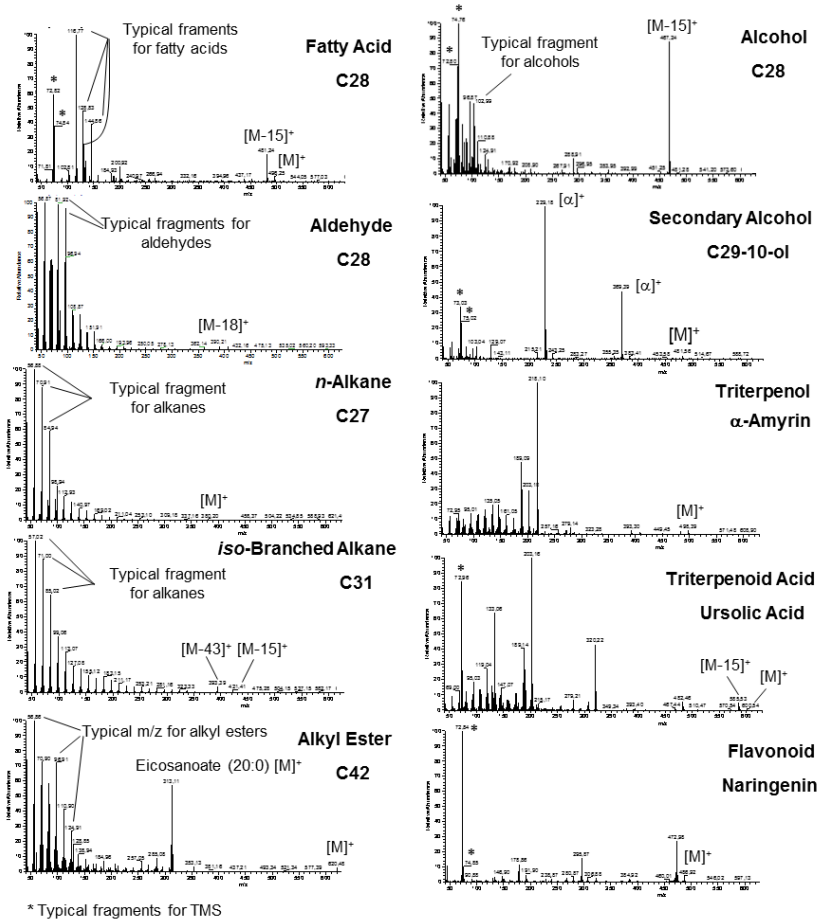


Figure I.S1 Typical mass spectra obtained for the different wax classes observed. *The typical fragments are shown for a fatty acid, aldehyde, alkane, alkyl-ester and alcohol wax classes. Also the fragments m/z 73 and 75 originating from the trimethyl silyl (TMS) group are shown. Molecular mass in the positive ionization mode* $[M]^+$ *and specific m/z fragments (e.g.* $[M-15]^+$ *or* $[M-18]^+$*) are also shown.*

Additional Tables

You can find these additional tables in the excel file "*CHAPTER I_Additional Tables I.S1-I.S5*" into the CD attached to this thesis.

Table I.S1 Cuticular wax composition for tomato fruit (Solanum lycopersicum) cv. M82 across years 2010, 2011 and 2012. *Specific fragments are shown. For typical fragmentation patterns of each wax class see mass spectra in Online Resource 1.*

Table I.S2 Cuticular wax composition for tomato fruit (S. lycopersicum) cv. MicroTom. *Specific m/z fragments are showed. Specific fragments are shown. For typical fragmentation patterns of each wax class see mass spectra in Online Resource 1.*

Table I.S3 Cuticular wax composition for tomato leaf (Solanum lycopersicum) cv. M82. *Specific m/z fragments are showed. Specific fragments are shown. For typical fragmentation patterns of each wax class see mass spectra in Online Resource 1.*

Table I.S4 Cuticular wax composition for apple fruit (Malus domestica) cultivars Golden Delicious (GD) and Granny Smith (GS). *Specific fragments are shown. For typical fragmentation patterns of each wax class see mass spectra in Online Resource 1.*

Table I.S5 Cuticular wax composition for hybrid aspen leaf (Populus tremula x P. tremuloides). *Specific fragments are shown. For typical fragmentation patterns of each wax class see mass spectra in Online Resource 1.*

CHAPTER II

QUANTITATIVE-TRAIT-LOCI ANALYSIS FOR TOMATO FRUIT CUTICLE COMPOSITION USING THE *Solanum pennellii* INTROGRESSION LINE POPULATION

II.1 INTRODUCTION

The modern tomato (*Solanum lycopersicum*) cultivars have been obtained after several rounds of domestication and improvement (Gur et al., 2011; Frary et al., 2010). Wild tomato relatives have been often used in modern breeding as a source of genes for stress tolerance (Pineda et al., 2012), disease resistance (Sandbrink et al., 1995) and, more recently, also for increasing yield (Gur and Zamir, 2004). The wild species *Solanum pennellii* is specially adapted to extreme stress conditions (Yeats, et al., 2013; Bolger et al., 2014) and has been an important germplasm donor for improvement of the cultivated tomato *S. lycopersicum*. Furthermore, the leaf cuticle of *S. pennellii* shows special characteristics for adaption to such conditions, including those involved in the regulation of cuticle related genes (Bolger et al., 2014). These characteristics probably make this species an ideal source of genes for cuticle reinforcement. One of the most used genetic resources to identify gene regions associated to quantitative traits of interest is the *Solanum. pennellii* X *S. lycopersicum* cv. M82 interspecific population. In such set of lines the full genome of the wild species is represented as small introgressed fragments into 75 nearly isogenic lines of the cultivar M82 genetic background (Eshed and Zamir, 1995). Out of the more than 3,000 QTLs reported (Alseekh et al., 2013), close to 2,000 QTLs are controlling fruit quality traits (i.e. Brix, sugar content, volatiles, firmness or flavonoids) (Chapman et al., 2012; Alseekh et al., 2013). To date, however, the *S.*

pennellii population has not been used to study QTL involved in fruit cuticle composition.

The cuticle is the outermost layer covering all aerial plant organs and is largely composed by acyl-lipids, polysaccharides and proteins (Heredia and Domínguez, 2009; Riereder and Müller, 2006). The cuticle is structured in two main layers, the cuticle proper and the cuticular layer (Riereder and Müller, 2006). The cuticle proper is formed by a cutin matrix in which, cuticular waxes are embedded (Riereder and Müller, 2006, Yeats and Rose, 2013). Cutin is a non-soluble three-dimensional cross-linked structure composed mainly by C_{16} and C_{18} hydroxy- and epoxy- fatty acids (Pollard et al., 2008; Beisson and Pollard 2012). The cuticular layer placed below the cuticle proper and made of cutin polymers linked to polysaccharides of pectin, hemicellulose and cellulose from the secondary cell wall (Heredia and Domínguez, 2009; López-Casado et al., 2007; Riederer and Müller, 2006). The cuticular waxes are a heterogeneous mix composed of ubiquitous aliphatic compounds and other secondary metabolites (e.g. triterpenoids or flavonoids) that are embedded and deposited onto the cutin matrix (Samuels and Jetter, 2008; Buschhaus and Jetter, 2011). The cuticle has an important role in the physiology, ecology and development of the plant (Hen-Aviv et al., 2014; Martin and Rose, 2014) and its appropriate functionality is apparently crucial for plant survival (Riederer and Müller, 2006). Thus, it is important to investigate the cuticle in plant crops to ensure high yield (e.g. preventing water loss or pathogen infection) (Hovav et al., 2007; Isaacson et al., 2009; Buxdorf et al., 2014), but also to improve quality traits (e.g. visual or industrial qualities) (Domíngez et al., 2011; Chapman et al., 2012; Lara et al., 2014).

During the last decades, the tomato cuticle has received some attention with focus on several aspects associated with biomechanics (Domínguez and Heredia, 2011), water-loss (Hovat et al., 2007; Isaacson et al., 2007), fruit's cracking (Domínguez et al., 2012), pathogen infection,

specific gene expression programs (our lab's *unpublished* results) and glossiness (Kunst and Samuels, 2009). This was performed via the characterization of mutant and transgenic lines displaying various phenotypes. Yet, the biosynthetic routes to cuticular lipids biosynthesis and the especially the regulatory mechanisms controlling them still remain unclear (Yeats and Rose, 2013; Hen-Aviv et al., 2014; Martin and Rose, 2014). A recent study on tomato fruit cuticle from several wild species (Yeats et al., 2012a) has revealed a large variability in cuticle composition in genetically distant wild relatives, including *S. pennellii*. The availability of the *S. pennellii* X *S. lycopersicum* cv. M82 introgression line population and the recent release of the *S. pennellii* genome sequence represent an excellent opportunity to investigate the genetic basis for variability of tomato cuticle (Bolger et al., 2014). Therefore, the aim of this study was to identify QTL for cuticular wax and cutin monomer composition in the tomato fruit cuticle and moreover to propose putative candidate genes that could be involved in their molecular regulation.

II.2 MATERIALS & METHODS

Plant Material

We used the interspecific introgression line (IL) population derived from the cross between the cultivated tomato *Solanum lycopersicum* cv. M82 and the wild tomato *Solanum pennellii*, LA716. From the 75 ILs (Eshed and Zamir, 1995), 73 ILs were analysed for cutin monomer composition and 63 ILs for cuticular waxes. The population was grown in a single-block trial under standard growth conditions with natural light and controlled temperature (24ºC during day and 18ºC during night) in a greenhouse during 2010 summer season in Rehovot, Israel. Four plants per IL and ten for the parental M82 were assessed and tomato fruits were harvested at red stage. Four fruits per plant were harvested, except for lines producing small fruits where eight fruits were collected. For validation of the genetic effects, selected lines were grown in 2011 under the same conditions as in 2010.

Chemicals for cuticular wax analyses

Unless otherwise specified, all chemicals were purchased from Sigma Aldrich®: n-alkanes (C_8-C_{41}), fatty acids (C_{29}-C_{31}), α- and β-amyrins, BSTFA (bis-trimethylsilyltrifluoroacetamide), MetOH/BF_3 (10% w/w), pyridine, pectinase and cellulase enzymes (*Aspergillus niger*), sodium azide, sodium hydrocarbonate, sodium sulphate, sodium tetraborate decahydrate, except for sodium acetate (Merck®), glacial acetic acid (Furtarom®) and chloroform LC-pure (ACS, ISO, Reag. Ph Eur chloroform, EMSURE®). Alkanes and fatty acids were used as internal standards or as a standard for identification at individual concentrations of 0.2 mg/ml. The analysis of fruit cuticular waxes was performed using enzymatically isolated cuticular membranes (ICM) (Riederer & Müller, 2006). Further, both epi- and intracuticular waxes are extracted together and they are considered as cuticular waxes.

Tomato fruit cuticle membrane isolation

Tomato fruit cuticle membranes (CMs) were obtained as previously described (Hovav et al., 2007; Fernandez-Moreno et al., *under review*). Briefly, fruit skin discs (total surface area of ~19 cm^2) were dissected and CMs were isolated by enzymatic digestion in a three day process. The enzyme solution contained 1% pectinase and 0.5% cellulase in a 0.1M sodium acetate buffer at pH 3.8 and 1mM sodium azide to prevent microbial growth. The enzymatic incubation was performed at 37ºC for a total of two days, replacing the enzymatic solution once after 24h. Isolated cuticle membranes (ICMs) were then treated with sodium tetraborate decahydrated solution (10mM TTBS, pH 9-9.2) for one minute and washed twice in distilled water for one minute each time and then dried at 60ºC during 48 h. Finally, ICMs were weighted and stored in a non-oxidative atmosphere by flushing N_2 gas.

Cuticular lipids extraction

The cuticular wax extraction was performed as described previously (Fernandez-Moreno et al., *under review*). Briefly, ICM discs were dipped into 4 ml of chloroform at room temperature and shaken gently during 30 sec. twice. Immediately after the extraction, *n*-tetracosane (0.2 mg/ml) was added as internal standard (ISTD). The final volume was reduced to 1 ml using a stream of N_2 and samples were stored at -20ºC until further analysis. Next, dewaxed ICMs were exhaustively delipidated in MetOH/$CHCl_3$ (1:1 v/v) for 15 days and air-dried during 12 hours to proceed with the cutin extraction. Cutin matrix was then depolymerisated by using a portion of the sample surface (~13 cm^2 from the dewaxed area of ~19 cm^2), which was weighed and incubated in MetOH/BF_3 (2 ml) for 16 hours at 70ºC. The cutin monomers obtained after this methyl-esterification reaction were extracted as described previously (Franke et al., 2005; Shi, Jian-Xi et al., 2012). Briefly, samples were incubated with *n*-tetracosane (ISTD) and the esterification

reaction was stopped with 2 ml of saturated $NaHCO_3$–water solution. Next, cutin monomers were extracted with 2 ml $CHCl_3$ three times. In each extraction the organic fraction was transferred to a new vial. Finally, the pooled organic fraction was washed twice with 1 ml water and the remaining water was eliminated with anhydrous Na_2SO_4. The final extract was concentrated to 1ml using a stream of N_2 gas. Cutin extracts were stored at -20ºC until further analysis.

Cuticular lipids analysis

The cuticular lipid screening profiles was performed by Gas Chromatography/Flame Ionization Detector (GC-FID) using three replicates per line (Adato et al., 2009). Additionally, a reference sample was prepared containing a mixture of the three replicates per line to identify cuticular waxes and cutin monomers using Gas Chromatography/Mass Spectrometry (GC-MS) (Mintz-Oron et al., 2008; Fernandez-Moreno et al., *under review*). For the validation analysis of the best QTL candidates obtained during the initial screening, we quantified in a second year the cuticular lipids using a GC-MS (Mintz-Oron et al., 2008; Fernandez-Moreno et al., *under review*).

Before GC analysis, hydroxyl groups in the cuticular lipid extracts were transformed into the corresponding trimethylsilyl derivatives as follows. The volume of wax extracts was first reduced to 100 μl under a stream of N_2. Then, 40 μl of BSTFA and pyridine (1:1 v/v) were added and the solutions mixed by vortexing for 1 minute. Derivatization was allowed to proceed at 70ºC for 40 minutes. After cooling to room temperature, the whole reaction mixture was transferred to a vial and an aliquot injected into the GC column.

The screening for variability in cuticular lipids was done in a capillary GC with FID under the same GC conditions as described for the GC-MS analysis, but with H_2 carrier gas inlet pressure programmed for a constant flow of 2 ml min^{-1} (Adato et al., 2009). In parallel, cuticular lipid compounds were identified by a GC (Trace GC, Thermo Scientific),

connected to an electron impact MS detector (DSQ2; Thermo Scientific) set at 70 eV and m/z 40-850 Da (± 0.5 Da) in positive mode (Fernandez Moreno et al., *under review*). Samples were injected using a solvent split mode (50 ml/min of split flow) in a PTV injector, and compounds were separated in a Zebron DB-1 column (Phenomenex ZB-1MS, 30 m length, 0.25 mm I.D. and 0.50 μm film thickness). GC was carried out in an oven programmed for 0 min at 65ºC, 20ºC min-1 to 220ºC, 3 min at 220ºC, 3ºC min-1 to 300ºC, 0 min at 300ºC, 5ºC min-1 to 330ºC and 35 min at 330ºC using a helium flow of 1.2 ml/min (Fernandez Moreno et al., *under review*) for both, cuticular waxes and cutin monomers. During the second year the qualitative and quantitative composition in cuticular lipids was validated by the GC-MS using the same conditions above described in order to validate the results on a number of the lines selected from the first year screening.

Different wax constituents were identified in the total ion chromatogram using their Kovat's indices and by comparing their mass spectra with those of authentic standards as well as data from the literature (Fernandez Moreno et al., *under review*). The Avalon algorithm (Xcalibur software, Thermo Fisher Scientific) was used for the detection and peak integration of the GC-FID datasets, and the ICIS algorithm (Xcalibur software, Thermo Fisher Scientific) was used for the detection and peak integration on GC-MS datasets using the following parameters: an area noise factor of 5, a peak noise factor of 10 and a minimum peak height (s/n) of 3. The quantification of each wax compound was performed in the same way for all the wax classes, and quantities calculated based on the ratio between each peak area and the area of the ISTD. Peak integration was double checked manually. Finally, each ratio was normalized based on the total amount of ISTD (5 μg per mixture) and the total area of extraction (18.84 cm^2 for cuticular waxes, and 12.56 cm^2 for cutin monomers) to produce the final dataset (μg/cm^2). The ICM weight (ICMW) was also normalized according to the area extracted (mg/cm^2). Before statistical analysis, data

were pretreated by centering ($\tilde{x}ij = xij - \bar{x}i$) and scaling using *range scaling* ($\tilde{x}ij = \frac{xij - \bar{x}i}{Max(xj) - Min(xj)}$) (van den Berg et al., 2006).

Hierarchical clustering analysis

The phenotypic variation among the *S. pennellii* IL population for both cuticular wax and cutin monomer metabolites was represented by a two-way hierarchical clustering analysis (HCA). A normalized and pre-treated (centered and range-scaled) dataset were used for the HCA. JMP® 10.0.0 software was used for the analysis (SAS Institute Inc., 2012).

QTL analysis and mapping

Cuticular wax, cutin monomers and ICM weight QTL analysis was carried out by a one-way ANOVA and the Tukey's HSD test as post-hoc analysis in conjunction with the ANOVA. A set of 187 genetic markers, including RFLPs and CAPSs (EXPEN 2000 map, http://www.solgenomics.net/), distributed among the 73 ILs (**Figure II.S1**) was selected for cutin monomer analysis. The statistical threshold to declare significant effects was (α = *0.05/73bins*), being 73 the number of bins defined by the recombination points in the IL collection. Further, we also performed a pooled *t*-test analysis on the different metabolite traits, and we kept the same threshold as for the ANOVA in order to declare significant effects. Additionally, ICM weight trait and the validation analysis for specific metabolites in a selected set of lines during the second year were studied comparing each IL against parental M82 for each metabolite by a Student's *t*-test analysis ($\alpha < 0.05$). Every statistical analysis herein was done using JMP® 10.0.0 software (SAS Institute Inc., 2012).

Microscopy imaging

For transmission electron microscopy (TEM) observations, 9 cm^2 fruit exocarp pieces were fixed (using paraformaldehyde, glutaraldehyde and

osmium tetroxide solutions) and embedded in EPON resin following the protocol developed at the Electron Microscopy Unit at Weizmann Institute of Science (Rehovot, Israel) and available online: *http://www.weizmann.ac.il/Chemical_Research_Support/EMUnit/links*. Ultrathin sections were also used for fluorescent microscopy observations (Nikon fluorescence microscope, Nikon Eclipse e-800, Japan) and to measure the cuticle thickness using IMAGEJ software.

II.3 RESULTS & DISCUSSION

Cuticular lipids identified in the IL population

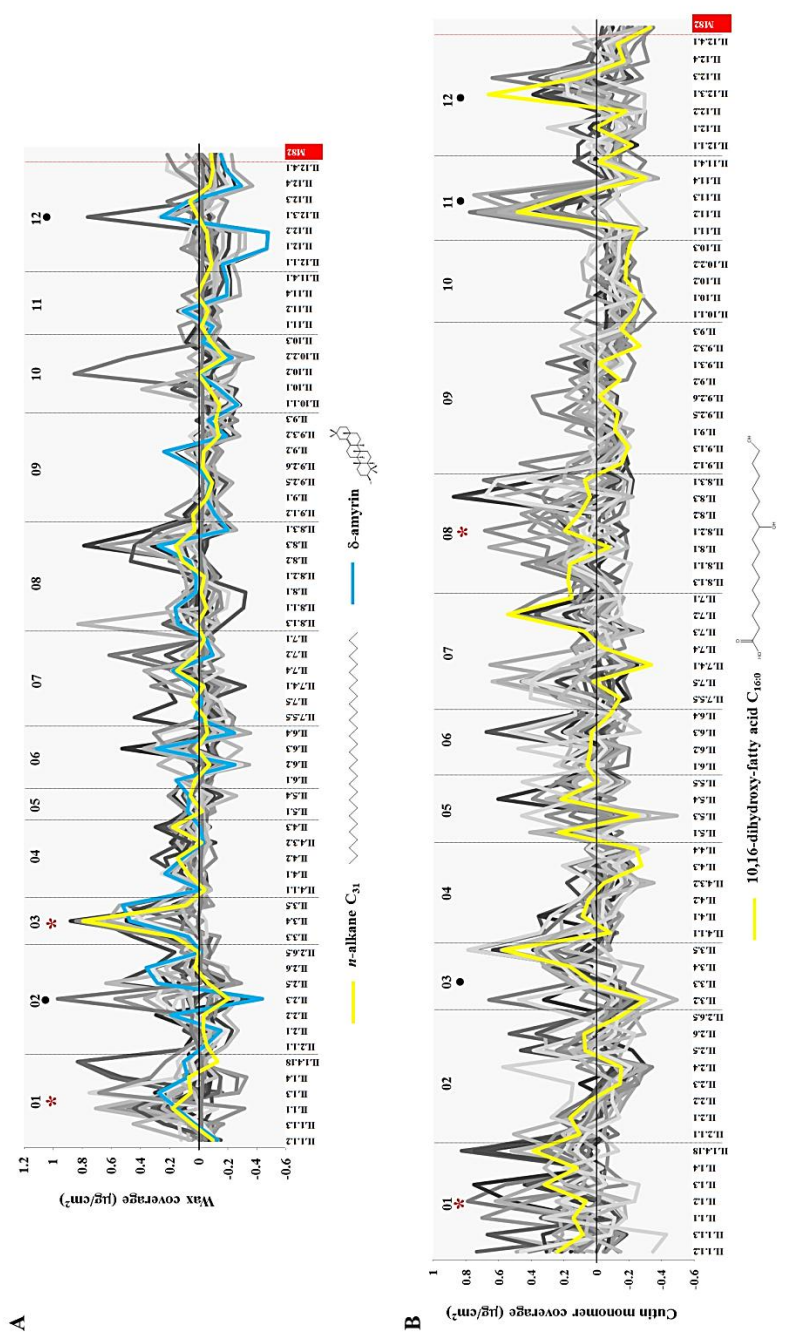
An adaptation of the methodology described by Hovav et al. (2007) was used herein to study the variability of cuticular waxes in tomato fruit IL population (Fernandez-Moreno et al., *under review*). We used this method first to screen cuticular wax and cutin monomer compositions in tomato fruit across the *Solanum pennellii* x *S. lycopersicum* cv. M82 IL population (Eshed and Zamir, 2005). Next, the major cuticular lipids in the IL population were successfully identified, 26 cuticular waxes and 26 cutin monomers, by GC-MS (**Figure II.S2**, **Tables II.S1** and **II.S2**). Cuticular waxes identified included (**Figure II.S2A**, **Table II.S1**): even-numbered fatty acids C_{22}, C_{24}, C_{26} and C_{30}; *n*-aldehydes C_{24}; odd-numbered *n*-alkanes C_{27}-C_{33} and even-numbered *n*-alkanes C_{28}-C_{32}; *iso*-alkanes C_{30} and C_{31}; *n*-alcohols C_{22}, C_{23}, and C_{32}; three pentacyclic triterpenols, α-, β-and δ-amyrins; and the flavonoid naringenin. In addition, we could identify two waxes with similar spectrum to *n*-aldehydes (i.e., $unknown_1$ and $unknown_2$), other two waxes with a typical spectrum for pentacyclic triterpenoids (i.e., $unknown_3$ and $unknwon_4$) and one additional unidentified wax ($unknown_5$). Into the cutin monomer set, we could identify 12 cutin monomers (**Figures II.S2C** and **II.S2D**, **Table II.S2**): seven $C_{16:X}$ monomers, ω-hydroxylated fatty acid $C_{16:0}$, ω-hydroxylated fatty acid $C_{16:1}$, 10,16-ω-dihydroxylated fatty acid $C_{16:0}$, 9,16-ω-dihydroxylated fatty acid $C_{16:0}$, 8(9/10),16-ω-dihydroxylated fatty acid $C_{16:0}$, α,ω-dicarboxilic acid $C_{16:0}$, 9(10)-hydroxylated α,ω-dicarboxilic acid $C_{16:0}$; and four $C_{18:X}$ monomers, 9,10-dihydroxy fatty acid $C_{18:1}$, 9,10,18-trihydroxy fatty acid $C_{18:1}$, 9,10-epoxy 18-ω-hydroxy fatty acid $C_{18:0}$ and 9,10-epoxy 18-ω-hydroxy fatty acid $C_{18:1}$. The remaining 14 cutin monomers detected (**Figure II.S2D**, **Table II.S2**) were characterized by some structural features such as the presence of 9(10)-midchain-hydroxyl groups of some saturation. Additional work is required to obtain a complete identification for these 14

unknown cutin monomers. In summary we screened a set of 52 total cuticular lipid components (i.e., 26 cuticular waxes and 26 cutin monomers) over the IL population.

Phenotypic variation in cuticular lipid composition across the *S. pennellii* IL population

Out of the 75 ILs comprising the population 73 ILs were studied (**Figure II.S1**). We obtained the cutin monomer profile for the 73 ILs (**Figure II.1B**) but 10 lines could not be processed for cuticular wax analysis, and therefore the cuticular wax profile was obtained for 63 of the 73 initial ILs (**Figure II.1A**). To determine the variability among ILs and the parental M82 for the 52 metabolites we performed two independent two-way hierarchical clustering analyses (HCA), one for ILs *vs.* cuticular waxes (**Figure II.2**) and a second one for ILs *vs.* cutin monomers (**Figure II.3**).

Figure II.1 Quantitative variation of cuticular metabolites among the S. pennellii x S. lycopersicum cv. M82 introgression line population. ***A****. Cuticular wax variation. The major cuticular wax accumulating in tomato fruit (n-alkane C_{31}) is highlighted in yellow, and the second most abundant wax (δ-amyrin) in blue.* ***B****. Cutin monomer variation. The major cutin monomer accumulating in tomato fruit (10,16-ω-dihydroxy $C_{16:0}$ fatty acid) is highlighted in yellow. ILs are placed bin the corresponding chromosomes. Metabolite coverage ($\mu g/cm^2$) for the pretreated dataset is represented. *, hotspot chromosomes for metabolite accumulation. ●, other chromosomes containing interesting ILs for metabolite accumulation. [next page].*


A
Wax coverage (µg/cm²)
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
01
02
03
04
05
06
07
08
09
10
11
12
M182
n-alkane C31
δ-amyrin
B
Cutin monomer coverage (µg/cm²)
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
10,16-dihydroxy-fatty acid C16:0

Cuticular wax phenotyping

In the ILs *vs.* cuticular waxes HCA, we observed seven Introgression Lines (IL) clusters (see 1-7 in **Figure II.2**) and two wax clusters (see I and II in **Figure II.2**). The wax-cluster I contained the major cuticular waxes accumulated in tomato fruit: the three amyrins and the *n*-alkane C_{31} while, the wax-cluster II contained the second most accumulated cuticular wax, *n*-alkane C_{29}. In wax-cluster I, we find three sub-clusters: (a) the wax-cluster Ia, including only *n*-alkanes with $\geq C_{30}$ carbons; (b) wax-cluster Ib, which contained the two *iso*-alkanes C_{30} and C_{31} quantified in the population, the α-amyrin and other several waxes (i.e., fatty acids, *n*-alcohols and *n*-aldehydes); and (c) the wax-cluster Ic, containing both δ- and β-amyrins. Similarly, in wax-cluster II, we find two sub-clusters: (a) the wax-cluster IIa, which contained the fatty acid C_{30}; and the wax-cluster IIb, containing *n*-alkanes $\leq C_{30}$ carbons and other waxes (i.e., fatty acids, *n*-alcohols and naringenin). On the other hand, the ILs were clustered in seven IL-clusters based on their levels for those waxes belonging to the wax-cluster I. The IL-cluster 1 thus contained ILs with the lowest levels of waxes belonging to the wax-cluster I, and the IL-cluster 7, which consist only in IL3.4 (see asterisk in **Figure II.2**), showed the highest levels for them. Further, we found that the IL with the most different cuticular wax composition relative to the parental M82 (see # in **Figure II.2**) is the IL3.4 (IL-cluster 7, see asterisk in **Figure II.2**). This IL is the one in the population with the highest values for *iso*- and *n*-alkanes, as well as for other waxes in wax-clusters Ia, Ib, IIa and IIb. In contrast, the parental cv. M82 was placed in IL-cluster 1, exhibiting lower levels of cuticular waxes than ~76% (48 ILs) of the population.

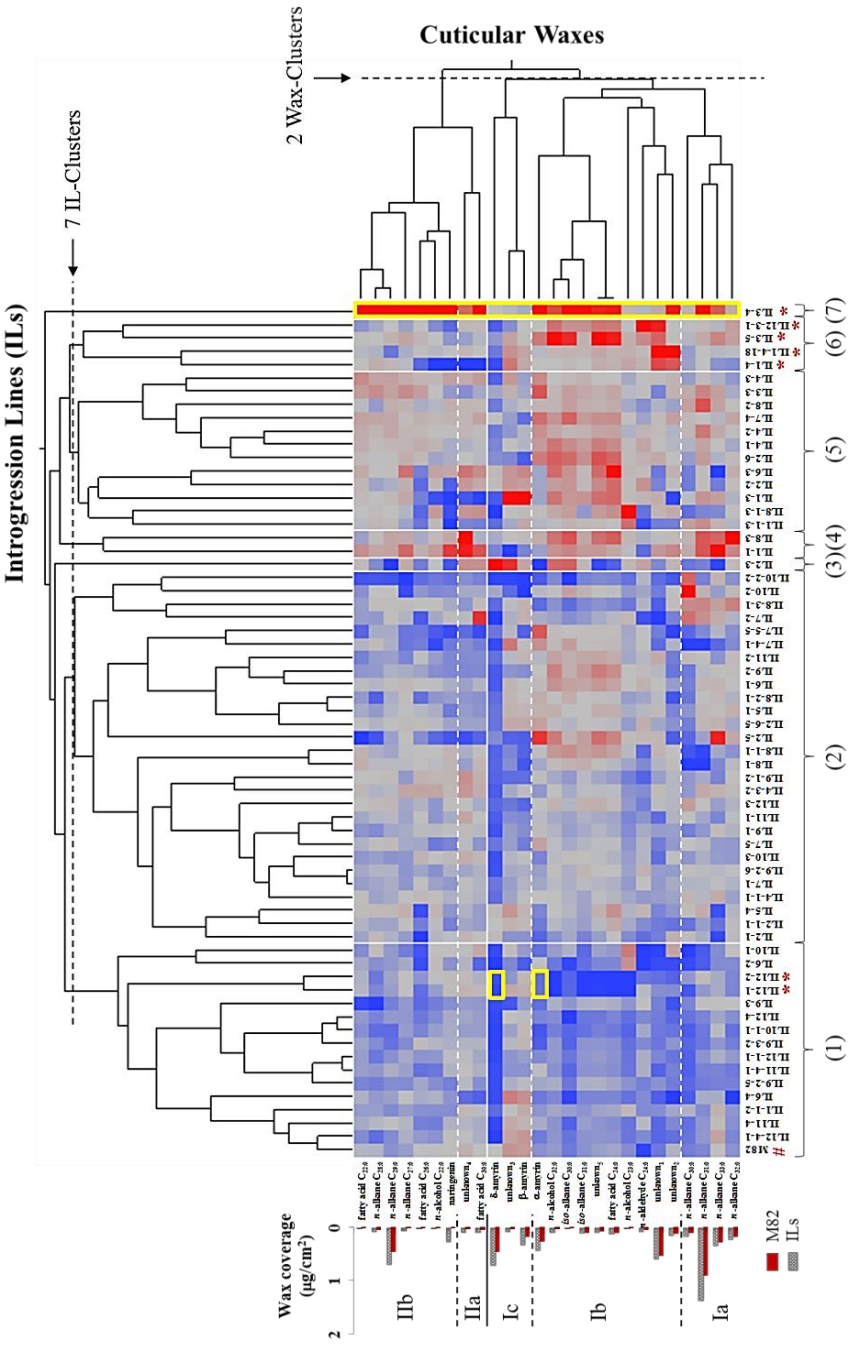
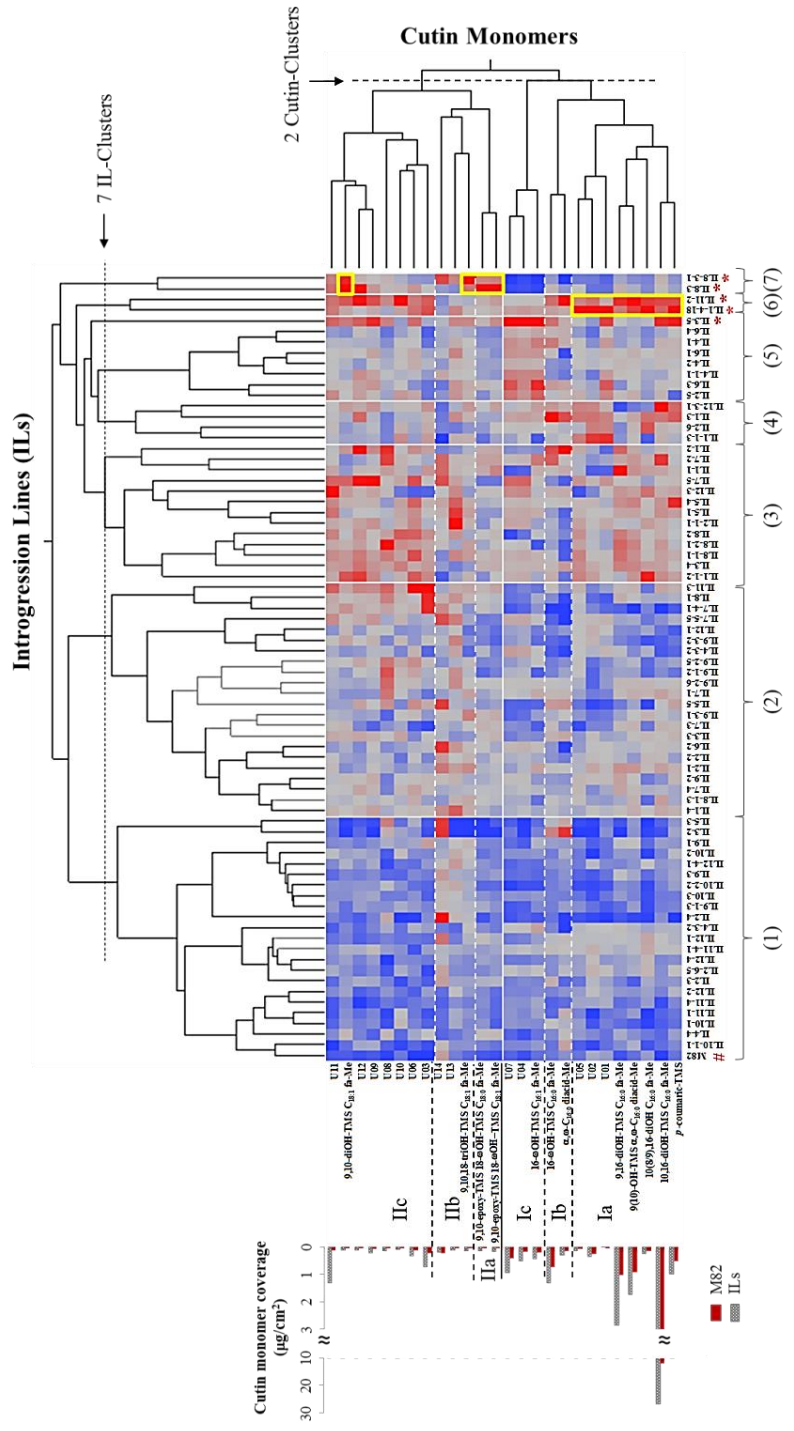

Cuticular Waxes
2 Wax-Clusters
7 IL-Clusters
Introgression Lines (ILs)
(1)
(2)
(3)(4)
(5)
(6) (7)
M82
ILs
Wax coverage (µg/cm²)
IIb
IIa
Ic
Ib
Ia

Figure II.2 Screening for cuticular wax composition in the S. pennellii x S. lycopersicum cv. M82 introgression line population. *A two-way hierarchical clustering analysis for 31 cuticular waxes in 63 ILs is presented. Seven IL-clusters (1-7) and two wax-clusters (I and II) were obtained from the HCA. Metabolite sub-clusters are represented by a Romanic number followed by a lower case letter (e.g., Ia, IIb, etc.). . Color legend: lower values are in blue and higher values in red for each cuticular wax. The cuticular wax coverage ($\mu g/cm^2$) for parental M82 (solid red bars) and for the IL population-mean (dotted grey bars) are also shown on the left side of the heatmap. Asterisks represent the most different ILs respect parental line M82 clustered in IL-clusters 6 and 7 (i.e., IL1.4.18, IL IL3.5, IL12.3.1 and IL3.4), and also represent the ILs with the lowest amount of amyrins (i.e., IL12.1 and IL12.2). #, represent the parental line M82. [previous page].*

Figure II.3 Screening for cutin monomer composition in the S. pennellii x S. lycopersicum cv. M82 introgression line population. *A two-way hierarchical clustering analysis for 27 cutin monomers in 73 ILs is shown. Seven IL-clusters (1-7) and two cutin-clusters (I and II) resulted from the HCA. Metabolite sub-clusters were represented by a Romanic number followed by a lower case letter(e.g., Ia, IIb, etc.). Representative lines containing the lowest or the highest values for epoxy-hydroxy and trihydroxy fatty acids (IIa and IIb), as for major cutin monomers (Ia) are marked with asterisks. Color legend: lower values are in blue and higher values in red for each cutin monomer. Cutin monomer coverage ($\mu g/cm^2$) for the parental M82 cultivar (solid red bars) and for the IL population-mean (dotted grey bars) for each cutin monomer is showed at the left side of the heatmap. TMS, trimethylsilyl-eter; Me, methyl-ester; OH, hydroxyl; fa, fatty acid; diacid, dicarboxylic acid; U, unidentified monomer. #, represent the parental line M82. [next page].*


Cutin Monomers
2 Cutin-Clusters
7 IL-Clusters
Introgression Lines (ILs)
(1)
(2)
(3)
(4)
(5)
(6)(7)
IIc
IIb
IIa
Ic
Ib
Ia
M82
ILs
Cutin monomer coverage (µg/cm²)

Cutin monomer phenotyping

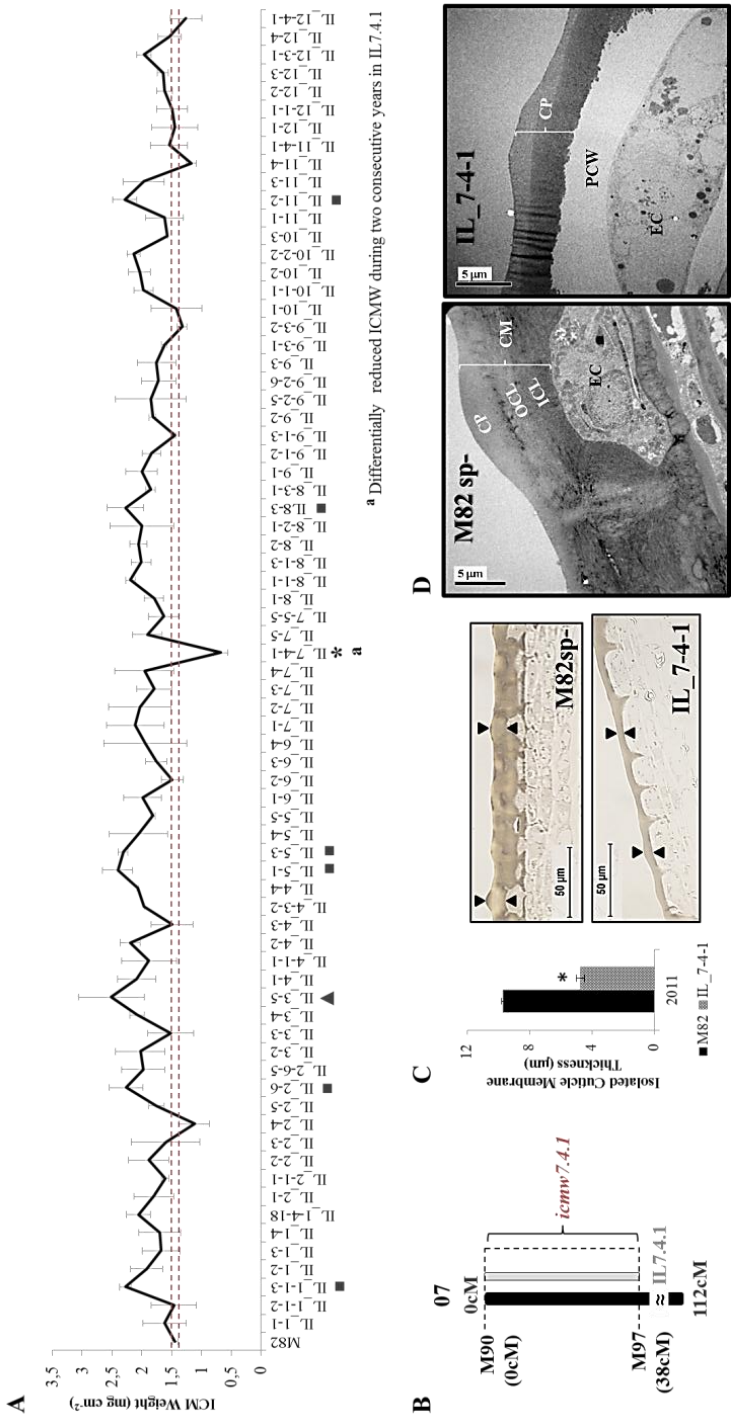
Further, we also found seven IL-clusters (see 1-7 in **Figure II.3**) in the IL *vs.* cutin monomers HCA, and two main cutin-clusters (see I and II in **Figure II.3**). Cutin monomers were splited into cutin-cluster I containing the $C_{16:X}$ cutin monomers (most abundant monomers; see I in **Figure II.3**) and cutin-cluster II containing the $C_{18:X}$ monomers (less abundant monomers; see II in **Figure II.3**). Both cutin-clusters I and II could be divided in three sub-cluster each: (a) the cutin-cluster Ia, containing the mid-chain hydroxylated $C_{16:X}$ monomers and the *p*-coumaric acid; (b) the cutin-cluster Ib, including the ω-hydroxylated fatty acid $C_{16:0}$ and the α,ω-dicarboxilic acid $C_{16:0}$; (c) the cutin-cluster Ic, containing only the ω-hydroxylated fatty acid $C_{16:1}$; (d) the cutin-cluster IIa, which showed the two epoxy-hydroxylated fatty acids $C_{18:X}$; (e) the cutin-cluster IIb, containing the trihydroxy fatty acid $C_{18:1}$; and (f) the cutin-cluster IIc, including the mid-chain hydroxylated fatty acid $C_{18:X}$. The cutin composition in tomato fruit is typically dominated by the $C_{16:X}$ monomer 10,16-dihydroxy $C_{16:0}$ fatty acid and its isomers 9(8),16-dihydroxy $C_{16:0}$ fatty acid, followed by *p*-coumaric acid and 9(10)-hydroxy-α,ω-dicarboxylic acid $C_{16:0}$. These cutin monomers were clustered into the cutin-cluster Ia (**Figure II.3**). Based on the levels for these cutin monomers the ILs were clustered into seven IL-clusters. For example, ILs into cutin-cluster 1 showed lower levels for these three cutin monomers while ILs in cutin-cluster 6 exhibited the highest levels for them. The IL1.4.18 and IL11.2 in cutin-cluster 6 (see asterisks in **Figure II.3**) represented the most different ILs with respect to parental cv. M82 (see # in **Figure II.3**) when the major cutin monomers are considered. Furthermore, the levels of $C_{18:X}$ cutin monomers appear to determine the distribution of some ILs with otherwise invariable levels of the major $C_{16:X}$ cutin monomers. For example, this is the case of IL8.3 and IL8.3.1 both in cutin-cluster 7 (see asterisks in **Figure II.3**) exhibited higher levels than parental cv. M82 for cutin monomers which are characteristics of cutin-cluster II ($C_{18:X}$), but invariable levels for the major

cutin monomers. Finally, the parental cv. M82 displayed reduced cutin monomer accumulation than ~69% (50 ILs) of lines in the IL population in a similar manner as for cuticular waxes.

Comparing the HCA results with the distribution of the 52 metabolites in the IL population (**Figure II.1**) we observed that cutin monomer composition showed a wider variability than cuticular wax composition in the *S. pennelli* IL population. A closer inspection revealed chromosomes 01, 03 and 08 as most relevant for variability in cuticular lipid composition (see asterisks in **Figure II.1**) followed by chromosomes 02, 11 and 12 (see dots in **Figure II.1**). Lines in these chromosomes showed the largest variation in cuticular wax and/or cutin monomer composition with respect parental cv. M82 (e.g. IL1.4.18, IL2.4, IL2.5, IL3.4, IL3.5, IL8.3 and IL8.3.1, IL11.2, IL12.1, IL12.2 and IL12.3.1; see asterisk in **Figures II.2** and **II.3**).

Variation in the weight of isolated cuticle membranes across the pennelli IL population

The weight of the isolated cuticle membranes (ICMW) from the same tomato fruit cuticle samples used in the metabolic profiling was analyzed for the 73 ILs (**Figure II.4A**). Most of the lines showed similar or higher cuticle weight as compared to the parental M82 cultivar. IL3.5 showed the highest ICMW values (see triangle in **Figure II.4A**) as well as higher levels of triterpenoid waxes, followed by six ILs (see squares in **Figure II.4A**), including both IL8.3 and IL11.2 which also exhibited higher levels of cutin monomer accumulation. Interestingly, out of the entire IL population, only IL7.4.1 exhibited reduced ICMW with respect to the M82 cultivar (see asterisk in **Figure II.4A**).


A
ICM Weight (mg cm-2)
M82
IL_1-1
IL_1-1-2
IL_1-1-3
IL_1-2
IL_1-3
IL_1-4
IL_1-4-18
IL_2-1
IL_2-1-1
IL_2-2
IL_2-3
IL_2-4
IL_2-5
IL_2-6
IL_2-6-5
IL_3-2
IL_3-3
IL_3-4
IL_3-5
IL_4-1
IL_4-1-1
IL_4-2
IL_4-3
IL_4-3-2
IL_4-4
IL_5-1
IL_5-3
IL_5-4
IL_5-5
IL_6-1
IL_6-2
IL_6-3
IL_6-4
IL_7-1
IL_7-2
IL_7-3
IL_7-4
IL_7-4-1
IL_7-5
IL_7-5-5
IL_8-1
IL_8-1-1
IL_8-1-3
IL_8-2
IL_8-2-1
IL8-3
IL_8-3-1
IL_9-1
IL_9-1-2
IL_9-1-3
IL_9-2
IL_9-2-5
IL_9-2-6
IL_9-3
IL_9-3-1
IL_9-3-2
IL_10-1
IL_10-1-1
IL_10-2
IL_10-2-2
IL_10-3
IL_11-1
IL_11-2
IL_11-3
IL_11-4
IL_11-4-1
IL_12-1
IL_12-1-1
IL_12-2
IL_12-3
IL_12-3-1
IL_12-4
IL_12-4-1
a Differentially reduced ICMW during two consecutive years in IL7.4.1
B
07
0cM
M90
(0cM)
icmw7.4.1
M97
(38cM)
IL7.4.1
112cM
C
Isolated Cuticle Membrane Thickness (µm)
2011
M82 IL_7-4-1
D
M82 sp-
CP
CCL
ICL
CML
EC
IL_7-4-1
PCW
5 µm
50 µm

QTL[a]	IL	Chr	Marker interval (cM)[b]	Phenotype	Tomato gene name[c]	Arabidopsis protein ortholog[d]	Protein similarity	Reference
icmw7.4.1	IL7.4.1/IL7.4	7	M90-M92 (2)	Isolated cuticle membrane thickness showed half weight than parental cv. M82. Apparently, it is lacking the secondary cell wall.	Chitinase-like (*Solyc07g005100*)	Glycosyl hydrolase family 18 protein (*At4g19810*)	55%	*Not characterized*
					Glucan endo-1 3-β-glucosidase (*Solyc07g005330*)	β-1,3-glucanase, putative (*At3g55430*)	52%	*Not characterized*
			M95-M97 (29)		Chitinase-like (*Solyc07g009500*)	-	-	*Not characterized*
					Chitinase-like (*Solyc07g009510*)	-	-	*Not characterized*
					Chitinase-like (*Solyc07g009520*)	-	-	*Not characterized*
					SlXTH2 (*Solyc07g009380*)	AtXTR3 (*At5g57550*)	64%	Catalá et al., 2001
					Glucan endo-1 3-β-glucosidase 5 (*Solyc07g017730*)	Glycosyl hydrolase family 17 protein (*At4g31140*)	57%	*Not characterized*
					Glucan endo-1 3-β-glucosidase 3 (*Solyc07g047710*)	Glycosyl hydrolase family 17 protein (*At5g35740*)	58%	*Not characterized*
					Polygalacturonase 1 (Solyc07g015870)	Polygalacturonase/ pectinase, putative (*At1g60590*)	58%	*Not characterized*
					Pectinesterase (Solyc07g017600)	AtMPE44 (*At4g33220*)	59%	
					Pectinesterase inhibitor (*Solyc07g042390*)	Invertase/pectin methylesterase inhibitor family protein (*At5g20740*)	51%	*Not characterized*
					α,β-hydrolase domain-containing protein 5 (*Solyc07g032180*)	α,β-hydrolase family protein (*At4g24160*)	72%	*Not characterized*
					β-galactosidase (*Solyc07g038120*)	AtBGAL2 (*At3g52840*)	44%	*Not characterized*
					Cellulose synthase (*Solyc07g043390*)	AtCSLG1 (*At4g24010*)	33%	*Not characterized*
					SlbHLH140 transcription factor (*Solyc07g020960*)	ERF family protein (*At4g29100*)	47%	Sun et al., 2015
					SlbHLH051 transcription factor (*Solyc07g039570*)	bHLH family protein (*At5g56960*)	33%	Sun et al., 2015
					Transcription factor-like protein (*Solyc07g021550*)	Unknown protein (*At5g56780*)	39%	*Not characterized*

[a] QTLs are named by the abbreviation of "isolated cuticle membrane weight" (*icmw*), plus the IL number

[b] Genetic markers used in the table are described in Figure S1

[c] Gene annotations were obtained from SL2.40 database [SGN, http://www.solgenomics.net]

[d] Blastp for Protein ITAG realese 2.40 database [SGN, http://www.solgenomics.net]

Figure II.4 Identification of an isolated cuticular membrane weight-related QTL in IL7.4.1. A*. Screening for ICMW in 73 lines in the S. pennellii x S. lycopersicum cv. M82 IL population. Total tissue surface used for chemical analysis was weighed for the screening.* ***B****. Chromosomal location for icmw7.4.1 QTL.* ***C****. ICM thickness measurement (μm) and images from fluorescent microscopy (scale: 50 μm).* ***D****. TEM images from epidermal cells (EC) showing the cuticle membrane for both M82 (right) and IL7.4.1 (left) (scale: 5 μm).* ***E****. Candidate genes proposed as associated with icmw7.4.1 QTLs. CM, cuticle membrane. CP, cuticle proper. CL, cuticular layer. OCL, outer cuticular layer. ICL, inner cuticular layer. PCW, primary cell wall. Error bar: standard deviation (n = 3). *, IL showing the lowest value for ICMW. ►, IL showing the highest value for ICMW. ■, set of ILs showing higher ICMW values. [previous pages].*

QTL Mapping of cuticular lipids composition and isolated cuticle membrane weight

We used 187 DNA markers from the Tomato EXPEN 2000 map (SGN, http://www.solgenomics.net) covering the 73 ILs analyzed (**Figure II.S1**) to perform a one-way ANOVA analysis on the 53 traits. The analysis revealed 34 ILs containing *S. pennellii* genomic regions affecting cuticle wax, cutin monomer and/or ICMW phenotypes. **Table II.1** summarizes the ILs containing those genomics regions associated with differential accumulation of metabolites and the ICMW trait.

Wax QTL mapping reveal consistent regions regulating fruit cuticle wax composition

Out of the 26 cuticular waxes analyzed, 24 (~92%) were differentially accumulated in 13 of the 63 ILs (~21%; **Table II.1**). Most of the ILs showed one or a small number of cuticular waxes which were differentially accumulated when compared to the parental cv. M82, but we found only 2 ILs (i.e. IL3.4 and IL12.1) exhibiting alteration in three or more

cuticular waxes. Changes observed in those 2 ILs were grouped in three wax-related QTLs, namely *vlcfa3.4*, *ttp3.4* and *ttp12.1*, in chromosomes 3 and 12 (**Figure II.5A**).

The QTLs *ttp3.4* and *ttp12.1* showed an opposite triterpenoid accumulation profile: while *ttp12.1* caused a reduction in levels of triterpenoids, *ttp3.4* produced an increase of almost 2-fold (**Figure II.5B**). Thus, we found differentially reduced levels for δ- and α-amyrins but not for β-amyrin in the overlapped region IL12.1.1/IL12.1/IL12.2 (20 cM) and in the overlapped region IL12.1/IL12.2 (1 cM). The effect was stronger in the IL12.1/IL12.2 overlapped region (see **Table II.1**). Noteworthy, three triterpenoid synthases in tandem could be identified from the 77 genes included in this 1 cM region, *TTS1* (*Solyc12g006530*), *TTS2* (*Solyc12g006520*) and a *TTS-like* (*Solyc12g006510*) (**Figure II.5C**) suggesting that they are candidate genes to be involved in the genetic control of the phenotype observed. In fact, TTS2 is a general triterpenoid synthase able to produce the three amyrins, while TTS1 is a specific β-amyrin synthase (Wang et al., 2011). When we investigated the gene expression levels for these three candidate genes in the RNA-Seq dataset available in SGN website (http://www.solgenomics.net/), we found that both *TTS1* and *TTS2* genes showed in the corresponding ILs a reduced expression level when compared to the parental cv. M82 (**Figures II.S3A**).

Table II.1 A list of those ILs containing any trait (cuticular wax, cutin monomer and/or ICMW) showing differential accumulation in the S. pennellii x S. lycopersicum cv. M82 IL population. *Those lines containing reduced levels of cuticular waxes and ICMW are marked by ↓ to differentiate them from the remaining lines showing greater values than the M82 parental cultivar. The significance was considered as $\alpha = 0.05/$No. bins, being 73 the No. of bins analyzed (one-way ANOVA) for the metabolite dataset and $\alpha < 0.05$ (two-tailed t-test, n = 3) for ICMW data. [next page].*

TRAIT	INTROGRESSION REGION AFFECTED
Cuticular Wax	
Fatty acids	
C_{22}	IL8.2/IL8.3; IL8.3; IL8.3.1
C_{24}	IL3.4; IL7.2
C_{26}	IL1.4.18; IL3.3/IL3.4
C_{30}	IL3.4/IL3.5; IL12.1.1↓; IL12.1↓; IL12.2↓; IL12.1/IL12.2
Aldehydes	
C_{24}	IL1.1; IL3.3/3.4; IL3.4; IL3.4/IL3.5
$Unknown_1$	IL3.3/3.4; IL3.4; IL3.4/IL3.5
$Unknown_2$	IL3.3/3.4; IL3.4; IL3.4/IL3.5
n-Alkane	
C_{27}	IL10.2; IL10.2.2
C_{28}	IL1.1; IL2.5
C_{29}	IL3.4; IL8.3
C_{30}	IL3.4
C_{31}	IL3.4
C_{32}	IL3.4
C_{33}	IL3.3; IL3.4
iso-Alkanes	
C_{30}	IL1.3; IL1.3/IL1.4
C_{31}	IL1.3; IL1.3/IL1.4
Alcohols	
C_{22}	IL1.4.18
C_{23}	IL1.1; IL8.3
C_{32}	IL3.3/3.4; IL3.4; IL3.4/IL3.5
Pentacyclic triterpenoids	
δ-Amyrin	IL3.4; IL3.5; IL12.1.1↓; IL12.1↓; IL12.2↓
β-Amyrin	IL3.4; IL3.5
α-Amyrin	IL3.4; IL3.5; IL12.1.1↓; IL12.1↓; IL12.2↓

Unknown$_3$	IL3.3/3.4; IL3.4; IL3.4/IL3.5
Unknown waxes	
Unknown$_5$	IL2.2/IL2.3; IL2.3/IL2.5
Cutin Monomer	
Identified cutin monomers	
p-coumaric	IL3.5
16-ω-hydroxy fatty acid $C_{16:0}$	IL1.3
16-ω-hydroxy fatty acid $C_{16:1}$	IL3.5; IL6.3
9,16-dihydroxy $C_{16:0}$	IL1.1
α,β-dicarboxilic acid $C_{16:0}$	IL1.2; IL3.2; IL11.2
(9/10)-hydroxy α,β-dicarboxilic acid $C_{16:0}$	IL11.2
9,10-dihydroxy fatty acid $C_{18:1}$	IL8.3; IL8.3.1
9,10-epoxy 18-ω-hydroxy fatty acid $C_{18:0}$	IL8.3; IL8.3.1
9,10-epoxy 18-ω-hydroxy fatty acid $C_{18:1}$	IL8.3; IL8.3.1
9,10,18-trihydroxy fatty acid $C_{18:1}$	IL8.3; IL8.3.1
Unidentified cutin monomer	
Unknown$_2$	IL1.4.18
Unknown$_3$	IL11.2
Unknown$_4$	IL3.5
Unknown$_7$	IL3.5
Unknown$_{10}$	IL11.2
Isolated Cuticle Membrane Weight	
ICMW	IL1.1.3; IL2.6; IL2.6.5; IL3.4; IL3.5; IL4.1; IL4.2; IL4.3.2; IL5.1; IL6.1; IL7.4.1↓; IL8.1.1; IL8.3; IL10.1.1; IL11.2; IL12.3.1

This suggested that both genes could be involved in the biosynthesis of the amyrins. A similar conclusion was proposed for LA3917 in a study conducted in the *S. habrochaites* x *S. lycopersicum* cv. MoneyMaker IL population (Yeats et al., 2012a).

On the other hand, we found the three amyrins, including β-amyrin, as differentially accumulated in the overlapped lines IL3.4 and IL3.5 (**Figure II.5B**). The best candidates for the effect on these triterpenoids could be found in the overlapped region between IL3.4 and IL3.5 (11cM containing 112 genes). This region appears to have no *triterpenoid synthase* gene but a candidate *terpenoid cyclase* (*Solyc03g118570*) which activity could be acting upstream in the triterpenoid biosynthesis (Halá et al., 2010) (**Figure II.5C**). However, the gene expression level for this gene did not show differences between IL3.4, IL3.5 and M82 parental line (**Figure II.S3B**; RNA-Seq information from the SGN dataset). Additionally, several VLCFAs mRNAs were identified as differentially over-accumulating that mapped to the no-overlapped region of IL3.4 (7 cM), including fatty acids C_{24}, C_{26} and C_{30}, *n*-aldehyde C_{24},*n*-alkanes C_{29}, C_{30}, C_{31}, C_{32}, C_{33} and *n*-alcohol C_{32} (**Figure II.5B**). Further, IL3.4 was partially overlapped with IL3.3 (38 cM) and with IL3.5 (11 cM). The effects observed in the no-overlapped region of IL3.4 were also observed in IL3.3/IL3.4 and IL3.4/IL3.5 overlapped regions. However, these effects were not observed in the no-overlapped regions of IL3.3 nor IL3.5. Thus, the IL3.4 was considered as a QTL for very-long-chain fatty acid biosynthesis or regulation: *vlfa3.4*. We searched for candidate genes controlling the increase in VLCFAs levels in the no-overlapped specific region of IL3.4 (7cM) (**Figure II.5A**). We found, out of the 112 genes localized in such region, two transcription factors whose Arabidopsis orthologs have been described as cuticular wax regulators: *MYB96/MYB30* (*Solyc03117800*; Seo et al., 2011 / Raffaelle et al., 2008; Bolger et al., 2014) and *SHINE1/WAX INDUCER1* (*SHN1/WIN1*) (*Solyc03g116610*; Aharoni et al., 2004; Kannangara et al., 2007). In Arabidopsis, MYB96 regulates wax

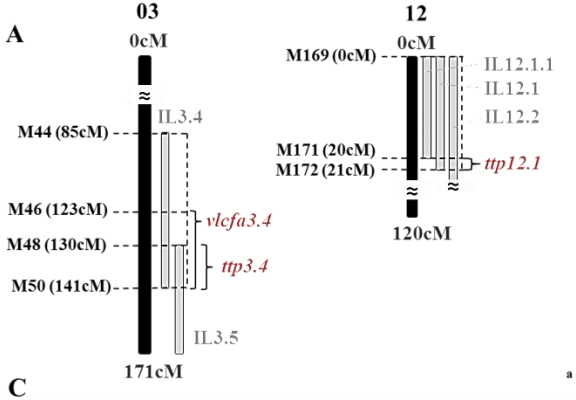
biosynthesis in leaves under drought stress conditions as response to ABA signaling (Seo et al., 2011) and MYB30 regulates very-long-chain fatty acid biosynthesis in response to pathogen infection (Raffaelle et al., 2008). MYB96 mostly regulates waxes through the decarbonylation pathway (alkanes and aldehydes) acting over a large set of biosynthetic genes (Seo et al., 2011); MYB30 controls the production of VLCFAs and *n*-alkanes acting on the fatty acid elongation complex. Both MYB96 and MYB30 transcription factors were suggested to regulate *CER3/WAX2*, involved in the formation of *n*-alkanes (Rowland et al., 2007). Interestingly, we found a tomato ortholog of *CER3/WAX2* (*Solyc03g117800*) in the IL3.4/IL3.5 overlapped region (**Figure II.5C**). Furthermore, *CER3/WAX2* is also a likely target for SHN1/WIN1 in Arabidopsis (Aharoni et al., 2004). However, SHN1/WIN1 is a major cutin biosynthetic regulator, and only a gain-of-function produces an increment in wax content, mainly in alkanes and aldehydes, as secondary effect (Aharoni et al., 2004). While *MYB96/MYB30* or *SHN1/WIN1* tomato orthologs are good candidates to be associated with the *vlcfa3.4* QTL this might not be the case as: (i) no evidence for differential accumulation of cutin monomers was found in IL3.4 which could be associated with the *SHN1/WIN1* transcription factor; (ii) no evidence for differential accumulation of several other VLCFAs was found in IL3.4 which would suggest *MYB30* is the gene controlling the phenotype; and (iii) no wax-esters were differentially accumulated in IL3.4 in contrast to what was described for the mutant *myb96* in Arabidopsis leaves (Seo et al., 2006). When we investigate the changes in gene expression levels for these two candidate genes in the RNA-Seq dataset available in SGN website, we found that *MYB96/MYB30* was up-regulated in IL3.4 and IL3.5 (there were not data available for IL3.3) while *SHN1/WIN1*, which was not expressed in M82 nor in IL3.5 but was expressed in IL3.4 (**Figure II.S3B**). The gene expression levels of *CER3/WAX2* were similar for IL3.4, IL3.5 and parental cv. M82 (**Figure II.S3B**). These results suggested that *SHN1/WIN1* could be responsible for the wax variation in the QTL *vlcfa3.4* as it is differentially

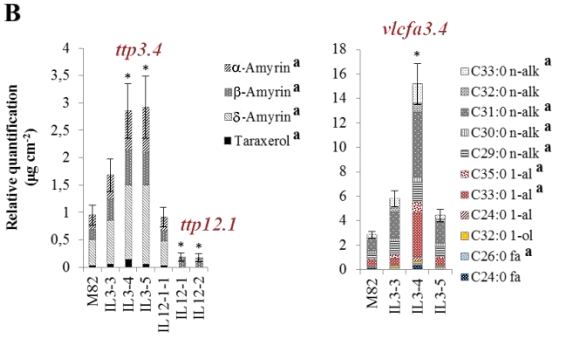
expressed in IL3.4 but did not in IL3.5. Interestingly, this IL3.4 had not any cutin monomer differentially accumulated, suggesting a putative new role for *SHN1/WIN1* in tomato fruit cuticular wax biosynthesis regulation rather than the already known role in the cutin biosynthesis regulation. To validate the changes in *ttp3.4* and *vlcfa3.4* QTLs, IL3.4 was analyzed and evaluated in a second year experiment. As in the first season, the three amyrins, *n*-alkanes C_{29}, C_{30}, C_{31}, C_{32}, C_{33}, *n*-aldehyde C_{24} and fatty acid C_{26} were accumulated to higher levels as compared with the M82 parental line (**Figure II.5B**), and therefore validated both wax-related QTLs.

Cutin QTL mapping revealed a limited peppered control of cutin monomer composition across the tomato genome

We found 15 (~58%) out of 26 cutin monomers monitored, differentially accumulated in 10 (~14%) of the 73 ILs analyzed (**Table II.1**). As in the case of cuticular waxes, mostly single cutin monomers per IL were found as differentially accumulated in the 10 ILs. To our surprise, we did not find lower levels of cutin monomers with comparison to the M82 parental line in any of these 10 ILs. Furthermore, we did not find in any of the 10 ILs significant changes in levels of the major tomato fruit cutin monomer (10,16-ω-dihydroxy fatty acid $C_{16:0}$). This result is consistent with the invariant level of this compound reported for *S. pennellii* as compared to the M82 cultivar (Yeats et al, 2012a). On the other hand, the cutin monomers 9,10-dihydroxy fatty acid $C_{18:1}$, 9,10-epoxy-18-ω-hydroxy fatty acid $C_{18:0}$, 9,10-epoxy-18-ω-hydroxy fatty acid $C_{18:1}$ and 9,10,18-trihydroxy fatty acid $C_{18:1}$ were found differentially accumulated in IL8.3 and IL8.3.1 (**Figure II.6B**).

Figure II.5 Identification of three cuticular wax-related QTLs in IL3.4 and IL12.1. ***A****. Linkage map between candidate genes and cuticular wax QTLs.* ***B****. Differentially accumulated cuticular waxes for ttp3.4, ttp12.1 and vlcfa3.4 QTLs.* ***C****. Candidate genes proposed as associated with ttp3.4, ttp12.1 and vlcfa3.4 QTLs. [next page].*





C

[a] Cuticular waxes differentially accumulated during two consecutive years for IL3.4

QTL[a]	IL	Chr	Marker interval (cM)[b]	Phenotype	Tomato gene name[c]	Arabidopsis protein orthologue[d]	Protein similarity	Reference
ttp3.4	IL3.4/IL3.5	3	M48-M50 (11)	Increased levels of taraxerol, δ-, β- and α-amyrins	Geranylgeranyl transferase II/ Terpenoid cyclase (Solyc03g118570)	At3g12070	80%	Hală et al., 2010
vlcfa3.4	IL3.4	3	M46-M48 (7)	Increased levels of VLCFAs (>C_{20} carbons): mainly *n*-aldehydes and *n*-alkanes, but also fatty acids and *n*-alcohols.	MYB transcription factor (Solyc03g116100)	MYB96 - At5g62470/ MYB30 - At3g28910*	51% 50%	Seo et al., 2011/ Raffaele et al., 2008
	IL3.4	3	M46-M48 (7)		SHN1/WIN1 (Solyc03g116610)	At1g15360	65%	Aharoni et al., 2004
	IL3.4/IL3.5	3	M48-M50 (11)		CER3/WAX2 (Solyc03g117800)	At5g57800	65%	Rowland et al., 2007
ttp12.1	IL12.1/IL12.2	12	M171-M172 (1)	Reduced levels of taraxerol, δ- and α-amyrins	TTS1 (Solyc12g006530) TTS2 (Solyc12g006520) TTS-*like* (Solyc12g006510)	*genes described in tomato*		Wang et al., 2011; Yeats et al., 2012

[a] QTLs are named by the abbreviation of the specific wax class, triterpenoid (*ttp*) or very-long-chain fatty acid (*vlcfa*), plus the IL number.

[b] Genetic markers used in the table are described in Figure S1.

[c] Gene annotations were obtained from SL2.40 database [SGN, http://www.solgenomics.net]

[d] Blastp for Protein ITAG realese 2.40 database [SGN, http://www.solgenomics.net]

* Annotations obtained from Bolger et al., (2014) when they were different from those obtained from SL2.40 database [SGN, http://www.solgenomics.net].

In a previous study, Yeats et al., (2012a) reported that fruit of *S. pennellii* contained high amounts of 9,10,18-trihydroxy fatty acid $C_{18:0}$ compared to the cultivated M82, and similar levels those found in both the *S. chmielewskii* and *S. habrochaites* wild species. Here, IL8.3 and IL8.3.1 contained the highest amount of epoxy-hydroxystearic acids and trihydroxystearic acid among fruit of the entire IL population. The most likely biosynthetic pathway for these epoxy-hydroxylated fatty acids is represented in **Figure II.6C** (adapted from Pollard et al., 2008). Thus, these two ILs could provide insight to the biosynthesis or possibly the regulation of this pathway. To contribute to that, we searched for candidate genes located in IL8.3 and IL8.3.1. Thus, two regions were investigated in this two ILs: (i) IL8.3.1 which is fully overlapped into IL8.3 (3 cM containing 101 genes); and (ii) the specific region of IL8.3 (17cM containing 258 genes) which is not overlapped with IL8.3.1 neither with IL8.2 (17cM) (**Figure II.S1**). The search for candidate genes in these two regions (i.e., 17 cM and 3 cM) resulted in five genes whose orthologous in Arabidopsis have already been described as involved either in cuticle assembly or in fatty acid biosynthesis. Four candidate genes were located in the specific 17 cM of the IL8.3 region and one in the 3 cM of the overlapped IL8.3/IL8.3.1 region (**Figure II.6D**): (i) *HOTHEAD-like* (*HTH-like*, *Solyc08g080190*), a choline dehydrogenase involved in cutin monomer biosynthesis (Kurdyukov et al., 2006a); (ii) *CYP86A8* (*LCR*, *Solyc08g081220*) (Wellensen et al., 2001) also similar to *CYP86A7* (Bolger et al., 2014), a ω-hydroxylase without activity on oxygenated epoxy fatty acids (Rupashingue et al., 2007); (iii) *LACS4-like* (*Solyc08g082280*), a protein putatively involved in fatty acid and VLCFAs biosynthesis, as well as in defense response to insect and wounding (Sockey et al 2002, Jessen et al., 2014). It is also similar to *CER8/LACS1* (Bolger et al., 2014), involved not only in cuticular wax and cutin biosynthesis (Lü et al., 2009) but also in seed oil biosynthesis in Arabidopsis (Zhao et al., 2010); (iv) *GPAT9-like* (*Solyc08g082340*), a glycerol-3-phosphate transferase likely involved in storage oil, diacylglycerol and triacylglycerol biosynthesis (Gidda et al.,

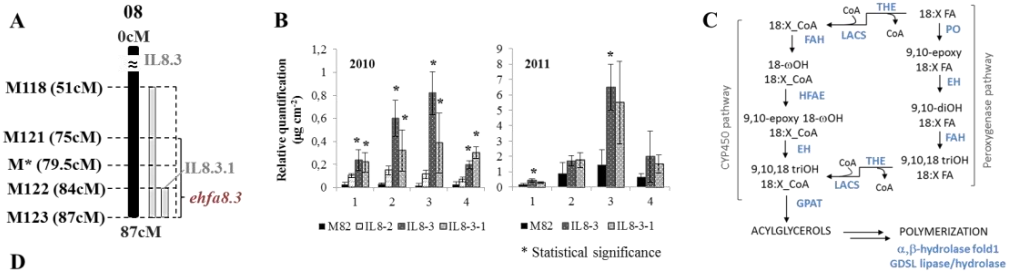
2009); and (v) *BDG1/BDG3* (*Solyc08g083190*) putatively encoding an α,β-hydrolase fold-1 family protein involved in cutin polymer biosynthesis (Kurdyukov et al., 2006b; Bolger et al., 2014). Furthermore, we found four more genes with hydrolase activity (see **Figure II.6D**), i.e. two encoding a protein with a GDSL-motif, *Solyc08g082490* and *Solyc08g082500* (both of them in the overlapped IL8.3/IL8.3.1 region, see **Figure II.6D**), and two additional genes encoding a putative α,β-hydrolase, *Solyc08g082150* (in the specific 17cM region into IL8.3) and *Solyc08g083140* (in the overlapped IL8.3/IL8.3.1 region, see **Figure II.6D**). None of these four genes has been previously characterized in the tomato fruit. We searched for the RNA-Seq data for these nine candidate genes (SGN website), and we found that *CYP86A8/LCR* (*Solyc08g081220*) and *α,β-hydrolase* (*Solyc08g082150*) were not expressed in fruits of IL8.3, IL8.3.1 nor in parental cv. M82. On the other hand the two genes encoding GDSL-motif proteins were expressed similarly in both IL8.3 and IL8.3.1 and parental line M82. In contrast the *HTH-like* (*Solyc08g080190*) was highly expressed in IL8.3 but not in IL8.3.1 nor the parental cv. M82 (**Figure II.S3C**) and GPAT9-like (*Solyc08g082340*) showed a slightly reduced expression level in IL8.3 with respect to both IL8.3.1 and the parental line M82 (**Figure II.S3C**). The *α,β-hydrolase* (*Solyc08g083140*) showed higher levels of expression in IL8.3.1 but not in IL8.3 and the parental cv. M82 (**Figure II.2C**). Finally, the two remaining candidate genes, *LACS4-like* (*Solyc08g082280*) and *BDG1/BDG3* (*Solyc08g083190*) showed higher expression levels in IL8.3.1 than in IL8.3, and both ILs showed also higher expression levels than the parental M82 (**Figure II.S3C**). All these results suggested that *HTH-like*, *LACS4-like*, *GPAT9-like*, *BDG1/BDG3* and one of the *α,β-hydrolase* (*Solyc08g083140*) could be involved in the genetic control of the metabolite variation observed in IL8.3 and IL8.3.1. In order to validate the effect observed in these two ILs, we analyzed them in a second year of study (**Figure II.6B**). The second season samples were more variable and statistical significant differences were only found for 9,10-epoxy-18-ω-hydroxy $C_{18:0}$ and 9,10-dihydroxy $C_{18:1}$ fatty

acids in IL8.3 (see asterisks in **Figure II.6B**). Changes in IL8.3.1 were not statistically significant even though this IL showed higher amounts than parental cv. M82 for the four monomers. Then, these results reinforce that the candidate genes located in IL8.3 but possibly not those in IL8.3/IL8.3.1 overlapped region, (i.e., *HTH-like*, *LACS4-like* and *GPAT9-like*) are more stable/ reliable as candidates for the metabolite phenotype observed in IL8.3. Therefore, we propose IL8.3 as a candidate QTL for epoxy-hydroxylated fatty acids (e*hfa8.3*QTL), but we cannot exclude the overlapped IL8.3/IL8.3.1 region as containing other putative polyhydroxylated fatty acid-related QTLs.

Isolated Cuticular Membrane Weight QTL mapping revealed a putative new secondary cell wall QTL

Although no statistically significant differences were detected in ICMW between M82 and any of the 73 ILs using a restrictive *p*-value (α = 0.05/73bins), significant differences for ICMW in 16 ILs (~22%) were found when using a less restrictive *p*-values and by *t*-test analysis (see **Table II.1**), as we observed previously in the phenotyping (**Figure II.4A**).

Figure II.6 Identification of a cutin monomer-related QTL in IL8.3. A*. Linkage map between candidate genes and cutin monomers QTLs.* ***B****. Differentially accumulated cutin monomers in ehfa8.3 QTLs.* ***C****. Oxygenated and polyhydroxylated fatty acid most likely biosynthetic pathway (adapted from Pollard et al., 2008).* ***D****. Candidate genes proposed as controlling ehfa8.3 QTLs.* ***1****, 9,10-dihydroxy* $C_{18:1}$ *fatty acid;* ***2****, 9,10-epoxy 18-ω-hydroxy* $C_{18:1}$ *fatty acid;* ***3****, 9,10-epoxy 18-ω-hydroxy* $C_{18:0}$ *fatty acid; and* ***4****, 9,10,18-trihydroxy* $C_{18:1}$ *fatty acid. LACS, Long-chain Acyl-CoA Synthetase; THE, Thioesterase; FAH, Fatty Acyl (ω-1) hydroxylase; PO, Peroxygenase; HFAE, ω-hydroxy fatty acyl epoxygenase; EH, epoxyde hydrolase; GPAT, Glycerol-3-Phosphate Acyl Transferase.[next page].*



D

QTL[a]	IL	Chr	Marker interval (cM)[b]	Phenotype	Tomato gene name[c]	Arabidopsis protein orthologue[d]	Protein similarity	Reference
ehfa8.3	IL8.3	8	M121-M122 (8)	Increased levels of 9,10-epoxy 18-ω-hydroxy $C_{18:X}$ fatty acids and 9,10,18-trihydroxy $C_{18:1}$ fatty acid.	Choline dehydrogenase (Solyc08g080190)	**HOTHEAD** - At1g72970*	57%	Kurdyukov et al., 2006a
					CYP450 (Solyc08g081220)	**CYP86A8/LCR** - At2g45970	72%	Wellesen et al., 2001
						CYP86A7 - At1g72970*	72%	Rupasinghe et al., 2007
					α,β-hydrolase (Solyc08g082150)	α,β-hydrolase (partial) - At5g41850	77%	*Not characterized*
					LACS (Solyc08g082280)	**LACS4** - At2g45970	76%	Shockey et al., 2002; Jessen et al., 2014
						CER8/LACS1 - At2g47240	50%	Lü et al., 2009; Zhao et al., 2010
					GPAT3 (Solyc08g082340)	**GPAT9** - At5g60620	86%	Gidda et al., 2009
	IL8.3/IL8.3.1	8	M122-M123 (3)		GDSL-esterase/lipase (Solyc08g082490)	GDSL-lipase/SGNH-esterase/hydrolase - At4g10950	59%	*Not characterized*
					GDSL-esterase/lipase (Solyc08g082500)	GDSL-lipase/SGNH-esterase/hydrolase - At4g10950	53%	*Not characterized*
					α,β-hydrolase fold1 (Solyc08g083140)	esterase/lipase/thioesterase (hydrolase activity) - At3g50790	61%	*Not characterized*
					α,β-hydrolase fold1 (Solyc08g083190)	**BDG1** - At1g64670	59%	Kurdyukov et al., 2006b
						BDG3 - At4g24140*	57%	Kurdyukov et al., 2006b

[a] QTLs are named by the abbreviation of the cutin monomer class, epoxyhydroxylated fatty acids (*ehfa*), plus the IL number

[b] Genetic markers used in the table are described in Figure S1

[c] Gene annotations were obtained from SL2.40 database [SGN, http://www.solgenomics.net]

[d] Blastp for Protein ITAG realese 2.40 database [SGN, http://www.solgenomics.net]

* Annotations obtained from Bolger et al., (2014) when they were different from those obtained from SL2.40 database [SGN, http://www.solgenomics.net]

Such significant ICMW variation included ILs with heavier fruit cuticles than the parental cv. M82. However, IL7.4.1 was the only IL exhibiting significantly reduced ICMW, almost half of the parental weight. To our surprise, IL7.4.1 did not show reduction in any of the cuticular lipids measured here. In a second year analysis, we observed the same ICMW phenotype in IL7.4.1 even though several cuticle metabolites were differentially accumulated (*data not shown*). These results suggest that the cuticular membrane weight in IL7.4.1 could be independent from cuticle composition. Therefore, we propose this region as a QTL for ICMW (*icmw7.4.1*) (**Figure II.4B**). Additionally, microscopy imaging showed that cuticle thickness of IL7.4.1 was half of the parental line M82 (**Figure II.4C**). Further, both ICM weight and thickness were correlated (R^2= 0.92826). This phenotype was already described for *S. pennellii* wild species (Yeats et al., 2012a), in which the ICM thickness was almost half of the parental line M82. To further elucidate the possible causes of the IL7.4.1 phenotype, we analyzed the ICM structure using transmission electron microscopy (TEM). TEM analysis revealed that the cutinized secondary cell wall in IL7.4.1 was missing and a primary cell wall was observed instead (no differences in the proper cuticle between IL7.4.1 and M82 were found; **Figure II.4D**). This observation suggests that secondary cell wall biosynthesis could be altered in IL7.4.1. However no previously reported cell wall-related QTL or texture QTL had been associated to this region or line (Graham Seymour, personal communication). We searched for cell wall-related genes into IL7.4.1 excluding the region overlapping with IL7.5 (7 cM), covering a total of 31cM and containing 1110 genes. IL7.4.1 was fully overlapped with IL7.4, however the phenotype observed for IL7.4.1 was not observed in IL7.4. So, we restricted our search to cell wall-related genes with different expression levels in IL7.4.1 with respect IL7.4 and parental cv. M82. We found 17 genes showing different gene expression levels in IL7.4.1 than parental line M82 (**Figure II.4E**, **Figure II.S3D**) and out of them, 11 genes also showed a different expression level in IL7.4.1 respect IL7.4 (**Figure II.S3D**). These

genes included four *chitinases* (*Solyc07g005100*; *Solyc07g009500*; *Solyc07g009510*; and *Solyc07g009530*) with reduced levels of expression in IL7.4.1 and IL7.4 respect parental cv. M82. In addition to the chitinases, six other cell wall-related genes showed higher expression levels in IL7.4.1 than in IL7.4.1 and M82: (i) *polygalacturonase / pectin lyase* (*Solyc07g015870*); (ii) *Glucan endo-1 3-β-glucosidase 5* (*Solyc07g017730*); (iii) *Glucan endo-1 3-β-glucosidase* (*Solyc07g005330*); (iv*) α,β-hydrolase domain-containing protein 5* (*Solyc07g032180*); (v) *cellulose synthase* (*Solyc07g043390*), a tomato ortholog of *AtCSLG1* ; and (vi) bHLH051 transcription factor (*Solyc07g039570*; Sun et al., 2015). Finally, seven genes showed lower expression levels in IL7.4.1 than in IL7.4 and parental cv. M82: (i) *SlXTH2* (*Solyc07g009380*), XYLOGLUCAN ENDOTRANSGLUCOSYLASE / HYDROLASE 2 (Catalá et al., 2001), a tomato ortholog of *AtXTR3* (*XYLOGLUCAN ENDOTRANSGLYCOSYLASE 3*); (ii) *pectinesterase* (*Solyc07g017600*), a tomato ortholog of *AtPME44* (an enzyme inhibitor/pectinesterase in Arabidopsis); (iii) *pectinesterase inhibitor* (*Solyc07g042390*); (iv) *Glucan endo-1 3-β-glucosidase 3* (*Solyc07g047710*); (v) *β-galactosidase* (*Solyc07g038120*), a tomato ortholog of *AtBGAL2* (*BETA-GALACTOSIDASE 2*); and (vi) bHLH140 transcription factor (*Solyc07g020960*; Sun et al., 2015); and (vii) unknown transcription factor (*Solyc07g021550*). Therefore, we have been able to identify a putative new QTL for secondary cell wall formation (*icmw7.4.1*) and proposed 11 candidate genes to be involved in the control of the phenotype observed in IL7.4.1. Future cell wall composition analyses could facilitate the identification of the best candidate genes controlling the phenotype observed in this *icmw7.4.1* QTL.

II.4 CONCLUDING REMARKS

In conclusion, our screening of the *S. pennellii* x *S. lycopersicum* cv. M82 IL population for cuticular lipids composition provided 37 QTLs in tomato fruit. Our work also provided a more detailed analysis for the five major QTLs. Thus, two QTLs were related with cuticular wax composition: *ttp3.4*, putatively involved in the biosynthesis of amyrins, and *vlcfa3.4*, putatively involved in the regulation of very-long-chain fatty acids and *n*-alkanes biosynthesis. A third QTL (*ehfa8.3*) related with cutin monomer composition, is proposed to be involved in epoxy-hydroxy fatty acid biosynthesis. A fourth QTL (*icmw7.4.1*) related to cuticle membrane weight is proposed to be involved in the biosynthesis of the secondary cell wall. Finally, we also identify a fifth QTL, the *ttp12.1*, whose phenotype and candidate genes involved in it were also described in a different IL population. Further, we propose 29 candidate genes for those five major QTLs. Thanks to the RNA-Seq data available in the tomato SGN website we could identify 19 of the 29 genes localized in their respective ILs genes that presented changes in the gene expression levels consistent with the phenotype observed in each case. Moreover, we proposed a new function for SHN1/WIN1 in the biosynthesis of cuticular waxes in tomato fruit. Therefore, we are providing new information for the study of the genetic control of the cuticular lipids biosynthesis and / or regulation in tomato fruit. This information could be on interest to be expoloited in tomato breeding programs.

II. 4 ANNEXES II

Additional Figures

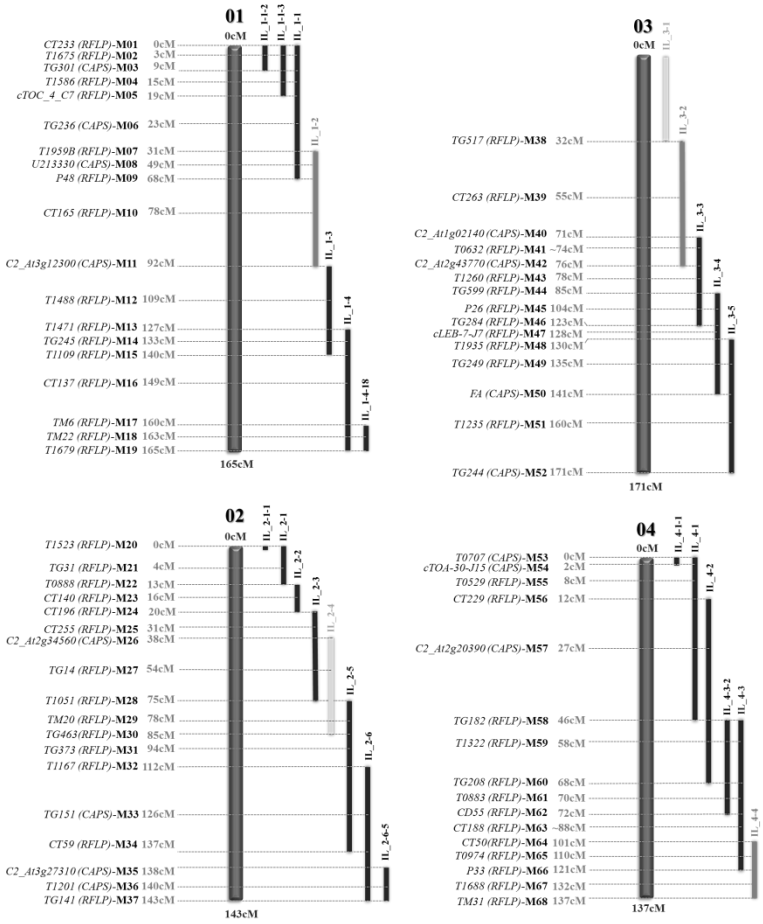


Figure II.S1 Genetic markers and ILs used for the screening of cuticular components. *The schematic representation of the introgression lines are placed to the right of each chromosome and genetic markers to the left. Those ILs analysed for cuticular waxes and cutin monomers are in dark grey and those analysed only for cutin monomers in grey. Missed ILs are in light grey. For genetic markers: the name is in italics, the type into brackets, the code in bold and the location (cM) in grey.*

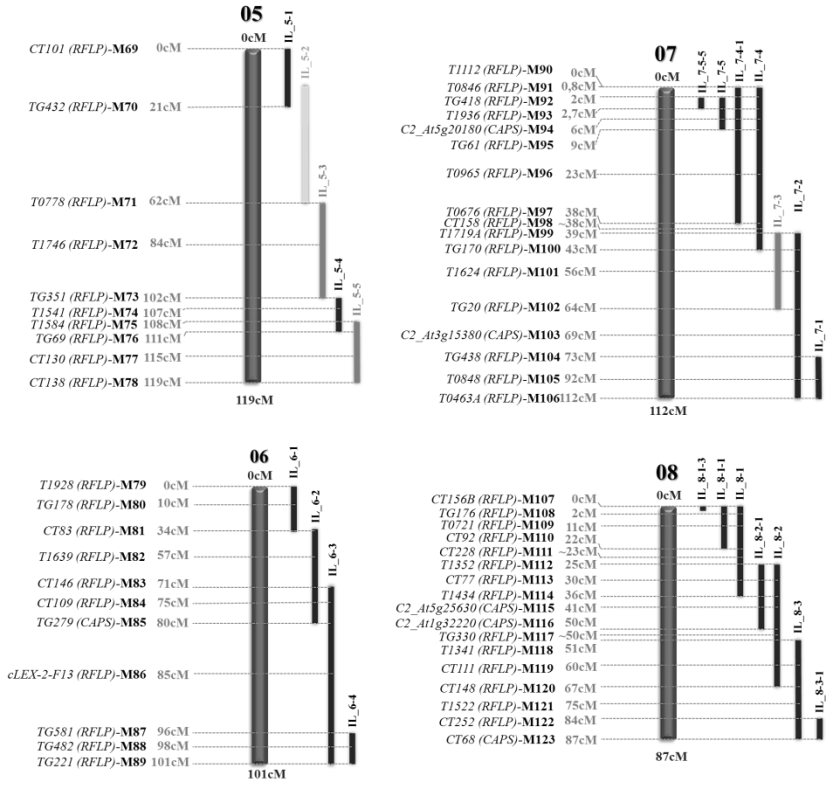


Figure II.S1 *(…continued)*

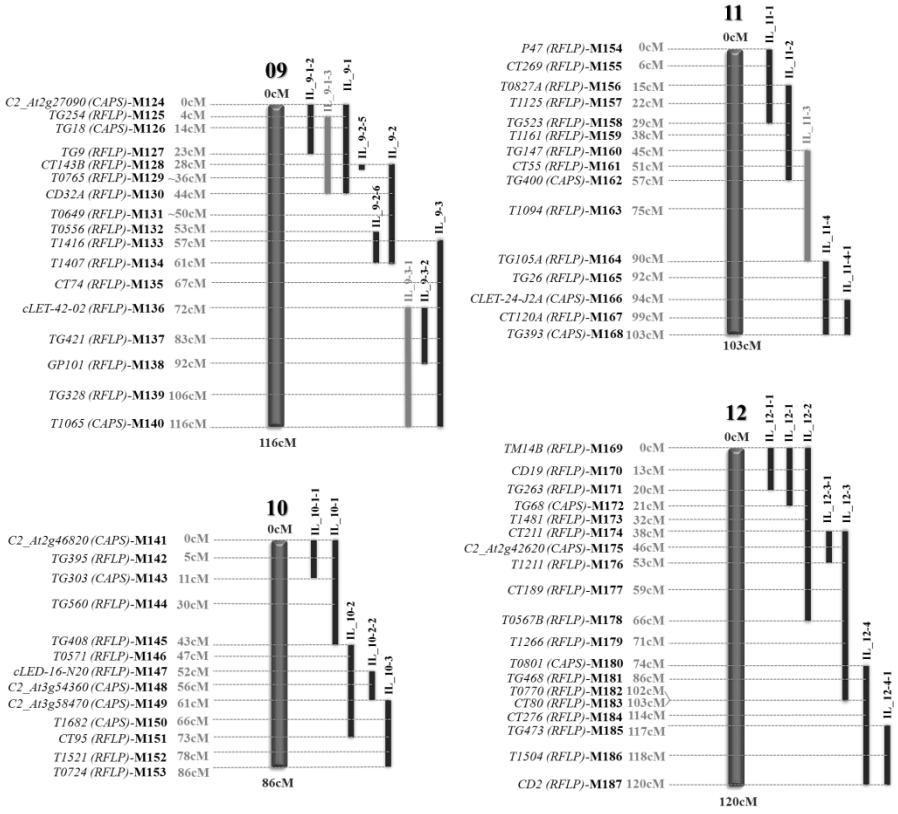


Figure II.S1 *(…continued)*

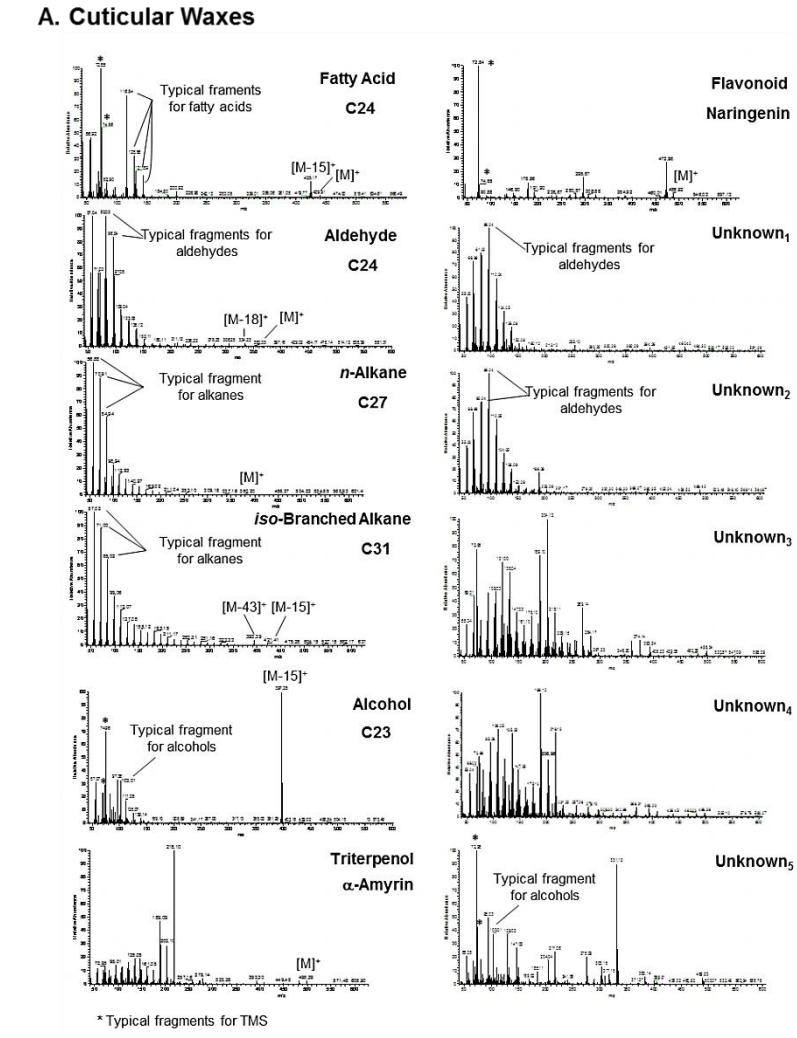


Figure II.S2 Typical mass spectra obtained for the different wax classes and different cutin monomers observed. A. *The typical fragments are shown for a fatty acid, aldehyde, alkane, and alcohol wax classes.* ***B****. Mass spectra for cutin monomers* $C_{16:x}$ *identified.* ***C****. Mass spectra for cutin monomers* $C_{18:X}$ *identified.* ***D****. Mass spectra for unknown cutin monomers detected and quantified herein. Also the fragments m/z 73 and 75 originating from the trimethyl silyl (TMS) group are shown. Molecular mass in the positive ionization mode* $[M]^+$ *and specific m/z fragments (e.g.* $[M-15]^+$ *or* $[M-47]^+$*) are also shown if they were identified in the spectrum.*

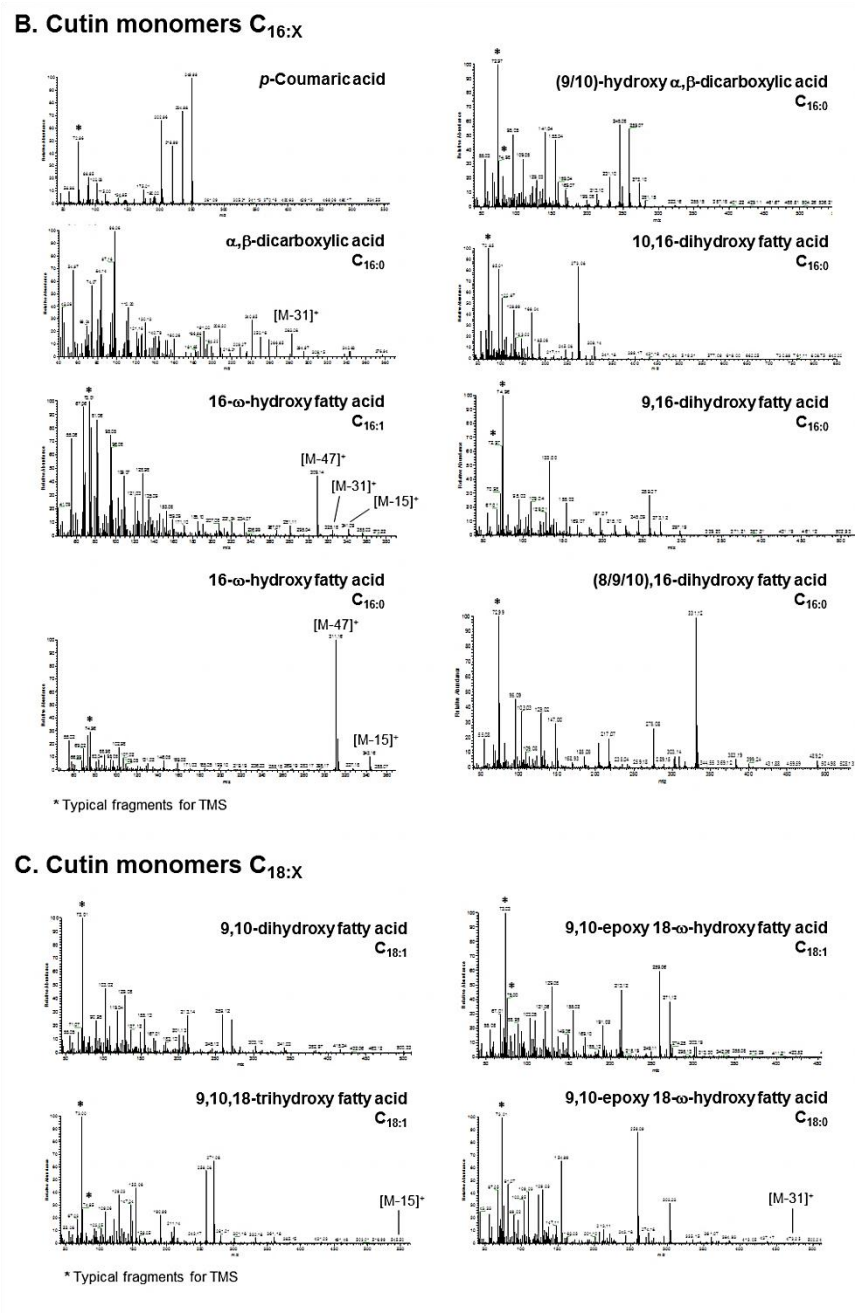


Figure II.S2 *(…continued)*

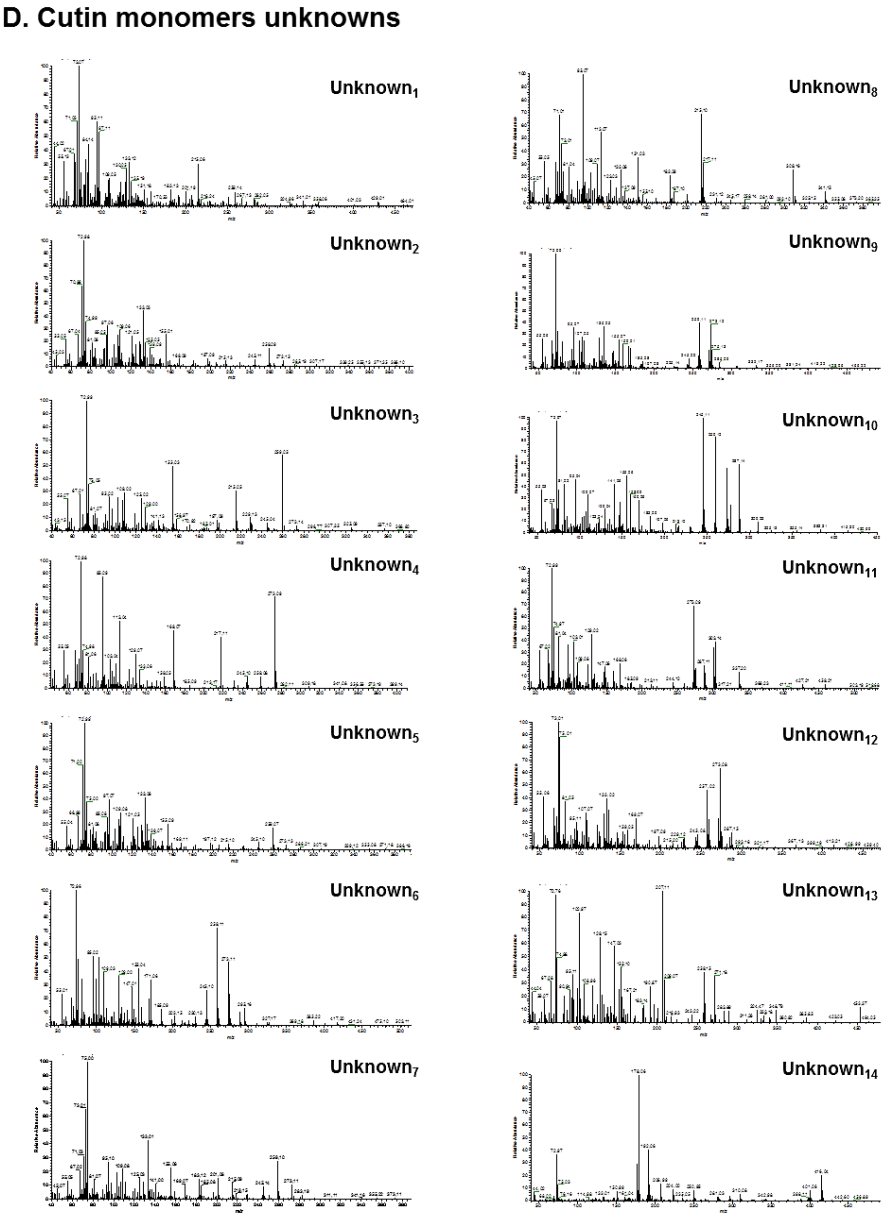


Figure II.S2 *(…continued)*

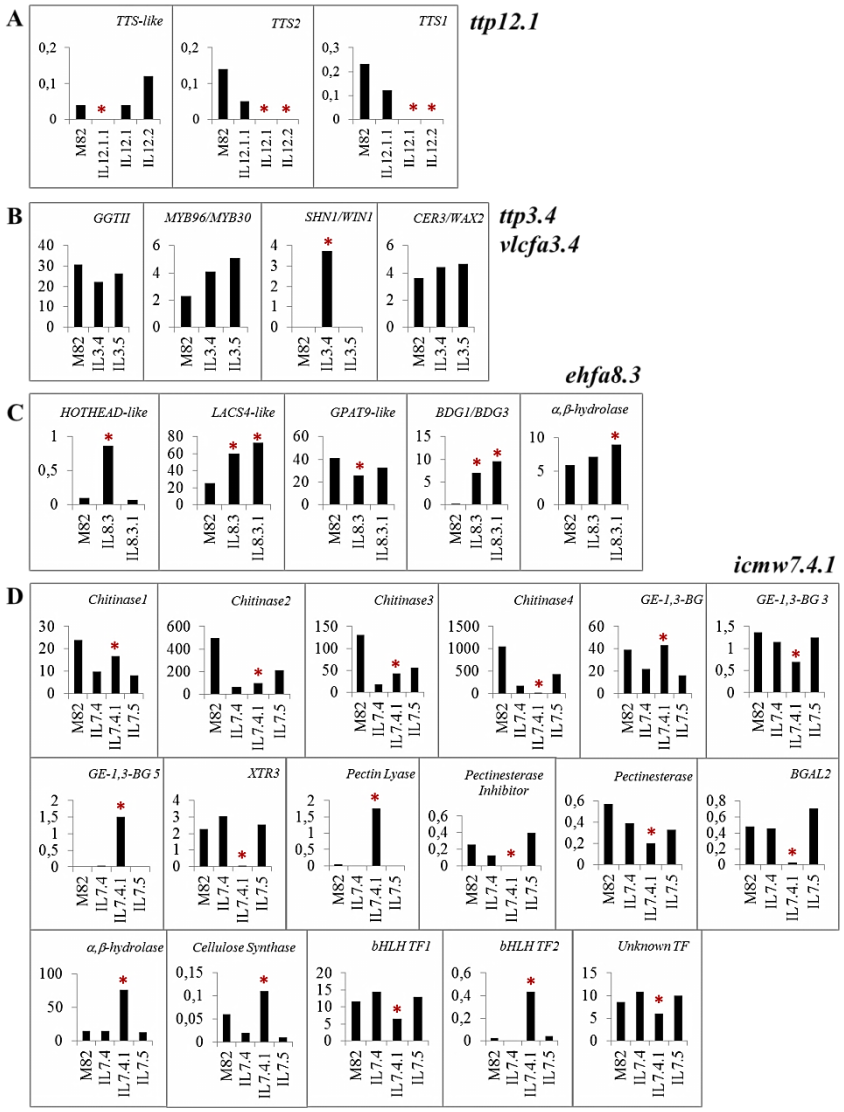


Figure II.S3 Gene expression levels for the 29 candidate genes localized in the four major QTLs (ttp12.1, ttp3.4, vlcfa3.4, ehfa8.3 and icmw7.4.1). *Values showed in the graphs are from the RNA-Seq dataset available in the tomato website SGN (http://www.solgenomics.net). Asterisks represent the IL containing the QTL of interest that show a differential gene expression level respect parental cv. M82.*

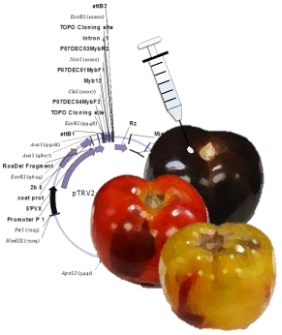
Additional Tables

You can find these additional tables in the excel file "*CHAPTER II_Additional_Tables_II.S1-II.S2*" into the CD attached to this thesis.

Table II.S1 Cuticular wax composition for tomato fruit (Solanum lycopersicum) cuticle for parental cv. M82 and for the Solanum pennellii x S. lycopersicum cv. M82 introgression lines.

Table II.S2 Cutin monomer composition for tomato fruit (Solanum lycopersicum) cuticle for parental cv. M82 and for the Solanum pennellii x S. lycopersicum cv. M82 introgression lines.

CHAPTER III



VIRUS-INDUCED GENE SILENCING, A TOOL TO STUDY FRUIT DEVELOPMENT IN *Solanum lycopersicum*

Josefina-Patricia Fernández-Moreno[1], Diego Orzáez[1] and Antonio Granell[1]

Adapted from the publication by the same authors in Methods in Molecular Biology. Springer Life Editorial. Volume 975, 2013, Chapter 14 (pp183-196).

[1] Fruit Genomics and Biotechnology Laboratory. Instituto de Biología Molecular y Celular de Plantas (CSIC-UPV). Valencia, Spain.

SUMMARY

A visually traceable system for fast analysis of gene functions based on Fruit-VIGS methodology is described. In our system, the anthocyanin accumulation from purple transgenic tomato lines provides the appropriate background for fruit-specific gene silencing. The tomato *Del/Ros1* background ectopically express *Delila* (*Del*) and *Rosea1* (*Ros1*) transgenes under the control of fruit ripening E8 promoter, activating specifically anthocyanin biosynthesis during tomato fruit ripening. The *Virus-Induced Gene Silencing* (VIGS) of *Delila* and *Rosea1* produces a colour change in the silenced area easily identifiable. *Del/Ros1* VIGS is achieved by agroinjection of an infective clone of *Tobacco Rattle Virus* (pTRV1 and pTRV2 binary plasmids) directly into the tomato fruit. The infective clone contains a small fragment of *Del* and *Ros1* coding regions (named DR module). The co-silencing of reporter *Del/Ros1* genes and a gene of interest (GOI) in the same region enables us to identify the precise region where silencing is occurring.

The function of the GOI is established by comparing silenced sectors of fruits where both GOI and reporter DR genes has been silenced with fruits in which only the reporter DR genes has been silenced. The Gateway vector pTRV2_DR_GW was developed to facilitate the cloning of different GOIs together with DR genes. Our tool is particularly useful to study genes involved in metabolic processes during fruit ripening, which by themselves would not produce a visual phenotype.

Key words: *Virus-Induced Gene Silencing*, tomato fruit, agroinjection, *Tobacco Rattle Virus*, Gateway, pTRV2*_Del/Ros1_*GW, co-silencing, anthocyanin, gene function.

III.1 INTRODUCTION

Virus-Induced Gene Silencing (VIGS) is a technique based on RNA-mediated antiviral plant defence mechanism that has been used to analyze gene function in plants (Brigneti et al., 2004; Fu et al., 2005; Lu et al., 2003; Ratcliff et al., 2001; Robertson et al., 2004; Shao et al., 2008). A fragment of a plant gene of interest (GOI) is inserted into the recombinant viral genome used for infection. The specific degradation of endogenous GOI's mRNA is the result of plant antiviral defence and produces the silencing of the endogenous GOI (Burch-Smith et al., 2004). VIGS presents multiple advantages when compared to other loss-of-function techniques (Burch-Smith et al., 2004; Fu et al., 2005; Orzaez et al., 2009a; Unver et al., 2009) and therefore qualifies as an advantageous technique for reverse genetic studies.

Genomics project are generating an overwhelming amount of information, thanks to a more powerful RNA sequencing and array technologies. This is also happening with tomato, which is a particularly important crop, not only because its fruit contributes importantly to the human diet, but also because it is becoming a model crop species. To improve both the nutritional value and organoleptic features of this crop, first it is necessary to understand the genetic basis of the metabolic pathways that operate during fruit ripening processes (Orzaez et al., 2009a). The easy way to obtain genomic sequence information contrasts with our current lack of understanding about the function of many genes in the genome. In many fruit crops, one way to investigate gene function is altering its expression by stable transformation. Different techniques can be used for that objective, but are often cumbersome and lengthy taking several months or years. A rapid and high-throughput method is required, which allows both to analyze the enormous amount of sequence data from genomic projects and to link sequences to gene functions to phenotypes (Orzaez et al., 2009a). VIGS technique can be successfully applied in tomato fruit for that purpose

(Ballester et al., 2010; Fu et al., 2005; Orzaez et al., 2006; 2009a and 2009b). The use of *Agrobacterium tumefaciens* as a vehicle for transfection is the common way to introduce effectively viral-modified vectors for VIGS approaches. We developed a new VIGS methodology in fruits, named '*Fruit Agroinjection*' (**Figure III.2**), which introduces *Agrobacterium* suspension, into tomato fruit tissues through the stylar apex (Orzaez et al., 2006). This method speeds up the experimental procedures and confine the VIGS signal into the fruit, allowing to increase the throughput of VIGS by 'one organ-one biological replicate' approaches (Orzaez et al., 2006).

Different viruses have been used as suitable VIGS vectors (Robertson et al., 2004). *Tobacco Rattle Virus* (TRV) was described as a VIGS vector one decade ago (Ratcliff et al., 2001), and since then it has been one of the most widely used (Brigneti et al., 2004; Shao et al., 2008). TRV-based vectors for VIGS approaches consist in pTRV1 and pTRV2 binary plasmids. GOIs are cloned into pTRV2 plasmid by digestion/ligation cloning or into pTRV2_GW vector by Gateway recombination (Liu et al., 2002). In tomato fruits, TRV-based vectors normally produce partial VIGS penetration and patchy tissue distribution as a result of partial and highly variable silencing from fruit to fruit (Orzaez et al., 2009a). This low efficiency causes serious limitation for VIGS use in the investigation of gene loss-of-function that yields non-visual phenotypes (Orzaez et al., 2009a). An internal reference that monitors the levels of silencing was developed to overcome these limitations and increase the sensitivity of downstream analysis, allowing the dissection of silenced from non-silenced tissues (Orzaez et al., 2009a). In our system, the anthocyanin accumulation in purple transgenic tomato lines provides the appropriate background for fruit-specific gene silencing. These lines were obtained in Dr. Cathie Martin group (Butelli et al., 2008) by ectopically expressing, under the control of tomato *E8* promoter, *Delila* (*Del*) and *Rosea1* (*Ros1*) genes (two transcription factors that activate the anthocyanin branch of flavonoid biosynthesis pathway in

Antirrhinum majus flowers (Kevin et al., 2006)). This resulted in the activation of anthocyanin biosynthesis specifically during tomato fruit ripening (Butelli et al., 2008; Orzaez et al., 2009a). The silencing of both *Del* and *Ros1* genes, using small fragments of their coding regions (named reporter DR genes) delivered by pTRV2_DR expression vector (developed in our laboratory, Orzaez et al., 2006 and 2009a) (**Figure III.1A**), results in the lack of anthocyanin production. As reporter, DR genes silencing involves the blockage of a pathway not normally active in tomato fruit, the lack of anthocyanin accumulation produces silenced red sectors that present similar characteristics in metabolism and development as 'wild type' tomato fruit (Orzaez et al., 2009a). To facilitate the dissection of silenced tissues and to increase the yield of silenced areas for downstream analysis, we transferred the *Del* and *Ros1* transgenes from cherry-type MicroTom (Butelli et al., 2008) to a large globe-type Money-Maker tomato background (**Figure III.3A**) by standard crossing and selection (Orzaez et al., 2009a) (see 2.1.1). The integration of DR-reporter module and GOI in the same viral genome (pTRV2_DR_GOI VIGS vector) is required for an efficient co-silencing of both the reporter module and target gene in the same tissue area (co-silencing in tandem, Orzaez et al., 2009a) (**Figure III.3B** and **III.3C**). Finally, in order to facilitate high-throughput tandem cloning of subsequent GOIs, we modified pTRV2_DR vector into pTRV2_DR_GW vector (**Figure III.1B**) by the introduction of a Gateway recombination cassette (Orzaez et al., 2009a). This system has proved to be particularly useful for the analysis of genes of unknown function involved in different stages of fruit ripening, specially genes associated with different branches of metabolism in fruit (Ballester et al., 2010; Orzaez et al., 2009a).

In the present chapter we describe the methodology of Fruit-VIGS based on anthocyanin accumulation in tomato fruit, as a tool for studying gene function during ripening stages especially useful as when it is related with quantitative characters, such as secondary metabolites in tomato fruit.

Details on plant cultivation and maintenance are explained. Some recommendations on silenced areas harvesting and their analysis are provided too. At the end of the chapter, we present two examples of the efficiency of the tool performed by us, and some future perspectives for using this tool with different reporter genes and different promoters aimed to study genes involved in different developmental stages of tomato fruit.

III.2 MATERIALS

2.1 Plant material

1. Globe-type purple tomatoes (*Solanum lycopersicum*) were obtained by crossing *Del/Ros1* MicroTom N line (T2 homozygous generation from Micro Tom plants transformed with *Delila* and *Rosea1* cDNAs under the control of the E8 ripening-specific promoter; (Butelli et al., 2008)) with wild-type Money Maker plants (Orzaez et al., 2009a). Segregating sibling lines were selfed and selected through to the F7 generation. Selection was based on globe-type fruit, smooth leaves, indeterminate growth, and best fruit VIGS response (Orzaez et al., 2009a).

2. Plants were grown in a greenhouse supplemented with artificial light from mercury vapour lamps (OSRAM) of 400ω (PHILILPS HDK) 400HPI®N (96μmoles $m^{-2}s^{-2}$ (Marti et al., 2007)).

3. Plants were irrigated four times per day with a HOAGLAND Nº1 nutritive solution supplemented with oligoelements by automatic dipping irrigation system (Marti et al., 2007).

2.2 Cloning procedures and vectors construction

Gateway technology (*http://www.invitrogen.com*) has been used to generate the different VIGS vectors following the manufacturer's instructions.

1. For amplification: Advantage® 2 DNA Polymerase Mix (Clontech, Mountain View, CA, USA[Note1]), specific forward and reverse primers (10 mM each one), dNTPs mixture (10 mM each dNTP) and sterilized water.

2. For DNA clean-up: QUIAquick® PCR Purification kit (Qiagen, Valencia, CA, USA) and sterilized water; or appropriate percentage of agarose in TAE 1X (40 mM Tris-Acetate and 1 mM Na_2EDTA) and QUIAEXII® Gel Extraction kit (Qiagen, Valencia, CA, USA). Nanodrop spectrophotometer (Nanodrop ND-100 Spectrophotometer. Thermo Fisher Scientific Inc., USA). Follow the manufacturer's instructions in each case.

3. For cloning in pDONOR to generate a pENTR clone: pCR®8/GW/TOPO® TA Cloning® Kit (Invitrogen, Carlsbad, CA, USA) and freshly purify PCR. Follow the manufacturer's instructions.

4. For final generation of a VIGS Vector (Expression clones): TRV-based silencing vectors pTRV1 and pTRV2 were provided by Prof. Dinesh Kumar (Liu et al., 2002); pTRV2_DR and pTRV2_DR_GW VIGS vectors were generated in our group (Orzaez et al., 2009a) (**Figure III.1**). Gateway® LR Clonase™ II Eznyme Mix (Invitrogen, Carlsbad, CA, USA).

5. For *E. coli* transformation: use One Shot®TOP10 or One Shot®Mach1™T1R chemically competent *E. coli* kit (Invitrogen, Carlsbad, CA, USA). Sterile LB liquid medium and solid LB agar plates containing 50 μg/mL Spectinomycin in case of Entry clones (TA-Cloning) or 50 μg/mL Kanamycin for Expression clones (LR reaction). Growing chamber set at 37ºC.

6. For *E. coli* DNA plasmid extraction: use Plasmid Mini Kit I E.Z.N.A. (Omega Bio-tek, Doraville, GA, USA).

7. For *E. coli* colony glycerol stocks: in a sterile Eppendorf tube mix 700 μL of fresh liquid *E. coli* culture and 300 μL of 50% sterile glycerol. Freeze it quickly in liquid N_2 and store it at -80ºC.

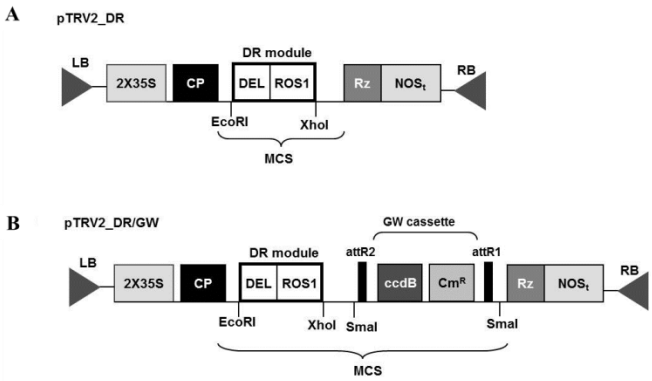


Figure III.1 Map of pTRV2_DR and pTRV2_DR/GW. *TRV cDNA clones are flanked between duplicated CaMV 35S promoter (2X35S) and nopaline synthase terminator (NOSt) (Liu et al., 2002).* ***A****. DR-module (284 bp) was introduced into RNA2 MCS between EcoRI and XhoI restriction enzymes sites to generate pTRV2_DR final expression vector.* ***B****. Gateway (GW) cassette was introduced using SmaI restriction site into pTRV2_DR vector to generate the pTRV2_DR_GW destination vector (Orzaez et al., 2009a). LB and RB, left and right borders of T-DNA; Rz, self-cleaving ribozyme; CP, coat protein; MCS, multiple cloning site; attR1 and attR2, GW recombination sites; DEL, 129 pb from Delila A. majus gene; ROS1, 155 pb from Rosea1 A. majus gene; ccdB, ccdB toxin gene; CmR, chloramphenicol resistance gene.*

2.3 Agrobacterium transformation and agroinjection suspension preparation

1. For *Agrobacterium* transformation: Eppendorf tubes containing 40 µL of electro-competent Agrobacterium C58 cells stored at -80ºC[Note2]. Electroporator (Bio-Rad, gene-pulser 165-2077), 1mm electroporation cuvettes (Bio-Rad Laboratories, CA, USA) and 15 mL plastic tubes containing 250 µL of S.O.C. medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4, 20 mM glucose).

Sterile LB liquid medium and solid LB agar plates containing 50 μg/mL Kanamycin and 50 μg/mL Rifampicin antibiotics. Growing chamber set at 28ºC.

2. For *Agrobacterium* DNA plasmid extraction: use QUIAprep®Miniprep Kit (Qiagen, Valencia, CA, USA).

3. For *Agrobacterium* glycerol stocks: in a sterile Eppendorf tube mix 700 μL of fresh liquid *Agrobacterium* culture and 300 μL of 50% sterile glycerol. Freeze it quickly in liquid N_2 and store it at -80ºC.

4. For *Agrobacterium* C58 cultures and sub-cultures for agroinjection: 15 mL plastic tubes.

5. MES infiltration buffer: 10 mM MES (Sigma-Aldrich, MO, USA[Note3], 10 mM MgCl[Note3], 200 μM acetosyringone (Sigma-Aldrich, MO, USA[Notes3,4]. Rotating and swaying mixer (CAT RM-5). Spectrophotometer (UV/VIS Spectrophotometer SP8001, DINKO) set at a wavelength of 600 nm and transparent plastic cuvettes.

2.4 Fruit agroinjection

1. Tomato *Del/Ros1* fruits at Mature Green (MG) stage, 30-35 days post-anthesis (dpa).

2. Sterile 1mL Plastipak needle syringes (25 GA 5/8 IN, needle: 0.5x16 mm, BD Plastipak™).

2.5 Dissection and collection of silenced sectors

1. For silenced sectors dissection: Glass board to dissect the fruit with sharp knife[Note5].

2. For silenced sectors collection: Screwed cap plastic tubes (25mL) to store the samples at -80ºC. Liquid Nitrogen (N_2) in a suitable container.

3. For silenced sectors crushing: Thermal-cover mortar, metallic little spoon, thermal gloves and protective glasses.

2.6 Evaluation of GOI silencing

1. A suitable RNA extraction method for tomato fruits (Bugos et al., 1995).

2. SuperScript™First-Strand Synthesis System for RT-PCR (Invitrogen, Carlsbad, CA, USA). Follow the manufacturer's instructions.

3. Power SYBR® Green PCR Master Mix and RT-PCR (Applied Biosystems, Madrid, CA, USA) and 7500 Fast Real-Time PCR system (Applied Biosystems, Madrid, CA, USA).

III.3 METHODS

3.1 Plant cultivation and maintenance

1. Estimation of the plant numbers required for the experiment. For that, find out: how many genes you will be silencing (Ratcliff et al., 2001); how many individual silencing constructions will be generated (Lu et al., 2003); and how much material will be necessary for downstream analysis[Note6] (Fu et al., 2005). For each construction we use three plants with twenty fruits each, ten of them for control DR-silencing and the other ten for GOI-DR silencing[Note7].

2. Sowing and seedling. Sow more seeds than plants will be required and keep some extra seedlings to allow selection of best-performing plants[Note8].

3. Pruning and labelling. Remove secondary buds to give structure to the plant (a main axis with first lateral branches only) and to increase reproductive vigour[Note9]. Throughout the flowering period, flower trimming and labelling are required. Keep six flowers per truss and label them by indicating the anthesis date. When small fruits develop, keep six or seven floral trusses and remove the rest. Later, the excess of fruit per truss should be removed. But before that, make sure that at least four of them have reached Immature Green (IMG) stage[Note10]. When the fruits get to Mature Green (MG) stage, select which ones will be used for DR control silencing and which others will be used for GOI-DR co-silencing. Then label them appropriately[Note11] before agroinjection.

3.2 Cloning procedures and vector construction

1. Design the primers for a specific GOI region. Optimal size of GOI ranges between 100 and 500 bp in length[Note12].

2. GOI sequence amplification by PCR reaction using Advantage® 2 DNA Polymerase Mix. PCR conditions will depend on fragment length and primers Tm.

3. Amplified DNA fragment purification. If the PCR yields a single product, purify the PCR reaction using QUIAquick® PCR Purificaton kit, and elute it in 50μL of volume. If the PCR yield gives several products, load the complete PCR reaction on an electrophoresis gel with appropriate percentage of agarose. Separate the PCR product of interest and cut the corresponding band from the gel. Purify the DNA using QUIAEXII® Gel Extraction Kit, and elute it in 30μL of volume. In both cases, quantify the DNA with a Nanodrop spectrophotometer.

4. Ligation to obtain the pCR8_GOI entry clone. Prepare as many 1,5 mL Eppendorf tubes as amplified GOIs. Prepare the TA-Cloning reaction following the pCR®8/GW/TOPO® TA Cloning® Kit manufacturer's instructions. Incubate the ligation 1 hour at room temperature.

5. Transformation of entry clone in E. coli. Take from -80ºC a tube of chemically competent E. coli cells per each Entry clone to be generated (One Shot®TOP10 or One Shot®Mach1™T1R) (see 2.2.5). Thaw competent cells on ice. Add 2 μL of pCR8-GOI ligation reaction to each tube, incubate without shaking and perform transformation following manufacturer's instruction[Note13]. Collect 50 μL from the bacterial culture and spread on a solid LB plate containing 50 μg/mL Spectinomycin[Note14]. Incubate the plates overnight at 37ºC.

6. Validation of entry clone. Pick 4 to 6 colonies into 3mL of liquid LB media with 50 μL/mL Spectinomycin using toothpicks in sterile conditions. Allow them to grow overnight at 37ºC. Isolate the plasmid DNA using the E.Z.N.A.® Plasmid Mini Kit I, and elute it in 50μL of volume. Validate the pCR8_GOI entry clone by restriction analysis[Note15] and by sequencing using M13 forward and reverse primers. Generate a glycerol stock (see 2.2.7) with a positive colony previously validated.

7. Generation of final construct in the expression vector. For each LR reaction, mix in a 1,5 mL sterile Eppendorf: 50-150 ng Entry clone (pCR8-GOI) (1 to 7 μL), 150 ng/μL (1 μL) of pTRV2_DR_GW destination vector and TE buffer pH 8.0 up to 8 μL. Follow the instructions from Gateway® LR Clonase™ II Enzyme Mix manufacturer's instructions.

8. Transformation of final expression vector. Proceed as 3.2.5 but using 1 μL of LR reaction and 50 μg/mL Kanamycin for selective LB plates.

9. Validation of the final expression construct. After overnight incubation, pick 4 to 6 colonies and proceed as 3.2.6. Validate the expression vector by restriction analysis and sequencing[Note16]. Finally, generate a glycerol stock (see 3.2.6).

3.3 Agrobacterium transformation and agroinjection suspension preparation

1. *Agrobacterium* transformation. For each GOI take a sterile Eppendorf containing 40μL of *Agrobacterium* C58 electrocompetent cells and thaw them on ice. Add to each tube 1μL of a positive *E. coli* plasmid miniprep from step 3.2.9. Electroporate the samples at 1,5 V. Add 250 mL of SOC medium and incubate them in 15 mL plastic cap tube for 2 hours shaking (150-200 rpm) at 28ºC. Then, collect them by spin at 13000 rpm in a micro-centrifuge. Remove the supernatant, leaving approximately 100μL. Resuspend the cells and spread them on selective LB plates containing 50μg/mL of both Kanamycin and Rifampicin. Incubate for at least 48 hours at 28ºC.

2. Validation of *Agrobacterium* clones. Pick 4 colonies from selective plates in sterile conditions (as 3.2.6) and allow them grow in 5 mL of liquid LB media containing Kanamycin and Rifampicin (50 μg/mL) at 28ºC for 48 hours. Isolate the plasmid DNA by miniprep and validate by digestion with the suitable restriction enzyme[Note17]. Generate a glycerol stock with a positive colony previously validated.

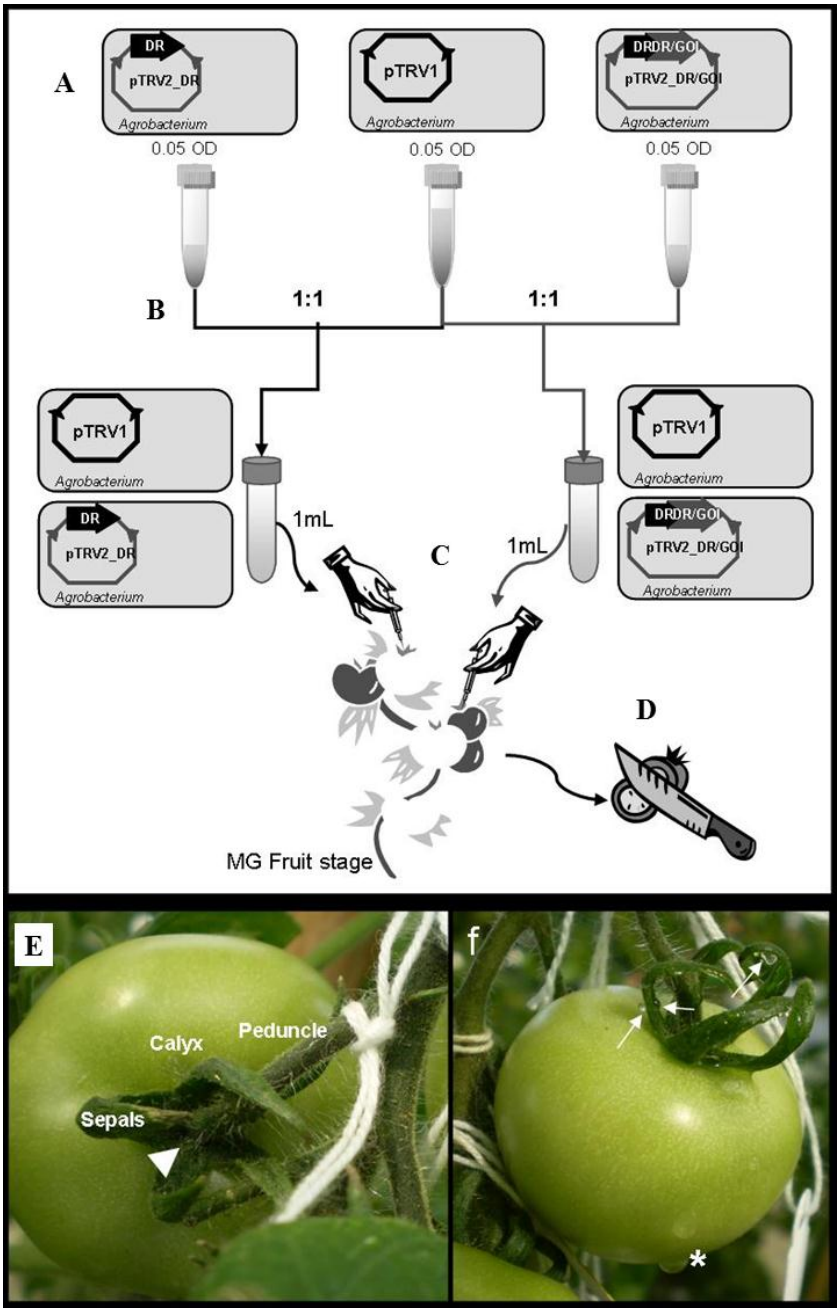
3. Culturing and sub-culturing of *Agrobacterium* clones. Grow pTRV1, pTRV2_DR and each pTRV2_DR_GOI construct from frozen stocks individually in Kanamycin and Rifampicin (50 µg/mL) selective LB plates. Pick a colony fromeach LB plate and put them into a 50 mL plastic tube containing 5 mL of LB medium with Kanamycin and Rifampicin (50 µg/mL). Grow them shaking 48 hours at 28 °C. Based on the number of labelled MG fruit (see 3.1.3), make an estimation of Agrobacterium suspension volume required for each final expression vector. The amount of agroinjection suspension mix (pTRV1: each pTRV2) varies depending on the fruit size, but 1mL should be enough to infiltrate one MG Del/Ros1 MM fruit. For small fruits, use 0,5 mL of agroinjection suspension. Prepare a fresh pre-culture of *Agrobacterium*. Take 100 µL of each pre-cultures and inoculate a 50 mL plastic tube containing 5 mL LB medium with Kanamycin and Rifampicin (50 µg/mL) (**Figure III.2A**). Grow overnight with shaking at 28 °C.

4. Agroinjection suspension preparation. Collect the *Agrobacterium* cells by centrifugation at 3000 rpm for 15 minutes. Discard the supernatant by inversion. Prepare acetosysiringone solution and add it to MES infiltration buffer. Protect it from light by wrapping the bottle with aluminium foil. Resuspend the pellet with the cells in 15mL MES infiltration buffer and vortex it[Note18] to produce the agroinjection suspension. Wrap the plastic tubes in aluminium foil. Incubate them at room temperature with gentle agitation (20 rpm) in a rotating and swaying mixer for at least 2 hours. Check the optical density (OD) at 600 nm wavelength of each suspension and dilute them adding more MES infiltration buffer to reach 0.05 OD[Note19]. Prepare the agroinjection suspension by mixing 1:1 (volume:volume) the pTRV1 suspension with each pTRV2 suspensions, including pTRV2_DR control (**Figure III.2A** and **III.2B**[Note20]).

5. Fruit agroinjection. Use different sterile 1 mL needled syringes for each agroinjection suspension (containing pTRV1 and pTRV2_DR or PTRV1 and each pTRV2_DR/GOI vectors) and agroinject them into MG

fruits (30-35 dpa) (**Figure III.2C** and **III.2E**). Agroinjection proceeds by inserting the needle about 3 to 4 mm into the fruit through the calyx region, between sepals and peduncle junction. Inject the cell suspension carefully[Note20]. The successful spread of agroinjection suspension into the fruit can be monitored by the colour change observed in the fruit tissues from light to dark green. Agroinjection is finished when the fruit is fully infiltrated, and a few drops "guttate" through the sepal hydatodes (**Figure III.2F**). Dry the drops on the fruit and keep the fruit surface clean. DR (and GOI) genes silencing can be observed ten days after agroinjection, when the breaker stage starts.

Figure III.2 The Fruit-VIGS experimental system. A. *Grow pTRV1, pTRV2_DR and pTRV2_DR/GOI viral vectors from individual Agrobacterium stocks in a suitable solid medium. Pick a colony of each one and grow them in liquid medium. Prepare a fresh subculture of each one and measure their OD.* ***B****. Dilute cell suspension to 0.05OD and combine 1:1 volumes pTRV1 vector with each pTRV2 vectors.* ***C****. Agroinject 1 mL of each agroinjection suspension mix into MG fruits, through the calyx region (****E****).* ***E****. 1 mL syringe needled should be introduced between peduncle and sepals junction (head narrow) to agroinject the cell suspension into vascular system.* ***F****. After agroinjection, some drops may appear by the hyatodes of sepals (arrows) and drip over the fruit surface (asterisk).* ***D*** *Harvest the silenced fruits, dissect the silenced sectors with a sharp knife and store them. DR, Delila/Rosea1 silencing module; GOI, gene of interest; MG, Mature Green tomato fruit stage; MM, Money Maker tomato cultivar. [next page].*


A
DR
pTRV2_DR
Agrobacterium
pTRV1
Agrobacterium
DRDR/GOI
pTRV2_DR/GOI
Agrobacterium
0.05 OD
0.05 OD
0.05 OD
B
1:1
1:1
pTRV1
Agrobacterium
DR
pTRV2_DR
Agrobacterium
pTRV1
Agrobacterium
DRDR/GOI
pTRV2_DR/GOI
Agrobacterium
1mL
1mL
C
D
MG Fruit stage
E
Calyx
Peduncle
Sepals
f
*

3.4 Dissection and collection silenced sectors

1. Harvesting. Fruits are harvested at different ripening stages depending on the particular interest of each study. Harvest DR- and DR/GOI-silenced fruits separately and keep the labels for identification.

2. Dissection and collection. Rinse fruits with tap water and dry them. Separate silenced from non-silenced areas by cutting them apart with a sharp knife (**Figure III.2D**). For pericarp tissue studies, slice the fruit and discard seeds and gel. Sort out silenced from non-silenced areas[Note21], transfer them quickly to conveniently labelled screw cap tubes (25 mL) and hold in liquid N_2. Store at -80ºC. Grind the samples up in a mortar with liquid N_2 to obtain a fine frozen powder. In **Figure III.3** we show an example of DR- and DR/GOI-silenced fruits in comparison with DR-non-silenced fruit. In the accompanying example we used the Phytoene Desaturase (PDS) gene as GOI because it provides a visual phenotype by itself, which give us the opportunity to evaluate the co-silencing of DR module and the PDS gene (Orzaez et al., 2009a).

3.5 Evaluation of GOI silencing

1. Extract the RNA from the silenced region by using a suitable RNA isolation protocol (Bugos et al., 1995).

2. Synthesize cDNA from isolated RNA with SuperScript™ First-Strand Synthesis System for RT-PCR following the manufacturer's instructions.

3. Perform a relative quantification of transcript abundance in 7500 Fast Real-Time PCR system using Power SYBR® Green PCR Master Mix and an established RT-PCR protocol.

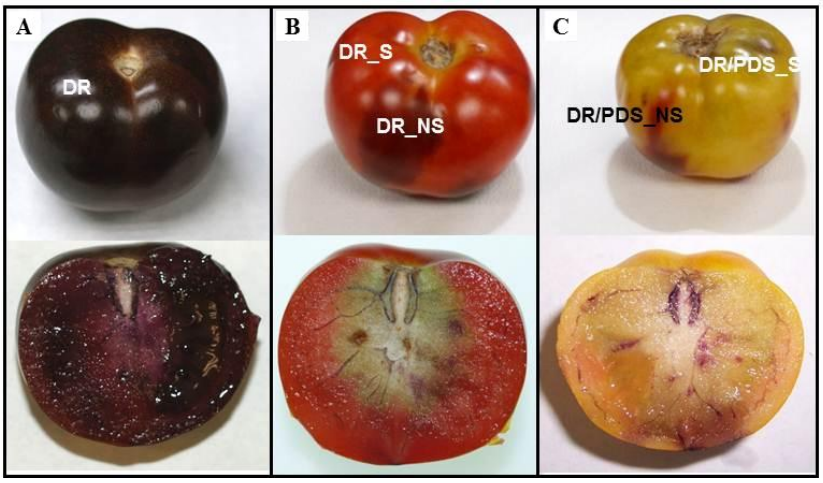


Figure III.3 DR silencing and DR/PDS co-silencing examples. A*. Del/Ros1 background presents a purple phenotype in tomato fruit (top). Anthocyanin accumulation happens in all tissues of the fruit (bottom).* ***B****. The silencing of DR-module (DR_S) prevents the anthocyanin accumulation in silenced tissues and the fruit presents a wt red colour (top and bottom). The purple areas represent regions with non-silenced DR-module (DR_NS).* ***C****. The co-silencing of a GOI with DR-module allows the identification of silenced area as shown, where PDS gene represents the co-silenced GOI (Orzaez et al., 2009a). In this case, the PDS gene silencing phenotype gives a yellow colour as results of the blockage of lycopene accumulation in the fruit. The yellow areas represent DR and PDS co-silencing in the same region. No areas were observed where only one of the two genes was silenced. Purple coloured sector represent DR/PDS-non-silenced areas (DR/PDS_NS). PDS, Phytoene desaturase.*

III.5 FUTURE PERSPECTIVES

The Fruit-VIGS tool described in this methodological report was successfully used by us in the identification of the gene function for two MYB-type transcription factors involved in the regulation of the secondary metabolism during tomato fruit ripening: SlODO1, which regulates the biosynthesis of volatiles derived from benzoic acid (Orzáez et al., 2009b) and SlMYB12, involved in the regulation of the phenylpropanoid and flavonol biosynthesis (Ballester et al., 2010).

The strategy used in purple *Del/Ros1* transgenic plants as background that provides a monitoring system for VIGS, can be adapted to other developmental stages of tomato fruit by appropriate genetic engineering. Different approaches are possible to achieve this objective: *first*, by changing the stage-specific promoter. The pENFRUIT collection developed in our laboratory can be used as promoter source for specific developmental stages (Estornell et al., 2009); s*econd*, by changing the reporter gene. Other visual reporters easily traceable can be used, e.g. fluorescent DsRed protein. The 35S-DsRed tomato cv. Money Maker transgenic line developed in our laboratory, works as a new and attractive monitoring VIGS strategy for early stage of fruit development (Orzaez and Granell, 2009b); and *third*, to combine both different specific-stage promoters with different reporter genes. In this way, complete developmental and ripening processes could be studied in tomato fruit.

III.6 NOTES

1. We have used Advantage^{TM} as a high fidelity polymerase which adds a single adenine overhang at 3'-ends of each amplicon. It is necessary for TA cloning reaction.

2. In tomato fruit (*Solanum lycopersicum*) we observed that *Agrobacterium* C58 strain is more infective and efficient in transient silencing assays than LBA4404.

3. Stocks for MES infiltration buffer: *first*, MES 100mM (10X) pH 5,6. Dissolve in sterile water and adjust the pH with KOH 1M. Require sterilization by autoclaving or filtration; *second*, Magnesium Chloride ($MgCl_2$) 1 M (100X). Dissolve in sterile water. Require sterilization by filtration; and *third*, 200mM acetosyringone solution. Dissolve 78.48 mg in 2 mL of dimethyl sulfoxide and filter-sterilize it. Divide in 200 µL alicuots and store at -20ºC. IMPORTANT: Is better to prepare the 200 mM stock acetosyringone on the same day you plan to use it.

4. Acetosyringone is photosensitive and it needs to be stored in the dark. Tubes wrapped with aluminium foil will work for that.

5. Sharp knife is important to avoid crushing the tissue.

6. Depending on many variables, such as the final expression vector size or environmental conditions, the yield of silenced sectors per fruit will be different. For example, in our experience we observed that expression vectors with larger sizes produce less silenced areas per fruit than vectors with smaller sizes. In these cases, we advise to use more plants per experiment in order to increase the amount of tissue from silenced samples.

7. Based on our experience, good results were obtained with 20 fruits per plant. We try to keep 5 trusses per plant, containing 4 fruits each. In each truss, 2 fruits are used for control silencing and other 2 fruits are used for

DR/GOI co-silencing. With 3 plants per construction, enough material is obtained for most downstream analysis (e.g. metabolomics).

8. Even though our *Del/Ros1* MM background is from a F7 generation, sometimes plants with some MT traits appear and should be removed.

9. If pruning can affect the expression of your trait, do not prune the plants. If this is not the case, prune them every week until the last fruit is collected.

10. Put special attention at plant pruning after first fruit reaches Mature Green (MG) stage: Remove every floral truss and secondary shoots.

11. Label the MG fruits indicating date, silencing vectors agroinjected and write a code that indicates fruit, truss and plant number. IMPORTANT: keep this code for downstream analysis because it represents a unique biological replicate.

12. Selection of a specific GOI region is particularly important when working with gene families. Overlapping primer designs can be used to clone several gene fragments in tandem inside the same vector.

13. We routinely use *E. coli* (One Shot®TOP10 or One Shot®Mach1TMT1R chemically competent *E. coli* kit) following the transformation conditions from pCR®8/GW manual.

14. When the transformation efficiency has not been optimal, we spread the rest of culture (150μL) on a different plate for recover some colonies.

15. We use Vector NTI 10.3.0 (Jul 31, 2006© Invitrogen, Carlsbad, CA, USA) program to choose the suitable restriction enzymes and predict the expected sizes after plasmid DNA digestion. Proper enzymatic reaction conditions can be found at the enzyme manufacturer website. The pCR®8/GW/TOPO® TA Cloning® Kit, recommend EcoRI restriction enzyme because it releases the cloned GOI from the plasmid. Be careful with

additional EcoRI digestion products which can be obtained as a result of internal EcoRI sites in your fragment.

16. We use Vector NTI 10.3.0 program too to select the best restriction enzyme. We commonly use EcoRV (Takara Bio Europe, France; Takara Bio Inc., Shiga, Japan).

17. To increase the yield of *Agrobacterium* plasmid DNA minipreps, collect the cells from 5mL liquid LB medium and follow QUIAprep® Miniprep Kit procedures. Elute with 20μL to obtain a more concentrated plasmid DNA preparation.

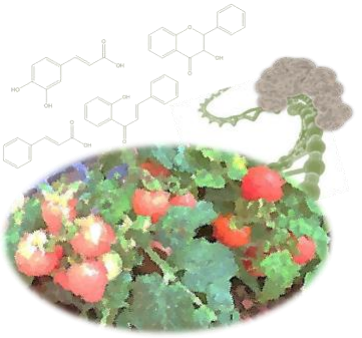
18. It is IMPORTANT to resuspend completely the cells in the MES infiltration buffer. Vortex them for 1-2 minutes.

19. Usually we make a 0.5 OD intermediate dilution for each agroinjection suspension, in order to equalize them before reaching the final 0.05 OD.

20. In some fruits you may find a higher initial resistance to agroinjection than normal. In those cases try to find a more suitable entry position by changing the depth of needle insertion in the fruit.

21. Sometimes you may find difficulties in distinguishing between silenced and non-silenced areas. Small patchy silencing or a gradual silencing make hard to dissect some sectors. If there are enough samples, discard those fruits. If not, try to select areas in patchy silenced fruit with high similarity to well-silenced fruit.

CHAPTER IV



CHARACTERIZATION OF A NEW '*pink fruit*' TOMATO MUTANT RESULTS IN THE IDENTIFICATION OF A NULL ALLELE OF THE *SLMYB12* TRANSCRIPTION FACTOR.

Josefina-Patricia Fernández-Moreno[1], Oren Tzfadia[2], Javier Forment[1], Ilana Rogachev[2], Sagit Meir[2], Diego Orzaez[1], Asaph Aharoni[2] and Antonio Granell[1]

(Adapted from the manuscript submitted to *Plant Physiology*)

[1] Fruit Genomics and Biotechnology Laboratory. Instituto de Biología Molecular y Celular de Plantas (CSIC-UPV). Valencia, Spain.

[2] Department of Plant Sciences and the Environment. Weizmann Institute of Science. Rehovot, Israel.

SUMMARY

The identification and characterization of new tomato (*Solanum lycopersicum*) mutants, affected in fruit pigmentation and nutritional content, are imperative for understanding the underlying biology, and also as a source of new alleles for breeding programs. To date, characterized tomato pink pigmented fruit mutants have been reported to be a result of low *SlMYB12* transcript levels in the fruit skin. Two new Ethyl Methyl Sulphonate mutant lines displaying a '*pink fruit*' phenotype (*pf1* and *pf2*) were characterized in this study. While *SlMYB12* transcript accumulates to wild type levels in these mutants, it exhibits a truncated transcript lacking the last exon encoding the essential MYB activation domain. Allelism and complementation tests performed with the *pf* mutants revealed that they were allelic to the *y* locus

and showed the same recessive *pf* allele in homozygosis. A set of molecular and metabolic effects, reminiscent of those observed in the arabidopsis *myb11myb12myb111* triple mutant, were found in the mutants *pf*, which have not been previously described, and support them being null mutants. These yet undescribed aspects include a reduction of several flavonol glycosides and the down-regulation of some associated glycosyl transferases. Transcriptome analysis revealed extended effects from the flavonoid pathway into the interface between primary and secondary metabolism. Finally, a screening for Myb-binding sites on candidate gene promoter sequences revealed that 116 of the 152 co-down-regulated genes were putative direct targets of SlMYB12 regulation activity. In contrast to the previous transcriptional hypomorphic pink fruit lines, the new *pf* mutants likely correspond to a null allele.

Keywords: EMS-mutants, tomato fruit skin, ESI-UPLC-QTOF, phenylpropanoids, flavonoids, RNA-Seq, R2R3-MYB transcription factor, Myb-Binding motif.

IV.1 INTRODUCTION

The consumption of pink-colored tomatoes in Asia is high (Ballester et al., 2010; Lin et al., 2014) and their popularity is increasing worldwide. The "*pink fruit*" trait was described by Lindstrom et al., in 1925 as a fruit with a transparent epidermis that lacks its natural yellow flavonoid compounds (Lindstrom, 1925). Fruits from homozygous plants that contain the recessive "*y*" allele in the *y* locus located in chromosome 1 have failed to accumulate the yellow flavonoid precursor naringenin chalcone in fruit peel, which causes the pink-colored fruit phenotype (Lindstrom, 1925; Rick and Butler, 1956; Adato et al., 2009; Ballester et al., 2010).

Flavonoids are a class of phenolic compounds that are widely distributed in the plant kingdom (Domínguez et al., 2011), and exhibit a wide range of ecological and physiological functions (Winkel-Shirely, 2001; Grotewold, 2006; Domínguez et al., 2009a, 2009b and 2011). The interest shown in flavonoids is because of their effect on fruit appearance, and also their nutritional and health-related properties (Rein et al., 2006; Butelli et al., 2008; Luo et al., 2008; Martin et al., 2011). In tomato fruit skin, flavonoids have been reported to contribute to correct cuticle functioning as they affect its rigidity and susceptibility to cracking and postharvest behavior (Domínguez et al., 2009a and 2011; Lara et al., 2014). Increasing the accumulation of polyphenol species, including flavonoids, or producing a wider range of them to improve the nutritional value and tolerance to environmental stresses, are currently important fruit biotechnology and breeding goals (Luo et al., 2008).

The biosynthesis of flavonoids has been well studied as a branch of the general phenylpropanoid biosynthetic pathway (Grotewold, 2006). Phenylpropanoids (PPs) are derived from the L-phenylalanine (Phe) an amino acid synthesized by the shikimic and chorismic acids biosynthetic pathways. In the first steps of the PP pathway, Phe is transformed into 4-

coumaroyl-CoA, the precursor of flavonoid biosynthesis, by the successive activities of PHENYLALANINE AMMONIA-LYASE (PAL), CINNAMATE 4-HYDROXYLASE (C4H) and 4-COUMARATE:COENZYME A LIGASE (4CL) enzymes. Next, CHALCONE SYNTHASE (CHS) produces naringenin chalcone by incorporating three molecules of malonyl-CoA to a molecule of 4-coumaroyl-CoA. Using naringenin chalcone as a substrate, a series of enzymatic steps produce different flavonoids (flavones, flavonones, isoflavones, dihydroflavonols and flavonols), proanthocyanidins and anthocyanins through distinct branches of the main pathway (**Figure IV.1**). More than 70 different flavonoids have been identified in tomato fruit, including glycosylated and other conjugated flavonoids (Moço et al., 2006; Mintz-oron et al., 2008; Iijima et al., 2008; Adato et al., 2009; Ballester et al., 2010). Studies on different species (arabidopsis, maize, snapdragon, rose, gentian, tomato or grape, among others) have revealed that R2R3-MYB-type transcription factors are able to regulate not only *CHS* and *CHALCONE ISOMERASE* (*CHI*) gene expression, but many other genes of the PP, flavonoid and anthocyanin biosynthetic pathways (Grotewold et al., 1994; Chopra et al., 1996; Kranz et al., 1998; Stracke et al., 2007; Adato et al., 2009; Czemmel et al., 2009; Ballester et al., 2010; Lin-Wang et al., 2010; Hichri et al., 2011; Du et al., 2012; Nakatsuka et al., 2012). Specific R2R3-MYB transcription factors (TFs) capable of regulating the expression of the genes in the flavonoid branch are characterized by the presence of the SG7 motif (Kranz et al., 1998; Stracke et al., 2007; Czemmel et al., 2009), which is positioned downstream of the R2R3 DNA-binding domain. Almost every gene on the PP biosynthetic pathway is regulated by an SG7-MYB TF (Luo et al., 2008; Czemmel et al., 2009).

Although overexpression of *CHI* and silencing of *CHS* result in a reduction in naringenin chalcone levels in fruit skin, which leads to a pink fruit phenotype (Muir et al., 2001; Schijlen et al., 2007; Bovy et al., 2008), all

currently characterized *pink fruit/y* mutants have been related with low expression levels of chromosome 1-located R2R3-MYB transcription factor *SlMYB12*. This is the case of modern cultivars that carry the "*y*" mutation, and of some introgression lines that derive from *Solanum chmielewskii* in the cv. Moneyberg background population (Ballester et al., 2010), the LA3189 line in the cv. Ailsa Craig background (Adato et al., 2009), and is probably the case of a large number of globally characterized pink accessions (Lin et al., 2014). To date, *y* mutant fruit has been reported to exhibit low levels of the *SlMYB12* transcript, which likely results in the down-regulation of *CHS*, *CHI*, *FLAVONOID 3-HYDROXYLASE* (*F3H*) and *FLAVONOL SYNTHASE* (*FLS*) expression, and in a partial block of the flavonol biosynthetic branch (**Figure IV.1**). In addition, Adato et al., (2009) provided a putative target gene set for the SlMYB12 TF, apart from the flavonol biosynthetic branch, and suggested that other MYB-type TFs in tomato (e.g. *SlTHM27*) would be primary targets of *SlMYB12* regulatory activity (**Figure IV.1**) (Adato et al., 2009).

The economical and nutritional relevance of flavonoids in tomato fruit suggests that identification of new mutants, studying the genetic variability associated with the pink fruit trait and deciphering their biosynthetic regulatory network are important research targets. In this paper, we characterized two pink fruit tomato mutants by performing: (i) a metabolomics study using liquid chromatography coupled to mass spectrometry (LC-MS) and (ii) a gene expression analysis by next generation sequencing (i.e., RNA-Seq). The aim of the study was to identify the metabolic and molecular changes in the *y* mutation, and to discover new candidate target genes of the SlMYB12 transcription factor.

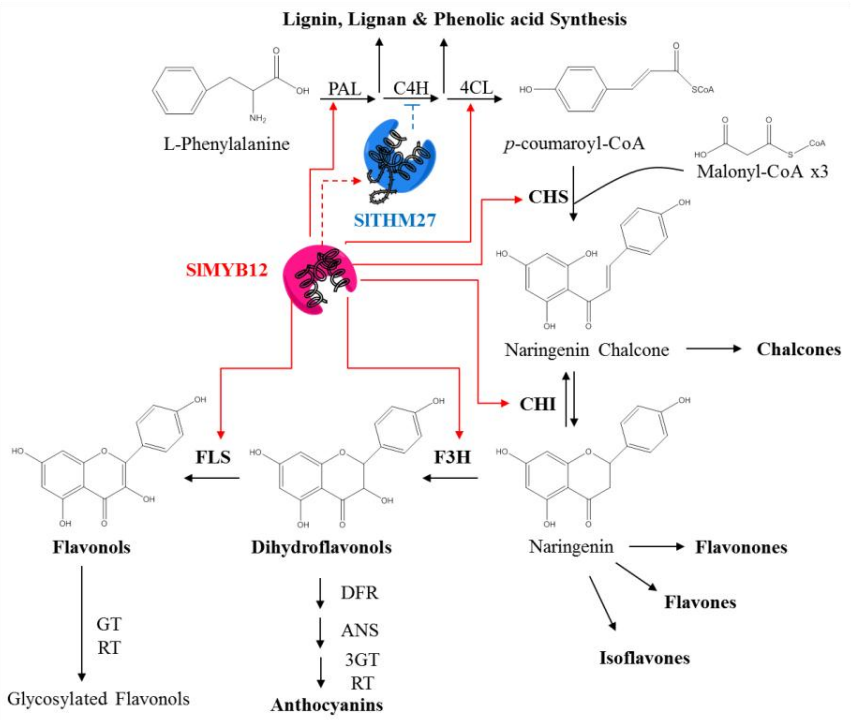


Figure IV.1 Schematic overview of the flavonoid biosynthetic pathway and regulation in tomato fruit. *The amino acid L-Phenylalanine is transformed into naringenin chalcone by the action of enzymes PAL, C4H and 4CL. Naringenin chalcone is the precursor of different flavonoid classes: chalcones, flavonones, flavones, isoflavones, dihydroflavonols and flavonols. SlMYB12 TF is the main activator of the flavonol biosynthetic branch to act on genes PAL, 4CL, CHS, CHI, F3H and FLS (Adato et al., 2009; Ballester et al., 2010). The SlTHM27 transcription repressor, putatively involved in the regulation of gene C4H, could also be regulated by SlMYB12 TF (Adato et al., 2009). Abbreviations: PAL, PHENYLALANINE AMONIO LYASE; C4H, CINNAMATE 4-HYDROXYLASE; 4CL, 4-COUMARATE CoA LIGASE; CHS, CHALCONE SYNTHASE; CHI, CHALCONE ISOMERASE; F3H, FLAVONONE 3 HYDROXYLASE, FLS, FLAVONOL SYNTHASE; DFR, DIHYDROFLAVONOL REDUCTASE; ANS, ANTHOCYANIDIN SYNTHASE; 3GT, 3-O-GLYCOSYL TRANSFERASE; RT, RHAMNOSYL TRANSFERASE.*

IV.2 MATERIALS & METHODS

Plant material

Mutant seeds were obtained from the EMS (Ethyl Methyl Sulphonate) mutagenized tomato population (*Solanum lycopersicum* cv. MT, 'MicroTom'), developed by Tsukuba University in Japan (TOMATOMA, http://www.tomatoma.nbrp.jp/; Saito et al., 2001; Watanabe et al., 2007). F_3 seeds from two mutant lines with a pink fruit phenotype, 4497 and 5663, were used. All the "*wild type*" (WT) and the *pf1* and *pf2* mutant plants were grown under natural light and controlled temperature conditions (24ºC during the day, 18ºC at night) in a greenhouse. Once fruit had reached the mature green stage, plants were checked daily to harvest fruits in the breaker (Br) stage. Peels from 5-8 Br fruits per genotype were manually separated and gently cleaned of excess mesocarp, as described by Adato et al. (2009), and were frozen in liquid nitrogen until further use. Three biological replicates, which represented the peel of 5-10 fruits, were collected from three independent plants, and were used for each experimental data point.

UPLC-QTOF-MS profiling of semi-polar compounds and data analysis

An analysis of semi-polar compounds, including flavonoids and other phenylpropanoids, was carried out using an UPLC-ESI-QTOF instrument (HDMS Synapt; Waters) with the UPLC column connected online to a photodiode array (PDA) detector and then to the MS detector, following the methodology described by Rogachev and Aharoni (2012). The analysis was done on methanolic extracts prepared from frozen tissues. A 100 x 2.1-mm i.d., 1.7-µm UPLCBEH C18 column (Waters) was used for the separation of metabolites. PDA detector was set to 210-500 nm. Masses of the eluted compounds were detected by a qTOF mass spectrometer, equipped with an electrospray ionization (ESI) source, in positive and negative

ionization modes, at the m/z range from 50 to 1500Da. The XCMS open-source LC-MS-based data analysis software (Smith et al., 2006) was used for peak picking and peak alignment, as described elsewhere (Rogachev and Aharoni, 2012). Compounds were identified using standards, injected at the same LCMS conditions, or putatively identified as previously described (Adato et al., 2009). Statistical analysis of the XCMS outputs was performed by a home-made computer program. Only the metabolites that changed more than 2-fold between the mutants and control samples were considered to differentially accumulate.

RNA-Seq gene expression analysis

For Whole Transcriptome Shotgun Sequencing (WTSS or RNA-Seq), total RNA was extracted, short cDNA libraries of 50 bp length were produced (Zhong et al., 2011), and was quality tested in an Agilent 2100 bioanalyzer (Faculties of Life Sciences, Weizmann Institute of Science, Israel) and sequenced by the Illumina technology (Crown Institute of Genomics, Weizmann Institute of Science, Israel). High-quality reads were filtered by the CASAVA 1.8 software. Clean reads were mapped to the tomato genome using TOPHAT (Trapnell et al., 2012). The resulting alignment files were used to generate a transcriptome assembly for both the *pf* mutants and the Micro Tom wild-type (MT-WT) with the help of CUFFLINKS (Trapnell et al., 2012). Those assemblies were merged together to calculate gene and transcript expressions with the CUFFMERGE utility of the CUFFLINKS package. CUFFDIFF calculated the expression levels and tested the statistical significance of the CUFFMERGE results. The differentially expressed (p-value < 0.05) transcripts between each mutant and control were aligned by the BWA software (Li et al., 2010) to create consensus sequences by the iAssembler software (Zheng et al., 2011). The free software IGV (Integrative Genomics Viewer) was used to view the consensus sequences (Robinson et al., 2011).

PCR and RT-PCR gene expression analyses

For the PCR analysis, genomic DNA was extracted by the CTAB-chloroform procedure (15g CTAB, 40.9g NaCl, 0.5M EDTA pH 8, 1M Tris-HCl pH 8, 0.2% β-mercaptoethanol). The aqueous phase was treated with RNAse (50 mg/mL), and DNA was precipitated with 3M sodium acetate and isopropanol. The pellet was washed with 70% and 80% of ethanol, and was resuspended in DDW. The PCR reaction was set as follows: 1 μl of diluted (1:100) genomic DNA, 12.5 μl GOTaq® Green Master Mix (Promega) and 0.5 μl of each *SlMYB12*-specific forward (F) and reverse (R) oligonucleotide primers in a 25 μl reaction. The PCR conditions were: 94ºC, 4 min; (94ºC, 30 s; 50-60ºC (depending on the primer sequence), 30 s; 72ºC, 1-4 min (depending on the primer length)) x 35 cycles; 72ºC, 5 min; 8ºC, ∞ min. The gene-specific oligonucleotide primers were designed using Vector NTI 10.3.0 (2006© Invitrogen, Carlsbad, CA, USA) (**Figure IV.S7**).

For the RT-PCR analysis, total RNA was extracted using Tri-reagent 100 mg/mL (38% water-saturated phenol, 0.8M guanidine thiocyanate, 0.4M ammonium thiocyanate, 0.1M sodium acetate pH 5, 5% glycerol). The aqueous phase was extracted twice with chloroform and precipitated with isopropanol. The RNA pellet was washed twice with 70% ethanol and resuspended in DDW. Isolated RNA was treated with DNase I to be reverse-transcribed using a high capacity complementary DNA reverse transcription kit (Applied Biosystems). Real-time PCR analysis was performed as previously described (Adato et al., 2009; Mintz-Oron et al., 2008). The *SlMYB12*-specific oligonucleotide primers were designed using Software Express (Applied Biosystems) (**Figure IV.S7**). The TIP41-like gene (*Solyc10g049850*) was used as an endogenous control (Exposito-Rodriguez et al., 2008).

Identification of transcription factor DNA recognition elements

The gene promoter region for a selection of differentially down-regulated genes in the *pf* mutants was extracted in the FASTA format (http://www.solgenomics.net/). The promoter region was defined as the region of 2000 base-pairs upstream of the 100 base-pairs downstream of the start codon (Veerla et al., 2006). Using the 2100 base-pair length promoter sequences, DNA motifs were found with the free online MEME tool, version 4.9.0 (Bailey et al., 1994). The MEME analysis was run to search for the 100 motifs of a width size between 6 and 50 base pairs using the default distribution model (Zero Or One Per Sequence, ZOOPS) and stopping if the motif E-value was above $1E^{-03}$. The analysis included those motifs that repeated between 2 and 600 times, and considered the maximum dataset size as 1,000,000 characters. Then the free online TOMTOM tool (Gupta et al., 2007) was used to identify the different transcription factor recognition elements in the significant DNA motifs obtained by the MEME analysis. The TOMTOM analysis was run with the default settings. The transcription factor recognition elements were searched using the JASPAR (Portales-Casamar et al., 2010) and UniPROBE (Newburger and Bulyk, 2009) databases. Only the plant DNA-binding elements in these databases were considered in our study.

In addition, the Myb-Binding Site (MBS) motif consensus sequences for the target genes of AtMYB111 in the *Arabidopsis thaliana myb111 myb11 myb12* triple mutant, as described by Franco-Zorrilla et al., (2014) ((T/C)AAC(T/A)A(A/C)C and ACCTA(A/C)C), were used as a string search query of the gene promoter regions. The free online FUZZNUC tool (http://www.emboss.bioinformatics.nl/cgi-bin/emboss/fuzznuc/) was used to search the MBS patterns in the promoter sequences of interest.

Gene ontology analysis for primary metabolism

The differentially expressed gene set shared by both mutant *pf1* and *pf2* was analyzed by the MAPMAN 3.6.0 RC1 software (http://mapman.gabipd.org/eb/guest/mapman/) to identify the genes involved in primary metabolism processes among the differentially expressed gene set. To map the genes into the primary metabolism processes, we used an updated tomato mapping file available in the GoMapMan open web-accessible resource for gene functional annotations (http://www.gomapman.org/).

IV.3 RESULTS & DISCUSSION

Characterization of the *pink fruit* phenotype in the new mutants

A phenotypic description of tomato EMS mutants (cv. MT), available in the TOMATOMA database (Kazusa, Japan) (http://www.tomatoma.nbrp.jp/), revealed two mutant lines with pink-colored fruits (database numbers 4498 and 5663), which were named *pink fruit 1* (*pf1*) and *pink fruit 2* (*pf2*), respectively. The fruit peel of both the *pf* mutants showed no yellowish coloration, which was confirmed in methanolic extracts; the *pf* mutant extracts were colorless, while the control extract was shiny yellow (**Figure IV.2A**). This pink fruit phenotype was similar to the phenotype previously described for the *y* mutant (Adato et al., 2009), the *y* lines (Ballester et al., 2010; Lin et al., 2014) and the *CHI* and *CHS* transgenic tomato lines (Muir et al., 2001; Bovy et al., 2007; Schijlen et al., 2007). We further defined the *pink fruit* mutation by both allelic and complementation diagnostic tests. The allelic test between the *pf* mutants and the WT resulted in F_1 individuals, which showed only red fruits. This is consistent with the *pink fruit* mutation caused by a recessive allele of a single gene in mutants *pf1* and *pf2* (**Figure IV.2B**). The complementation test done between both the *pf* mutants resulted in that F_1 individuals showed only pink fruits. Thus, both *pf1* and *pf2* mutants could be alleles of a single gene: pf^1 and pf^2 (**Figure IV.2B**). Similar tests with the *y* mutants (Ballester et al., 2010) confirmed that they were also alleles of the same *y* locus.

The *SlMYB12* gene is expressed in both new *pink fruit* mutants

The visual pink fruit phenotype in the *pf* mutants suggested that the function of *SlMYB12* might be altered, as previously reported for the *y* mutants (Adato et al., 2009; Ballester et al., 2010; Lin et al., 2014). To determine *SlMYB12* expression in the *pf* mutants, RT-PCR analysis experiment was conducted with the mRNA of fruit skin in the breaker stage.

In contrast to previous reports on *y* fruit from mutants or accessions (Adato et al., 2009; Ballester et al., 2010; Lin et al., 2014), the *SlMYB12* transcripts showed equal or slightly higher expression levels in the *pf* mutants compared to the WT fruit (**Figure IV.2C**).

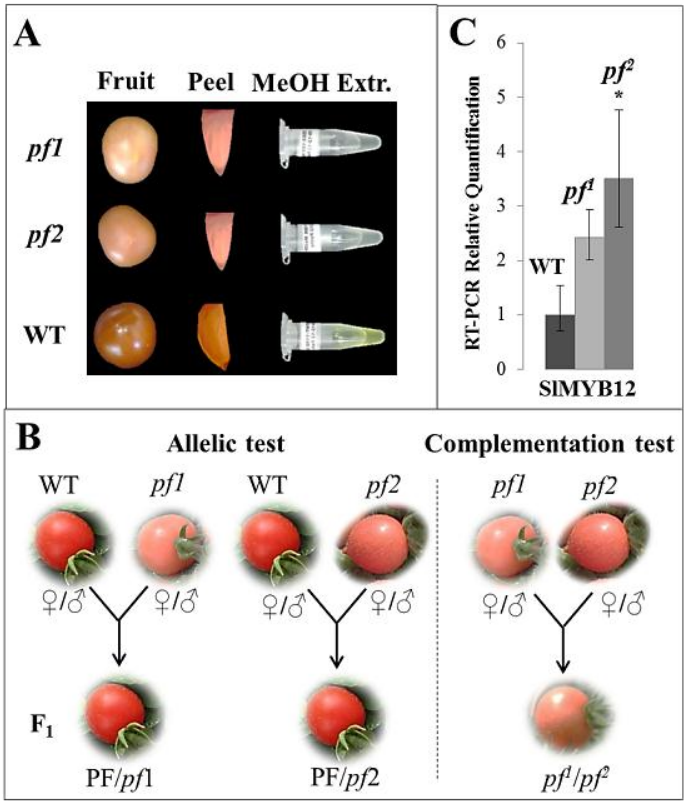


Figure IV.2 The pink fruit phenotype in mutants pf1 and pf2. ***A.****The phenotype for tomato fruit, fruit skin and methanolic extracts in both the pf mutants and WT.* ***B.*** *Allelic and complementation tests.* ***C.*** *The SlMYB12 gene expression level detected by the RT-PCR analysis in both the pf mutants and WT fruit skin. Abbreviations: PF, PINK FRUIT wild type allele; pf, pink fruit mutant allele.*

Accumulation of glycosylated flavonols is specifically diminished in both *pink fruit* mutants

Although skin color is indicative of similarity between the new and previously described pink fruit mutants, the normal *SlMYB12* transcript levels in both the *pf* mutants led us to investigate possible differences in their metabolite profiles. Thus, a biochemical analysis was performed on both the *pf* mutants and the WT fruit skin. The LC-MS analysis revealed 46 putative phenylpropanoid (PP) compounds that were differentially accumulated in both the *pf* mutants as compared to the WT fruit skin (**Table IV.S1**). Most of those metabolites (40 PPs) showed reduced accumulation levels in the *pf* mutants. These results mostly resembled the PP profile for the *y* mutant, as described by Adato et al. (2009). Similarly to that described for other *y* fruits, levels of naringenin, naringenin chalcone (Adato et al., 2009; Ballester et al., 2010), and their hydroxylated and glycosylated derivatives (Adato et al., 2009), were reduced in both *pf* mutants (**Figure IV.3**, **Table IV.S1**). We also observed reduced levels of dihydrokaempferol, quercetin and 13 glycosylated flavonol derivatives that contained one to four glycosyl groups (**Figure IV.3**). Only four of these glycosylated flavonols have been previously reported in *y* fruits (see a, b in **Figure IV.3**): rutin (Adato et al., 2009; Ballester et al., 2010), quercetin-dihexose-deoxyhexose, quercetin-hexose-deoxyhexose-pentose and kaempferol-glucose-rhamnose (Adato et al., 2009). The remaining 6 up-regulated metabolites belonged to the hydroxycinnamate phenolic acid biosynthetic branch (Liu et al., 2015) (see 1-7 in **Table IV.S1**).

Identification of a complex gene set associated with the new *pink fruit* mutant phenotype

Since we detected specific differences in the PP metabolite levels between the *pf* mutants and the previously described *y* mutant, a transcriptome analysis was performed to decipher the molecular basis that underlay the *pf* mutants metabolic phenotype.

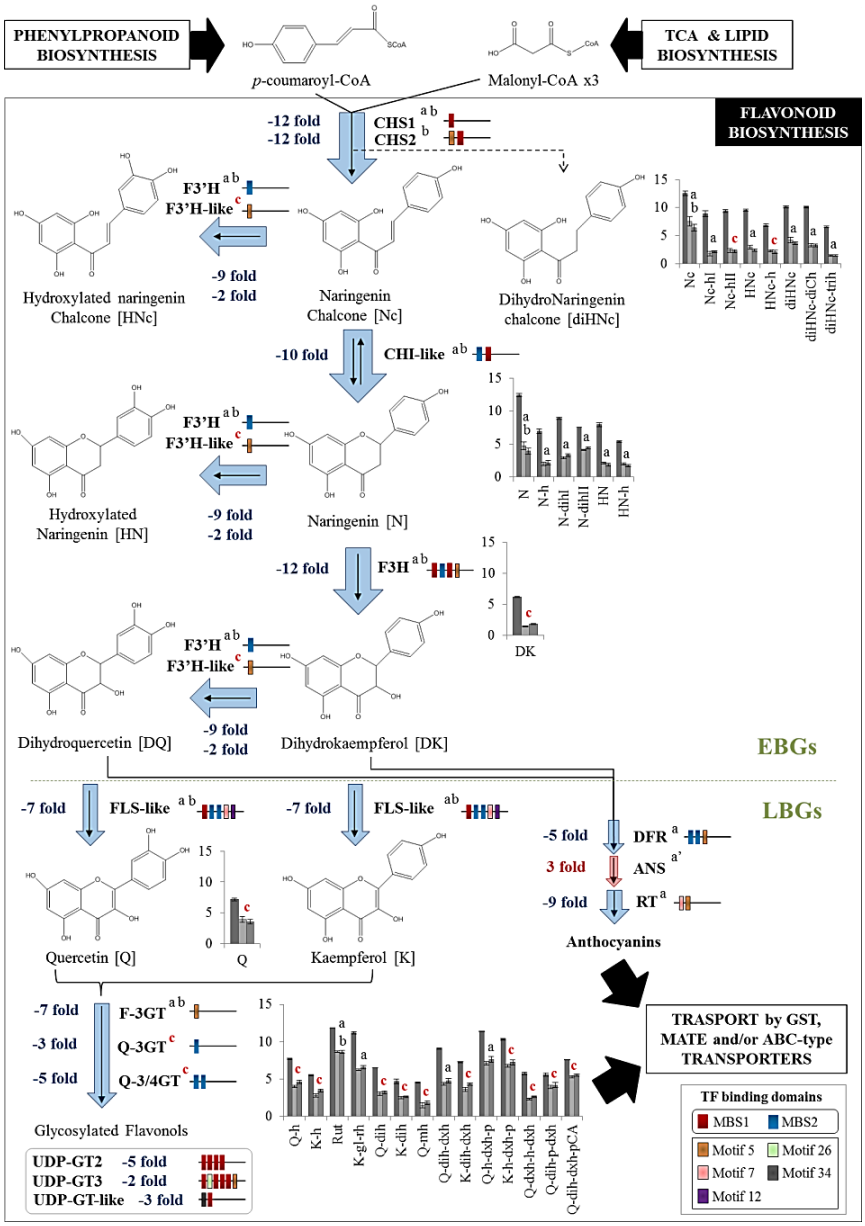

PHENYLPROPANOID BIOSYNTHESIS
TCA & LIPID BIOSYNTHESIS
p-coumaroyl-CoA
Malonyl-CoA x3
FLAVONOID BIOSYNTHESIS
-12 fold
-12 fold
CHS1
CHS2
F3'H
F3'H-like
-9 fold
-2 fold
Hydroxylated naringenin Chalcone [HNc]
Naringenin Chalcone [Nc]
DihydroNaringenin chalcone [diHNc]
Nc
Nc-hl
Nc-hll
HNc
HNc-h
diHNc
diHNc-diCh
diHNc-triH
-10 fold
CHI-like
Hydroxylated Naringenin [HN]
Naringenin [N]
N
N-h
N-dihl
N-dihll
HN
HN-h
-12 fold
F3H
DK
Dihydroquercetin [DQ]
Dihydrokaempferol [DK]
EBGs
LBGs
-7 fold
FLS-like
Quercetin [Q]
Kaempferol [K]
Q
-5 fold
DFR
3 fold
ANS
-9 fold
RT
Anthocyanins
TRASPORT by GST, MATE and/or ABC-type TRANSPORTERS
-7 fold
-3 fold
-5 fold
F-3GT
Q-3GT
Q-3/4GT
Glycosylated Flavonols
Q-h
K-h
Rut
K-gl-rh
Q-dih
K-dih
Q-mh
Q-dih-dxh
K-dih-dxh
Q-h-dxh-p
K-h-dxh-p
Q-dxh-h-dxh
Q-dih-p-dxh
Q-dih-dxh-pCA
TF binding domains
MBS1
MBS2
Motif 5
Motif 26
Motif 7
Motif 34
Motif 12
UDP-GT2 -5 fold
UDP-GT3 -2 fold
UDP-GT-like -3 fold

Figure IV.3 Flavonol biosynthesis is blocked in both the new pink fruit mutants. *Differentially expressed genes are shown by blue (down-regulated) or red (up-regulated) arrows (thicker arrows represent a higher fold change). The transcription factor (TF) binding domains found in the promoter region of the down-regulated genes are also included (in colored boxes). Bar graphs show the metabolite levels for flavonols. WT is represented in dark gray (left), the pf^1 mutant in light gray (middle) and the pf^2 mutant in gray (right). Error bars are standard deviations for n=3. Genes and metabolites differentially expressed also in the y mutant (Adato et al., 2009) and in the other y lines (Ballester et al., 2010) are marked a and b, respectively; and those differentially expressed only in the pf mutants are marked c (in red). Abbreviations: TCA, tricarboxylic acid cycle; EBGs, early biosynthetic genes; LBGs, late biosynthetic genes; h, hexose; dxh, deoxyhexose; p, pentose; gl, glucose; rh, rhamnose; mh, malonyl-hexose; pCA, p-Coumaric acid; Rut, rutin; a', the ANS gene was down-regulated in the y mutant (Adato et al., 2009).[previous page]*.

In addition, RNA-Seq analysis was conducted to provide better coverage of the transcriptomic changes that occurred in the *pink fruit* mutants, which might have not been detected in previous experiments done using similar mutants with array technologies. Consequently, the breaker fruit skins from the same samples used in the LC-MS analysis were also used for the transcriptome analysis. The differential gene expression analysis of the RNA-Seq outputs by the CUFFLINK package (Zhong et al., 2011) revealed 444 genes with significantly altered expression levels in the pf^1 mutant, and 442 in the pf^2 mutant. In all, 312 genes showed a similar altered expression levels in both the *pf* mutants compared to the wild type (**Figure IV.4**). Of the 312 genes, 159 were up-regulated and 152 were down-regulated. When comparing these results with the differential expression analysis performed in the *y* mutants by the *Affymetrix* microarray technology (273 total genes with an altered gene expression in the breaker stage) (Adato et al., 2009), RNA-Seq detected ~2-fold more genes that showed differential gene expression. Only about 14% (45 genes) of the 312 genes were also

shared by the *y* mutant (**Figure IV.4**, **Table IV.S2**). However among the remaining 267 genes, we detected some genes with a similar functional annotation, but with a different tomato gene ID (http://www.solgenomics.net/) to those found to be differentially expressed in the *y* mutant (Adato et al., 2009). This result can be explained by the different power of the profiling technologies, and also by the genetic backgrounds used in the *y* and *pf* mutant experiments (cultivars Ailsa Craig and MT, respectively). Next, we focused on the 312 genes whose differential expression was shared by both pf^1 and pf^2 mutants, including the 45 genes reported in previous *y* mutant studies.

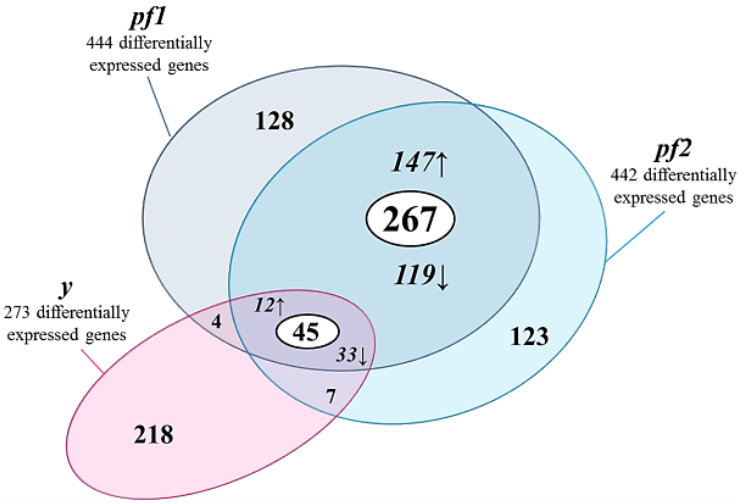


Figure IV.4 Genes differentially expressed in the new pink fruit mutants detected by the RNA-Seq approach. *A Venn diagram is presented that includes all the genes differentially expressed in mutant pf^1 and pf^2 (dark blue and light blue circles, respectively). The genes detected as being differentially expressed in the y mutant (pink circle) by the Affymetrix microarray approach (Adato et al., 2009) were also included. The genes shared between mutant pf^1 and pf^2, and between both the pf mutants and the y mutant, were represented in a white circle. The shared up-regulated genes (↑) and the shared down-regulated genes (↓) are also represented.*

The *SlMYB12* transcript is truncated at the same position in both new *pink fruit* mutants

RNA-Seq analysis confirmed that *SlMYB12* was expressed at equal or slightly higher levels in pf^1 and pf^2 compared to the WT fruit, respectively (**Figure IV.5A**). Mapping the RNA-Seq reads against the *SlMYB12* gene sequence (*Solyc01g079620*) (http://www.solgenomics.net/) showed a similar profile for the pf^1 and pf^2 versions, as well as for the WT, from the start codon to nucleotide 661. Notably downstream of that position, no RNA-Seq reads mapped in the *pf* mutants, unlike the WT reference which covered the full transcript length (1161 bp) (**Figure IV.5B**). The consensus sequence of all the RNA-Seq reads for each *pf* mutant revealed two nucleotide changes at the 661 (A:T > G:C) and 662 (T:A > A:T) transcript positions compared to the WT transcript consensus sequence (**Figure IV.5B**). We investigated the *SlMYB12* transcript structure to determine the effect of the nucleotide change observed in both the *pf* mutants. When comparing the *SlMYB12*-WT, *Slmyb12-*pf^1 and *Slmyb12-*pf^2transcript sequences with the cv. Heinz tomato genomic sequence (http://www.solgenomics.net/), we observed that *SlMYB12* was spliced into four exons and three introns (**Figure IV.5C**), which confirms the structure proposed by Ballester et al., (2010) and contrasts with the 3-exon structure of *AtMYB12* (*At2g47460*) in arabidopsis (http://www.arabidopsis.org/). Furthermore, the assembled *Slmyb12-pf* transcript sequences revealed that no reads matched the transcript sequence immediately after the sequence that corresponds to the end of exon-3, which indicates that the full exon-4 sequence was missing in the *pf* mutants (**Figure IV.5B**).

We subsequently sequenced the different genomics DNAs (gDNAs) derived from the breaker fruit skin samples of the various genotypes that covered the different introns and exons (**Figure IV.5C**). While we observed that the expected amplicon spanned from exon-1 to intron-3 in the WT and both the $pf^1$and pf^2 mutants, we failed to amplify exon-4 in the mutants

(**Figure IV.5D**). This scenario suggests genomic reorganization (e.g. which most likely involves a deletion) at the end of the gene in the *pf* mutants. These results further indicate that both the *pf* mutants correspond to a single allele (*pf*) of the same mutation of the *y*/*MYB12* locus.

The deduced amino acid sequence of the *SlMYB12*-WT and *Slmyb12-pf* mutant transcripts revealed that exon-1 and exon-2 corresponded to the R2 and R3 DNA-binding domains described for the MYB-type TF proteins (**Figure IV.5C**, **Figure IV.S1**). At the N-terminal, immediately upstream of the R2 motif, we found a well-conserved motif for the flavonol and proanthocyanidin regulators group (Czemmel et al., 2009) (**Figure IV.S1**). Downstream in the sequence, exon-3 contained a flavonol regulation motif, which was shared by R2R3 TFs that belonged to subgroup 7 (SG7), as described in arabidopsis (Kranz et al., 1998; Stracke et al., 2007; Czemmel et al., 2009) (**Figure IV.S1**). Finally, exon-4 of the WT version of *SlMYB12* contained a putative activation domain (defined by a negatively charged amino acid rich sequence) (Chopra et al., 1996) and a very well-conserved flavonol regulation motif shared by a subset of SG7 R2R3 TFs(SG7-2) (Czemmel et al., 2009) located at the C-terminus of the protein (**Figure IV.S1**). The *pf* mutants containing a putative truncated Slmyb12-*pf* protein exhibited the effect of a knock-out null mutant for *SlMYB12* in tomato fruit, even when the *Slmyb12-pf* truncated transcript was expressed at normal levels.

The new *pink fruit* mutant is a SlMYB12 null mutant

In order to produce a strong effect in flavonol biosynthesis in arabidopsis, a *myb11 myb12 myb111* triple null mutant was required as these transcription factors showed partly overlapping functions. One of the specific changes associated with this triple null mutant in arabidopsis is the reduced accumulation of several flavonol glycosides, together with the down-regulation of some flavonoid glycosyl transferases (Stracke et al., 2007).

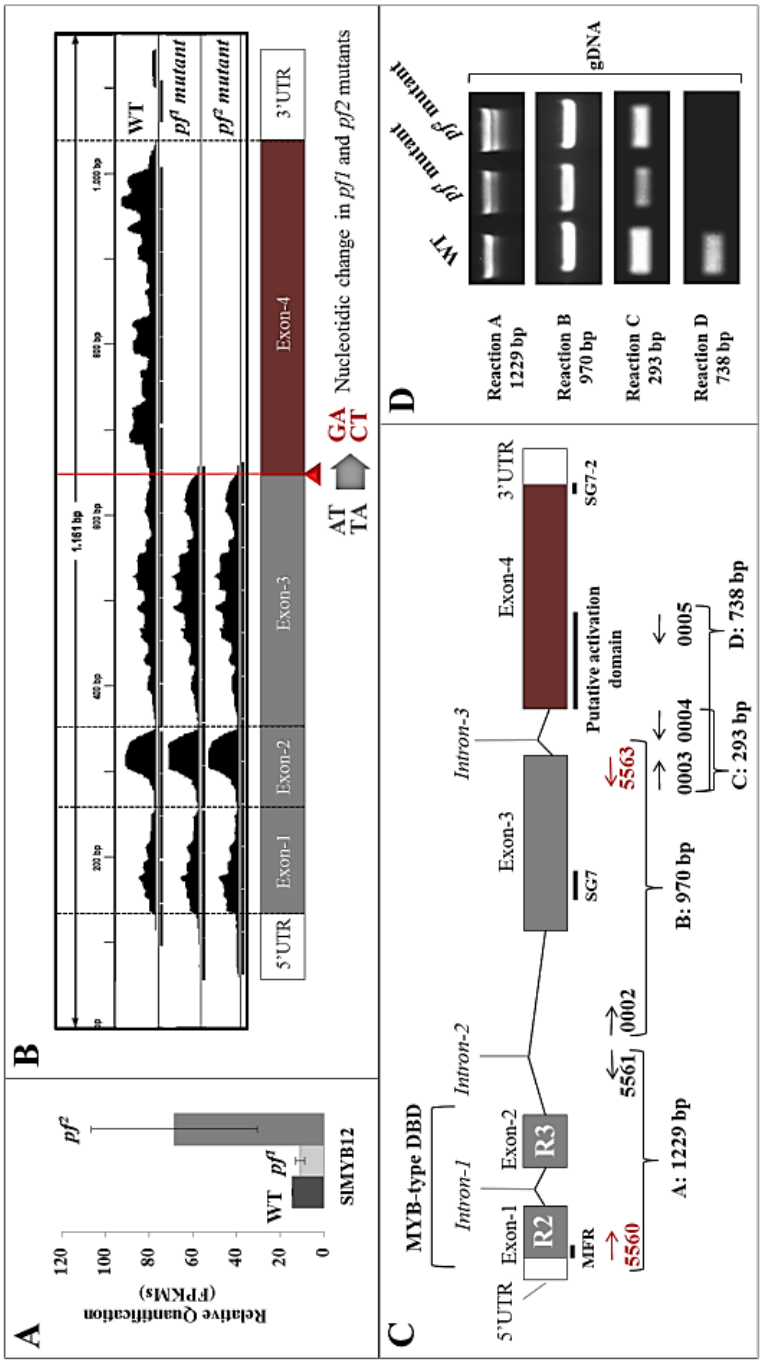

A
Relative Quantification (FPKMs)
120
100
80
60
40
20
0
pf2
WT pf1
SlMYB12
B
1,161 bp
WT
pf1 mutant
pf2 mutant
5'UTR
Exon-1
Exon-2
Exon-3
Exon-4
3'UTR
AT
TA
GA
CT
Nucleotidic change in pf1 and pf2 mutants
C
MYB-type DBD
Intron-1
Intron-2
Intron-3
5'UTR
Exon-1
Exon-2
Exon-3
Exon-4
3'UTR
R2
R3
MFR
SG7
SG7-2
Putative activation domain
5560
5561
0002
5563
0003
0004
0005
A: 1229 bp
B: 970 bp
C: 293 bp
D: 738 bp
D
VT
pf1 mutant
pf2 mutant
gDNA
Reaction A
1229 bp
Reaction B
970 bp
Reaction C
293 bp
Reaction D
738 bp

***Figure IV.5 Slmyb12 in the pf mutants. A**. The RNA-Seq-reads mapped among the SlMYB12 transcript sequence in mutants pf^1 and pf^2 (graphs from the IGV software; Robinson et al., 2011). Nucleotide changes (G:C → A:T) (position 661) and (A:T → T:A) (position 662) are also represented. **B**. The Slmyb12-pf gene expression levels by RNA-Seq-FPKM measurements. **C**. SlMYB12 genomic sequence in the tomato database (http://www.solgenomics.net/) and the PCR oligonucleotides distribution in the gene sequence. **D**. SlMYB12 PCR amplicons using genomic DNA (gDNA) templates from the WT, pf^1 and pf^2 mutants obtained from the amplification of the gene region, which spanned exon-1 to exon-4. Abbreviations: DBD, DNA-binding domain; MFR, Myb-like flavonol regulation motif (Czemmel et al., 2009); SG7, subgroup 7 with the flavonol regulation motif (Kranz et al., 1998; Czemmel et al., 2009); SG7-2, subgroup 7-2 with the specific flavonol regulation motif (Czemmel et al., 2009). [previous page].*

For both the pf^1 and pf^2 mutants, we also observed a similar effect on flavonol glycosides accumulation and on several glycosyl transferases (**Figure IV.3**). These observations support the hypothesis that the truncated *Slmyb12-pf* transcript could produce a non functional protein and, therefore, the *pf* mutants represent a SlMYB12 knock-out mutant. The closest null mutant in SlMYB12 has been provided by Adato et al., (2009), who developed artificial microRNA to target *SlMYB12 (amiR-Slmyb12)* in the cv. MT background (Adato et al., 2009). The *amiR-Slmyb12* line showed low levels for three additional flavonol glycosides compared to the original *y* mutant, which only showed a single reduced flavonol glycoside (**Figure IV.3**) (Adato et al., 2009). However, neither the *y* mutant nor the *amiR-Slmyb12* transgenic line displayed the dramatic changes observed in the flavonol glycosides content detected herein for the *pf* mutants as described below.

The flavonoid biosynthetic pathway is severely altered in the *pf* mutants

Early flavonoid biosynthetic genes (EBGs), namely *CHS1*, *CHS2*, *CHI-like* and *F3H*, were considerably reduced in both the *pf* mutants. *F3'H*,

encoding FLAVONOID 3' HYDROXYLASE, was also down-regulated in both mutants. The F3'H enzyme was shown to hydroxylate EBGs products and to determine the quercetin/kaempferol flavonols ratio by hydroxylating dihydrokaempferol into dihydroquercetin (http://www.arabidopsis.org/) (**Figure IV.3**, **Table IV.S3A**). We also found that some late flavonoid biosynthetic genes (LBGs) were differentially expressed in the *pf* mutants (**Figure IV.3**, **Table IV.S3A**). These belonged to the flavonol biosynthetic branch (a strongly reduced *FLS-like* gene) and also to the anthocyanin biosynthetic branch (a strongly reduced *DIHYDROFLAVONOL REDUCTASE* (*DFR*) and a slightly increased *ANTHOCYANIDIN SYNTHASE* (*ANS*)). Finally, four flavonoid GLYCOSYL TRANSFERASES (GTs), likely involved in the decoration of the flavonoid aglycones, showed strongly reduced expression levels in the *pf* mutants (**Figure IV.3**, **Table IV.S3A**) including three FLAVONOL 3-GTs (*F-3GT*, QUERCETIN 3-GT (*Q-3GT)* and *Q-3/4GT*) and an anthocyanin 3GT (*RT*). We also found three additional GTs that considerably reduced in the *pf* mutants (*UDP-GT2*, *UDP-GT3* and *UDP-GT-like*) and could be related with the decoration of flavonoids before their transport to the vacuole (**Figure IV.3**, **Table IV.S3A**).

The above-described EBGs (*CHS1*, *CHS2*, *CHI-like*, *F3H)* and LBGs in the flavonol biosynthetic branch *(F3'H*, *FLS-like* and *F-3GT)* were also down-regulated in the previously described *y* lines (Adato et al., 2009; Ballester et al., 2010). The Arabidopsis *myb11 myb12 myb111* triple mutant, also showed a co-down-regulation of the homologous genes for *CHS2* (*At5G13930*), *CHI-like* (*At5g05270*), *F3H* (*At3g51240*), *FLS-like* (*At5g08640*) and *F3GT* (*At5g17040*) (Stracke et al., 2007). This finding supports the results observed in the tomato *pf* mutants. Moreover, we found that the *F3'H* (*At5g07990*) and the *UDP-GT2* (*At2g22590*) genes were co-down-regulated in Arabidopsis *myb11 myb12 myb111* triple mutant (Stracke et al., 2007) in the same way as in the *pf* mutants, but not in the *y* lines (Adato et al., 2009; Ballester et al., 2010). Thus we propose that F-3GT could

be the GT involved in the decoration of EBG products as it was differentially expressed in both the *pf* mutant and *y* lines in which the EBGs glycosylated products were also found to be reduced (Adato et al., 2009; Ballester et al., 2010). Moreover, Q-3GT, Q-3/4GT, UDP-GT2 may possibly be the GTs involved in the decoration of LBG products in the flavonol biosynthetic branch as they were specifically down-regulated in the *pf* mutants.

Interestingly, the anthocyanin LBGs, DFR and RT, have also been reported as differentially expressed (co-down-regulated) in the *y* mutant and in the *pf* mutants (Adato et al., 2009). Yet *DFR* (*Solyc02g089770*) was annotated as a *CCR-like* given its similarity with a *CCR-related* gene in arabidopsis, and it was only recently annotated as *DFR* in tomato databases (http://www.solgenomics.net/; **Figure IV.S2**). It is noteworthy that no anthocyanin LBG was co-down-regulated in the Arabidopsis *myb11 myb12 myb111* triple mutant (Stracke et al., 2007).

These results support SlMYB12 TF as the main regulator to control flavonol biosynthesis and accumulation in tomato fruit skin by regulating both EBGs and LBGs, which thus confirms and extends previous reports (Luo et al., 2008; Adato et al., 2009; Ballester et al., 2010). Furthermore, since flavonols accumulate as glycosylated forms, the detection of several down-regulated GTs in the *pf* mutants suggests that this step is also regulated by SlMYB12 in tomato fruit.

Genes that act up- and down-stream of the flavonoid biosynthetic branch are also affected in the new *pink fruit* mutant

As described previously for *y* mutant fruit (Adato et al., 2009; Ballester et al., 2010), we observed alterations in the expression of the genes that belong to the PP pathway upstream of the flavonoid biosynthetic branch in the *pf* mutants. Thus we found nine *PAL* homologous genes and a *4CL3* gene, which were all down-regulated and belonged to the core PP biosynthetic pathway (**Figure IV.6**, **Table IV.S3A**). However, the *C4H* gene was not differentially expressed in either the *pf* mutants or the *y* mutant fruit

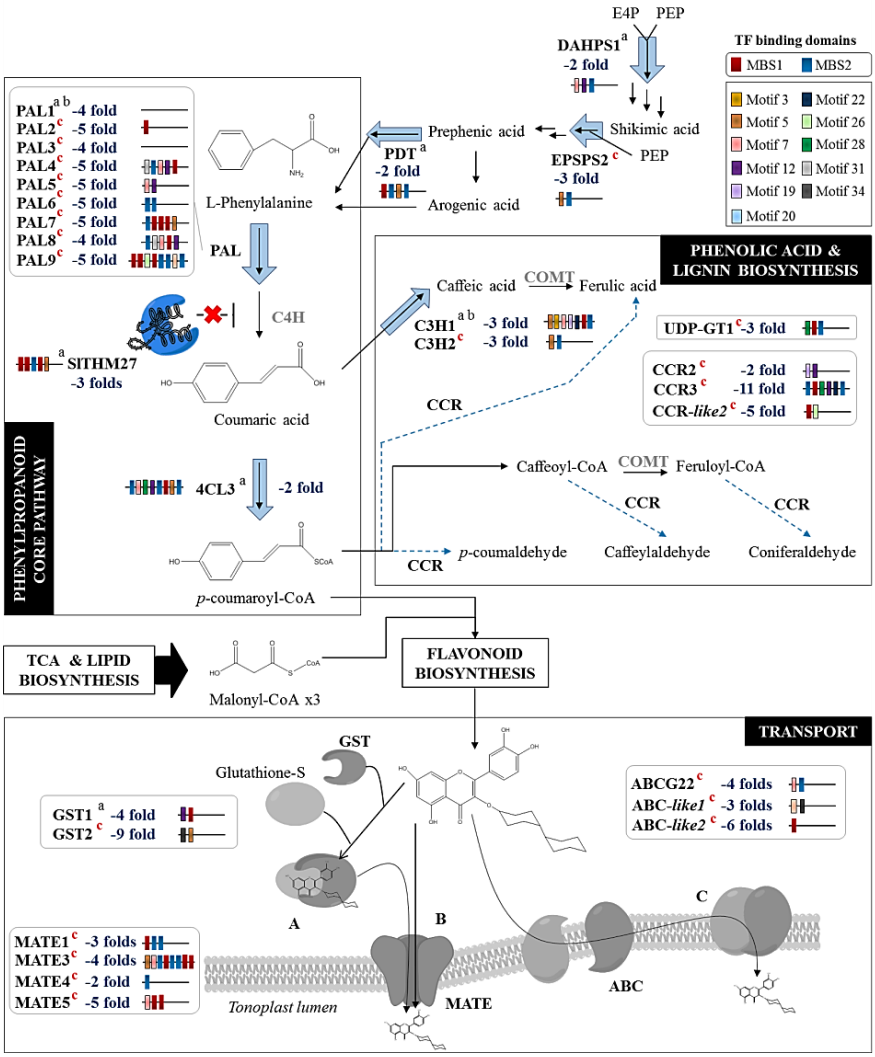
(Adato et al., 2009). In Arabidopsis, *C4H* gene expression is repressed by the AtMYB4 transcriptional repressor (Hemm et al., 2001). Adato et al., (2009) described AtMYB4 as the orthologe of the tomato SlTHM27 (*Solyc10g055410*) TF, and observed that *SlTHM27* was down-regulated in the *y* mutant and proposed that it is a SlMYB12 target gene. In support of this hypothesis, we also found *SlTHM27* to be down-regulated in the *pf* mutants (**Figure IV.6**, **IV.S3D**). However, the down-regulation of *SlTHM27* in both *pf* mutants had not effect over *C4H* gene expression and suggested other role for this regulator in tomato fruit.

We also found some genes that belonged to the hydroxycinnamate phenolic acid biosynthetic branch (Liu et al., 2015) to be differentially expressed in the *pf* mutants (**Table IV.S3A**). Most of these genes were down-regulated (**Figure IV.6**), but we also found a single gene to be up-regulated, namely *CCR-like4* (**Table IV.S3A**). AtMYB12 was proposed by Luo et al. (2008) as the regulator of the gene expression for genes involved in the production of caffeoylquinic acids and other phenolic compounds derived from hydroxycinnamic acid (Luo et al., 2008). However, our results do not support that SlMYB12 regulates the caffeoylquinic acid biosynthetic branch in tomato fruit. In both the pf^{1} and pf^{2} mutants, we only found a few polyphenols that are derived from the hydroxycinnamate phenolic acid biosynthetic branch. This could reveal a different specificity between the AtMYB12 and SlMYB12 TFs beyond the flavonol biosynthetic branch, or a consequence of the different approaches used in each case (knock out vs. overexpression).

We also detected reduced levels of *trans*-resveratrol, an additional polyphenol derived from the PP pathway. The enzymes *STILBENE SYNTHSE* (*STS*) and *CHS* are two members of the *CHS* superfamily of the III Polyketide synthases that catalyze similar condensation reactions, but different cyclizations, to produce resveratrol or naringenin chalcone, respectively (**Figure IV.S3B**) (Yamaguchi et al., 1999; Yu et al., 2012;

Stewart Jr. et al., 2013). These enzymes appear to be promiscuous in a way that STS can produce small amounts of naringenin chalcone while CHS can generate small amounts of resveratrol (Yamaguchi et al., 1999). Possibly, resveratrol production in tomato fruit can be catalyzed by a protein product of one of the down-regulated *CHS* homologous genes (**Figure IV.S3C**). As *trans*-resveratrol was also reduced in the *y* mutant, where only *CHS1* was found to be down-regulated (Adato et al., 2009), we propose *CHS1* to be the enzyme involved in *trans*-resveratrol production in tomato fruit. However, further studies are required to confirm this hypothesis.

Figure IV.6 The genes upstream and downstream of flavonol biosynthesis were affected in both the new pink fruit mutants. *The figure represents the shikimate, arogenate, core phenylpropanoid, phenolic acid and lignin pathways upstream of flavonol biosynthesis, and flavonoid transport. Differentially down-regulated (blue) genes are shown by arrows (the thicker the arrow, the higher the fold change). The transcription factor (TF) binding domains found in their promoters were also included (in color boxes). The genes in gray were not differentially expressed in the pf mutants. Genes and metabolites differentially expressed also in the y mutant (Adato et al., 2009) and in other y-lines (Ballester et al., 2010), are marked a and b, respectively; and those differentially expressed only in the pf mutants are marked c (in red). Abbreviations: E4P, ERITHROSE 4-PHOSPHATE; PEP, PHOSPHOENOL PYRUVATE; DAHPS, 3-DEOXY-D-ARABINOHEPTULOSONATE 7-PHOSPHATE SYNTHETASE; EPSPS, 5-ENOLPYRUVYLSHIKIMATE 3-PHOSPHATE SYNTHETASE; PDT, PREPHRENATE DEHYDRATASE; PAL, PHENYLALANINE AMONIO LYASE; C4H, CINNAMATE 4-HYDROXYLASE; 4CL, 4-COUMARATE CoA LIGASE; C3H, P-COUMAROYL 3-HYDROXYLASE; CCR, CINNAMOYL CoA REDUCTASE; COMT, CAFFEIC ACID O-METHYLTRANSFERASE; HCT, CINNAMOYL CoA TRANSFERASE; HQT, HYDROXYCINNAMOYL CoA QUINATE TRANSFERASE; TCA, TRICARBOXYLIC ACID CYCLE; UDP-GT, GLUCURONOSYL TRANSFERASE. [next page].*


E4P
PEP
DAHPS1 a
-2 fold
TF binding domains
MBS1
MBS2
Motif 3
Motif 22
Motif 5
Motif 26
Motif 7
Motif 28
Motif 12
Motif 31
Motif 19
Motif 34
Motif 20
PAL1 a b -4 fold
PAL2 c -5 fold
PAL3 c -4 fold
PAL4 c -5 fold
PAL5 c -5 fold
PAL6 c -5 fold
PAL7 c -5 fold
PAL8 c -4 fold
PAL9 c -5 fold
L-Phenylalanine
Prephenic acid
PDT a
-2 fold
Arogenic acid
Shikimic acid
EPSPS2 c
-3 fold
PEP
PAL
C4H
a
SITHM27
-3 folds
Coumaric acid
PHENOLIC ACID & LIGNIN BIOSYNTHESIS
Caffeic acid
COMT
Ferulic acid
C3H1 a b -3 fold
C3H2 c -3 fold
UDP-GT1 c -3 fold
CCR2 c -2 fold
CCR3 c -11 fold
CCR-like2 c -5 fold
CCR
PHENYLPROPANOID CORE PATHWAY
4CL3 a
-2 fold
Caffeoyl-CoA
COMT
Feruloyl-CoA
CCR
CCR
p-coumaldehyde
Caffeylaldehyde
Coniferaldehyde
CCR
p-coumaroyl-CoA
TCA & LIPID BIOSYNTHESIS
Malonyl-CoA x3
FLAVONOID BIOSYNTHESIS
TRANSPORT
GST
Glutathione-S
GST1 a -4 fold
GST2 c -9 fold
ABCG22 c -4 folds
ABC-like1 c -3 folds
ABC-like2 c -6 folds
A
B
C
MATE1 c -3 folds
MATE3 c -4 folds
MATE4 c -2 fold
MATE5 c -5 fold
Tonoplast lumen
MATE
ABC

Upstream of the PP biosynthetic pathway, we also detected differentially expressed genes in the shikimate pathway upstream of PP biosynthesis (**Figure IV.6**, **Table IV.S3A**). The two steps of the shikimate pathway downstream of phosphoenol pyruvate (PEP) were down-regulated in the *pf* mutants (steps catalyzed by the 3-DEOXY-D-ARABINOHEPTULOSONATE 7-PHOSPHATE (DAHP) and 5-ENOLPYRUVYLSHIKIMATE 3-PHOSPHATE (EPSP) synthases). Our results suggest that the alternative synthesis of L-Phe by the phenylpyruvate pathway (Tzin et al., 2009) in the *pf* mutants could be partially blocked, while the main biosynthetic pathway through arogenate (Maeda et al., 2011) is active and produces L-Phe (**Figure IV.6**, **Table IV.S3A**). Moreover, the down-regulation of the 9 *PAL* genes in the *pf* mutants (more than half of the 16 annotated *PAL* genes in tomato; http://www.solgenomics.net/) could explain why we found higher levels of L-Phe in the *pf* tomato skin as compared to wild-type fruit. This suggests that in tomato fruit, SlMYB12 likely regulates PP upstream pathways to ensure the supply of precursors to feed the flavonol biosynthetic pathway.

We found that several of the differentially expressed genes are likely involved in flavonoid glycosides transport to the vacuole (**Figure IV.6**, **Table IV.S3A**). These included several *GSTs* (*GLUTATHIONE S-TRANSFERASES*), *MATE-type* (*MULTI-ANTIMICROBIAL EXTRUSION*) and *ABC-type* (*ATP-BINDING CASSETTE*) transporter genes, which were mostly down-regulated (**Figure IV.6**, **Table IV.S3A**). We also detected a differentially down-regulated gene that could also be involved in the regulation of flavonoid accumulation (**Table IV1**), a homolog to *SlMYB12* which we termed *SlMYB12-like2* (*SlMYB12L2*, *Solyc12g04930*). The gene structure of *SlMYB12L2* is closer to the Arabidopsis *AtMYB12* than to *SlMYB12* as it contains three and not four exons (http://www.solgenomics.net/) (**Figure IV.S4B**).

Additional processes likely affected in the *pink fruit* mutant

Of the 312 differentially expressed genes in the *pf* mutants, 47 genes are likely involved in PP biosynthesis. The remaining 265 genes related to different processes, which include primary metabolism (i.e. photosynthesis, carbohydrate or energy metabolism) (**Figure IV.S5** and **Table IV.S3B**), secondary metabolism other than PP (**Table IV.S3B**), stress response (**Table IV.S3C**), transcriptional and hormonal regulations (**Table IV.S3D**), and additional cellular processes (i.e. signal transduction, DNA or protein metabolism) (**Table IV.S3E**). The genes involved in similar processes, other than flavonoid biosynthesis, have already been described for the tomato *y* mutant (Adato et al., 2009) and for the Arabidopsis *myb11 myb12 myb111* triple mutant (Stracke et al., 2007). However, the molecular characterization of the *pf* mutants by RNA-Seq allowed us to characterize the gene expression program with a higher resolution (Zhao et al., 2014), as compared to the previously used gene expression analysis platforms (Stracke et al., 2007; Adato et al., 2009). For example, we observed that genes associated with sugar and amino acid production and cell wall biosynthesis and developmental altered in the *pf* mutants could be related to fruit ripening (Giovannoni, 2004). Thus the *pf* mutants might prove to be a new opportunity to study the role of SlMYB12 TF during fruit ripening, as well as in the homeostasis between primary and secondary metabolisms.

Since flavonols are important for fruit cuticle biomechanics and post-harvest behavior (Domínguez et al., 2009a and 2011; Lara et al., 2014), we investigated if there were any genes differentially expressed in the *pf* mutants that were related with cuticle development. Among these, we identified a set of 18 genes which are involved in the biosynthesis of cutin monomers, cuticular waxes (i.e., alkanes and esters) and triterpenoids (especially in amyrin biosynthesis) (**Table IV.1**, **Table IV.S3B**). Interestingly, most of these genes were up-regulated in the *pf* mutants. These results support the phenotype observed in the breaker stage for the *y* mutant

(Adato et al., 2009) which exhibited altered wax accumulation patterns (i.e., increased levels of alkanes and amyrins), as well as some quantitative changes in cutin composition.

A key role of *MYB12* in biotic (i.e., fungi and bacterial infections) and abiotic (i.e., phosphate, nitrogen, light, temperature, UV or drought) stress responses in Arabidopsis has been recently reported (Nakabayashi et al., 2014; Schenke et al., 2014; Zoratti et al., 2014). We identified a set of 19 genes that were differentially expressed in the *pf* mutants and could be involved in these processes (**Table IV.1**, **Table IV.S3C**). Furthermore, the down-regulated genes (e.g., that encode defensin, α-dioxygenase, peroxidases, etc) (**Table IV.1**) are more likely to be involved in biotic stress responses, while the up-regulated genes (e.g., that encode heat shock proteins, dehydrins, cold-induced proteins, etc.) (**Table IV.1**) are possibly involved in abiotic stress responses. In addition, it has been described that flavonoid glycosides themselves are involved in several stress response mechanisms (Nakabayashi et al., 2014), and we observed that these compounds were reduced in the *pf* mutants (**Figure IV.3**, **Table IV.S1**). We thus suggest that SlMYB12 could be playing a role in the tomato fruit stress response.

Identification of the putative target genes of SlMYB12 regulator

Two different "*in silico*" methods were used to scan the upstream promoter sequence regions of the 152 down-regulated genes in the *pf* mutants to identify the possible direct targets of the SlMYB12 transcription factor.

We first scanned the promoter sequences for two different MYB Binding Sites (MBS) motifs that have been previously reported for arabidopsis *AtMYB12* , and for its functionally redundant TFs *AtMYB11* and *AtMYB111* (Hartmann et al., 2005; Mehrtens et al., 2005; Franco-Zorrilla et al., 2014). These MBSs were: (i) MBS1, (T/C)AAC(T/A)A(A/C)C; and (ii) MBS2, ACCTA(A/C). We observed similar nucleotide frequencies for the

MBS1 and MBS2 consensus sequences to the ones described for the arabidopsis *myb11 myb12 myb111* triple mutant (Franco-Zorrilla et al., 2014). Out of the 152 down-regulated genes, 88 genes containing the MBS1 motif (~66%) in their promoter sequences, showed the following consensus sequence [**C**(73%)/**T**(67%)]**AAC**[**A**(81%)/**T**(59%)]**A**[**A**(84%)/**C**(56%)]**C**. While the remaining 69 differentially down-regulated genes whose promoters contained the MBS2 motif, showed the consensus sequence **ACCTA**[**C**(47%)/**A**(46%)]. Thus, the consensus sequences obtained for the MBS1 and MBS2 motif were (**T**/**C**)**AACAAAC,** [(**T**≈**C**)**AAC**(**A**>**T**)**A**(**A**>>>**C**)**C**], and **ACCTA**(**A**/**C**)**C**, [**ACCTA**(**A**≈**C**)**C**], respectively (**Figure IV.7B**).

Most of the promoters of the co-down-regulated genes in the arabidopsis *myb11 myb 12 myb111* triple mutant were reported to contain an MBS2 motif, while a small set of genes contained the MBS1 motif (Franco-Zorrilla et al., 2014). In tomato fruit skin, 125 of the 152 down-regulated genes in the *pf* mutants contained at least one of the two MBS motifs (**Figure IV.7A**). Yet unlike arabidopsis, the number of down-regulated genes whose promoters contained the MBS1 motif was higher than those that contained the MBS2 motif (88 and 69, respectively). Furthermore, the promoters of 42 of these 125 genes contained both the MBS1 and MBS2 motifs (**Figure IV.7A**).

It has been proposed that the presence of several MBS2 motifs in promoters indicates primary target genes of the MYB11, MYB12, MYB111 TF in arabidopsis (Franco-Zorrilla et al., 2014). Moreover, those genes with a single MBS2 or MBS1 motif could represent secondary targets of these TFs (Franco-Zorrilla et al., 2014). However in the *pf* mutant tomato fruits, we found that most of the 125 genes that contained an MBS motif exhibited one or two MBS1, and/or the MBS2 motif in their promoters, and very few genes (less than 20%) showed several MBS1 and/or the MBS2 motifs (**Figure IV.8A**). It is possible that SlMYB12 TF in tomato fruit does not require

several MBS motifs to regulate the gene expression of its primary target genes. Another possibility is that other MBS motifs that differ from those being investigated are recognized by SlMYB12, or that other functional homologs of SlMYB12 could act coordinately with SlMYB12 (Grotewold, 2004).

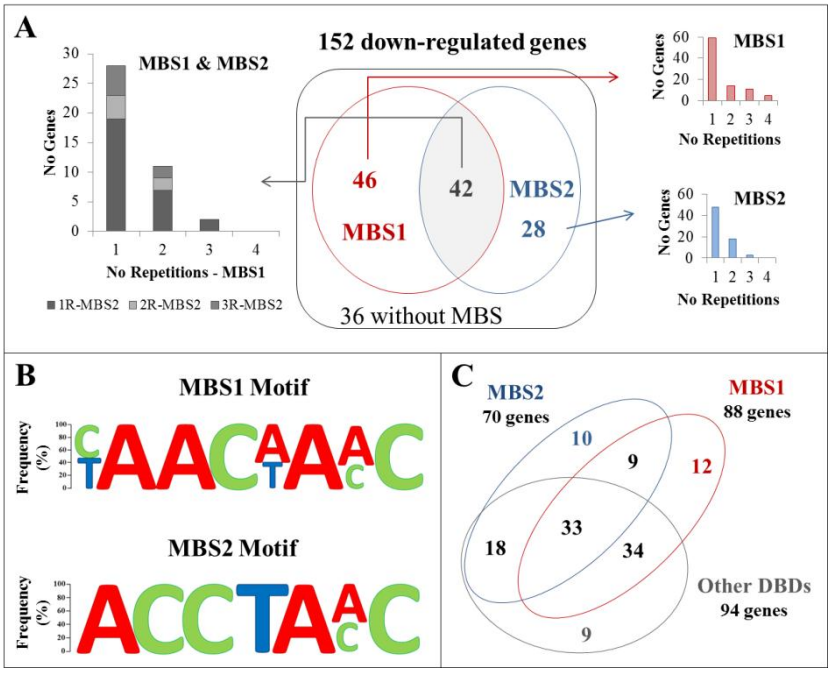


Figure IV.7 Transcription factor recognition elements for the SlMYB12 candidate target genes in both the new pink fruit mutants detected by an 'in silico' analysis. A. *Venn diagram representing the Myb-binding sites recognized by SlMYB12 TF in the 152 down-regulated genes set in the pf mutants. 125 of these 152 genes contained Myb-binding sites (MBS1 and MBS2). Bar graphs represent the amount of genes whose promoters contained one or more than one MBS1 (in red), MBS2 (in blue), or both MBS (in gray), motifs.* ***B.*** *Consensus sequences MBS1 and MBS2 for the candidate genes in the pf tomato fruit mutants.* ***C.*** *Venn diagram showing the down-regulated genes that contained different DNA-binding domains (DBDs) recognized by distinct transcription factors, obtained by tools MEME (Bailey et al., 1994) and TOMTOM (Gupta et al., 2007).*

Table IV.1 Genes differentially expressed putatively involved in PP regulation, cuticle specialized metabolism biosynthesis and stress response in the pf mutants. *Down-regulated genes are highlighted in gray. The number of MBS motifs in the gene promoters was represented by "xNo" (e.g. $_{x1}$ or $_{x4}$). MEME-motifs were represented by "MNo" (e.g. M5 or M7) (**Table IV.2**, **Table IV.S4**).* † - Shared by *pf*[1], *pf*[2] and *y* mutants.

Tomato Annotation	Short name	Gene ID	Fold-Change *pf1* $\log_2$	Fold-Change *pf2* $\log_2$	SlMYB12 MBS	MEME motif
Transcription factors putatively involved in phenylpropanoid biosynthesis						
MYB	THM27 [AtMYB4]†	*Solyc10g055410*	-3.72954	-2.65141	$MBS1_{x3}$ $MBS2_{x1}$	M5
MYB12	MYB12L2 [AtMYB12/ ZmMYBP]	*Solyc12g049350*	-1.798 e+308	-1.798 e+308	$MBS2_{x1}$	-
Fatty acid synthesis and elongation (plastid)						
3-oxoacyl-(ACP) synthase2	**KAS 2** [AtKASI]	*Solyc02g070790*	2,21624	2,36354	-	-
3-oxoacyl-ACP-reductase	**KAR**	*Solyc06g071910*	2,03983	2,0442	-	-
(3R)-hydroxymyristoyl, β-hydroxyacyl-(ACP) dehydratase FabZ	**HAD**	*Solyc01g105060*	2,20293	2,08279	-	-
Enoyl-[ACP] reductase (NADH)	**ER** [AtENR1]†	*Solyc10g078740*	4,33897	4,45364	-	-
Long-chain-fatty-acid CoA ligase (AMP-dependent)	**LACS**	*Solyc09g092450*	-2,14828	-1,48707	$MBS1_{x1}$	M7
Very-Long-Chain fatty acid synthesis and elongation (endoplasmic reticulum)						
Very-long-chain-fatty acid 3-ketoacyl CoA synthase	[At KCS3]	*Solyc11g072990*	2,07499	2,01566	-	-
Very-long-chain-fatty acid 3-ketoacyl CoA synthase	[At KCS11]	*Solyc09g083050*	2,67897	4,44723	-	-
Very-long-chain-fatty acid 3-ketoacyl CoA synthase	[At KCS11]	*Solyc08g067410*	2,48575	2,33948	-	-
Cuticular wax n-alkanes biosynthesis (endoplasmic reticulum)						
CER1 protein Fragment (Fatty acid hydroxylase, FAH)	**CER1** [At CER3/ WAX2]	*Solyc03g117800*	2,41359	3,4569	-	-
Cuticular wax esters biosynthesis (endoplasmic reticulum)						
O-acyltransferase WSD1	**WSD1**	*Solyc07g053890*	2,55808	3,47231	-	-
Cutin monomer biosynthesis (endoplasmic reticulum)						
Glycerol-3-phosphate dehydrogenase	[AtGPDHC1] †	*Solyc04g016330*	2,5422	2,26198	-	-
ER glycerol-phosphate (phospholipid/glycerol) acyltransferase	[AtGPAT6]	*Solyc02g087500*	2,83586	2,45514	-	-
Glycerol-3-phosphate acyltransferase 4	[AtGPAT4]	*Solyc01g094700*	2,264	2,09214	-	-
Fatty acid desaturation (endoplasmic reticulum)						
Stearoyl-acyl carrier protein (ACP) desaturase (Fatty acid desaturase 2)	**FAD2**	*Solyc03g063110*	1,67668	2,00451	-	-
Pentacyclic triterpenoid biosynthesis (cytosol)						
Squalene monooxygenase/ epoxidase (triterpenoid biosynthesis)	[AtXF1]	*Solyc04g077440*	2,72066	2,91864	-	-
General pathogen infection						
Major allergen Mal d 1, Bet v I allergen; Pathogenesis-related protein	[StSTH-2]†	*Solyc09g090980*	-2,36295	-3,03614	$MBS1_{x1}$ $MBS2_{x1}$	M5 M7
Ammonium transporter	[AtAMT2]	*Solyc10g076480*	-3,56048	-3,39429	$MBS1_{x2}$ $MBS2_{x1}$	M5 M12 M20
Defensin protein, γ Purothionin		*Solyc07g007750*	-6,16815	-6,78755	$MBS1_{x1}$; $MBS2_{x1}$	-
Defensin protein, γ		*Solyc07g007760*	-7,4657	-6,52064	-	-

…continues

Fungi infection						
Mildew resistance locus O-like protein 4	**MLO-like4**	*Solyc10g044510*	-3,9483	-3,4251	$MBS1_{x1}$ $MBS2_{x1}$	M12
putative resistance receptor like protein with an antifungal domain	**RLK-like**	*Solyc02g080080*	-2,35775	-1,7833	-	-
Bacterial infection						
α-dioxygenase (*oxylipin synthesis*)	[AtDOX1]	*Solyc02g087110*	-7,62299	-7,59258	$MBS1_{x2}$	-
General abiotic stress response						
Heat shock protein DnaJ chaperone		*Solyc01g090550*	3,52025	3,44625	-	-
Maleylacetoacetate isomerase/glutathione S-transferase	**GST1**†	*Solyc01g091330*	-4,63089	-4,25519	$MBS1_{x1}$	-
Glutathione S-transferase, C-terminal	**GST7**	*Solyc01g081250*	2,56747	3,12369	-	-
Hight Light Stress						
high light intensity/heat/hydrogen peroxide stresses	**CYPb561** [AtACYB-2]	*Solyc03g025840*	-2,62729	-3,27487	$MBS2_{x1}$	M5
RLK, receptor like kinase	[AtLight repressible RLK]	*Solyc03g121230*	-2,10248	-4,37655	-	-
RLK, Receptor like kinase		*Solyc12g014350*	-1,75973	-1,98744	$MBS1_{x1}$	M5
Drought & salt stress						
Dehydrin		*Solyc01g109920*	2,34412	2,86069	-	-
Glutathione S-transferase 3, C-terminal	**GST2** [At ERD9]	*Solyc05g006730*	-9,08275	-1.798e+308	-	M7 M4
ABC transporter G family member 22	**ABCG22**	*Solyc04g070970*	-4,11519	-5,17841	$MBS2_{x1}$	M7
Cold stress						
Cold shock protein-1	[AtCSDP1]	*Solyc03g033500*	3,03166	3,1725		
RNA recognition motif, glycine rich protein	[AtGR-RBP2]	*Solyc10g081180*	2,60452	2,71959		
RNA recognition motif, glycine rich protein	[AtGR-RBP2]	*Solyc05g053780*	3,8869	3,98303		
Oxidative stress						
Peroxidase, Suberization-associated anionic peroxidase precursor	**PERX**	*Solyc02g079500*	3,42092	3,69597		
Peroxidase, Suberization-associated anionic peroxidase precursor	**PERX**	*Solyc02g079510*	4,33888	5,62518		
Peroxidase	**PER** [AtCationic peroxidase]	*Solyc04g064690*	2,22658	3,18769		
Peroxidase 4	**PER4** [AtPER12]	*Solyc04g071890*	-2,53767	-2,3596	$MBS1_{x2}$	M12
Detoxification						
Heavy metal transport/detoxification protein		*Solyc09g008200*	4,57863	4,12865		
Multidrug resistance protein	**MATE1**	*Solyc03g025200*	-2,82269	-2,60382	$MBS1_{x1}$ $MBS2_{x2}$	-
Multidrug resistance protein [At ripening-responsive protein]	**MATE3**	*Solyc01g109320*	-4,40027	-3,25208	$MBS1_{x3}$ $MBS2_{x3}$	M5 M7
Multidrug resistance protein	**MATE4** [AtALF5]	*Solyc10g007100*	-2,55314	-2,31086	$MBS2_{x1}$	-
Multidrug resistance protein [At Ripening-responsive protein]	**MATE5**	*Solyc04g074840*	-2,00162	-4,20145	$MBS1_{x2}$	M7

However in tomato fruit, no gene was previously reported whose function could be functionally redundant with *SlMYB12* has been described, which supports the notion that SlMYB12 TF, as the gene that underlies the *y* and *Slmyb12-pf* mutant phenotypes by regulating EBGs and LBGs, is involved in the flavonol biosynthetic branch. A *SlMYB12* gene homolog, *SlMYB12-like* (*Solyc06g009710*), has been described earlier (Adato et al., 2009), whose expression levels were neither altered in the *y* mutant (Adato et al., 2009) nor in the new *pf* mutants. In contrast, we were able to identify a differentially expressed *SlMYB12* homolog, which we named *SlMYB12-like2* (*SlMYB12L2*, *Solyc12g04930*) (**Table IV.1**). The promoter of this TF contained an MBS2 motif, which suggests that it could be a target gene of SlMYB12. A co-expression analysis, done with the TOMEXPRESS database (Mohamed, Z. and Bouzayen, M., *unpublished*), suggested that SlMYB12L2 activity could relate more to cutin biosynthesis rather than to flavonol production.

In the second "*in silico"* method used to investigate other possible MBS motifs that could be recognized by SlMYB12 TF, we performed a non targeted search for DNA-binding domains (DBDs), including MYB and other different TF families binding sites. Using the online MEME (Bailey et al., 1994) and TOMTOM (Gupta et al., 2007) tools we examined the 2 Kb upstream sequences of the 152 down-regulated genes in the *pf* mutants. Overall, we detected 23 different TOMTOM plant-DBDs in 11 MEME-motifs (**Figure IV.S6**) in 94 genes (**Table IVS4**). These plant-DBDs were recognized by 10 different TF families that belonged to six TF classes (**Table IV.2**). When comparing both the MBS-containing and DBD-containing gene sets, we observed three different gene groups: (i) 85 genes that shared the MBS and DBD motifs; (ii) 31 genes that contained only the MBS motifs; and (iii) 9 genes that contained only the DBD motifs (**Figure IV.7C**). In **Figures IV.3** and **IV.6**, we represent MBS motifs and DBDs for the PP biosynthetic genes. Interestingly, we found three sets of genes to be putatively involved in

three different PP biosynthetic pathways: (i) genes that contained only MBS1/MBS2 motifs, and are likely involved in the biosynthesis of EBG glycoside products (**Figure IV.8A**); (ii) genes that contained MBS1/MBS2 motifs, together with the *MADS-like* repressor and/or the *SQUAMOSA-like* activator DBDs in MEME-motifs 7 and 12, putatively involved in the production of phenolic acids and LBG flavonols (**Figure IV.8B**); and (iii) genes that contained MBS1/MBS2 motifs, as well as the *MADS-like* repressor DBD into MEME-motif 5, putatively involved in the production of phenolic acids and anthocyanins (**Figure IV.8C**).

Finally, we also detected MBS motifs and DBDs in the genes that belonged to other metabolic and cellular processes (**Table IV.S5**).

Table IV.2 Transcription factor-binding elements by the MEME-TOMTOM combined analysis. *A combined MEME-TOMTOM analysis performed on 152 genes down-regulated in the pf mutant revealed 11 motifs that contained 23 plant DNA-binding domains recognized by 10 different transcription factors classes (**Figure IV.S6** and **Table IV.S4**).*

TF FAMILY	TOMTOM-DBD ANNOTATION	MEME MOTIF
β-Hairpin-Ribbon TF Class		
AP2/MBD-like	ABI4	26
Zinc-Coordinating TF Class		
β,β,α-Zinc finger	id1	3
Squamosa-Binding Protein [SBP]	SPL8	3, 7, 12, 19
Zipper-type TF Class		
Helix-Loop-Helix [HLH]	BES1 (repressor)	31
	BZR1 (repressor)	31
	MYC3	31
	MYC4	31
	PIF3	3
	PIL5 [PIF3-like]	3, 31
	PIF4	3, 31
	PIF5	3
Leucine Zipper	ABF1	3
	HY5	31
Helix-Turn-Helix TF Class		
Homeo	HB1	19
	HAT5	19
MYB	MYB84 [SlBLIND]	26
Other α-helix TF Class		
High Movility Group [HMG]	HMG-I/Y	7, 12, 19, 20, 34
MADS	FLC	3
	PI	3
	SEP3	3
	SVP [AGL22] (repressor)	3
	SOC1 [AGL20]	3, 5, 7, 12, 19, 34
AT-rich TF Class		
Arid-Bright	ARID3A	19, 22, 28, 34

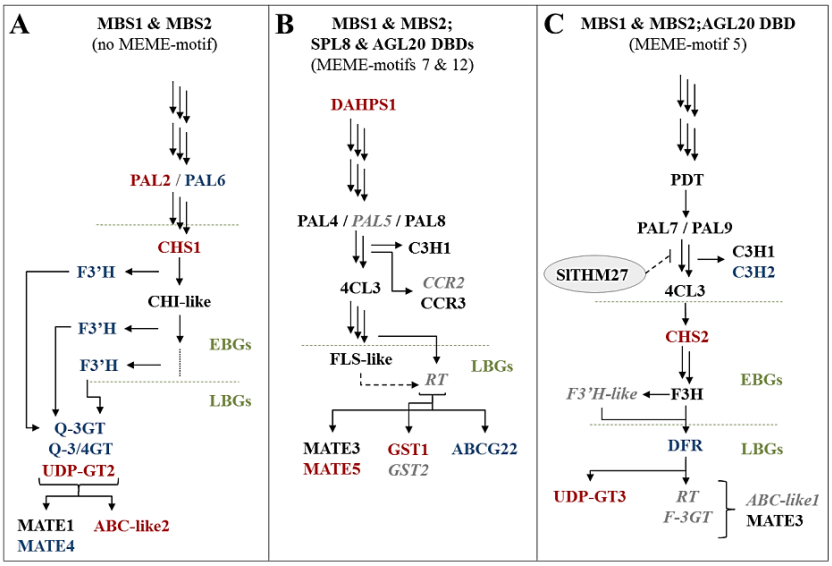


Figure IV.8 Three proposed mechanisms for controlling tomato fruit skin flavonoid biosynthesis based on the putative target genes of the SlMYB12 regulator. A. *Proposed biosynthetic regulation of EBG glycoside products. In this branch, the promoter sequences of genes contained only Myb-binding sites (MBS) are included.* ***B.*** *Proposed biosynthetic regulation of lignins and flavonols. Here genes contained MBS motifs and DNA-binding domains (DBDs) for AGAMOUS-like20 (AG-like20) and/or SQUAMOSA protein like 8 (SPL8) are included.* ***C.*** *Proposed biosynthetic regulation of coumarins and anthocyanins. Genes contained MBS motifs and AG-like20 DBD are incorporated. In A, B and C, the genes that contained the MBS1 motif are in red, MBS2 in blue, both the MBS motifs in black, and none in gray italics. EBG, early biosynthetic genes; LBG, late biosynthetic genes.*

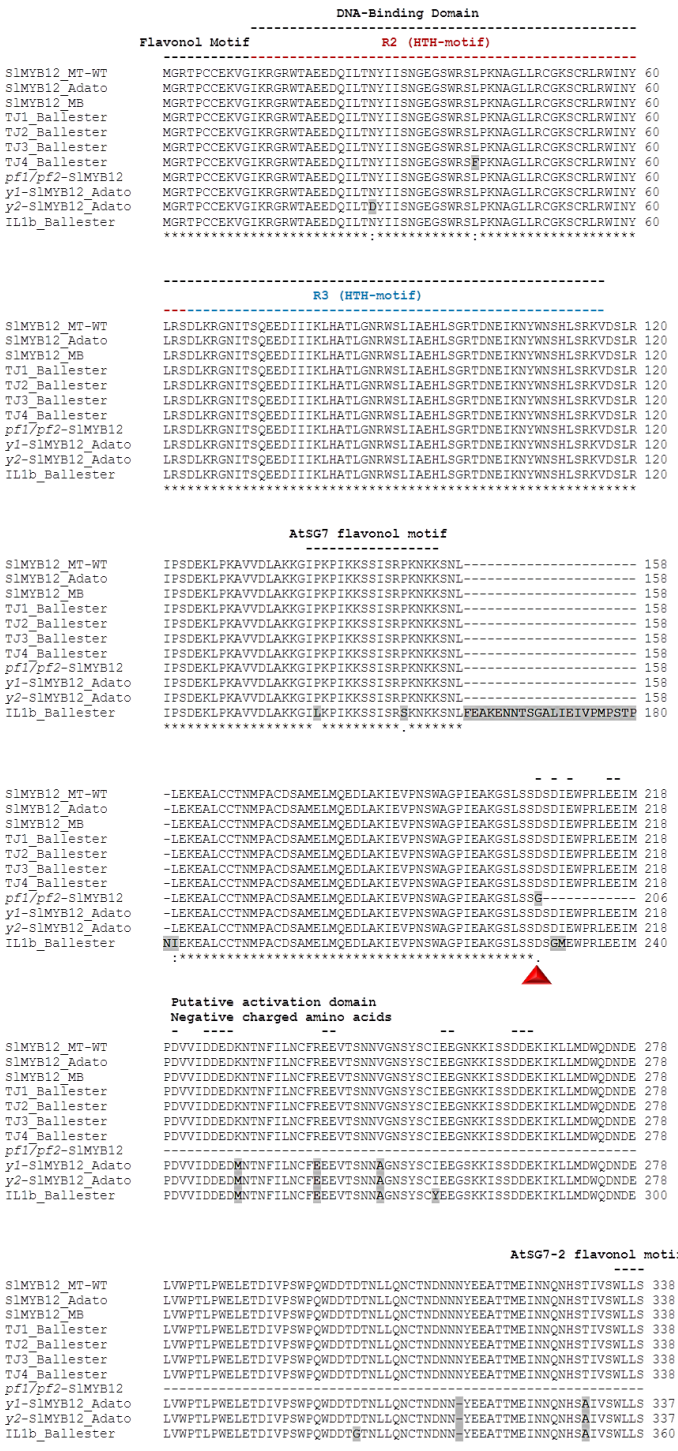
IV.4 CONCLUDING REMARKS

In this study we have characterized a null allele (*pf*) associated with the *pink fruit* phenotype in two new EMS tomato mutants. In contrast to the previously characterized *y* lines, the *pf* mutants accumulated normal levels of a truncated version of the *Slmyb12* transcript and did not accumulate flavonol glycosides in their fruit skin. Transcriptome analysis, allowed us to identify a larger set of genes that were differentially expressed in the *pf* mutants than those reported in the previously characterized *y* lines, and to examine the molecular program that underlies the *pf* mutation with higher resolution. These analyses provided us with new insights into the role of SlMYB12 in primary and secondary metabolism interaction/homoestasis. Finally, characterization of the *pf* mutants provided us with a set of candidate target genes of the SlMYB12 transcription factor even beyond the flavonol biosynthetic branch which will be subject for future in–depth characterization.

IV.5 ANNEXES IV

Additional Figures

Figure IV.S1 SlMYB12 protein alignment for the pf and the y mutants backgrounds. *The SlMYB12 protein was deduced from the consensus sequence obtained from the RNA-Seq-reads in the WT, pf^1 and pf^2 mutants. The Slmyb12 protein sequences were aligned with the consensus sequences obtained for the WT and y1-y2 alleles by Adato et al. (2009), together with the Moneyberg-WT (MB), TJ1-4 lines and IL-1b consensus sequences by Ballester et al., (2010). The amino acid changes are in gray. The red head arrow points to the truncation in the pf mutants. The three flavonol regulation motifs are marked over-lined in the picture: flavonol, AtSG7 and At SG7-2 motifs (Kranz et al., 1998; Stracke et al., 2007; Czemmel et al., 2009). The putative activation domain is also over-lined in the picture, and indicates the negatively charged amino acids. Alignment was performed with the free online EMBL-EBI software ClustalW 2.1 (http://www.ebi.ac.uk/Tools/msa/clustalw2/). [next page].*


DNA-Binding Domain
Flavonol Motif
R2 (HTH-motif)
R3 (HTH-motif)
AtSG7 flavonol motif
Putative activation domain
Negative charged amino acids
AtSG7-2 flavonol motif
SlMYB12_MT-WT
SlMYB12_Adato
SlMYB12_MB
TJ1_Ballester
TJ2_Ballester
TJ3_Ballester
TJ4_Ballester
pf1/pf2-SlMYB12
y1-SlMYB12_Adato
y2-SlMYB12_Adato
IL1b_Ballester

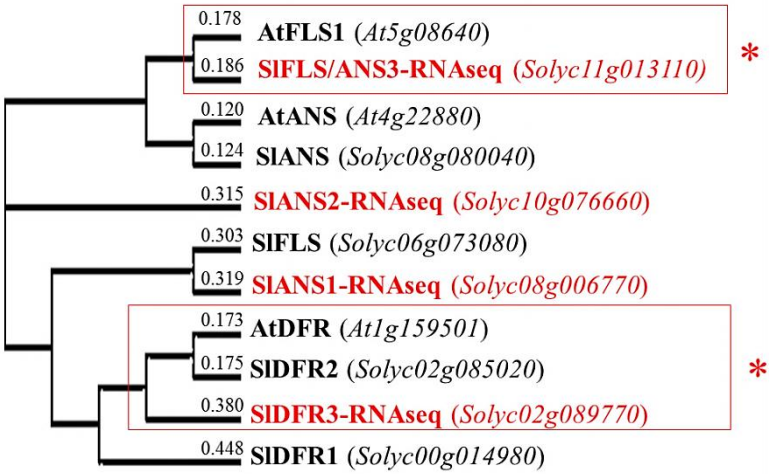


Figure IV.S2 FLS and DFR phylogenetic comparison for tomato (Solanum lycopersicum) and Arabidopsis thaliana. *The genes annotated as FLS, ANS or DFR in tomato (http://www.solgenomics.net/) and arabidopsis (http://www.arabidopsis.org/) are in black. The genes differentially expressed by RNA-Seq in the pf mutants with an ambiguous annotation are in red. The red squares, marked by an asterisk (*), highlight those selected functional annotations for ambiguous genes. Alignment was performed with the free online EMBL-EBI software ClustalW 2.1 (http://www.ebi.ac.uk/Tools/msa/clustalw2/).*

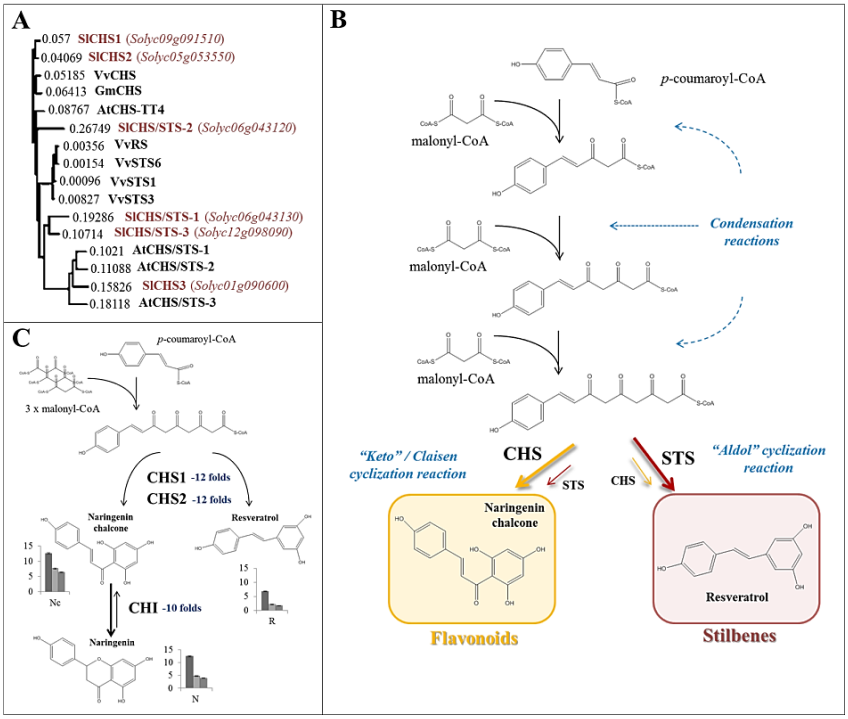


Figure IV.S3 Similarities between enzymes stilbene synthase and chalcone synthase. A. *Phylogenetic tree for enzymes stilbene synthase (STS) and chalcone synthase (CHS) from tomato (Solanum lycopersicum, highlighted in red), grapevine (Vitis vinifera), soybean (Glycine max) and Arabidopsis thaliana. The real distances of alignment re represented in the figure.* ***B.*** *Enzymes CHS and STS catalyze similar condensation reactions by a molecule of p-coumaroyl-CoA and three molecules of malonyl-CoA, but different cyclization reactions, to produce naringenin chalcone by a "keto" cyclization reaction or resveratrol by an "aldol" cyclization reaction, respectively.* ***C.*** *In the pf mutants, naringenin chalcone and resveratrol levels lowered at the same time as two CHS alleles (CHS1 and CHS2) were down-regulated. No STS was down-regulated in the pf mutants. Fold change per gene is provided. Levels of trans-resveratrol (R), naringenin chalcone (Nc) and naringenin (N) are expressed as the area under the peak (LC-MS approach). In the graphs, WT is represented in dark gray (left), the pf^1 mutant in light gray (middle) and the pf^2 mutant in gray (right). CHI, chalcone isomerase. Alignment was performed with the free online EMBL-EBI software ClustalW 2.1 (http://www.ebi.ac.uk/Tools/msa/clustalw2/).*

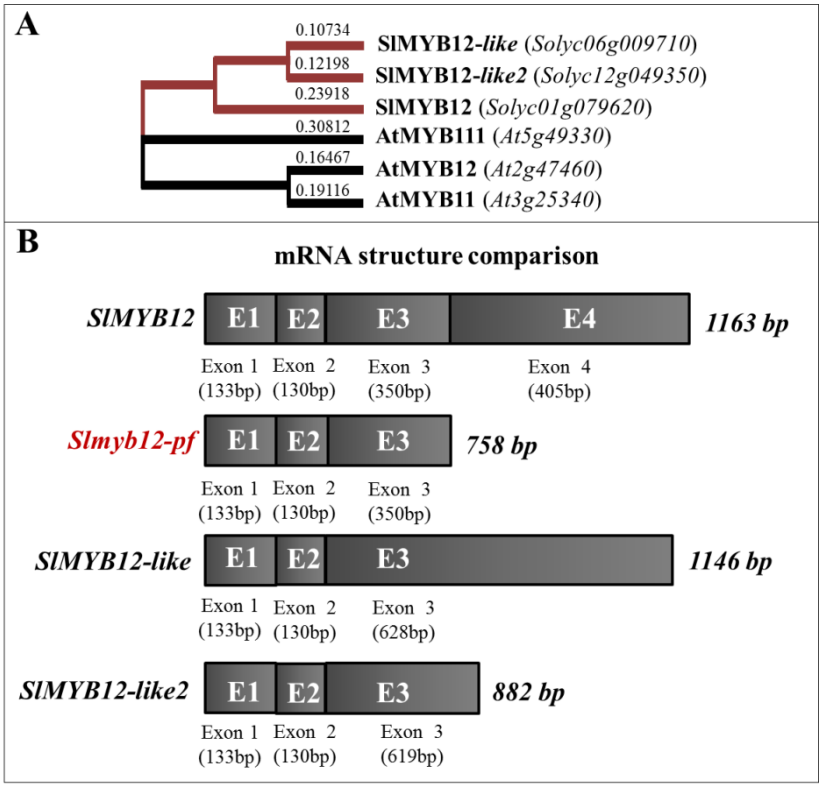


Figure IV.S4 SlMYB12, Slmyb12-pf, SlMYB12-like and SlMYB12-like2 proteins alignment. A. *Phylogenetic tree for the MYB12 proteins from tomato (Solanum lycopersicum) and arabidopsis (Arabidopsis thaliana).* ***B.*** *mRNA structures comparison between transcripts SlMYB12, Slmyb12- pf (in red), SlMYB12-like and SlMYB12-like2.*

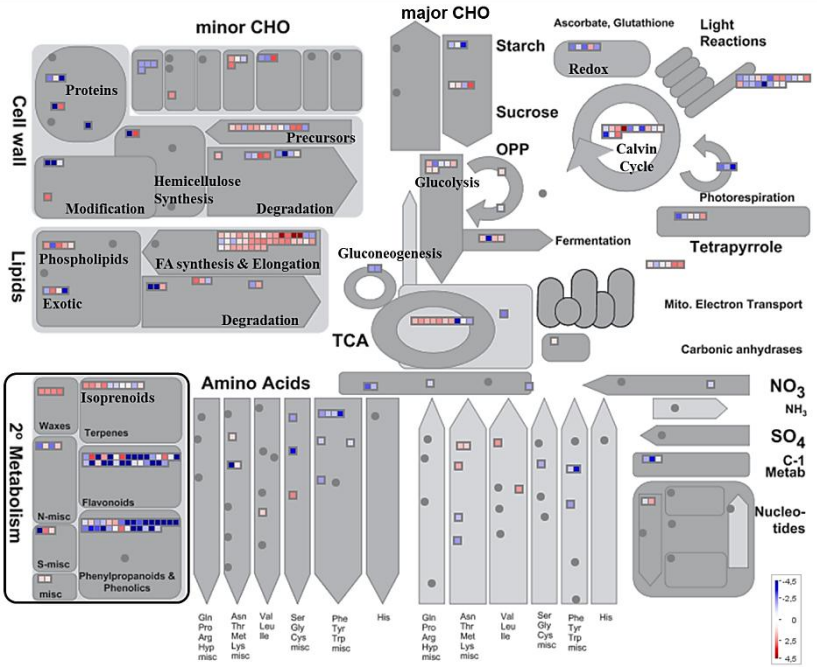


Figure IV.S5 General overview of the metabolic processes affected in the pf mutants. *Several genes in the primary and secondary metabolism (in the black square) were differentially expressed in the pf mutants. OPP, oxidative pentose phosphate; TCA, tricarboxylic acid cycle; CHO, carbohydrate metabolism. MapMan figure (http://mapman.gabipd.org/eb/guest/mapman/;http://www.gomapman.org/).*

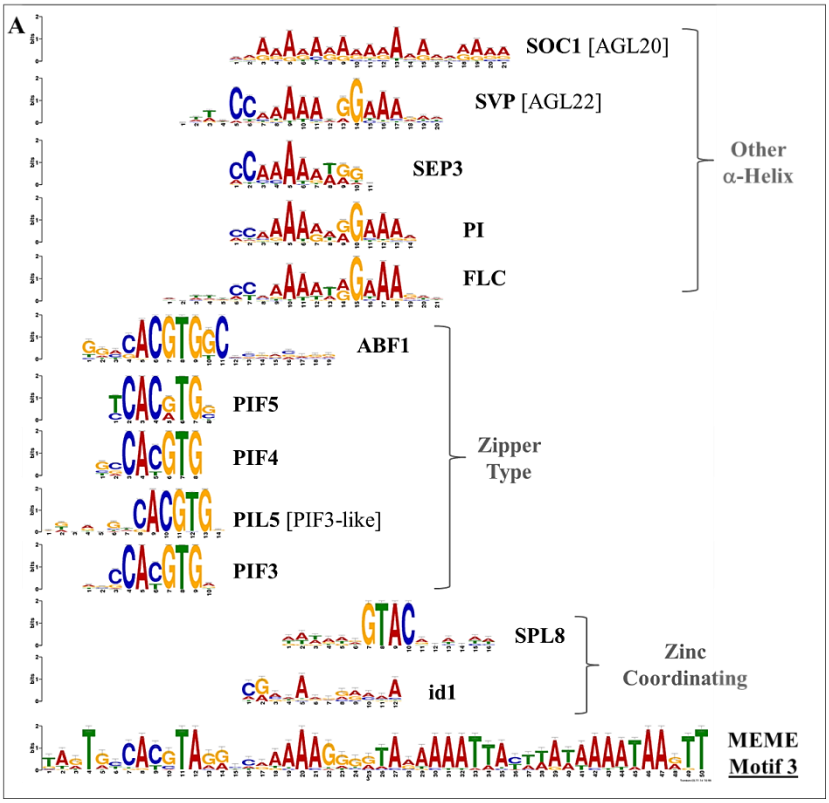


Figure IV.S6 Transcription factor-binding element consensus sequences. *The MEME-motif and TOMTOM plant-DNA-binding domain consensus sequences are provided. The location of each plant-DBD in each motif is shown. DBDs were grouped by classes in each motif. Those DBD found into MEME-motif 3 are in* ***A****, those into MEME-motifs 5, 20, 22, 26 and 28 in* ***B****, those into MEME-motifs 7, 12 and 34 in* ***C****, and those into MEME-motifs 19 and 31 in* ***D****.*

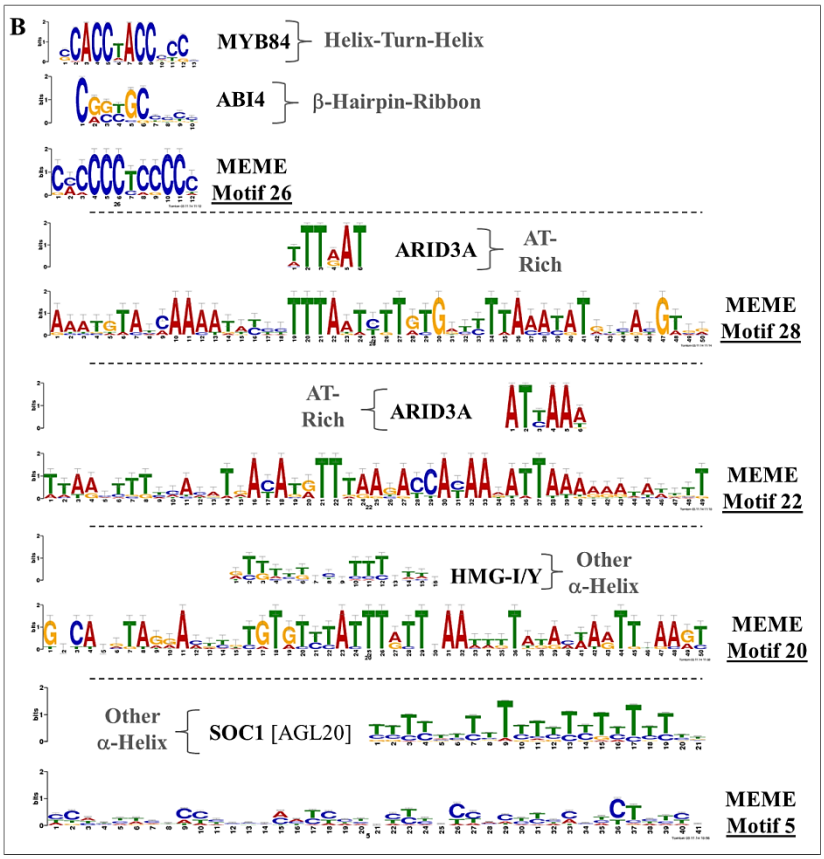


Figure IV.S6 *(…continued)*

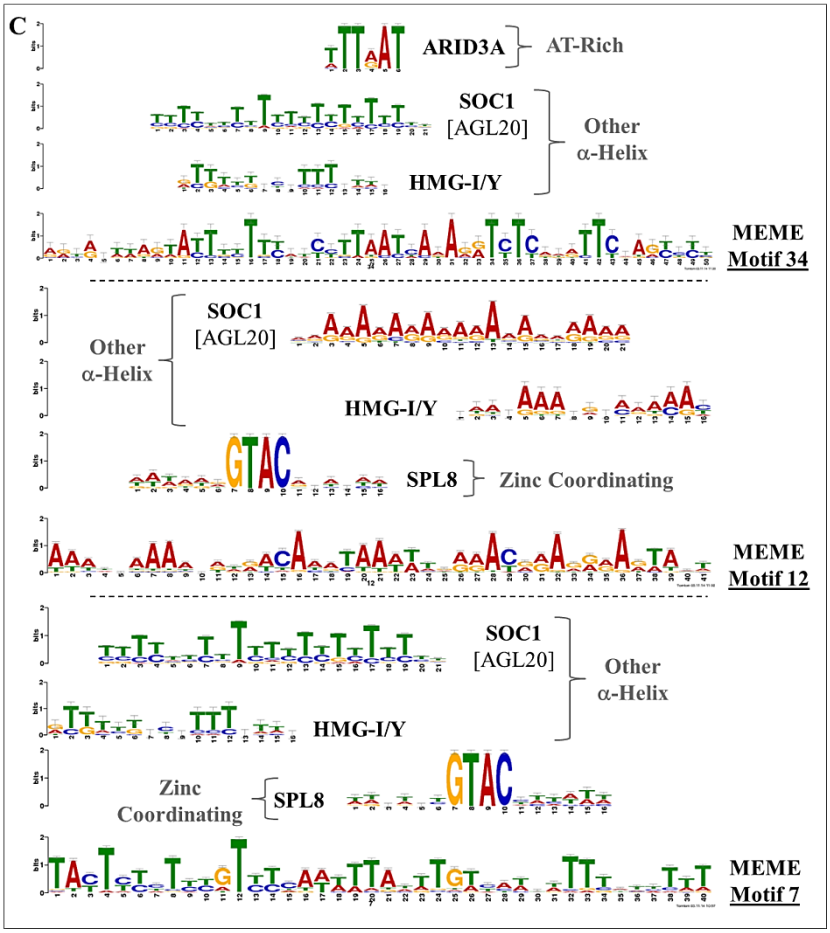


Figure IV.S6 *(…continued)*

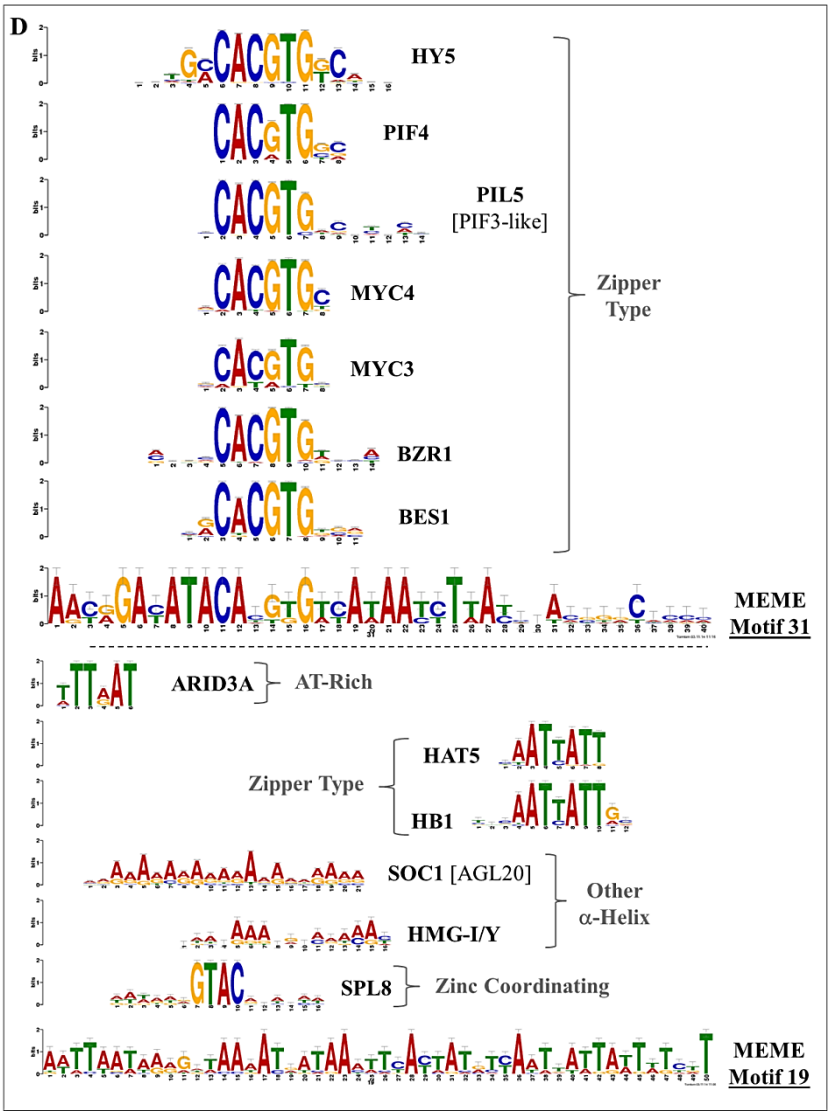
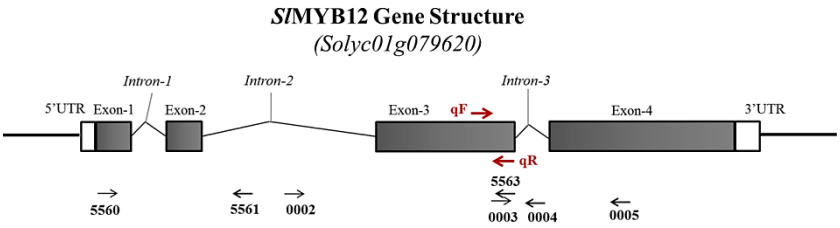


Figure IV.S6 *(…continued)*



Oligo ID	Sequence (5' → 3')
PCR Analysis	
5560 (F)	ATGGGAAGAACACCTTGTTGTAAAAAG
5561 (R)	GCTCGAATCCATTACACTATGTTATCATG
0002 (F)	CTACTTCAATCAATTATACCATTGACCTAG
5563 (R)	CTGAACTAAGGCTTCCCTTGGCCTCTATAG
0003 (F)	GAGGCCAAGGGAAGCCTTAGTTC
0004 (R)	CTATTAGAGGGTGTTTGATCTTTTTCGC
0005 (R)	CATTATTGCTCGTTACTTCTTCTCTG
RT-PCR Analysis	
qF	CCAGCTTGTGATAGTGCCAT
qR	AGGCTTCCCTTGGCCTCTA

Figure IV.S7 Primers for the SlMYB12 molecular analysis. *Oligonucleotide primer sequences for the PCR and RT-PCR analyses, and their location in the SlMYB12 gene structure.*

Additional Tables

You can find these additional tables in the excel file "*CHAPTER IV_Additional_Tables_IV.S1-IV.S5*" into the CD attached to this thesis.

Table IV.S1 UPLC-QTOF-MS-detected metabolites differentially accumulated in the skin of breaker fruits for both mutant pf^1 and pf^2.

Table IV.S2 Differentially expressed genes in breaker fruit skin from the pf mutants and shared with the y mutant*. The tomato annotation for each gene was included. A short name for the tomato and Arabidopsis annotations was provided if available. RNA-Seq FPKM averages (n=3) per sample and corresponding fold changes provided by CUFFDIFF (Trapnell et al., 2012) are included.*

Table IV.S3 Common genes differentially expressed in the pf mutants. *Data from mutant pf^1 and pf^2 were obtained by an RNA-Seq approach using tomato fruit skin in the breaker stage. The tomato annotation for each gene was included. A short name for tomato and Arabidopsis was provided if available. The RNA-Seq FPKM averages (n=3) per sample and corresponding fold changes provided by CUFFDIFF (Trapnell et al., 2012) are included.* ***A.*** *Genes that belong to the phenylpropanoid biosynthetic pathway.* ***B.*** *Genes that belong to primary and other secondary metabolisms.* ***C.*** *Genes that could be involved in stress response.* ***D.*** *Genes that could be involved in transcriptional and hormonal regulations.* ***E.*** *Genes that belong to other cellular processes.*

Table IV.S4 Transcription factor-binding elements by an MEME-TOMTOM analysis. *A combined MEME-TOMTOM analysis revealed 11 motifs that contained*

23 plant DNA-binding domains recognized by 10 different transcription factors classes. Motif information and the genes that contained each motif are provided.

Table IV.S5 Down-regulated genes in the pf mutants that contained transcription factor DNA-binding domains. *Only the genes that contained the SlMYB12 Myb-binding sites (MBS1 and MBS2) (Franco-Zorrilla et al., 2014) and MEME-TOMTOM TF DNA-binding domains are shown.* ***A.*** *Genes that belong to the phenylpropanoid biosynthetic pathway.* ***B.*** *Genes that belong to primary and other secondary metabolisms.* ***C.*** *Genes that could be involved in stress response.* ***D.*** *Genes that could be involved in transcriptional and hormonal regulations.* ***E.*** *Genes that belong to other cellular processes.*

GENERAL DISCUSSION

GENERAL DISCUSSION

The lack of a proper methodology for biochemical analysis and the lack of knowledge about the genetic basis controlling those biochemical changes, have supposed important limitations to plant breeding. Especially for more scientific, knowledge-based approaches, mainly when aimed to improve fruit quality. The establishment of new analytical methodologies applied to mid/high-throughput analysis allowed to assess the levels of many metabolites at once, e.g., metabolic profiling, (Fernie et al., 2006; Osorio et al., 2009) in mapping populations. And also provided additional basis for the genetic improvement of quality traits related to primary metabolites (e.g., sugar content; Causse et al., 2002), or secondary metabolites accumulations (e.g., carotenoids, flavonoids and volatiles; Fulton et al., 2003; Liu et al., 2003; Rosseaux et al., 2005). In contrast, the analytical methodologies to study fruit quality traits associated to cuticular composition in mid/high-throughput experiments were very limited. In *chapter I* of this thesis, the adaptation of an already described method (Hovav et al., 2007) for the screening of cuticular waxes in mid/high-throughput analysis of large breeding populations is described (Fernandez-Moreno et al., *under review*) and shown to be useful for a number of crops and organ surfaces. The adaptated method required smaller surface areas for analysis and shorter extraction periods using a uniform and versatile solvent. This protocol thus provided a useful methodology to screen cuticular waxes in tomato and an improved methodology to study cuticular wax composition in apple fruit and hybrid aspen leaf from those described in the literature (Belding et al., 1998; Cameron et al., 2002). Using the methodology described in the *chapter I*, the cuticular wax together with the cutin monomer compositions were analyzed throughout *Solanum pennellii* x *S. lycopersicum* cv. M82 introgression line population as is shown in *chapter II*. The biochemical analysis was combined with a QTL analysis using 187 genetic markers distributed across the IL population. Out of the 34 ILs containing differentially accumulated cuticle

metabolites, we validated five QTLs related with cuticular wax regulation (*vlcfa3.4*), triterpenoid biosynthesis (*ttp3.4*, *ttp12.1*), oxygenated epoxy-hydroxy fatty acids biosynthesis and/or regulation (*ehfa8.3*) and cuticle membrane weight (*icmw7.4*.1). Into these regions, nineteen candidate genes were identified, from which, seventeen still remain uncharacterized in tomato fruit. Only *TTS*s and *SHN1/WIN1* were characterized in tomato previously (Wang *et al.*, 2011; Shi et al., 2013). However herein our results suggest a new function, other rather than cutin monomer biosynthesis, for *SHN1/WIN1* in the regulation of cuticular wax biosynthesis in tomato fruit. Thus, we provided new clues about the genetic control of cuticle biosynthesis and regulation in tomato fruit. The *chapter III*, describes a detailed protocol for the use of the Fruit-VIGS methodology, a visually traceable system based on reverse genetics for fast analysis of gene functions (Fernández-Moreno et al., 2013). This protocol was used to study the function of two Myb-type transcription factors: SlODO1, the tomato homologue of Petunia PhODO1 involved in the regulation of aroma (Orzáez et al., 2009b), and SlMYB12, the tomato homologous of Arabidopsis AtMYB12, a key regulator of the flavonoid biosynthetic pathway (Ballester et al., 2009). The flavonoids have important benefits for the human health (Luo et al., 2008) and possess an important role in the correct functioning of the tomato fruit cuticle (Domínguez et al., 2011). Therefore, the comprehension of the complex network regulating the flavonoid biosynthesis is of interest for breeding programs. Thus, in the *chapter IV*, a new regulatory role of SlMYB12 in the biosynthesis of glycosylated flavonoids was revealed by identifying and characterizing the first null allele of this gene (*Slmyb12-pf*) which was obtained from an EMS mutant collection (Watanabe et al., 2007). Tomato mutants containing *Slmyb12-pf* allele were characterized metabolically and transcriptomically using high resolution mass spectrometry and new generation of sequencing methodologies, respectively. As discussed in *chapter IV*, the new allele provided insights into the regulation network as well as new target genes for SlMYB12, similarly to the effect observed for

the *myb11 myb12, myb111* triple mutant in Arabidopsis (Strake et al., 2007; Franco-Zorrilla et al., 2014)

Therefore, this thesis has covered the study of the tomato fruit peel following a multidisciplinary approach: *metabolomics*, using LC-MS for flavonoids and GC-MS for cuticular lipids analyses; *reverse genetics*, in helping to define the function of SlMYB12 and SlODO1 transcription factors, by developing and using Fruit-VIGS technology; *transcriptomics*, using RNA-Seq methodology to characterize the first null allele for SlMYB12, *Slmyb12-pf*, belonging to *pink fruit* EMS tomato mutants; *genomics*, for identification of genomics regions associated to cuticle metabolite changes, using a QTL analysis in the *Solanum pennellii* introgression line population; and *bioinformatics*, performing a candidate gene approach on the genomic regions of interests and by searching transcription factor binding domains into the promoters of the genes of interest.

The *future perspectives* arising from this thesis include: (**i**) to generate databases for cuticular wax and cutin monomer biochemical analysis, not only in tomato but also in other crop surfaces; (**ii**) to investigate additional QTLs associated to cuticle composition and to study the genomic regions of interest; (**iii**) to characterize the gene function for the candidate genes obtained in this thesis by conducting gene expression and reverse genetic approaches; (**iv**) to study the protein-protein interactions between SlMYB12 transcription factor and other regulators of the flavonoid biosynthetic pathway, as also the protein-promoter interactions between SlMYB12 and its target genes; (**v**) to study the gene function for the two SlMYB12-like transcription factors, their role in the pathway and their interaction with SlMYB12, not only for tomato fruit, but also in other organs as leaves.

CONCLUSIONS

CONCLUSIONS

The main conclusions of this thesis are:

First, we provide a general method for cuticular wax analysis method in mid/high-througput screenings easily adaptable to different species and organs.

Second, we present the first QTL analysis for cuticular wax and cutin monomer compositions on tomato fruit cuticle using the *Solanum pennellii* X *S. lycopersicum* cv. M82 introgression line population. Moreover, five QTLs were validated and nineteen genes are proposed as candidate for the genetic control of the cuticle composition.

Third, we standardize the Fruit-VIGS methodology for functional genomics in tomato fruit.

And ***fourth***, we characterize two new '*pink fruit*' tomato mutants metabolically and transcriptomically. Further, a new role in the regulation of the glycosylated flavonoids biosynthesis in tomato fruit was proposed for SlMYB12 transcription factor in response to the phenotype observed in the mutants containing a null allele for it.

REFERENCES

REFERENCES

Adato, A., Mandel, T., Mintz-Oron, S., Venger, I., Levy, D., Yativ, M., Domínguez, E., Wang, Z., De Vos, R.C.H., Jetter, R., Schreiber, L., Heredia, A., Rogachev, I. and Aharoni, A. (2009). Fruit-surface flavonoid accumulation in tomato is controlled by a SlMYB12-regulated transcriptional network. *PLoS Genet*. **5**, 1-23.

Aharoni, A., Dixit, S., Jetter, R., Thoenes, E., van Arkel, G. and Pereira, A. (2004). The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in Arabidopsis. *Plant Cell*. **16**, 2463-2480.

Albert, Z., Ivanics, B., Molnár, A., Miskó, A., Tóth, M. and Papp, I. (2013). Candidate genes of cuticle formation show characteristic expression in the fruit skin of apple. *Plant Growth Regul.* **70**, 71-78.

Alseekh, S., Ofner, I., Pleban, T., Tripodi, P., Di Dato, F., Cammareri, M., Mohammad, A., Grandillo, S., Fernie, A.R. and Zamir, D. (2013). Resolution by recombination: breaking up *Solanum pennellii* introgressions. *Trends Plant Sci*. **18**, 536-538.

Bailey, T.L. and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, pp. 28-36, AAAI Press, Menlo Park, California.

Ballester, A.R., Molthoff, J., de Vos, R., Hekkert, B.L., Orzaez, D., Fernandez-Moreno, J.P., Tripodi, P., Grandillo, S., Martin, C., Heldens, J., Ykema, M., Granell, A., Bovy, A. (2010). Biochemical and molecular analysis of pink tomatoes: deregulated expression of the gene encoding transcription factor SlMYB12 leads to pink tomato fruit colour. *Plant Physiol*. **152**, 71-84.

Bauer, S., Schulte, E. and Their, H.-P. (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 Their, H.-P. (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.

Beisson, F., Li-Beisson, Y. and Pollard, M. (2012). Solving the puzzles of cutin and suberin polymer biosynthesis. *Curr Opin Plant Biol*. **15**, 329-337.

Belding, R.D., Blankenship, S.M., Young, E. and Leidy, R.B. (1998). Composition and variability of epicuticular waxes in apple cultivars. *J Amer Soc Hort Scl*. **123**, 348-356.

Belding, R.D., Sutton, T.B., Blankenship, S.M. and Young, E. (2000). Relationship between apple fruit epicuticular wax and growth of Peltaster fructicola and Leptodontidi umelatius, two fungi that cause sooty blotch disease. *Plant Disease*. **8**, 767-772.

Bermudez, L., Urias, U., Milstein, D., Kamenetzky, L., Asis, R., Fernie, A.R., Van Sluys, M.A., Carrari, F. and Rossi, M. (2008). A candidate gene survey of quantitative trait loci affecting chemical composition in tomato fruit. *J Exp Bot*. **59**, 2875-2890.

Bernacchi, D., Beck-Bunn, T., Eshed, Y., Lopez, J., Petiard, V., Uhlig, J., Zamir, D., and Tanksley, S. (1998). Advanced backcross QTL analysis in tomato. I. Identification of QTLs for traits of agronomic importance from *Lycopersicum hirsutum*. *Theor Appl Genet*. **97**, 381-397.

Bolger, A., Scossa, F., Bolger, M.E., Lanz, C., Maumus, Florian., Tohge, T., Quesneville, H., Alseekh, S., SØrensen, I., Lichtenstein, G., Fich, E.A., Conte, M., Keller, H., Schneeberger, K., Schwacke, R., Ofner, I., Vrebalov, J., Xu, Y., Osorio, S., Aflitos, S.A., Schijlen, E., Jiménez-Gómez, J.M., Ryngajllo, M., Kimura, S., Kumar, R., Koeing, D., Headland, L.R., Maloof, J.N., Sinha, N., van Ham, R.C.H.J., Lankhorst, R.K., Mao, L., Vogel, A., Arsova, B., Panstruga, R., Fei, Z., Rose, J.K.C., Zamir, D., Carrari, F., Giovannoni, J.J., Weigel, D., Uasadel, B. and Fernie, A.R. (2014). The genome of the stress-tolerant wild tomato species Solanum pennellii. *Nat Genet*. **46**, 1034-1039.

Bonaventure, G. and Orhlrogge, J.B. (2002). Differential regulation of mRNA levels of acyl carrier protein isoforms in Arabidopsis. *Plant Physiol*. **128**, 223-235.

Bovy, A., Schijlen, E. and Hall, R.D. (2007). Metabolic engineering of flavonoids in tomato (*Solanum lycopersicum*): The potential for metabolomics. *Metabolomics*. **3**, 399-412.

Brodsky, L., Moussaieff, A., Shahaf, N., Aharoni, A. and Rogachev, I. (2010). Evaluation of peak picking quality in LC-MS metabolomics data. *Anal Chem*. **82**, 9177-9187.

Bugos R.C., Chiang, V.L., Zhang X.-H., Campbell E.R., Podila, G.K. and Campbell, W.H. (1995). RNA isolation from plant tissues recalcitrant to extraction in guanidine. *Biotechniques*. **19**, 734-744.

Burch-Smith, T.M., Anderson, J.C., Martin, G.B. and Dinesh-Kumar, S.P. (2004). Applications and advantages of virus-induced gene silencing for gene function studies in plants. *Plant J*. **39**, 734-746.

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.

Butelli, E., Titta, L., Giorgio, M., Mock, H.-P., Matros, A., Peterek, S., Schijlen, E.G.W.M., Hall, R.D., Bovy, A.G., Luo, J. and Martin, C. (2008). Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat Biothecnol*. **26**, 1301-1308.

Buxdorf, K., Rubinsky, G., Barda, O., Burdman, S., Aharoni, A. and Levy M. (2014). The transcription factor SlSHINE3 modulates defense response in tomato plants. *Plant Mol Biol*. **84**, 37-47.

Caligiani, A., Malavasi, G., Palla, G., Marseglia, A., Tgnolini, M. and Bruni, R. (2013). A simple GC-MS method for the screening of betulinic, corosolic, maslinic, oleanolic and ursolic acid contents in commercial botanicals used as food supplement ingredients. *Food Chem*. **136**, 735-741.

Cámara, M. (2006). Calidad nutricional y salud. Mejora genética de la calidad en las plantas. Llacer, G., Díez, M.J., Carrillo, J.M., Badenes, M.L. Editors, UPV Editorial, pp. 43-65.

Cameron, K.D., Teece, M. A., Bevilacqua, E. and Smart, B. (2002) Diversity of cuticular wax among Salix species and Populus species hybrids. *Phytochem*. **60**, 715-725.

Casañas, F. and Costell, E. (2006). Calidad organoléptica. Mejora genética de la calidad en las plantas. Llacer, G., Díez, M.J., Carrillo, J.M., Badenes, M.L. Editors, UPV Editorial, pp. 19-41.

Catalá, C., Rose, J.K.C., York, W.S., Albersheim, P., Darvill, A.G. and Bennet, A.B. (2001) Characterization of a tomato xyloglucan endotransglycosylase gene that is down-regulated by auxin in etiolated hypocotyls. *Plant Phys*. **127**:1180-1192.

Causse, M., Saliba-Colombani, V., Lecomte, L., Duffé, P., Rouselle, P. and Buret, M. (2002). QTL analysis of fruit quality in fresh market tomato: a few chromosome regions control the variation of sensory and instrumental traits. *J Exp Bot*. **53**, 2089-2098.

Chapman, N.H., Bonnet, J., Grivet, L., Lynn, J., Graham, N., Smith, R., Sun, G., Walley, P.G., Poole, M., Causse, M., King, G.J., Baxter, G. and Seymour, G.B. (2012). High resolution mapping of a fruit firmness-related QTL in tomato reveals epistatic interactions associated with a complex combinatorial locus. *Plant Physisol*. **159**, 1644-1657.

Chen, F.Q., Foolad, M.R., Hyman, J., St Clair, D.A. and Beeleman, R.B. (1999). Mapping QTLs for lycopene and other fruit traits in a *Lycopersicon esculentum* X *L. pimpinellifolium* cross and comparison of QTLs across tomato species. *Mol Breeding*. **5**, 283-299.

Chopra, S., Athma, P., Peterson, T. (1996). Alleles of the maize P gene with distinct tissue specificities encode Myb-homologous proteins with C-terminal replacements. *Plant Cell*. **8**, 1149-1158.

Czemmel, S., Stracke, R., Wisshaar, B., Cordon, N., Harris, N.N., Walker, A.R., Robinson, S.P. and Bogs, J. (2009). The grapevine R2R3-MYB transcription factor VvMYBF1 regulates flavonol synthesis in developing grape berries. *Plant Physiol*. **151**, 1513-1530.

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

Davies, K.M., Marshall, G.B., Bradley, J.M., Schwinn, K.E., Bloor, S.J., Winefield, C.S. and Martin, C.R. (2006). Characterisation of aurone biosynthesis in Antirrhinum majus. *Physiologia Plantarum*. **128**, 593–603.

Dobson, G., Vasukuttan, V. and Alexander, C.J. (2012). Evaluation of different protocols for the analysis of lipophilic plant metabolites by gas chromatography-mass spectrometry using potato as a model. *Metabolomics*. **8**, 880-893.

Domínguez, E., España, L., López-Casado, G., Cuartero, J., Heredia, A. (2009a). Biomechanics of isolated tomato (*Solanum lycopersicum*) fruit cuticles during ripening: the role of flavonoids. *Functional Plant Biol*. **36**, 613-620.

Domínguez, E., Luque, P., Heredia, A. (2009b). Sorption and interaction of the flavonoid naringenin on tomato fruit cuticles. *J Agric Food Chem*. **57**, 7560-7564.

Domínguez, E., Cuartero, J. and Heredia, A. (2011). An overview on plant cuticle biomechanics. *Plant Sci*. **181**, 77-84.

Domínguez, E., España, L., Lípez-Casado, G., Cuartero, J. and Heredia A. (2009). Biomechanics of isolated tomato (*Solanum lycopersicum*) fruit cuticles during ripening: the role of flavonoids. *Funct Plant Biol*. **36**, 613-620.

Domínguez, E., Fernandez, M.D., Hernández, J.C., Parra, J., España, L., Heredia, A. and Cuartero, J. (2012). Tomato fruit continues growing while ripening, affecting cuticle properties and cracking. *Physiol Plantarum*. **146**, 473-486.

Du, H., Feng, B.R., Yang, S.S., Huang, Y.B. and Tang, Y.X. (2012). The R2R3-MYB transcription factor gene family in Maize. *PLoS One*. **7**: e37463 (doi: 10.1371/journal.pone.0037463).

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.

España, L., Heredia-Guerrero, J.A., Segado, P., Benítez, J.J., Heredia, A. and Domínguez, E. (2014). Biomechanical properties of the tomato (*Solanum lycopersicum*) fruit cuticle during development are modulatedd by changes in the relative amount of its compounents. *New Phytol*. **202**, 790-802.

Estornell, L. H., Orzaez, D., Lopez-Pena, L., Pineda, B., Anton, M.T., Moreno, V. and Granell, A. (2009). A multisite gateway-based toolkit for targeted gene expression and hairpin RNA silencing in tomato fruits. *Plant Biotechnol J*. **7**, 298-309.

Expósito-Rodríguez, M., Borges, A.A., Borges-Pérez, A. and Pérez, J.A. (2008). Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. *BMC Plant Biol*. **8**, 131-143.

Fernández-Moreno J.-P., Orzáez, D. and Granell, A. (2013). Virus-Induced Gene Silencing: a tool to study fruit development in Solanum lycopersicum. Methods in Molecular Biology. Springer Life Editorial. Volume 975, Chapter 14 (pp183-196).

Fernie, A.F., Tadmor, Y., Zamir, D. (2006). Natural genetic variation for improving crop quality. *Curr Opin Plant Biol*. **9**, 196-202.

Franco-Zorrilla, J.M., Lopez-Vidriero, .I, Carrasco, J.L., Godoy, M., Vera, P. and Solano, R. (2014). DNA-binding specificities of plant transcription factors and their potential to define target genes. *PNAS*. **111**, 2367-2372.

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

Frary, A., Göl, D., Keleş, D., Ökmen, B., Pinar, H., Ö Şiğva, H., Yemicioğlu, A. and Doğanlar, S. (2010). Salt tolerance in *Solanum pennellii*: antioxidant response and related QTL. *BCM Plant Biol*. **10**, 58-73.

Fu, D.-Q., Zhu, B.-Z., Zhu, H.-L., Jiang W.-B. and Luo, Y.-B. (2005). Virus-induced gene silencing in tomato fruit, *Plant J*. **43**, 299-308.

Fulton, T.M., Buchell, P., Voirol, E., Lopez, J., Petiard, V. and Tanksley, S.D. (2002). Quantitative-trait-loci (QTL) affecting sugars, organic acids and other biochemical properties possible contributing to flavour, identified in four advanced backcross populations of tomato. *Euphytica*. **127**: 163-177.

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.

Giovannoni, J.J. (2004). Genetic regulation of fruit development and ripening. *Plant Cell*. **16**, S170-S180.

Giovannoni, J.J. (2001). Molecular Biology of fruit maturation and ripening. *Annu Rev Plant Physiol Plant Mol Biol*. **52**, 725-749.

Goldman, I.L., Paran, I. and Zamir, D. (1995). Quantitative locus analysis of a recombinant inbred line population derived from a Lycopersicon esculetum X Lycopersicum cheesmanii cross. *Theor Appl Genet*. **90**, 925-932.

Grandillo, S., Ku, H.-M. and Tanksley, S.D. (1999). Identifying the loci responsible for natueral variation in fruit size and shape in tomato. *Theor Appl Genet*. **99**, 978-987.

Grotewold, E., Drummond, B.J., Bowen, B. and Peterson, T. (1994). The myb-homologous P gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. *Cell*. **76**, 543-553.

Grotewold, E. (2005). Plant metabolic diversity: a regulatory perspective. *Trends Plant Sci*. **10**, 57-62

Grotewold, E. (2006). The science of flavonoids. Columbus, Ohio (USA): Erich Grotewold, Eds.

Gupta, S., Stamatoyannopolous, J.A., Bailey, T. and Noble, W.S. (2007). Quantifying similarity between motifs. *Genome Biol*. **8**, R24 (doi:10.1186/gb-2007-8-2-r24).

Gur, A., Semel, Y., Osorio, S., Friedmann, M., Seekh, S., Ghareeb, B., Mohammad, A., Pleban, T., Gera, G., Fernie, A.R. and Zamir, D. (2011). Yield quantitative trait loci from wild tomato are predominately expressed by the shoot. *Theor Appl Genet*. **122**, 405-420.

Gur, A., and Zamir, D. (2004). Unused natrual variation can lift yield barriers in plant breeding. *PLoS Biol*. **2**: e245 (doi:10.1371/journal.pbio.0020245).

Halá, M., Soukupová, H., Synek, L. and Zárský, V. (2010). Arabidopsis RAB geranylgeranyl transferase beta subunit mutant is constitutively photomorphogenic and has shoot growth and gravitropic defects. *Plant* J. **62**, 615-627.

Han, Y., Vimolmangkang, S., Soria-Guerra, S.E., Rosales-Mendoza, S., Zheng, D., Lygin, A.V. and Korban, S.S. (2010). Ectopic expression of apple F3'H gene contributes to anthocyanin accumulation in the Arabidopsis tt7 mutant grown under nitrogen stress. *Plant Physiol*. **153**, 806-820.

Hartmann, U., Sagasse, M., Mehrtens, F., Stracke, R. and Weisshaar, B. (2005). Differential combinatorial interactions of cis-elements recognized by R2R3-MYB, BZIP, and BHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. *Plant Mol Biol*. **57**, 155-171.

Hemm, M.R., Herrmann, K.M. and Chapple, C. (2001). AtMYB4: a transcription factor general in the battle against UV. *Trends Plant Sci*. **6**, 135-136.

Hen-Aviv, S., Lashbrooke, J., Costa, F. and Aharoni, A. (2014) Scratching the surface: genetic regulation of cuticle assembly in fleshy fruit. *J Exp Bot*. **65**, 4653-4664.

Heredia, A. (2003). Biophysical and biochemical characteristics of cutin, a plant barrier biopolymer. *Biochim Biophys Acta*. **1620**, 1-7.

Heredia, A. and Domínguez, E. (2009). The plant cuticle: a complex lipid barrier between the plant and the environment. An overview. In *Counteraction to Chemical and Bilogical Terrorism in East European Countries*; Dishovsky, C.; Pivovarov, A., Eds.; Springer: Dordrecht, The Netherlands, pp 109-116.

Hichri, I., Barrieu, F., Bogs, J., Kappel, C., Delrot, S. and Lauvergeat, V. (2011). Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *J Exp Bot*. (doi:10.1093/jxb/erq442).

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 microfisuring and dehydration when expressed in tomato fruit. *Plant J*. **52**, 627-639.

Iijima, Y., Nakamura, Y., Ogata, Y., Tanaka, K., Sakuri, N., Suda, K., Suzuki, T., Suzuki, H., Okazaki, K., Kitayama, M., Kanaya, S., Aoki, K. and Shibata, D. (2008). Metabolite annotations based on the integration of mass spectral information. *Plant J*. **54**, 949-962.

Isaacson, T., Kosma, D.K., Matas, A.J., Buda, G.J., He, Y., Yu, B., Pravitasari, A., Batteas J.D., Stark, R.E., Jenks, M.A. and Rose, J.K.C. (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.

Jessen, S., Roth, C., Wiermer, M. and Fulda, M. (2014). Two activities of long-chain acyl-CoA synthetase are involved in lipid trafficking between the endoplasmic reticulum and the plastid in Arabidopsis. *Plant Physiol*. **167**, 351-366.

Jones, H.G. (1992) Plants and microclimate: a quantitative approach to environmental plant physiology. 2nd Ed. Cambridge University Press, New York, NY.

Kandel, S., Sauveplane, V., Olry, A., Diss, L., Benvenist, I. and Pinot, F. (2006). Cytochrome p450-dependent fatty acid hydroxylases in plants. *Phytochem Rev*. **5**, 359-372.

Kannangara, R., Branigan, C., Liu, Y., Penfield, T., Rao, V., Mouille, G., Höfte, 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.

Kimbara, J., Yoshida M., Ito, H., Hosoi, K., Kusano, M., Kobayashi, M., Ariizumi, T., Asamizu, E. and Ezura, H. (2012). A novel class of sticky peel and light green mutations causes cuticle deficiency in leaves and fruits of tomato (*Solanum lycopersicum*). *Planta*. **236**, 1559-1570.

Koeing, D., Jiménez-Gómez, J.M., Kimura, S., Fulop, D., Chitwood, D.H., Headland, L.R., Kumar, R., Covington, M.F., Kumar-Devisetty, U., Tat, A.V., Tohge, A., Schneeberger, K., Ossowski, S., Lanz, C., Xiong. G., Tylor-Teeples, M., Brady, S.M., Pauly, M., Weigel, D., Usadel, B., Fernie, A.R., Peng, J., Sinha, N.R. and Maloof, J.N. (2013). Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. *PNAS*. **110**, E2655-E2662.

Kolattukudy, P.E. (1970). Biosynthesis of cuticular lipids. *Annu Rev Plant Physiol*. **21**, 163-192.

Kranz, H.D., Denekamp, M., Greco, R., Jin, H., Leyva, A., Meissner, R.C., Petroni, K., Urzainqui, A., Bevan, M., Martin, C., Smeekens, S., Tonelli, C., Paz-Ares, J. and Wisshaar, B. (1998). Towards functional characterization of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. *Plant J*. **16**, 263-276.

Kunst, L. and Samuels, A.L. (2002). Biosynthesis and secretion of plant cuticular wax. *Prog Lipid Res*. **42**, 51-80.

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

Kurdyukov, S., Faust, A., Nawrath, C., Bär, S., Voisin, D., Efremova, N., Franke, R., Schreiber, L., Saedler, H., Métraux, J.P. and Yephremov, A. (2006a). The epidermis-specific extracellular BODYGUARD controls cuticle development and morphogenesis in Arabidopsis. *Plant Cell*. **18**, 321-339.

Kurdyukov, S., Faust, A.,Trenkamp, S., Bär, 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*. **22**4, 315-329.

Lara, I., Belge, B. and Goulao, L.F. (2014). The fruit cuticle as a modulator of postharvest quality. *Postharvest Biol Technol*. **87**, 103-112.

Lashbrooke, J., Aharoni, A. and Costa, F. (2015). Genome investigation suggests MdSHN3, and APETALLA2-domain transcription factor gene, to be a positive regulator of apple fruit cuticle formation and an inhibitor of russet development. *J Exp Bot*. (doi:10.1093/jxb/erv366).

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 β-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 fruit. *J Plant Physiol*. **168**, 871-877.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. **15**, 1754-1760

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., Zhang, Z., Lun, Y., Li, S., Wang, X., Huang, Z., Li, J., Zhang, C., Wang, X., Wan, T., Zhang, Y., Wang, A., Zhang, Y., Lin, K., Li, C., Xiong, G., Xue, Y., Mazzucato, A., Causse, M., Fei, Z., Giovannoni, J., Chetelat, R.T., Zamir, D., Ständler, T., Li, J., Ye, Z., Du, Y. and Huang, S. (2014). Genomic analyses provide insights into the history of tomato breeding. *Nat Genet*. **46**, 1220-1226.

Lindstrom, E.W. (1925). Inheritance in tomatoes. *Genetics*. **10**, 305-317.

Lin-Wang, K., Bolitho, K., Grafton, K., Kortstee, A., Karunairetnam, S., McGhie, T.K., Espley, R.V., Hellens, R.P. and Allan, A.C. (2010). An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. *BMC Plant Biol*. **10**, 50-66.

Liu, Y.S., Gur, A., Ronen, G., Causse, M., Damidaux, R., Buret, M., Hirschberg, J. and Zamir, D. (2003). There is more to tomato fruit colour than candidate carotenoid genes. *Plant Biotechnol J*. **1**, 195-207.

Liu, J., Xu, X. and Deng, X. (2005). Intergeneric somatic hybridization and its application to crop genetic improvement. *Plant Cell Tissue Organ Cult*. **82**, 19-44

Liu, Y., Schiff, M., and Dinesh-Kumar, S.P. (2002). Virus-induced gene silencing in tomato. *Plant J*. **31**, 777-786.

Llácer, G., Badenes. M.L., Díez. M.J. and Carrillo, J.M. (2006). La calidad de los productos agrícolas en el marco de la mejora genética actual. Mejora genética de la calidad en las plantas. Llacer, G., Díez, M.J., Carrillo, J.M., Badenes, M.L. Editors, UPV Editorial, pp. 5-17.

López-Casado, G.,Matas, A.J., Domínguez, 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, R., Martin-Hernandez, A. M., Peart, J. R., Malcuit, I. and Baulcombe, D.C. (2003). Virus-induced gene silencing in plants. *Methods*. **30**, 296-303.

Lü, S., Song, T., Kosma, D.K., Parson, F.P., Rowland, O. and Jenks, M.A. (2009). Arabidopsis CER8 encodes a long-chain acyl-CoA synthetase 1 (LACS1) and has overlapping functions with LACS2 in plant wax and cutin synthesis. *Plant J*. **59**, 553-564.

Luo, J., Butelli, E., Hill, L., Parr, A., Niggeweg, R., Bailey, P., Weisshaar, B. and Martin, C. (2008). AtMYB12 regulates caffeoylquinic acid and flavonol synthesis in tomato: expression in fruit results in very high levels of both types of polyphenol. *Plant J*. **56**, 316-326.

Maeda, H., Yoo, H. and Dudareva, N. (2011). Prephenate aminotransferase directs plant phenylalanine biosynthesis via arogenate. Brief communication, *Nat Chem Biol*. **7**, 19-21 (doi:10.1038/NCHEMBIO.485).

Marti, C., Orzaez, D., Ellul, P., Moreno, V., Carbonell, J. and Granell, A. (2007). Silencing of DELLA induces facultative parthenocarpy in tomato fruits. *Plant J*. **52**, 865-876.

Martin, C., Butelli, E., Petroni, K., Tonelli, C. (2011). How can research on plants contribute to promoting human health? *Plant Cell*. **23**, 1685-1699.

Martin, L.B. and Rose, J.K. (2014). There's more than one way to skin a fruit: formation and functions of fruit cuticles. *J Exp Bot*. **65**, 4639-4651

Matas, A.J, Yeats, T.H., Buda, G.J., Zheng, Y., Chatterjee, S.,Tohge, T., Ponnala, L., Adato, A., Aharoni, A., Stark, R., Fernie, AR., Fei, Z., Giovannoni, J.J. and Rose, J.K.C. (2011). Tissue- andcell-type specific transcriptome profiling of expanding tomato fruit provides insightsinto metabolic and regulatory specialization and cuticle formation. *Plant Cell*. **23**, 3893-3910.

Mehrtens, F., Kranz, H., Bednarek, P. and Weissharr, B. (2005). The Arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiol*. **138**, 1083-1096.

Mintz-Oron, S., Mandel, T., Rogachev, I., Feldberg, L., Lotan, O., Yativ, M., Wang, Z., Jetter, R., Venger, I., Adato, A. and Aharoni, A. (2008). Gene expression and metabolism in tomato fruit surface tissues. *Plant Physisol*. **47**, 823-851.

Moço, S., Bino, R.J., Vorst, O., Verhoeven, H.A., De Groot, J., van Beek, T.A., Vervoot, J., De Vos, C.H.R. (2006). A liquid chromatography-mass spectrometry-based metabolome database for tomato. *Plant Physiol*. **141**, 1205-1218.

Muir, S.R., Collins, G.J., Robinson, S., Hughes, S., Bovy, A., Ric de Vos, C.H., van Tunen, A.J. and Verhoeyen, M.E. (2001). Overexpression of petunia

chalcone isomerase in tomato results in fruit containing increased levels of flavonols. *Nature*. **19**, 470-4741

Nakabayashi, R., Yonekura-Sakakibara, K., Urano, K., Suzuki, M., Yamada, Y., Nishizawa, T., Matsuda, F., Kojima, M., Sakakibara, H., Shinozaki, K., Michael, A.J., Tohge, T., Yamazaki, M., Saito, K. (2014). Enhancement of oxidative and drought tolerance in Arabidopsis by overaccumulation of antioxidant flavonoids. *Plant J*. **77**, 367-379

Nakatsuka, T., Saito, M., Yamada, E., Fujita, K., Kakizaki, Y. and Nishihara, M. (2012). Isolation and characterization of GtMYBP3 and GtMYBP4, orthologues of R2R3-MYB transcription factors that regulate early flavonoid biosynthesis, in gentian flowers. *J Exp Bot*. **63**, 6505-6517.

Newburger, D.E. and Bulyk, M.L. (2009). UniPROBE: an online database of protein binding microarray data on protein-DNA interactions. *Nucleic Acids Res*. **37**, D77-D82.

Orzaez, D. and Granell, A. (2009a). Reverse genetics and transient gene expression in fleshy fruits: overcoming plant stable transformation. *Plant signaling & behavior*. **4**, 864-867.

Orzaez, D., Medina, A., Torre, S. Fernandez-Moreno, J.-P., Rambla, J.L., Fernandez-del-Carmen, A., Butelli, E., Marin, C. and Granell, A. (2009b). A visual reporter system for virus-induced gene silencing in tomato fruit based on anthocyanin accumulation. *Plant physiology*. **150**, 1122-1134.

Orzaez, D., Mirabel, S., Wieland, W. H. and Granell, A. (2006). Agroinjection of tomato fruits. A tool for rapid functional analysis of transgenes directly in fruit. *Plant physiology*. **140**, 3-11.

Osorio, S., Tohge, T. and Fernie, A.R. (2009). Application of metabolomics profiling for identifying valuable traits in tomato. *Perspect Agric Vet Sci Nutr Nat Res*. **4**, 1-9.

Overy, S.A., Walker H.J., Malone, S., Howard, T.P., Baxter, C.J., Sweetlove, L.J., Hill, S.A. (2004). Application of metabolite profiling to the identification of traits in a population of tomato introgression lines. *J Exp Bot*. **56**, 287-296.

Paterson, A.H., Lander, E.S., Hewitt, J.D., Peterson, S., Lincoln, S.E. and Tanksley, S.D. (1998). Resolution of quantitative traits into mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature*. **335**, 721-726.

Petit, J., Bres, C., Just, D., Garcia, V., Mauxion, J.P., Marion, D., Bakan, B., Joubès, J., Domergue, F. and Rothan, C. (2014). Analyses of tomato fruit brightness mutants uncover both cutin-deficient and cutin-abundant mutants and a new hypomorphic allele of GDSL lipase. *Plant Physiol*. **164**, 888-906.

Pineda, B., García-Abellán, J.O., Antón, T., Pérez, F., Moyano, E., García Sogo, B., Campos, J.F., Angosto, F., Morales, B., Capel, J., Moreno, V., Lozano, R., Bolarín, M.C. and Atarés, A. (2012). Tomato: genomic approaches for salt and drought stress tolerance, in improving crop resistance to abiotic stress, Volume 1 & Volume 2 (eds. N. Tuteja, S.S. GII, A.F. Tiburcio and R. Tuteja), Wiley-VCH Verlog GmbH & Co. KGaA, Winheim, Germany. doi: 10.100219783527632930. Ch43.

Pollard, M., Beisson, F., Li, Y. and Ohlrogge, J.B. (2008). Building lipid barriers: biosynthesis of cutin and suberin. *Cell Review*. **13**, 1360-1385.

Portales-Casamar, E., Thongjuea, S., Kwon, A.T., Arenillas, D., Zhao, X., Valen, E., Yusuf, D., Lenhard, B., Wasserman, W.W. and Sandellin, A. (2010). JASPAR 2010: the greately expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res*. **38**, D105-D110.

Raffaele, S., Vailleau, F., Joubès, J., Miersch, O., Huard, C., Blée, E., Mongrand, S., Domergue, F. and Roby, D. (2008). A MYB transcription factor regulates very-long-chain fatty acid biosynthesis for activation of the hypersensitive cell death response in Arabidopsis. *Plant Cell*. **20**, 752-767.

Rambla, J.L., Tikunov, Y.M., Monforte, A.J., Bovy, A.G. and Granell, A. (2014). The expanded tomato fruit volatile landscape. *J Exp Bot*. (doi:10.1093/jxb/eru128).

Ratcliff, F., Martin-Hernandez, A.M., and Baulcombe, D.C. (2001). Technical Advance. Tobacco rattle virus as a vector for analysis of gene function by silencing. *Plant J*. **25**, 237-245.

Rein, D., Schijlen, E., Kooistra, T., Herbers, K., Verschuren, L., Hall, R., Sonnewald, U., Bovy, A. and Kleemann, R. (2006). Transgenic flavonoid tomato intake reduces C-reactive protein in human C-reactive protein transgenic mice more than wild-type tomato. *J Nutr*. **136**, 2331-2337.

Rick, C.M. and Butler, L. (1956). Cytogenetics of the tomato. *Advan Gent Incorporating Mol Gen Med*. **8**, 267-382

Riederer, M. and Müller, C. (2006). Biology of the plant cuticle. Oxford, Blackwell Pub.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P. (2011). Integrative Genomic Viewer. *Nat Biotechnol*. **29**, 24-26

Rogachev, I. and Aharoni, A. (2012). UPLC-MS-based metabolite analysis in tomato. *Methods Mol Biol*. **860**, 129-144.

Rowland, O., Lee, R., Franke, R., Schreiber, L. and Kunst, L. (2007). The CER3 wax biosynthetic gene from *Arabidopsis thaliana* is allelic to WAX2/YRE/FLP1. *FEBS Lett*. **581**, 3538-3544.

Rupasinghe, SG.,Duan, H. and Schuler, M.A. (2007). Molecular definitions of fatty acid hydroxylases in Arabidopsis thaliana. *Proteins*. **68**, 279-293.

Saito, T., Ariizumi, T., Okabe, Y., Asamizu, E., Hiwasa-Tanase, K., Fukuda, N., Mizoguchi, T., Yamazaki, Y., Aoki, K. and Ezura, H. (2001). TOMATOMA: a novel tomato mutant database distributing Micro-Tom mutant collections. *Plant Cell Physiol*. **52**, 283-296.

Saladié, M., Matas, A.J., Isaacson, T., Jenks, M.A., Goodwin, S.M., Niklas, K.J., Xiaolin, R., Labavitch, J.M., Shackel, K.A., Fernie, A.R., Lytovchenko, A., O'Neill, M.A., Watkins, C.B. and Rose, J.K.C. (2007). A reevaluation of the key factors that influence tomato fruit softening and integrity. *Plant Physiol*. **144**, 1012-1028.

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

Sandbrink, J.M., van Ooijen, J.W., Purimahua, C.C., Vrielink, M., Verkerk, R., Zabel, P. and Lindout, P. (1995). Localization of genes for bacterial canker resistance in *Lycopersicon peruvianum* using RFLPs. *Theor Appl Genet*. **90**, 444-450.

Schauer, N., Semel, Y., Balbo, I., Steinfath, M., Repsilber, D., Selbig, J., Pleban, T., Zamir, D. and Fernie, A.R. (2008). Mode of inheritance of primary metabolic traits in tomato. *Plant Cell*. **20**, 509-523.

Schenke, D. and Cai, D. (2014). The interplay of transcription factors in suppression of UV-B induced flavonol accumulation by flg22. *Plant Signal Behav*. **9**: e28745 (doi: 10.4161/pbs.28745).

Schijlen, E.G.W.M., De Vos, C.H.R., Martens, S., Jonker, H.H., Rosin, F.M., Molthoff, J.W., Tikunov, Y.M., Angenent, G.C., Van Tunen, A.J., Bovy, A.G. (2007). RNA interference silencing of chalcone synthase, the first step in the flavonoid biosynthesis pathway, leads to parthenocarpic tomato fruits. *Plant Physiol*. **144**, 1520-1530.

Seo, P.J. and Park, C.M. (2011). Cuticular wax biosynthesis as a way of inducing drought resistance. *Plant Signaling and Behavior*. **6**, 1043-1045.

Seo, P.J., Xiang, F., Qiao, M., Park, J.Y., Lee, Y.N., Kim, S.G., Lee, Y.H., Park, W.J. and Park, C.M. (2009). The MYB96 transcription factor mediates abcisic acid signaling during drought stress response in Arabidopsis. *Plant Physiol*. **151**, 275-289.

Shi, J.X., Adato, A., Alkan, A., He, Y., Lashbrooke, J., Matas, A.J., Meir, S., Malitsky, S., Isaacson, T., Prusky, D., Leshkowitz, D., Schreiber, L., Granell, A., Widemann, E., Grausem, B., Pinot, F., Rose, J.K.C, Rogachev, I., Rothan, C. and Aharoni, A. (2012). The tomato SlSHINE3 transcription factor regulates fruit cuticle formation and epidermal patterning. *New phytol*. **197**, 468-480.

Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R. and Siuzdak, G. (2006). Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Analytical Chemistry*. **78**, 779-787.

Smith, R.M., Marshall, J.A., Davey, M.R., Lowe, K.C. and Power, B. (1996). Comparison of volatiles and waxes in leaves of genetically engineered tomatoes. *Phytochem*. **43**, 753-758.

Sockey, J.M., Fulda, M.S. and Browse, J.A. (2002). Arabidopsis contains nine long-chain acyl-coenzyme a synthase genes that participate in fatty acid and glycerolipid metabolism. *Plant Physiol*. **129**, 1710-1722.

Stewart Jr, C., Vickery, C.R., Burkart, M.D. and Noel, J.P. (2013). Confluence of structural and chemical biology: plant polyketide synthases as biocatalysts for a bio-based future. *Curr Opin Plant Biol*. **16**, 1-8.

Stracke, R., Ishihara, H., Huep, G., Barsch, A., Mehrtens, F., Niehaus, K. and Weisshaar, B. (2007). Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant J*. **50**, 660-677.

Sun, H. Fan, H.J. and Ling, H.Q. (2015). Genome-wide identification and characterization of the bHLH gene family in tomato. *BMC genomics*. **16**: 9. DOI 10,1186/s12864-014-1209-2.

Szakiel, A., Pączkowski, C., Pensec, F. and Bertsch, C. (2012). Fruit cuticular waxes as a source of biologically active triterpenoids. *Phytochem Rev*. **11**, 263-284.

Tadmor, Y., Fridman, E., Gur, A., Larkov, O., Lastochkin, E., Ravid, U., Zamir, D. and Lewinsohn, E. (2002). Identification of malodorous, a wild species allelel

affecting tomato aroma that was aelected against during domestication. *J Agric Food Chem*. **50**, 2005-2009.

Taksley, S.D., Grandillo, S., Fulton, T.M., Zamir, D., Eshed, Y., Petiard, V., Lopez, J. and Beck-Bunn, T. (1996). Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *Theor Appl Genet*. **92**, 213-224.

The Tomato Genome Consortium (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*. **485**, 635-641.

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

To, A., Joubès, J., Barthole, G., Lécureuil, A., Scagnelli, A., Jasinski, S., Lepiniec, L. and Baud, S. (2012). WRINKLED transcription factors orchestrate tissue-specific regulation of fatty acid biosynthesis in Arabidopsis. *Plant Cell*. **24**, 5007-5023.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L. and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. **7**, 562-578.

Tzin, V., Malitsky, S., Aharoni, A. and Galili, G. (2009). Expression of a bacteria bi-functional chorismate mutase/prephenate dehydratase modulates primary and secondary metabolism associated with aromatic amino acid in Arabidopsis. *Plant J*. **60**, 159-167.

Unver, T., and Budak, H. (2009). Virus-induced gene silencing, a post transcriptional gene silencing method. *Internat J Plant Genom*. ID:198680 (doi: 10.1155/2009/198680).

Van den Berg, R.A., Hoefsloot, H.C.J., Westerhuis J.A., Smilde A.K. and van der Werf, M.J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*. **7**, 142-157

Veerla, S. and Hoglund, M. (2006). Analysis of promoter regions of co-expressed genes identified by microarray analysis. *BMC Bioinformatics*. **7**, 384 (doi: 10.1186/1471-2105-384).

Veraverbeke, E.A., Lammertyn, J., Saevels, S. and Nicalï, B.M. (2012). Changes in chemical wax composition of three different apple (*Malus domestica* Borkh.) cultivars during storage. *Postharvest Biol Technol*. **23**, 197-208.

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 β-ketoacyl-CoA synthase. *J Exp Bot*. **55**, 1401-1410.

Watanabe, S., Mizoguchi, T., Aoki, K., Kubo, Y., Mori, H., Imanish, S., Yamazaki, Y., Shibata, D. and Ezura, H. (2007). Ethylmethanesulfonate (EMS) mutagenesis of *Solanum lycopersicum* cv. Micro-Tom for large-scale mutant screens. *Plant Biotechnol*. **2**4, 33-38.

Wellensen, K., Durst, F., Pinot, F., Benveniste, I., Nettesheim, K., Wisman, E., Steiner-Lange, S., Saedler, H. and Yephremov A. (2001). Functional analysis of the LACERATA gene of Arabidopsis provides evidence for different roles of fatty acid omega-hydroxylation in development. *Proc Natl Acad Sci U.S.A.* **98**, 9694-9699.

Winkel-Shirely, B. (2001). Flavonoid biosynthesis. A colourful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol*. **16**, 485-493.

Wu, G.Z. and Xue H.W. (2010). Arabidopsis b-ketoacyl-[acyl carrier protein] synthase i is crucial for fatty acid synthesis and plays a role in chloroplast division and embryo development. *Plant Cell*. **22**, 3726-3744.

Shao, Y., Zhu, H.L., Tian, H.Q., Wang, X.G., Lin, X.J., Zhu, B.Z., Xie, Y.H. and Luo, Y.B. (2008). Virus-Induce Gene Silencing in Plant Species. *Russian J Plant Physiol*. **55**, 168-174.

Yamaguchi, T., Kurosaki, F., Suh, D.-Y., Sankawa, U., Nishioka, M., Akiyama, T., Shibuya, M. and Ebizuka, Y. (1999). Cross-reaction of chalcone synthase

and stilbene synthase overexpressed in *Escherichia coli*. *FEBS Letters*. **460**, 457-461.

Yeats, T.H. and Rose, J.K.C. (2013). The formation and function of plant cuticles. *Plant Physiol*. **163**, 5-20.

Yeats, T.H., Buda, G.J., Wang, Z., Chehanovsky, N., Moyle, L.C., Jetter, R., Schaffer, A.A. and Rose, J.K.C. (2012). The fruit cuticles of wild tomato species exhibit architectural and chemical diversity, providing a new model for studying the evolution of cuticle function. *Plant J*. **69**, 655-666.

Yu, D., Xu, F., Zheng, J. and Zhan, J. (2012). Type III polyketide synthases in natural product biosynthesis. *IUMBM Life*. **64**, 258-295.

Zamir, D. (2001). Improving plant breeding with exotic genetic libraries. *Nat Rev Genet*. **2**, 983-989.

Zhao, S., Fung-Leng, W.-P., Bittner, A., Ngo, K. and Liu, X. (2014). Comparison of RNA-Seq and microarray in transcriptomic profiling of activated T cells. *PLoS One*. **9**: e78644 (doi: 10.1371/journal.pone.0078644).

Zhao, L., Katavic, V., Li, F., Haughn, G.W. and Kunst, L. (2009). Insertional mutant analysis reveals that long-chain acyl-CoA synthetase 1 (LACS1), but not LACS8, functionally overlaps with LACS9 in Arabidopsis seed oil biosynthesis. *Plant J*. **64**, 1048-1058.

Zheng, Y., Zhao, L., Gao, J. and Fei, Z. (2011). iAssembler: a package for de novo assembly of Roche-454/Sanger transcriptome sequence. *BMC Bioinformatics*. **12**, 453-461.

Zhong, S., Joung, J.G., Zheng, Y., Chen, Y.R., Liu, B., Shao, Y., Xiang, J.Z., Fei, Z. and Giovannoni, J.J. (2011). High-throughput ilumina strand-specific RNA sequencing library preparation. *Cold Spring Harb Protoc*. **8**, 940-949.

Zoratti, L., Karppinen, K., Escobar, A.L., Häggman, H. and Jaakola, L. (2014). Light-controlled flavonoid biosynthesis in fruits. *Plant Sci*. **5**: 534 (doi: 10.3389/flps.2014.00534).

WEBSITES

CLUSTALW 2.1: free online EMBL-EBI software for sequence alignments [http://www.ebi.ac.uk/Tools/msa/clustalw2/]

Example of commercial tomato varieties including pink tomatoes [http://www.tomatofest.com/]

FUZZNUC: free online tool to search for patterns in nucleotide sequences [http://emboss.bioinformatics.nl/cgi-bin/emboss/fuzznuc/]

GoMapMan: open web-accessible resource for gene functional annotations in the plant science [http://www.gomapman.org/]

MapMan: user-driven tool to display large datasets onto diagrams of metabolic pathways or other processes [http://mapman.gabipd.org/eb/guest/mapman/]

TAIR: the Arabidopsis Information Resource [http://www.arabidopsis.org/]

SGN: the SOL Genomics Network [http://solgenomics.net/]

TOMATOMA, tomato mutant archive [http://www.tomatoma.nbrp.jp/]