



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

A Multilevel, Developmentally Controlled Gene Engineering Strategy for Tomato Fortification and Protection

María Florencia Cocaliadis Caisson

Advisor

Dr. Antonio Granell Richart

Co- Advisor

Dr. Diego Orzáez Calatayud

Valencia, July 2017

Agradecimientos

El laboratorio 2.10 me ha cambiado por completo. Estos intensos años vividos constituyen un capítulo entero de mi vida, que concluye en mi fortalecimiento personal, así como en la generación de vivencias y nuevos amigos que siempre guardaré en mi corazón.

Toni te agradezco las oportunidades que me has dado durante estos años, esas estancias y congresos que me han enseñado tanto. A tu lado he aprendido a ser fuerte, independiente y capaz de tomar mis propias decisiones, serás un jefe que echaré de menos haya donde vaya y siempre recordaré. **Diego**, ha sido un placer compartir contigo tantos cafés y lab-meetings. Agradezco tus conversaciones, buen rollo y grandes ideas.

Silvia, mi amiga, maestra de buenas prácticas de labo y compañera de aventuras de montaña. Desde mi primer día supe que siempre podría contar contigo, gracias. Espero que nos queden muchas más ferratas juntas! Y ten mucha paciencia por aquí... que sin ti imperaría el caos!! **Asun**, mi amiga, profesora, profesora... persona 100% imprescindible en el 2.10! Gracias por tus buenos consejos y el saber escuchar. Espero tenerte siempre, allá donde vaya.



Agradecimientos

Joan, la última incorporación en mi vida y la más importante. Gracias por tu compañía, tus consejos, tu amor y tu amistad. Sin mencionar tu paciencia incondicional en este último periodo de escritura...Eres el mejor premio de tesis que se pueda pedir, te amo.

Marta, amiga. Las aventuras de mi doctorado las he vivido contigo. Has sido color en este capítulo de mi vida. Espero que siempre estés a mi lado y compartir muchos más viajes y experiencias contigo. **Alfredo**, tachinno, te echo de menos! Nuestras aventuras han sido breves pero intensas. Gracias por esas risas, consejos y buena compañía. Estoy deseando presentar la tesis para poder visitarte en EEUU.

Mis amigos de fuera pero no menos importantes. Habéis estado en mi vida de forma fugaz pero han sido grandes vuestras aportaciones. **Silvia** siempre recordaré nuestro viaje a Turin, no puedo decirte con palabras lo mucho que significo para mí. **Alessandro**, empezamos juntos y nos dimos mutuo apoyo. Fuiste mi primer amigo de labo y siempre serás especial. **Gabri**, gracias por tu bondad y tu felicidad, gracias por abrirme las puertas de tu casa en Viterbo y aquellos momentos en Italia con Alessandro. **Camilla**, cuantas risas, fiestas y que bonitos momentos por el sur!! **Ana Margarida** gracias por las comidas en el tranvía. **Diana** fue genial conocerte a ti y a tus bizcochos... pronto nos dejaremos caer por Portugal. Gracias a todos



Agradecimientos

vosotros, siempre formareis parte de mí y espero que nunca perdamos el contacto. Os quiero!

Julve, Sara y Estefania, mis compis 2.10!!!! Ha sido agradable y divertido pasar mí día a día a vuestro lado. Risas diarias, cerves, viajes varios y excursiones. Os echaré de menos a todos!

Fuera del 2.10 también hay mucha gente a la que agradecer y que han aportado mucho en mi vida estos años.

Empezando por ti, **María Ángeles**. Ha sido un placer conocerte y me da pena que las nuevas incorporaciones no tengan el gusto de verte a diario. He estado mis cuatro años pidiéndote favores y por eso sé que tampoco me olvidarás... gracias por tu sonrisa diaria y abrazos. **Juan** ha sido muy agradable compartir comidas/ horchatas contigo y tantas conversaciones interesantes. Sin ti ya no existe 4 labos y se te echa mucho en falta. Os deseo a los dos mucha felicidad en los años dorados que os vienen por delante, os la merecéis!

Raquel, lo nuestro fue breve pero muy agradable. Espero que por ahí fuera te esté yendo todo bien. Gracias por esas comidas en el tranvía, siempre las recordaré. **Marco**, si no fuese por ti habría muerto de inanición, siempre asegurándote de no saltarnos la hora de la comida. Es un placer haberte conocido, sabes que te aprecio mucho y te deseo lo mejor. Me he cansado de decírtelo... pero por última vez y para que quede registro... haz el favor de vivir un poquito!. **Noel** lo hemos pasado bien juntos. Gracias por saber escuchar y estar siempre ahí, por



Agradecimientos

nuestras fiestas y viajes. Espero que nos queden muchos más. Y en general, a todo el 2.07, gracias porque sois una gran familia.

Los servicios del IBMCP también merecen reconocimiento. **Eugenio**, que loco que estás! Cuanto me he reído contigo... te agradezco tus consejos varios y la buena calidad de tus secuencias. **Rafa y Primi**, para mí siempre ha supuesto felicidad ir de excursión al inver porque sabía que ibais a estar allí para mí, sois los dos geniales, gracias! **Marisol**, gracias por ser tan simpática e intentar ayudar siempre con tus microscopios. Y **María**, la de la limpieza, a ti decirte que ese matutino "buenos días guapa" lo echaré de menos...

Fuera del IBMCP, más bien fuera de esta ciudad, también hay mucha gente a la que agradecer. **Rafa (Fali)**, jefe y amigo. He disfrutado en cada una de nuestras conversaciones y cada minuto en la Mayora rodeada de tus gatos. Nunca te olvidaré y espero que jamás perdamos el contacto. Gracias de verdad por tu compañía, consejos y ayuda. En Toulouse, agradecer al laboratorio de genómica y biotecnología del fruto. **Mondher, Simo, Isabelle, Pierre, Brigit, Benoit** y mis compis de aventuras **Baowen, Ibtihaj, June y Cougant**. Esos tres meses allí han marcado una diferencia en mi vida y siempre os recordaré. Quiero dar las gracias a **Dani Zamir** por abrirme las puertas de su laboratorio en Israel, a toda su gente y especialmente a **Joseph** por ser mi guía y compañero en aquellos invernaderos.



Agradecimientos

También agradecer a **Gianfranco** por nuestra colaboración y por todos esos dulces! Ha sido un placer conocerte y trabajar contigo, espero que sigamos en contacto.

Gracias a mis amigas **Andrea** y **Ángela** por su apoyo incondicional y por estar siempre a mi lado tanto para los buenos como malos momentos.

Por último y más importante, gracias a **mamá y papá, Agu y Bele** porque me habéis dado todo en la vida y me considero la persona más afortunada del mundo por vuestro amor y apoyo. La familia no se escoge, pero si tuviera la opción nunca cambiaría la mía por nada ni nadie.



Summary

Plastids are the cellular organelles where many of the visual, health and flavor-related metabolites are produced and stored in the fruit, and therefore are valuable components for consumers and breeders. The more sugar and flavor the fruit has, the more appreciated is for the consumer and industry. Thus, one of the breeder's goals is to obtain new varieties with fruits improved in these aspects. Paradoxically, fruits with a high content in chloroplasts have been avoided by the breeders because it usually suffers of oxidative stress disorders; such yellow shoulder impairment and fruit cracking when the light intensity increases. For this reason breeding efforts has been focused mainly on avoiding fruit losses and organoleptic characteristics have been neglected.

This thesis aims to improve tomato fruit quality by engineering plants to produce fruits with enhanced fruit chloroplast functions and improved tolerance to oxidative stress, using cisgenic/ intragenic approaches. *SIGLK1*, *SIGLK2* and *SIAPRR2* transcription factors have been suggested to be involved in chloroplast development. Tomato MoneyMaker plants were engineered to express *SIGLKs* and *SIAPRR2* either singly or in combination early in development. Those lines provide fruits which accumulate more sugars, carotenoids and specific volatiles than WT. The fruit chloroplast enhanced lines were characterized at the structural, metabolic, proteomics and transcriptomics. A novel additive effect in the chloroplast regulation network resulted when both transcription factors were coexpressed and a hypothesis for this effect is presented.

In addition, two tomato traditional varieties (Muchamiel and Pera) expressing tomato genes for BMW anthocyanin regulatory complex under the control of the light inducible promoter (PLI) were produced and characterized. Engineered tomato plants showed large accumulation of anthocyanin specifically in the fruit peel and in Type VI trichomes. Characterization of those tissues indicated specific alterations of the flavonoid pathway that were highly dependent on the light conditions. These tomato lines could be of high interest to protect the fruit chloroplast enhancement lines from eventual stresses involving ROS, and also to assess the effect on plant growth under high light stress and in plant-pest interaction studies.



Resumen

Los plastidos son orgánulos celulares donde se producen y almacenan muchos de los metabolitos relacionados con atributos organolépticos y compuestos beneficiosos para la salud, por lo tanto se consideran componentes de alto valor añadido para consumidores y mejoradores vegetales. Cuanto mayor contenido en azúcares solubles y sabor presente el fruto más se valoran por los consumidores y la industria. Por lo tanto uno de los objetivos actuales de los mejoradores de tomate es mejorar el fruto en estos aspectos. Paradójicamente, se ha seleccionado en contra de frutos con alto contenido en cloroplastos porque este carácter, bajo alta intensidad lumínica, suele estar asociado a daños en el fruto por estrés oxidativo; como los hombros amarillos del tomate o el agrietado del fruto. Por este motivo los esfuerzos en mejora se han orientado principalmente a evitar pérdidas y consecuentemente la calidad organoléptica se ha visto reducida.

El objetivo de esta tesis es mejorar la calidad del fruto de tomate mediante el empleo de técnicas de ingeniería genética orientadas a incrementar los cloroplastos en fruto y mejorar la tolerancia al estrés oxidativo con una aproximación *cis/ intragenica*. Los factores de transcripción *SIGLK1*, *SIGLK2* y *SIAPRR2* han sido estudiados por influir en el desarrollo del cloroplasto. Plantas de tomate de variedad MoneyMaker fueron mejoradas genéticamente para expresar de forma individual o conjunta *SIGLKs* y *SIAPRR2* en estadios tempranos de desarrollo. Estas líneas proveen frutos con mayor acumulo de azúcares, carotenos y volátiles que el control MoneyMaker. Las líneas potenciadas en desarrollo de cloroplastos se caracterizaron a nivel estructural, metabólico, proteómico y transcriptómico. Se descubrió un novedoso efecto aditivo en la regulación génica del cloroplasto cuando ambos factores de transcripción se expresan simultáneamente y se presentó una hipótesis para dicho efecto.

Además se caracterizaron dos variedades tradicionales de tomate (Muchamiel and Pera) diseñadas para expresar genes pertenecientes al complejo de regulación de antocianinas BMW, bajo el control del promotor inducible por luz (PLI). Las plantas mejoradas genéticamente presentan una gran acumulación de antocianos, especialmente en piel de fruto y en tricomas tipo VI. Caracterización de estos tejidos indican alteraciones específicas en la ruta de flavonoides y una alta dependencia a condiciones de luz. Estas plantas podrían ser de gran interés para proteger frutos con altos niveles de cloroplastos frente al estrés oxidativo generado por ROS, para evaluar el efecto en el crecimiento de la planta bajo condiciones de alta luz y en futuros estudios de interacción planta-patógenos.



Resum

Els plastidis són orgànuls cel·lulars on es produeixen i emmagatzemen molts dels metabòlits relacionats amb atributs organolèptics i composts beneficiosos per a la salut, per tant es consideren components d'alt valor afegit per a consumidors i milloradors vegetals. Quant major contingut en sucres solubles i sabor presenta el fruit, més serà valorat per part dels consumidors i la indústria. Paradoxalment, s'ha seleccionat en contra dels fruits amb alt contingut en cloroplasts perquè aquest caràcter, davall alta intensitat lumínica, sol estar associat amb danys en el fruit per estrès oxidatiu; com muscles groguencs de la tomata o clevitjament del fruit. Per aquest motiu, l'esforç en millora s'ha orientat principalment a evitar pèrdues de manera que la qualitat organolèptica s'ha vist reduïda.

L'objectiu d'aquesta tesi es millorar la qualitat del fruit de tomata mitjançant l'ús de tècniques d'enginyeria genètica orientades a incrementar els cloroplasts al fruit i millorar la tolerància a l'estrès oxidatiu amb una aproximació cis/intragènica. Plantes de tomata de la varietat MoneyMaker foren millorades genèticament per expressar de manera individual o conjunta *SIGLK1*, *SIGLK2* y *SIAPRR2* als moments inicials del desenvolupament. Aquestes línies donen fruits amb major acumulació de sucres, carotens i volàtils que el control MoneyMaker. Les línies potenciades amb el desenvolupament de cloroplasts es caracteritzaren a nivell estructural, metabòlic, proteòmic i transcriptòmic. Es va descobrir un nou efecte additiu en la regulació gènica del cloroplast quan ambdós factors de transcripció s'expressen de manera simultània i es va presentar una hipòtesi per a dit efecte.

A més, es van caracteritzar dos varietats tradicionals de tomata (Muchamiel i Pera) dissenyades per a expressar gens que pertanyen al complex de regulació d'antocians BMW, davall el control del promotor induïble per llum (PLI). Les plantes millorades genèticament presentaren una gran acumulació d'antocians, especialment a la pell del fruit i en tricomes de tipus VI. La caracterització d'aquest teixit indica alteracions específiques en la ruta dels flavonoides i una altra dependència a condicions de llum. Aquestes plantes podrien ser de gran interès per a protegir els fruits d'alts nivells de cloroplasts front a l'estrès oxidatiu generat pels ROS, i per a avaluar l'efecte en el creixement de la planta davall condicions d'alta llum i en futurs estudis d'interacció planta-patògen.



Table de Contents

INTRODUCTION	1
1. Fruit plastids and their contribution to fruit quality	3
1.1. Role of photosynthesis in the fruit.....	4
1.2. Chloroplast are remodeled to chromoplast during ripening.....	9
1.3. Genetic regulation of genes involved in chloroplast development.	10
1.4. Fruit chloroplast reinforcement, a double-edged sword?.....	17
1.5. Walking on the wild side: oxidative stress during normal fruit ripening and under stressful environmental conditions.....	19
2. Photoprotection approach.....	20
2.1 Anthocyanins: natural source of photoprotection.....	20
2.2 Tomato fruit are capable of accumulating anthocyanins.....	21
2.3 Anthocyanin biosynthetic pathway.....	22
2.4 Regulation of anthocyanin Pathway.....	25
2.5 Tomato engineering for high content in anthocyanins.....	27
OBJECTIVES	31
CHAPTER 1: Plastid Fortified Tomatoes Obtained by Combined over-expression of <i>GLK</i> and <i>APRR2</i> Transcription Factors during Early Fruit Development.....	35
3. INTRODUCTION	37
4. RESULTS.....	40
2.1 Engineering tomato plants expressing <i>GLKs</i> and <i>APRR2</i> TFs under the control of the TFM5 promoter	40
2.2 <i>A dark green fruit phenotype is observed when GLKs and APRR2 are over-expressed simultaneously.....</i>	42
2.3 The largest increase in chlorophyll and carotenoid levels is observed in fruit at IMG stage from double engineered <i>GLKs-APRR2</i> plants.....	46
2.4 Primary metabolism also reflects the additive effect between <i>GLKs</i> and <i>APRR2</i> at IMG stage.....	49



2.5	<i>GLKs-APRR2</i> increases fruit plastid number, compartment size and grana stacking	51
2.6	Engineering <i>GLK</i> and <i>APRR2</i> early in ovary/ fruit development have consequences on fruit quality at the ripe stage.....	55
2.6.1	<i>dCEF</i> lines accumulate higher levels of total soluble solids....	55
2.6.2	<i>dCEF</i> lines show higher contents of nutritional and healthy compounds.....	56
2.6.3	Volatile profiles modified in <i>dCEF</i>	57
5.	DISCUSSION	60
6.	MATERIALS AND METHODS	66
4.1	Plant materials and growth conditions.....	66
4.2	Plasmid construction and generation of transgenic plants	68
4.3	Transgenic plants evaluation.....	69
4.4	RNA extraction and qRT PCR expression analysis.....	70
4.5	Carotenoid and chlorophyll analysis.....	71
4.6	Light and transmission electron microscopy	72
4.7	Volatile compounds profiling.....	74
4.8	Primary metabolite profiling.....	74
4.9	Data analysis and visualization	75
CHAPTER 2: Molecular mechanisms underlying the fruit phenotypes driven by early expression of <i>GLK</i> and <i>APRR2</i> . 77		
7.	INTRODUCTION	79
8.	RESULTS	82
2.1	Transcriptome data overview shows changes consistent with the metabolomic fruit phenotype	82
2.2	Early activation of <i>GLK2</i> or <i>APRR2</i> under the control of a fruit specific promoter affects a common set of genes that are related to fruit photosynthesis and chloroplast development.....	85
2.3	Nuclear genes implicated in chloroplast retrograde signaling are DEG in CEF	88
2.4	Joint expression of <i>APRR2</i> and <i>GLK2</i> early in development results in further reinforcement of chlorophyll synthesis and photosynthesis machinery in the fruit.....	89



2.5	The boost in photosystem machinery in double engineered <i>GLK-APRR2</i> fruit is maximized when using <i>GLK1</i>	91
2.6	Early expression of <i>GLK2</i> and <i>APRR2</i> still has a quantitatively large effect on fruit transcriptome at MG stage	94
2.7	Fruit chloroplast enhancement in <i>GLK1-APRR2</i> lines results in a higher total protein content.....	97
2.8	Proteomics analysis of fortified fruit revealed that mitochondria and chloroplast are the most affected cellular components	101
2.9	Chloroplast biogenesis increases in CEF genotypes.....	104
2.10	Largest differentially protein portion analyzed belong to chloroplast proteome	106
2.11	Relationship between proteome and transcriptome	108
2.12	Over-represented motives in common upregulated genes.....	109
9.	DISCUSSION.....	110
3.1	GLKs and <i>APRR2</i> affecting an array of chloroplast biogenesis and chloroplast functions including retrograde signaling program.....	111
3.2	Additive effect resulting from joint expression of <i>GLK</i> and <i>APRR2</i> early in fruit development and the possible underlying mechanism.....	114
3.3	Increased in plastid development is well-correlated with the upregulation of genes and proteins process-associated.....	117
3.4	Fruit photosynthesis enhancement supported by upregulation of carbon fixation enzymes.....	119
10.	Materials and Methods.....	120
4.1	Plant materials and growth conditions	120
4.2	Protein analysis	121
4.3	RNA extraction	121
4.4	Global transcriptomics profiling by RNAseq.....	122
4.5	RNA Data Analysis.....	123
CHAPTER 3: Expression of tomato complex BHLH:MYB:WD40 under the light inducible promoter PLI results in high levels of anthocyanin accumulation in the fruit peel and Type VI glandular trichomes.....		
11.	INTRODUCTION.....	127



12. RESULTS.....	132
1.1 Generation of tomato plants overexpressing tomato BHLH:MYB:WD40 complex in fruit peel.....	132
1.2 Engineered fruit display intense purple pigment accumulation in fruit peel.....	134
1.3 Biochemical analysis revealed the detailed composition of anthocyanin-enriched fruits.....	137
1.4 Unexpected anthocyanin accumulation in type VI trichomes.....	140
1.5 Trichome-type Specificity	142
1.6 Anthocyanin trichome expression appears to require BM transcriptional complex to be driven the PLI promoter.....	145
1.7 Specificities of the alteration of the flavonoid pathway in type VI glandular trichomes driven by the PLI::BMW.....	147
1.8 AEP phenotype is dependent on light.....	149
1.9 Differences in light intensity produces dramatic changes in the flavonoid profile of the fruit peel.....	154
13. DISCUSSION	157
1.10 Purple pigment in our AEP is restricted to peel similarly as it happens in fruit of wild species.....	157
1.11 Unexpected novel localization of anthocyanin in tomato trichomes	158
1.12 Engineering trichomes as biofactories.....	161
1.13 Enriched fruit flavonoid profile.....	162
1.14 Light dependent phenotype	164
1.15 Photo-oxidative Stress Protection.....	165
1.16 Conclusion and future perspectives	166
14. MATERIALS AND METHODS	167
1.17 Plant materials and growth conditions.....	167
1.18 Anthocyanin module construct and plant transformation.....	168
1.19 Transgene expression determination by quantitative real time PCR	169
1.20 Sampling method.....	170
1.21 Microscopy characterization.....	170



1.22 LC-ESI(+/-)-MS analysis of phenylpropanoids and anthocyanins in tomato fruits and trichomes 171

GENERAL DISCUSSION 175

15. Fruit Photosynthesis..... 177

16. Early-fruit Specific Overexpression of *GLKs* and *APRR2* TFs: Chloroplast and Photosynthesis Boost with Additive Effect..... 179

16.1 A Multi-level Analysis Confirmed the Chloroplast Boost..... 180

16.2 Beyond Chloroplasts- Mitochondria and other Related Functions 182

16.3 Molecular Basis of the *GLK::APRR2* additive Effect..... 183

16.4 Fruit Photosynthesis Boost Demonstrated through Metabolites Associated..... 184

16.5 An Early Chloroplast Boost Impacts on Red Fruit Quality..... 185

17. A Possible Solution for Tomato Fruit Photooxidative Stress. 186

17.1 Novel Anthocyanin Trichome Phenotype..... 188

18. Future perspectives for CEF and AEP 189

CONCLUSION..... 191

SUPPLEMENTARY FIGURES..... 195

SUPPLEMENTARY TABLES.....207

REFERENCES.....213



Abbreviations

A2 APRR2 lines

AEP Anthocyanin Engineered Plant

BMW bHLH/MYB /WD40

Br breaker stage

Car Carotenoids

CDS Coding Sequence

CEF Chloroplast Engineered Fruit

Chl Chlorophyll

dCEF Double Chloroplast Engineered Fruit

DEG Differentially Expressed Gene

DEP Differentially Expressed Protein

DPA post-anthesis days

DW Dry Weight

EBGs Early Biosynthetic Genes

Fe Fruit epidermis

FW Fresh Weight

G1A2 GLK1-APRR2 lines

G2 GLK2 lines

G2A2 GLK2-APRR2 lines

GB Golden Braid

GC-MS Gas Chromatography Mass Spectrometry



Abbreviations

GO Gene Ontology

GS Grana Stacking

HPLC High-PerformanceLiquid Chromatography

IMG Immature Green

La Locular area

L Luminosity

LBGs Late Biosynthetic Genes

LR Light Red

MG Mature Green

MM MoneyMaker

OE Over-Expression

Pc central Pericarp

PCA Principal Component Analysis

ROS Reactive Oxygen Species

RR Red Ripe

sCEF Simple Chloroplast Engineered Fruit

SP Strong Phenotype

TF Transcription factor

TSS Total Soluble Solids

WP Weak Phenotype

WT Wild Type



INTRODUCTION



1. Fruit plastids and their contribution to fruit quality

Tomato fruit is a model system for fleshy fruit ripening and at the same time is one of the most widely consumed vegetables worldwide, this making it an ideal system to assess genes, processes and environmental conditions affecting fruit ripening and quality ¹. The conspicuous presence of chloroplasts in the highly vacuolated cells of tomato fruit pericarp is responsible of the green color of the tomato fruit before ripening, but the contribution of fruit photosynthesis to fruit metabolite composition is not well understood ². Fruit ripening involves the acquisition of a series of fruit traits that makes it attractive and palatable ³. The conversion from chloroplast to chromoplast is an important part of this ripening process and it is normally associated with the dismantling of the photosynthetic machinery and the accumulation of carotenoids in chromoplasts, which occurs concomitantly with the accumulation of sugars, organic acids and volatile aroma compounds in the fruit cells ^{1,4}. Fruits, as photosynthate sinks, rely on the energy provided by sugars transported from leaves to conduct the highly demanding processes of development and ripening, and eventually these photosynthates contribute to the fruit organoleptic properties.

Recent investigations ⁵⁻⁷ suggest important links between the levels of fruit photosynthetic activity and the final organoleptic quality of the ripe fruit. Furthermore, the activation of genes encoding transcriptional factors that enhance chloroplast development in fruit result in higher sugar contents next to enrichment in tomato fruit-specialized metabolites ^{5,7,8}. These recent results suggest new ways to improve fruit quality by fortifying fruit chloroplasts and plastids, and prompted us to reevaluate the importance of chloroplast/ photosynthesis contribution to fruit development and ripening.

1.6 Role of photosynthesis in the fruit

It is generally assumed that leaves are powerhouses of the plant as they provide photo-assimilates, while fruits and roots are basically heterotrophic as they rely on nutrients transport from leaves to grow and develop to their final size and composition. If this is the case, what is the need for fruits of tomato to be green before ripening? Which is the role of fruit chloroplasts? Do they have any effect on the final fruit quality?

Our understanding of fruit chloroplast formation and of the role of photosynthesis in fruit at the different stages of development is rather incomplete when compared to leaves. It is known that fruit chloroplast formation, chlorophyll biosynthesis and the assembly of the photosynthetic apparatus require exposure to light and activation of a series of developmental cues. Chloroplast proteins involved in light-harvesting complexes, electron transfer and in CO₂ fixation are all expressed in the fruit cells and they are regulated by transcription factors, in a similar manner as it occurs in leaves^{9,10}. There is, however, some fruit-specific regulation of nuclear-encoded photosynthetic genes¹¹⁻¹⁵ in a similar manner as it has been described for trichomes^{16,17}; the purpose of this specific regulation is not clear, but is probably related to function optimization in the context of the tissue. Therefore, a number of critical questions remain to be answer: are these fruit chloroplasts capable of photosynthesis? Are fruit chloroplast net contributors or is the bulk of carbon accumulation in fruit development simply reliant on photo-assimilates imported from leaves? What is the contribution of fruit chloroplasts and fruit photosynthesis to the fruit metabolism before ripening? What role do chloroplast and fruit photosynthesis have on chloroplast conversion to chromoplast and in fruit quality at the red ripe stage?

The role of fruit photosynthesis in fruit metabolism and development has been extensively discussed^{13,14,18-20} but even now, information about its importance is controversial. One of the issues is whether the fruits are or are not net producers of fixed carbon. There are reports indicating that tomato fruits are not net assimilators of carbon dioxide despite the high expression of photosynthetic genes in this organ^{10,21}. High expression of genes associated with photosynthesis occurs in specific fruit tissues with difficult access to light, such as the locule²², which although capable of photosynthesis²³, is also likely to display higher rates of respiration. Moreover, the triose-phosphate and glucose-phosphate transporters are both active in tomato chloroplasts, indicating that they could, in principle, both import and export phosphor-esters.

A number of studies in tomato support that the vast majority of photo-assimilates in the fruit are supplied by the leaves rather than produced *de novo* in the fruit²⁴⁻²⁷. Consistent with this idea, it has been shown that the correct development and the appropriate sugar composition of fruit largely depend on the size of the photo-synthate pool available in leaves and also on the sink strength of the fruit^{28,29}. Genetic analyses of fruit growth and composition have also confirmed the importance of both photo-synthate pool²⁶ and sink strength^{30,31}. Furthermore, a series of studies involving reverse genetics and QTL analysis revealed that the main genetic factors contributing to fruit size and sugar content are those affecting (I) cell division/ number of cells³², (II) auxin signaling³³ in the initial stages of development, and (III) the ability to convert imported sucrose into glucose and fructose in the final developmental stages³⁰; all three factors are consistent with the importance of developing of strong sink signal. Further support to the importance of leaves vs fruit photosynthesis comes from a number of studies where elevation of tomato leaf photosynthesis results in a proportional increase in fruit yield^{34,35}.

Carbohydrate pool size in leaves and its partitioning between leaves and fruit are affected by a variety of environmental conditions, including those cultural practices which are known to determine fruit growth and quality³⁶⁻³⁸. For instance, the exposition to salinity, a cultural practice to obtain better quality tomatoes, causes an increase in soluble sugars in ripe tomato fruit apparently as a consequence of the upregulation of sucrose transport in leaves and the increased activity of ADP-glucose pyrophosphorylase in early fruit development³⁹.

To apparently close the issue for good, a recent study² indicated that fruit photosynthesis is not required for correct fruit development, or for the accumulation of photosynthates, including those impacting taste. In this study, transgenic MoneyMaker plants which repressed in fruits the expression of the chlorophyll biosynthesis gene glutamate 1-semialdehyde aminotransferase (GSA) were not affected in size or metabolite accumulation, despite showing a reduced photosynthetic rate, as determined by both CO₂ exchange and by the levels of intermediates of the Calvin-Benson cycle. These results suggest that transport from leaves can compensate for loss of fruit photosynthesis. The only phenotype observed was a delay in seed development, suggesting that fruit photosynthesis may be important for timely seed development. These results seem to indicate that fruit photosynthesis is dispensable in fruit energy metabolism, although it may be relevant under specific environmental conditions.

Despite all the above, it remains the case that as much as 20% of the total carbon of the fruit has been found to result from photosynthetic activity in the fruit itself⁹. All three stages in which fruit development is traditionally described including cell division, cell enlargement and ripening contribute to final sugar accumulation in the fruit. Particularly, the second stage is accompanied with the degradation of starch into soluble sugars^{20,40} with early studies already indicating that level of soluble

solids in ripe tomato fruit is related to level of the starch in immature and mature green fruit ⁴⁰. Fruit photosynthesis contribution to fruit growth and net sugar accumulation has been supported by early fruit shading experiments which analyzed the rate of fruit growth to conclude the fruit contributes by its own fixed carbon between 10% and 15% of the total ⁴¹.

There have been criticisms about these experiments as the bagging procedure may impact light receptors which are required for normal fruit development ^{42,43}. Additional support to the contribution and importance of fruit photosynthesis comes from experiments showing 15-20% negative effect on fruit development by depleting photosynthesis in the fruit following the antisense inhibition of the fruit chloroplastic Fru-1,6-bisphosphatase (FBPase) ⁴⁴. The importance of fruit photosynthesis in early fruit development is also suggested by the results of combined metabolomic and transcriptomic analyses ⁴⁵.

Important to mention is fruit are able to fix carbon (refixation), via malate, with the CO₂ supply for this derived from mitochondrial respiration of imported carbon ²¹ rather than from intercellular/ stomata/ diffusible CO₂, whose accessibility could be in part compromised in bulky organs like the tomato fruit with a thick cuticle and near absence of stomata. The possibility also exists that CO₂ generated by the oxidative pentose pathway could be re-assimilated by RuBisCO in green fruits as shown in green seeds ⁴⁶ thus providing further efficiency to the system.

All these reports taken together indicate that fruit photosynthesis contributes to fruit development and carbon economy. This contribution can be dispensable under normal growth but may become important under certain, limiting environmental conditions. The use of different genetic backgrounds and different environmental growth conditions (light) in the studies may also have contributed to the sometimes

contradictory results. Interestingly, we now know that the genetic background is determinant in the degree of contribution of fruit photosynthesis to fruit growth and fruit carbon economy.

This conundrum about the importance of fruit chloroplasts and fruit photosynthesis has received an unexpected twist with the recent identification of the genetic nature of the “uniform ripening” *u* mutation⁵. Tomato *U* gene mutants exhibit defective fruits in that chloroplast number and thylakoid grana are dramatically reduced as compared to wild type, and yet fruit develops to normal size and ripens at the same time as wild type, although the accumulation of sugars in red fruit is repressed by 10-15%. Consistent with this, the so called uniform ripening tomato mutants show a pale green fruit phenotype at the mature green stage, which contrasts with the darker-green shouldered phenotype of the wild type fruit⁴⁷⁻⁴⁹.

The introgression of the wild type *U* locus into fruits of (*u/u*) converts tomato pale immature fruit into dark-green-shouldered fruits with higher starch accumulation. The uniform ripening gene has been recently identified by positional cloning and the *u* mutants revealed to carry an additional *A* in a small *A*-repeat region in exon I of the *GLK2* gene. This introduces a premature termination codon in the encoded protein. Since *GLK2* is the predominant *GLK* form expressed in the fruit, and since *GLKs* are members of the GARP family of transcription factors that regulate the expression of chloroplast genes⁵⁰⁻⁵², *u* mutants qualify as fruit chloroplast mutants, the rest of the plant's needs being satisfied by the *GLK1* gene, with redundant *GLK* roles in vegetative tissues.

Adding wild type alleles of *GLK*, either by crossing or by genetic transformation under different promoter sequences, produces fruit with more chloroplasts that have more grana /thylakoids, and higher accumulation of chlorophyll and starch at the mature green stage⁵. Fruits with activated *GLKs* express

higher levels of transcripts for PSII and PSI components, as well as of other genes involved in sugar (*GLK2*) or other aspects of metabolism (*GLK1*)⁵ (Granell unpublished). Furthermore, *GLK*-overexpressing lines accumulate more sugars and lycopene in the red ripe stage, opening a way to increase fruit quality by acting at the *GLK* level.

Most of the tomato cultivated varieties that fill the stalls of large supermarkets, including the variety used for the GSA experiment², carry the *u* mutation and therefore have defective fruit chloroplasts with the associated effects of lower sugars and lycopene levels that they could potentially have.

1.7 Chloroplast are remodeled to chromoplast during ripening

No matter what the contribution of fruit chloroplasts is to net photosynthesis in green fruits, tomato fruits clearly undergo during ripening a physiological transition associated with the differentiation of chromoplasts from photosynthetically-active chloroplasts^{53,54}. This transition appears to be coupled with a decline in both nuclear- and plastid-encoded gene expression^{13,14,18,54} and in enzymatic activities^{19,20} that are associated with carbon assimilation and chloroplast components. Furthermore, most nuclear gene expression for chloroplast proteins, and all plastid-encoded photosynthesis gene expression, are developmentally regulated, decreasing already during the late stages of green fruit development prior to ripening⁵⁴ in anticipation to the partially-phototrophic to completely-heterotrophic shift occurring in the fruit at ripening. Only in some mutants this process is partially blocked, as is the case in the stay-green mutants⁵⁵.

It is noteworthy that most of the proteins of the Calvin-Benson cycle, including RuBisCO, and of the oxidative pentose pathway were identified in the proteome of the tomato chromoplast⁵⁶. If all these pathways were active, the CO₂ generated by the oxidative pentose pathway could be re-assimilated by RuBisCO⁴⁶ to satisfy the specific metabolic needs of the ripe fruit. Therefore, as fruit ripen, their chloroplasts are remodeled into chromoplasts that no longer contain chlorophyll but synthesize and accumulate lycopene, β-carotene and other metabolites important for ripe fruit sensory and nutritional attributes^{1,9,56}.

How can then fruit chloroplasts and photosynthesis in green fruit affect the metabolite constitution of ripen fruit? Early studies have related the content of soluble solids in ripe tomato fruit to the starch level in the immature and mature green fruit stages^{20,57} but this is probably only part of the story. Genes encoding components of fruit chloroplast development or light signaling represent genetic tools for manipulation of fruit color and nutritional value and support the importance of fruit photosynthesis.

1.8 Genetic regulation of genes involved in chloroplast development

There is evidence that tomato fruit, either selected from the available gene pool or engineered to have an increased number of highly active chloroplast at the green stage, will develop ripe tomatoes with more active chromoplasts, producing higher amounts of metabolites associated to organoleptic and nutritional quality⁵⁸⁻⁶⁰. For example, the *high pigment 1, 2 and 3* (*hp1*, *hp2* and *hp3*) tomato mutants possess altered chloroplast number and ultrastructure, resulting in fruit with elevated levels of chlorophyll, carotenoids and flavonoids, and altered patterns of aroma volatile production^{59,61-63}. These mutants have been of limited use in tomato breeding programs due to pleiotropic

effects of the light signal perception phenotype that affect other aspects of plant growth, making it also difficult to infer the contribution of the fruit photosynthesis to the fruit quality phenotype ⁶³.

Fortunately, a number of studies using molecular genetics approaches on these mutants resulted in the identification of the corresponding mutations as components of the light signal transduction pathway and opened the way to a more targeted approach with less pleiotropic effects. A series of *hp2* alleles in tomato were revealed ⁶⁴ as mutations in *DE-ETIOLATED1 (DET1)*, a negative regulator of light signal transduction and *hp1* has also been identified as a mutation in a tomato *UV-DAMAGED DNA-BINDING PROTEIN 1 (DDB1)* homolog whose Arabidopsis counterpart interacts with *DET1*, and therefore both genes are part of the light signaling pathway.

The identification of the genes underlying the *hp* mutations allowed fruit targeted engineering ⁶⁵ using RNAi driven by fruit specific promoters resulting in plants that do not present pleiotropic effects. The engineered fruit contained a higher content of quality metabolites ⁶⁶.

Metabolomics and transcriptomic analysis of fruit with a specific downregulation of *DET1* failed to pinpoint the mechanism underlying the phenotype of enhanced plastids in the fruit, and although the authors identified an increased plastid biogenesis per se, they could not conclude this to be the primary effect ⁶⁷. Consistent with this, they failed to detect in the *DET*-modified plants any effect on the expression of either transcription factors involved in plastid biogenesis (e.g. *GLK*), structural proteins involved in plastid biogenesis (e.g. ribosomal proteins), or plastid division (*FtsZ* and *MinD*).

Collectively, these data could suggest that in the case of *hp* mutants it is the initiation of core metabolic processes which

drives subsequent plastid biogenesis to provide a defined cell compartment for these processes ⁶². In contrast, *GLK*-reinforced plants contain more and bigger chloroplasts with more and denser grana, in fruit tissues that have activated expression of transcripts for PSI and PSII components. Some differential specificities of expression, with consequences in fruit metabolism, can be anticipated from the expression pattern in *GLK1*- and *GLK2*- overexpression lines ⁵. Interestingly, no changes in the expression levels for the main light transduction genes are detected in fruit overexpressing *GLKs* ⁵, suggesting that at least at the transcriptional level no interaction between the core light signal transduction and *GLK* exists in the fruit. However, a number of evidences support that the interactions occur at the protein level. Thus (1) DDB1 and DET1 are known to essential components of CULLIN4-RING ubiquitin ligases (CRL4s), that target specific substrate proteins for 26S proteasome degradation ⁶⁸, (2) DDB1-CUL4-DET1 were found to interact in tomato according to the yeast two-hybrid assay ⁶⁹ and co-immunoprecipitation assays ⁶⁸; and (3) tomato *GLK2* protein appears to be a substrate for degradation mediated by the CUL4-DDB1-DET1 as *SLGLK2* protein is a substrate of the CUL4-DDB1-DET1 ubiquitin ligase ⁶⁸. Therefore the mechanisms by which genetically impairing *CUL4*, *DDB1* or *DET1* (as in *hp* ^{64,70} or in engineered lines ⁶⁷ results in a chloroplast/plastid fortification may be due to an impairment in *SLGLK2* degradation by the CUL4-DDB1-DET1 ubiquitin ligase/ 26S proteasome.

The combination of *HY5* and *GLKs*, functioning downstream of light and auxin/ cytokinin signaling pathways, is responsible for coordinated expression of the key genes in chloroplast biogenesis during root greening ⁷¹ and it is possible that this can be extended to aerial organs including fruits. In this respect, it is possible that the activity of *GLKs* on fruit chloroplast enhancement is not only affected by light but also by the changing hormonal status of the fruit ^{1,59}. A large body of experimental data supports the participation of phytohormones

in regulating fruit chloroplast formation/ differentiation and evolution to chromoplast during ripening⁵⁶. For instance, down-regulation of the tomato auxin response factor *ARF4* resulted in a fruit specific dark-green phenotype and blotchy ripening of tomato fruit⁷² with cells in the outer pericarp tissue displaying more chloroplasts with more grana per cell. More recent work by Sagar et al 2013 revealed that fruits of those plants accumulated higher amounts of transient starch at the immature stages and higher sugar accumulation at the red ripe stage. The possibility that *ARF4* action of plastids in the fruit maybe mediated by regulating the expression of *SIGLKs* in the fruit, was shown as the levels of both the fruit *SIGLK2* (inactive protein in Microtom background) and ectopic expression in the fruit or the leaf *SIGLK2* transcripts increased in *SlARF4-ASL* fruit tissues. The hypothesis that down-regulation of *ARF4* expression results in the up-regulation of *SIGLK1* and consequently in activation of chloroplast functions was also supported by the presence for perfectly conserved canonical ARF binding (TGTCTC box) in the promoter region of the *SIGLK1* gene⁷.

Later study suggest the role not only the auxins but also the brassinosteroids to regulate chloroplast development. Tomato transgenic lines which overexpress Brassinazole resistant 1 (*BZR1*) provides fruits with dark green shoulder because of the up-regulation of *GLK2*⁷³. The observation that the fruit phenotype can also be produced by ectopic expression in tomato of the cytokinin biosynthesis gene *ipt* from *Agrobacterium tumefaciens*⁷⁴ and that exogenous treatment with cytokinin can mimic the *hp* mutant phenotype⁶⁴ indicates the possibility of cytokinins to positively influence fruit chloroplasts. All these examples support a role for hormones in fruit chloroplast development and open new avenues of research for fruit photosynthesis modulation.

Others putative light signal transduction genes have proven useful to modify fruit traits. RNA interference (RNAi)-mediated

gene repression of two light signal transduction genes, *LeHY5* and *LeCOP1LIKE* were found to antagonistically regulate fruit plastid pigmentation in tomato ⁷⁰. Although the results obtained are consistent with their role as positive and negative elements of the light signaling pathway in Arabidopsis, the mechanism by which they alter fruit pigmentation and quality in tomato fruit is unknown.

More fruit chloroplast mutants have also proven useful to alter fruit quality. Thus, the green-flesh and chlorophyll retainer mutations of tomato and pepper showed an inhibition of chlorophyll degradation caused by dysfunction of the corresponding homologs of the chloroplast targeted STAY-GREEN protein of rice. Again, these mutations influenced fruit chloroplast development by delaying the degradation of chloroplast components during ripening, and impacted fruit quality in ripe fruit very positively ^{55,75}. Recently, the modification of other chloroplast proteins was found to impact fruit ripening and quality. Thus, the *Orange ripening (OrrDS)* mutant that encodes the M subunit of the plastidial NADH dehydrogenase complex ⁶⁰ is affected in multiple aspects of the fruit ripening process, including the ethylene climacteric and carotenogenesis. In contrast, fruit of *lutescent 1 (l1)* and *lutescent 2 (l2)* ⁷⁶ mutants of tomato show an early and progressive loss of chlorophyll producing fruit with a whitish-yellow color prior to the onset of ripening and a delay in the onset of fruit ripening itself. Both *l1* and *l2* have been identified as tomato homologs of a chloroplast-targeted zinc metalloprotease ⁷⁶ that is required for chloroplast development (*EGY1* in Arabidopsis, ⁷⁷). This supports the importance of a chloroplast signal that stimulates ripening and that eventually affects fruit quality.

Moreover, Pan et al (2013) using a network analysis approach identified transcription factor related to *ARABIDOPSIS PSEUDO RESPONSE REGULATOR2-LIKE* gene (*APRR2-Like*) that influences pigmentation and ripening in tomato fruit. Most

interesting overexpression (OE) in transgenic lines under 35s promoter the *APRR2-like* produced fruits with larger and more numerous plastids and consequently with higher pigment chlorophyll levels in immature unripe fruits and carotenoids in red ripe fruits. In Arabidopsis, this gene seems to be a transcriptional activator with a GARP DNA-binding domain and a receiver-like domain similar to ARR-B genes⁷⁸. Interestingly, this *APRR2-like* showed its closest similarity in tomato with the *GLK*⁵ but it lacks the AREAEAA motif conserved in the *GLK* genes⁷⁸ suggesting a new specificity. Although the function of the *APRR2* in Arabidopsis has yet to be determined the central position in the regulatory hub in tomato, its regulation during ripening and the effect on *RIN*, *ACO* and *PG* expression indicated it has an important role in the ripening process. On the other hand, the similarity of the ectopic OE phenotype with that of *GLK-like* genes, together with the alteration of *APRR* expression in *hp* mutant plants⁷⁹ indicates a link between the functional deficiency of the *DDB1* light signal transduction gene in the *hp1* mutant and *APRR2-Like* gene expression suggested and a possible explanation for its implication in plastid development and metabolism (Fig. 1). Recent reports indicating that pepper⁸ and cucumber⁸⁰ mutant lines which provide albino fruits are due to mutations in *APRR2*, which suggest that in addition to its role during fruit ripening *APRR2* is required for green fruit phenotype and therefore for fruit chloroplast development.

More recently, a tomato mutation called *uniform gray-green* (*ug*)⁴⁷ described by the lack of the green shoulder surrounding the calyx similarly to the "*u*" mutation was characterized by positional cloning. *Ug* was mapped to the long-arm of chromosome 1 and assigned to *TKN4* gene which encodes a type KNOX TF. *TKN4* was suggested to positively regulate *GLK2* expression levels. Higher levels of *TKN2* were also associated to increased chloroplast development and to the up-regulation of *SIAPRR2* and *SIGLK2* genes⁸¹.

All these results consolidate further the idea that genes encoding relevant fruit chloroplast biogenesis genes, or components of light signaling, can be the targets of genetic and biotech approaches for enhancing fruit composition and nutritional value.

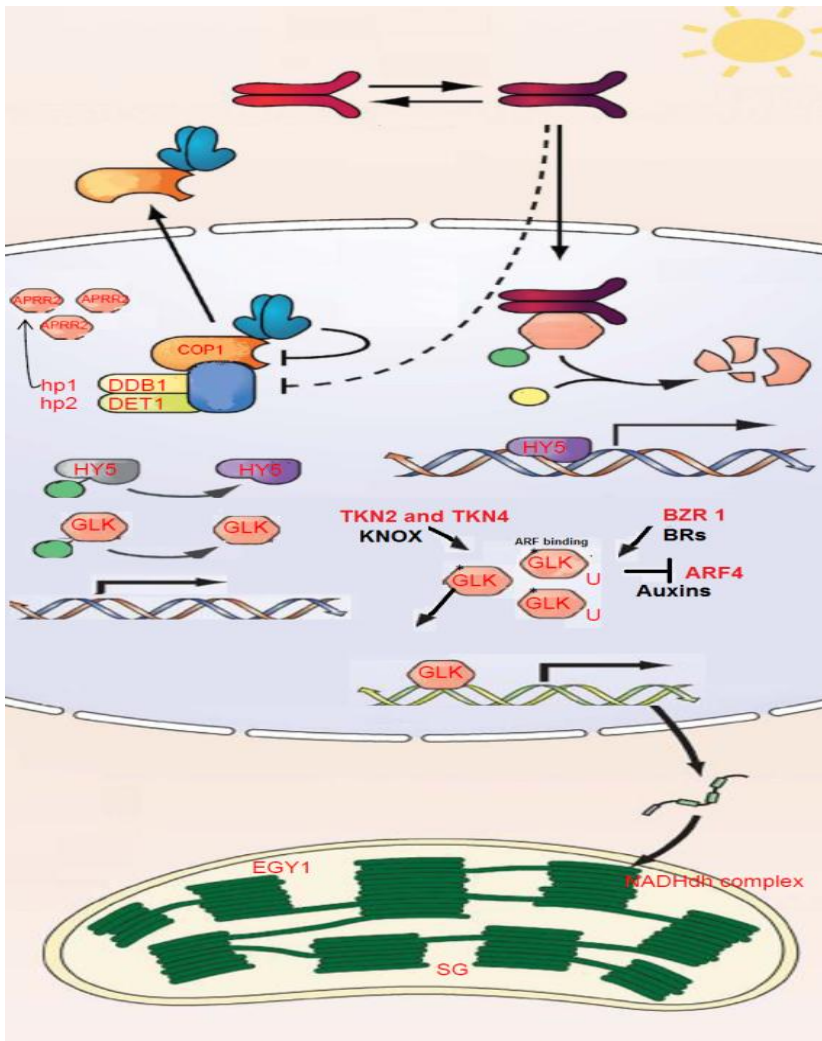


Figure 1: Light signalling and chloroplast development factors affecting fruit quality. In red are those components that in tomato have been reported to affect fruit chloroplast development and fruit quality (continued).

Figure 1 (continued): *COP1*, *DE-ETIOLATED 1 (DET1/hp2)* and *UV-DAMAGED DNA-BINDING PROTEIN 1 (DDB1/hp1)* cooperate in the dark to promote the ubiquitination and degradation (CUL4-based E3 ligase) of photomorphogenesis promoting transcription factors such as the transcription factor *HY5* or *GLK2*, whereas other light regulators probably act by repressing transcription of photomorphogenesis and pigment genes. Down-regulation of *LeCOP1-like*, *DET1/hp2*, and *DDB1/hp1* leads to higher chlorophyll/plastid and carotene content in green and mature fruit, respectively, while *HY5* or *GLK2* down-regulation causes the opposite effects, as observed in *hp1*, *hp2*, *HY5*, and *LeCOP1-like* mutants. Exposure to light causes *COP1* to exit from the nucleus and leads to the repression of *DET1/hp2* and *DDB1/hp1* and to active *HY5* and *GLK2* levels increasing. This enables G-box binding and transcriptional activation of photomorphogenesis and pigment genes. In a *DDB1/hp1*-independent regulatory pathway, *Golden 2-like* MYB transcription factors (*GLK1* and *GLK2*) participate; these are also major drivers of photosynthetic genes, with *GLK2* having a predominant role in fruits. Knox genes are upstream regulating *GLK2* expression. Alteration of phytohormones such as auxins or BRs also affects *GLK2* levels. Overexpression of *GLKs* increased chlorophyll content, chloroplast number, and thylakoid grana stacks in green fruit, which results later in higher levels of carotenoids and sugars in ripe fruit. [Adapted from Macmillan Publishers Ltd: The EMBO Journal. Waters MT, Langdale JA. The making of a chloroplast. EMBO J. 28, 2861-2873, figure 1, copyright (2009).]

1.9 Fruit chloroplast reinforcement, a double-edged sword?

Reinforcement of chloroplast in the fruit results in better quality fruit, so why was the *u* mutation that resulted in weak fruit chloroplasts introduced in 1940-50s in the first place? The reason appears to be that in addition to providing a fruit that ripens more uniformly, with no green shoulders, it facilitated maturity determinations crucial for the processing industry to avoid losses. Moreover, tomato breeders know well that *u* fruits are less prone to develop cracking and yellow shoulder at the stem end, two disorders often found under certain stressful environmental conditions^{82,83}. Absorbed solar energy may be harmful when it exceeds the capacity of photosynthesis to use it for assimilation, resulting then in photooxidative stress.

Photooxidative stress therefore exceeds the coping capacity of the plethora of mechanisms that plants have developed to manage excess of absorbed energy⁸⁴. Thus, part of the U fruit problem could be due to oxidative stress running out of control in very active fruit chloroplast as they become overheated, or unable to cope with high irradiation operating in many industrial facilities (which use completely different conditions to those at the centers of origin and diversification where wild relatives of tomato first evolved). Fruits of wild tomato species are inherently small as an adaptation to the limitations imposed by water and other environmental stresses in nature, and limited fruit growth is favored by survival and reproductive strategies during evolution over other fruit traits associated to growth and carbon gain, which are the focus of modern agriculture based on domesticated species with larger fruits.

Lhc/cab genes are among the most sensitive to photooxidative stress and their transcription is turned off soon after plants are subjected to photooxidative conditions^{85,86}. The decline of *cab* mRNA in maturing tomato fruit delayed chloroplast deterioration in a mutant tomato fruit⁸⁷ and coincided with the increase in superoxide dismutase⁸⁸. This strongly suggests that the decline in *cab* mRNA is in response to the increasingly photooxidative environment in maturing fruit.

On the other hand, control of redox chemistry is central to photosynthesis in the chloroplast and to cell survival and therefore is not surprising that plants go long ways in assuring that naturally. This control is exerted even when there is an imbalance between the two photosystems and is governed by the redox stage of the plastoquinone providing the plant with a short, rapid and direct control on photosynthesis which is encoded universally within the chloroplast⁸⁹.

1.10 Walking on the wild side: oxidative stress during normal fruit ripening and under stressful environmental conditions

Oxidative stress is induced in tomato fruit upon exposure to high irradiance and high temperature and, presumably, upon rising of photosynthetically-produced O₂ within a tissue in which gas exchange may be reduced because of near-absence of stomata. This may have consequences of fruit plastid bleaching and later lead to impairment in ripening that affects mainly the shoulders of the fruit. Yellow shoulder is a color and ripening disorder⁹⁰ of mature fruits (although probably induced when they are immature, Fernández-Muñoz, unpublished) injuring all types of tomatoes under high-light/ temperature conditions. It will, nevertheless, affect especially, severely fruits from genotypes lacking the uniform ripening gene⁹¹. It is long known that non-uniform green U tomatoes are far more sensitive to cuticle cracking⁹²⁻⁹⁴. Similarly to the yellow shoulder disorder, fruit cracking is also more frequent when developing fruits are directly exposed to high irradiance⁸³. It is therefore not surprising that to protect fruit from the deleterious, quality-reducing effect of oxygen radicals produced from electron transfer going out of control under high-irradiance/ temperature, breeders selected fruits with fewer number of inferior chloroplasts.

During fruit ripening, chloroplasts develop into chromoplasts accumulating high amounts of liposoluble antioxidant compounds, such as lycopene and beta carotene, which presumably will protect the chromoplast and fruit cells, in addition to serve as a visual cue for frugivores to eat the fruit and disperse the seed. Despite this, during the last stages of normal tomato fruit ripening oxidative stress in the fruit is known to increase markedly; this has been proposed to be part of the ripening program aimed at facilitating metabolic changes

associated to ripening and fruit softening required for seed release ⁹⁵. Consistent with this, short shelf-life tomato cultivars exhibit increased oxidative stress and reduced radical scavenging activity ⁹⁶. Engineering tomato fruit for higher antioxidant levels resulted in firmer fruit with extended fruit shelf-life ^{5,97,98}.

2. Photoprotection approach

Enhancing photoprotection has been proposed and demonstrated as a useful strategy to improve photosynthesis ⁸⁴. These approaches should be finely tuned and their beneficial effects should be weighed against the evidence that such actions may have a cost that could limit carbon availability, i.e. yield ⁸⁴. All this indicates that it should be possible to increase the fruit own supply of healthy nutrients by enhancing fruit plastids at the different stages of fruit development (fortification of both chloroplasts and chromoplasts) while keeping oxidative stress under control.

Understanding the mechanisms to achieve a fine redox control and manipulation of photo/ thermo-protection mechanisms in the plastid under real growth conditions will require further research to take full advantage of chloroplast fortification.

2.1 Anthocyanins: natural source of photoprotection

Under certain stress conditions plants react and produce anthocyanins compounds as a protective mechanism. Many environmental cues have been reported to stimulate the pigment synthesis such as low temperature, UV irradiance, drought, osmotic stress (sucrose, glucose, manitol), nutrients

deficiency (nitrogen, phosphorus), hormones (auxins and cytokinins) or wounding^{99,100}.

Anthocyanin synthesis is considered to be a light-dependent process; therefore, photoinduction is one of the most studied of the induced mechanisms. It has been demonstrated that is a phytochrome-mediated process where expression of enzymes in the pathway (CHS) is strictly dependent on light¹⁰¹⁻¹⁰³.

Anthocyanin not only acts as UV-b filter against high energetic irradiance (absorbing blue light and reflecting red wavelengths), but also is consider a sunscreen protector of PSII^{9,104,105}. The mechanism underlying flavonoid protection in ROS reduction is not well understood. However, it has been proved that is possible to stimulate flavonoid synthesis in low irradiance (kaempferol and quercetin glycosides) when other stresses, such as lack of nitrogen or salinity, are present. These conditions lead to ROS generation that could be mitigated by flavonoids¹⁰⁶.

In addition, the antioxidant capacity of those pigments makes them also a relevant trait for fruit breeding with impact in different stakeholder interests. Thus, it has been reported the capacity of flavonoids/ anthocyanins to extend the shelf-life of the fruits - important for industry, as well as the benefits associated to high levels of flavonoids in a regular diet¹⁰⁷⁻¹⁰⁹.

2.2 Tomato fruit are capable of accumulating anthocyanins

Modern tomato varieties do not accumulate anthocyanins in fruit. Nevertheless, is possible to find wild relatives of tomato which fruits store anthocyanins in vacuoles of the sub-epidermal layer¹¹⁰.

Many efforts have been oriented to the introgression of this trait in tomato with limited success and mostly restricted to the fruit peel. Anthocyanin fruit production from *S. chilense* was introduced in domesticated tomato through interspecific crosses, and this phenotype was associated to *Aft* gene¹¹¹ mapped in chromosome 10¹¹². In the same locus *Aubergine* (*Abg*) was identified and introgressed in tomato from *S. lycopersicoides*^{113,114}. Both are dominant genes and require high-light to express the phenotype. The trait has been also identified in *S. cheesmaniae* and mapped in chromosome 7 as a recessive gene called *atviolaceum* (*atv*). This gene is able to produce anthocyanin accumulation not only in fruits but also in leaves¹¹⁵. Introgression of *atv* has been partially successful because of the weak phenotype in tomato. *Aft* and *Abg* can provide strong purple pigmentation but *Abg* plants can give blotchy fruit when the light source is not uniform. Despite differences in tissue-specific expression of genes, both fruit and leaves have similar anthocyanin profile with petunidin, malvidin and delphinidin accumulating to the highest levels^{116,117}.

Moreover, the double homozygous mutants *Aft-atv* present higher levels of fruit anthocyanins and lower light requirements for their production than individual lines indicating a synergistic effect between the two genes¹¹⁸. Similar results were obtained from the interaction of *atv* with *Aft* or *Abg* with anthocyanin, achieving in the pyramided lines higher levels than in any of them separately¹¹⁹.

2.3 Anthocyanin biosynthetic pathway

Although the above results do not require the understanding underlying molecular and biochemical pathways, the important progress made in identifying the components of the Flavonoids/Anthocyanins (FLV/ANT) pathways and their regulators have

facilitated more targeted approaches to increase FLV/ANT accumulation in the fruit.

Anthocyanins are water-soluble pigments derived from flavonoids through phenylpropanoid pathway (Figure 2). Phenylalanine is the amino acid precursor located at the beginning of the phenylpropanoid pathway. In addition, 3 enzymes -phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H) and 4-coumaroyl CoA ligase (4CL), are implicated in generating the intermediate, p-coumaroyl CoA, necessary in downstream steps of the phenylpropanoid metabolism. Condensation of one molecule of 4-coumaroyl CoA with three molecules of malonyl-CoA (Chalcona synthase, CHS) is the first committed step for the flavonoid pathway and produces naringenin chalcone, which is the precursor of all flavonoids.

Downstream anthocyanin synthesis continues from the three different kinds of classes of dihydroflavonols (Dihydroquercetin, Dihydrokaempferol, Dihydromyrecetin) where, a Flavonol synthase (FLS), diverts the pathway to the major flavonol glycosides in tomato (rutin and kaempferol-3-rutinoside), or otherwise the pathway continues through Dihydroflavonol 4-reductase (DFR). This enzyme catalyzes the conversion of dihydroflavonol into flavan-3,4-diols and subsequently, the transformation into anthocyanidins is catalyzed by anthocyanidin synthase (ANS). After that, anthocyanidin glycosides are formed by the addition of sugar ^{120,121}. Three kinds of anthocyanin compounds derived from the three different dihydroflavonols: delphinidin (dihydromyricetin), cyanidin (dihydroquercetin) and pelargonidin (dihydrokaempferol).

The pathway continues then to form other downstream anthocyanins, such as petunidin or malvidin (Fig. 2). Phenylpropanoids and flavonoids are normally stored in

vacuoles of guard and epidermal cells, as well as in sub-epidermal layers ¹²².

Expression levels of the flavonoid pathway genes in pericarp tissue are normally below detection levels, but fruit peel presents all necessary transcripts and enzymes for the precursors of anthocyanin biosynthesis ^{121,123}. Naringenin chalcone represents over 1% of the cuticle dry weight, and also significant transcript levels of CHS, hydroxylase (F39H), and flavonol synthase (FLS) are detected in the fruit peel. CHI enzyme is the only one in low expression. In contrast to that, most of the flavonoid pathway seems to be turned off in fruit flesh tissue ¹²⁴.

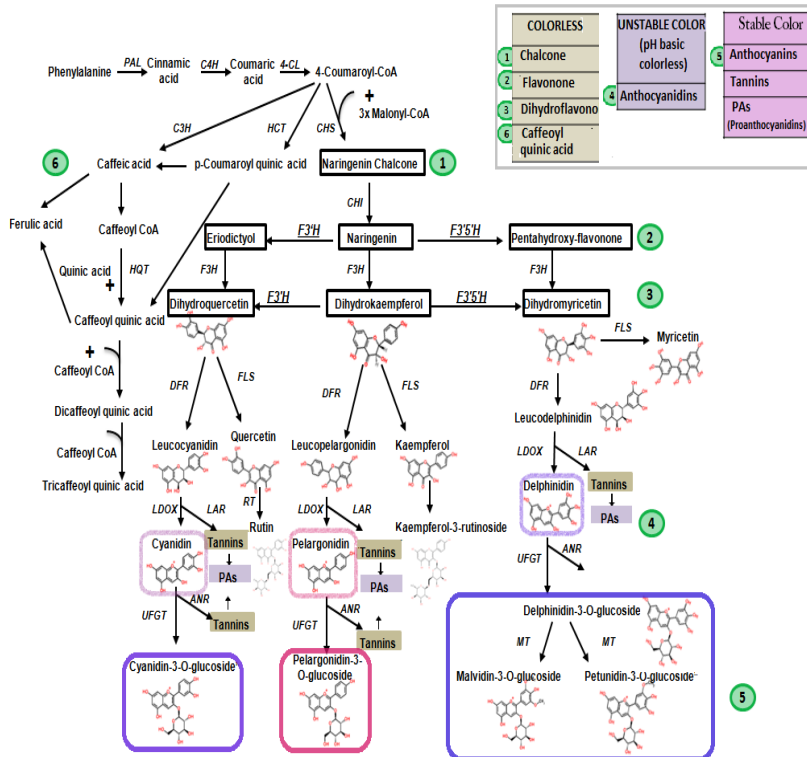


Figure 2: Schematic representation of flavonoid biosynthesis pathway. Upstream in the phenylpropanoid pathway phenylalanine is the precursor of a series of reactions catalyzed by PAL (ammonia lyase), C4H (cinnamate 4-hydroxylase) and 4CL (4-coumaroyl CoA ligase) (continued)

Figure 2 (continued): to generate the 4-coumaroyl CoA, precursor of the chalcones. From Naringenin chalcone pathway are synthesized flavonones, dihydroflavonols, anthocyanidins (alternative tannins or Proanthocyanidins) and anthocyanins. 4-coumaroyl CoA can also be a substrate for caffeoylquinic acids though *C3H* (p-coumaroyl ester 3-hydroxylase) activity, otherwise could be substrate for *HCT* (cinnamoyl CoA shikimate/quinate transferase) or *HQT* (hydroxycinnamoyl CoA quinate transferase). Enzymes represented in flavonoid pathway: *CHS* (chalcone synthase), *CHI* (chalcone isomerase), *F3H* (flavanone 3-hydroxylase), *F3'H* (flavonoid 3'-hydroxylase), *F3'5'H* (flavonoid 3',5'-hydroxylase), *DFR* (dihydroflavonol reductase), *LDOX* (leucoanthocyanidin oxidase), *UFGT* (UDP-glucose flavonoid 3-O-glucosyl transferase), *MT* (methyltransferase). Alternative branched from Leucoanthocyanidins: *LAR*, (leucoanthocyanidin reductase), *ANR* (anthocyanidin reductase). Major flavonoids groups of interest for our work are numbered and the reference is given in the upper right part of the chart. Pathway info was extracted from KEGGs (<http://www.genome.jp/kegg/>). Chemical structures from dihydroflavonols compounds to anthocyanins included (<https://pubchem.ncbi.nlm.nih.gov/>).

2.4 Regulation of anthocyanin Pathway

Regulation of anthocyanin biosynthetic pathway can be divided in three main blocks (Fig. 3). Two blocks are defined from phenylalanine to 4-coumaroyl CoA, and from this compound to the synthesis of non-anthocyanin flavonoids controlled by Early Biosynthetic Genes (EBGs). Finally, the third block is specific for the anthocyanin pathway and characterized by the regulation of Late Biosynthetic Genes (LBGs)^{125,126}.

Anthocyanin production at LBGs level has been described to be transcriptional regulated by a complex bHLH-MYB-WDR (BMW) described in several plant species and necessary for anthocyanin synthesis. In Arabidopsis, an R2R3-MYB protein called *Production of Anthocyanin Pigments 1 and 2* (*PAP1* and *PAP2*) has been identified, as well as one bHLH transcription factor *GLABROUS3* (*GL30*) (Gou et., 2011). The physical interaction between R2R3-MYB/bHLH was first described in maize, as occurring between R3 repeat of the MYB domain and N-terminal region of the bHLH factor¹²⁷. In petunia, two bHLH proteins were identified: *PhAN1* and *PhJAF13*, and both interact

with R2R3-MYBs (*PhAN2*); in addition a WD-repeat (WDR) called *AN11* (*PhAN11*)^{128,129} has been described. *AN11* encodes a small protein with five to six conserved WDR that modulate the activity of the R2R3-MYB or bHLH factors post transcriptionally¹²⁸⁻¹³⁰. A number of studies reported these TFs (BMW: bHLH, MYB, and WDR) as responsible for the anthocyanin pigmentation of many fleshy fruits. Some examples, including plant model and economically important crops, are shown in figure 3.

Tomato orthologues for the BMW Petunia complex have been identified. Two *PhAN2* (MYB) homologues were found in tomato, *SIAN2* and *SIAN1*, both encoded by two paralog genes mapped in chromosome 10. In addition, two bHLH of the petunia *JAF13* and *an1* were found in the tomato genome and mapped to *SJAF13* (Chr. 8) and *AN1* (Chr. 9). Finally, the corresponding WDR gene from tomato was located in chromosome 3 based on its homology with petunia *AN11* and later it was called *SIAN11*^{112,131}.

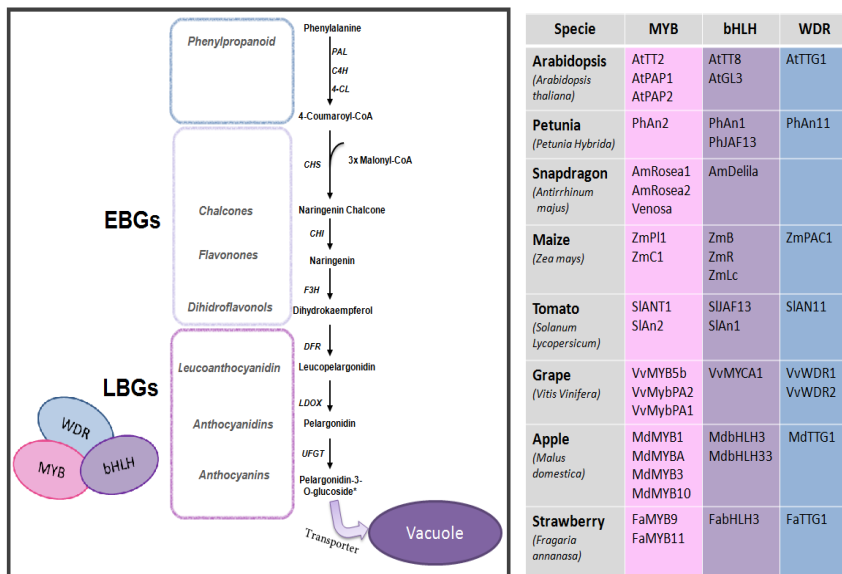


Figure 3: Regulation of the flavonoid biosynthesis pathway. Anthocyanin pathway is regulated at three main steps (continued)

Figure 3 (continued): (each step is highlighted by a color square). First, from Phenylalanine to 4-coumaroyl-CoA (blue square) and subsequently, regulation is controlled by two groups of genes: Early Biosynthetic Genes (EBGs, lilac square) and Late Biosynthetic Genes (LBGs, Purple Square). The complex BMW (bHLH/MYB /WD40 has been described as activator of LBGs, being the major regulator of anthocyanin synthesis¹³². Ternary complex is composed by three transcription factors described as well-conserved among species. On the right, a review table shows examples of MBW in plant models and economically important species. *Arabidopsis* (Borevitz et al., 2000; Nessi et al., 2001, Nessi et al.,2000; Walker et al., 1999), *Petunia* (Quattrocchio et al., 1999; Quattrocchio et al., 1993; De vetten et al., 1997), *Snapdragon* (Schwinn et al., 2007; Schwinn et al., 2006; Martin et al., 1991), *Maize* (Paz-Ares et al., 1987; Chandler et al., 1989; Ludwig et al., 1989; Cone et al., 1993; Golf et al., 1990; Carey et al., 2004), *Tomato* (De Jong et al., 2004; Kiferle et al., 2015), *Grape* (Walker et al. 2007; Bogs et al. 2007; Terrier et al. 2009; Matus et al., 2009), *Apple* (Talos et al., 2006; Ban et al., 2007 ; Vimolmangkang et al., 2013; Espley et al. 2007; An et al., 2012) and *Strawberry* (Schaart et al., 2012)

2.5 Tomato engineering for high content in anthocyanins

Before we had a detailed knowledge of the FLV/ANT pathway and its regulation, the strategies for breeding anthocyanin traits in tomatoes were limited to interspecific crosses and resulted in introgression lines carrying the corresponding trait^{113,133}. The identification of LBGs and key regulators opened new avenues for genetic engineered strategies.

The *Delila* gene (*del*), a BHLH factor of *Antirrhinum majus*¹³⁴ cloned under a constitutive promoter and introduced in Micro-Tom plants, resulted in enhanced anthocyanin synthesis in vegetative tissues but not in fruit (Mooney et al., 1995). Similar results were obtained when *SlANT1* (MYB) was overexpressed under 35s promoter in tomato: only small spots of anthocyanins were detected in fruit. Later Micro-Tom plants, transformed with *p35s:AtPAP1* (MYB), accumulated anthocyanins in leaves, stems, roots and flowers, whereas just small pigmented areas were found in fruit¹³⁵.

Moreover, fruit-ripening specific overexpression was attempted by using a strong specific promoter (E8). *LC* (MYB-type) and *C1* (MYC-type) transcription factors, described as regulators of anthocyanin synthesis in maize, were placed under the control of the E8 promoter and transformed in tomato. Fruit from those plants showed increased levels in total flavonols (up to 20 fold) but not in anthocyanins which remained undetectable in red ripe fruit¹²⁴. However, when the same E8 tomato promoter was used to control the overexpression of two transcription factors described in the BMW complex successful accumulation of anthocyanins in the fruit was achieved. More precisely *Rosea1* (*Ros1*) and *del* from *A. majus*, MYB and bHLH respectively, were cloned under the control of E8 promoter and transformed in Micro-Tom. Both TFs upregulated LBGs (*F3'5'H*, *ANS*, *3-GT*), but also activated genes upstream in the pathway acting on the two previous levels of regulation (first: PAL, C4H, 4CL and second: EBGs) (Fig. 3). Transgenic plants exhibited purple tomatoes resulting from the induced accumulation of anthocyanin not just in the peel, but also in pericarp and inner flesh¹³⁶. Those tomatoes showed an extended shelf-life due to the antioxidant capacity of anthocyanins¹⁰⁸.

The most recent paper aimed to restrict the anthocyanin accumulation to the peel chose Plastidial Light Inducible (PLI) promoter to drive the expression of the MYB and 35s promoter to control bHLH (plants also showed expression in the stems, hypocotyl and roots). Moreover, this study also incorporated the evaluation of *Aft/atv* genotype and analyzed the shelf-life of both transgenic and introgression lines. Results suggested that accumulation of anthocyanins in the skin is sufficient to delay post-harvest over-ripening and to extend the fruit life span^{137,138}.

To summarize, it has been demonstrated that anthocyanin accumulation provides fruit with antioxidant protection extending shelf-life (*Aft/atv* genotype). However, to date, no engineered tomato has been reported with strict skin

accumulation of anthocyanins. On the other hand, so far an intragenesis approach has never been used to enhance anthocyanin accumulation in tomato. Previous studies only used the tomato MYB overexpression and did not success.

All the data exposed above support tomato as an excellent candidate for increasing anthocyanin accumulation. The enhancement of tomato fruit chloroplast functions, combined with photo-protection strategy provided by anthocyanin engineering, define the aim of this work.

OBJECTIVES



All the information above reviewed lead us to propose a research plan directed to improving tomato fruit quality through the enhancement fruit chloroplast and increasing anthocyanins in the fruit peel using a cisgenic/ intragenic approach. To develop this general objective a set of specific objectives were planned:

1. To overexpress tomato *GLKs* and *APPR2* genes early in fruit development and analyze the effect on fruit at different stages of development (Chapter I).
2. To study de molecular basis of the phenotypes obtained from single or combined overexpression of *GLK* and *APPR2*, using different omics technologies (Chapter I and II).
3. To generate and characterize tomato lines that accumulate anthocyanins in the fruit peel by overexpressing tomato BHLH-MYB-WD40 under the control of a fruit peel specific promoter (Chapter III)

**CHAPTER 1: Plastid Fortified
Tomatoes Obtained by Combined
Over-expression of *GLK* and *APRR2*
Transcription Factors during Early
Fruit Development**



1. INTRODUCTION

Plastids are the plant specific cellular organelles where many of the visual-, health- and flavor-related metabolites are produced and stored in the fruit ¹³⁹⁻¹⁴¹, and therefore are valuable components for consumers and breeders. Before they turn into chromoplasts during ripening ¹⁴², chloroplasts contribute to fruit's carbon economy ¹⁴³. Both chloroplast and chromoplasts are the synthesis site of important metabolites and growth regulators ¹⁴⁴. The number and the functionality of fruit plastids at the different developmental stages determine several metabolic and developmental processes.

Although fruits are photosynthetic sinks, and rely mainly on carbon transported from leaves to support development and ripening ²⁵, plastid contribution to fixed-C have been estimated to reach around 15-20% ⁹ or even higher ⁷. Furthermore, high chlorophyll content and enhanced chloroplast content in green fruits was shown to result in increased photosynthesis rates, and more soluble solids and carotenoids levels in Red Ripe (RR) fruit ^{5,7,8,61,63,70,145}.

Important characteristics of the ripe fruit such as its shape ¹⁴⁶, size ¹⁴⁷ or weight ¹⁴⁸ are known to be determined early in development. Therefore, acting at the early developmental stages should provide the possibility of changing final fruit characteristics ¹⁴⁹, including some of the most important for fruit quality like the sugar content and the nutritional composition ⁵⁷. Plastids number appears to be defined also early in organ development ^{59,62}, but their functionalities change during fruit development and ripening. Thus, starch accumulated in the fruit chloroplast may reach up to 20% of the total dry weight in the young fruit but is negligible in ripe fruits ²⁰, indicating that there is a transient accumulation of starch in the Immature Green (IMG) and Mature Green (MG) fruits which is mobilized and converted into soluble sugars at the ripening stage ¹⁵⁰. Variability in the

accumulation of sugars stored as starch in the chloroplast of immature fruits is in part responsible for the differences in sugar content at the red ripe stage ^{39,151}. Carotenoids, on the other hand, increase during fruit ripening concomitantly or as a consequence of chlorophyll degradation ^{152,153}. Lycopene ¹³⁹ is mostly accumulated and their levels are in part also determined by genes affecting plastids early in development ^{69,70}.

The importance of fruit photosynthesis for starch accumulation, and subsequently for the sugar content of the ripe fruit, has received additional support with the discovery of several Transcription factors (TFs) and regulators that, when mutated or overexpressed, result in greener fruit at the IMG and MG, as well as fortified fruit at the ripe stage ^{5,8,64,66,81,154}. *Golden like* (*GLK*) belongs to GARP family of MYB transcription factors as defined by *GOLDEN2* in maize (*Zea mays*) ¹⁵⁵, and had previously been identified in several plants such as *Arabidopsis thaliana*, *Oryza sativa*, *Sorghum bicolor* and the moss *Physcomitrella* as TFs implicated in plastid development ^{50,78,156}.

GLKs are normally found in pairs; the ancestral of flowering plants was a single gene and the duplication occurred in some species ¹⁵⁷. Gene pairs appear to be functionally redundant. In moss and *Arabidopsis*, only double *GLK* mutants exhibit a characteristic pale green phenotype in vegetative tissues ⁵². Although both *GLK* genes are expressed in leaves (*GLK1* and *GLK2*), only *GLK2* is expressed in fruit. In tomato and *Arabidopsis*, single *GLK2* mutants produce pale green phenotype in fruit and siliques respectively ^{78,141}. It has been shown that *SIGLK2* influences photosynthesis in developing fruit by inducing transcription of photosystem I (PSI) and photosystem II (PSII) and other photosynthesis-related components. *GLK2* has been recently identified by positional cloning as the gene underlying the uniform (*u*) ripening mutation in tomato that produces a pale green fruit phenotype. The mutation was characterized as an additional Adenine (A) nucleotide in a region

of six consecutive As in exon I of *SlGLK2* gene, that introduces a premature stop codon and generates a non-functional protein⁵.

Most modern commercial tomato varieties are *u/u*, such as MoneyMaker (MM), showing pale green fruit at IMG and MG stages and lower sugars and carotenoid contents in RR fruit. Moreover, *GLK*-overexpressing lines accumulate more soluble solids in the red ripe stage, opening the way to increase fruit quality biotechnologically by acting at the *GLK* level. It has been studied that overexpression of either *GLK1*- or *GLK2*-, under the control of 35s promoter, exhibits impact in fruit quality. Total Soluble Solids (TSS) increase from 20 to 40% and 40 to 60%, and total carotenoid presents levels from 25 to 40% higher than the control 'M82' (*u/u*) line¹⁴¹. Although a functional equivalence of the *GLK1* and *GLK2* peptides has been reported in fruit, transgenic p35s::*SlGLK2* lines showed RR fruit with higher TSS and carotenoids values than p35s::*SlGLK1* plants¹⁴¹.

More recently, *ARABIDOPSIS PSEUDO RESPONSE REGULATOR2-LIKE* gene *APRR2-like* was identified as a transcriptional activator with a GARP DNA-binding domain and a receiver-like domain similar to ARR-B. *APRR2-like* has certain sequence similarity with *GLK2* but lacks the AREA₂EAA conserved motif⁷⁸. A *SlAPRR2-like* tomato gene shows ripening-associated expression pattern, with a strong induction at the breaker stage (Br)⁸. Ectopic overexpression (OE) of *SlAPRR2*, under the control of the 35s promoter, resulted in fruit with more plastids and higher chlorophyll amounts in immature stages and, eventually, higher sugar and carotenoid levels in the red ripen stage⁸.

Previous reports provide indirect evidence that early developmental expression of regulatory TFs is required to obtain enhanced chloroplast functions. *AtGLKs* expressed under the control of tomato ripening promoter (PDS) had no effect on fruit "greenness"⁵, whereas the *SlGLK2* or *SlAPRR2* cDNAs expressed under constitutive 35s promoter resulted in a clear chloroplast

enhancement^{5,8}. In our hands, 35s *SIGLK2* and *SIAPRR2* OE lines produced small and often seedless fruit, probably as a result of an excessive expression of regulatory factors (own observations). Therefore, to understand the role of chloroplasts in fruit development and their influence in the final fruit composition it is important to dissect their functional roles during early fruit development, measuring the effects of restricting chloroplast over-accumulation to IMG stages exclusively.

We report here the characterization of MoneyMaker tomato lines (*u/u*) which overexpress *SIAPRR2* and *SIGLK2* TFs under the control of tomato regulatory sequences conferring ovary-early-fruit specific gene expression. In addition, *SIGLK1* and *SIGLK2* fruit functionality was compared in a background where *SIAPRR2* was also OE. We found that the co-expression of the two TFs displays an additive effect ranged from fruit color to accumulation of primary metabolites and isoprenoid compounds, as well as modifications in chloroplast structure. Results indicate that intragenesis-like approaches are viable alternatives to transgenesis or conventional breeding for enhancing plastid functions in the fruit and consequently fruit quality.

2. RESULTS

2.1 Engineering tomato plants expressing *GLKs* and *APRR2* TFs under the control of the TFM5 promoter

To study the effect that *SIGLK2* and *SIAPRR2* early expression in the ovary/ fruit may have on fruit development and quality, the coding sequences of the two TFs were inserted between the TFM5 early fruit specific promoter and terminator regions^{2,158}. We chose *GLK2* instead of *GLK1* because, although a functional

equivalence has been described due to the OE (35s promoter) of both TFs, *GLK2* showed better results in the improvement of tomato fruit quality than *GLK1*¹⁴¹. This approach differs from previous ones in that it follows an all-tomato DNA strategy, aiming to specifically and transiently activate TFs at early fruit development (designed in part to overcome that 35s OE lines produced small and often seedless fruit, in our hands). Our approach targets the fruit stage when normally chloroplasts are dividing and developing, a stage associated with activation of all chloroplast transcriptional machinery (<http://gbf.toulouse.inra.fr/tomexpress17b>).

In addition, we constructed a series of transcriptional units either combining pTFM5::*GLK1* or pTFM5::*GLK2* with pTFM5::*APRR2* in the same binary construct or *GLK2* and *APRR2* cloned separately (Fig. 1a). Although *GLK1* and *GLK2* OE was previously compared¹⁴¹, it has never tested before the effect of those TFs in a background where *APRR2* TF is expressed.

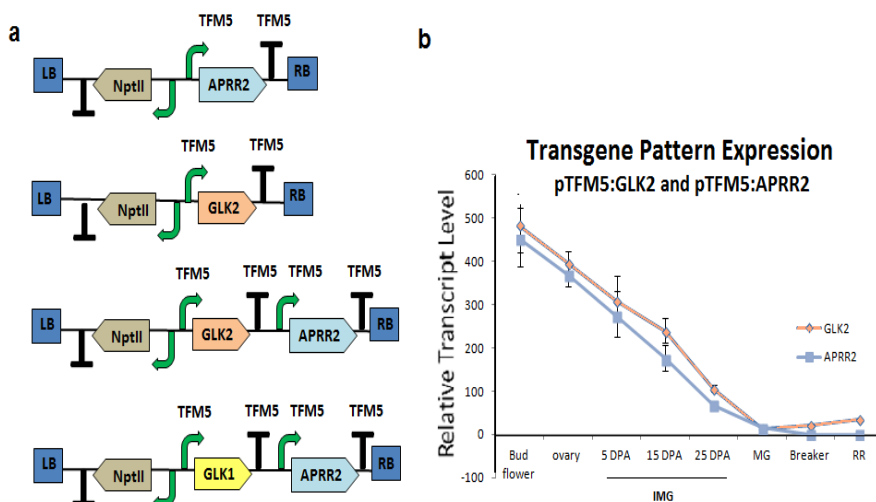


Figure 1: Fruit Chloroplast enhancing modules. (a) Modules for simple Chloroplast Engineered Fruits consisted in (from left border to right border): Kanamycin resistance (*NptII* CDS controlled under nopaline synthase promoter and terminator), TFM5 promoter, *APRR2*-CDS or *GLK2*-CDS followed by TFM5 terminator (continued)

Figure 1 (continued): Single constructs were used to build double fortified modules: *GLK2-APRR2* and *GLK1-APRR2* and used to generate double Chloroplast Engineered Fruits. **(b)** Expression levels of *GLK2* and *APRR2* driven by TFM5 promoter (*GLK2* and *APRR2* mRNA levels) analyzed by Real time PCR in *GLK2-APRR2* lines (T3 generation). Tissue sample included bud flowers, ovaries, IMG (immature green), MG (mature green), Breaker (Br) to RR (Red Ripe) fruits see M&M (4.1). The $2^{-\Delta\Delta CT}$ method was used to determine the relative mRNA abundance see M&M (4.4). Transgene expression in fruit tissues is expressed in relation to the sample with the least expression. Data are shown as mean \pm SD of three replicates in each independent experiment.

Fifteen engineered tomato lines for *GLK2*, fifteen for *APRR2*, forty lines for the combination of *GLK2-APRR2* and thirty lines for *GLK1-APRR2* were chosen at random among the KamR lines and cultivated in the greenhouse. Initial visual inspection clearly indicated that the double transgene lines showed darker green fruit color than the single gene lines, which were nevertheless darker than WT. No evident phenotypes other than fruit color were observed in any other part of the plant. A quantitative real-time PCR experiment revealed the expression pattern of the transgenes driven by the TFM5 promoter sequence at different stages of fruit development. As can be seen in Figure 1b both transgenes showed maximum expression in the ovary at anthesis, and decreased later with fruit development, showing almost undetectable levels during ripening.

2.2 A dark green fruit phenotype is observed when GLKs and APRR2 are over-expressed simultaneously

A detailed analysis of fruit color at different developmental stages was performed next. Compared to the wild-Type moneymaker fruit, which has very light green color (typical of *u/u* varieties) at IMG stage all Chloroplast Engineered Fruits (CEF) displayed a darker green color (Fig. 2). The fruit phenotype reached its major intensity from 0 DPA to IMG stage (10-15DPA) consistent with the transgene expression pattern described

above (Fig. 1b). Expression of functional *SlGLK2* under TFM5 promoter resulted in fruit that are homogeneously dark green (Fig. 2).



Figure 2: Fruit from CEF lines are greener than WT fruit. Comparison between WT fruits (upper row) and engineered fruits as follows: *APRR2*, *GLK2*, *GLK1-APRR2* and *GLK2-APRR2*. Fruits were characterized in green stages: two in immature green IMG stage of development (10 DPA and 20 DPA) and subsequently in mature green MG (35 DPA), as well as ripening stages of development: Breaker *Br*, Light red *LR(B+5)* and Red Ripe *RR(B+7)*.

Visual inspection indicated that double gene engineered (dCEF) plants showed darker color than in the single gene plants (sCEF), suggesting an additive/ synergistic effect of *GLKs* and *APRR2*. To better define the color phenotype variation associated with each construct further analyses were conducted on a population of selected T2 plants. In total, fifteen independent transformed lines were selected for each of the four types of engineered plants: *GLK2* (G2), *APRR2* (A2), *GLK1-APRR2* (G1A2) and *GLK2-APRR2* (G2A2), then between ten and fifteen plants per line were analyzed for fruit color phenotype.

Fruit surface color measures were conducted on IMG and MG fruits from a population of approximately 500 plants representing the different engineering events (Supplementary Fig. 1). A total of five lines for each construct were chosen based on surface color measures for their detailed analysis.

Three of them represent the average values of green intensity for each transgenic construct population and the other two correspond to the higher and lower extremes of the fruit phenotype for that construct population.

Figure 3 reveals the internal and external fruit color distribution for those selected subpopulation of plants. *APRR2* lines displayed fruits with the weakest color phenotype, which was almost similar to WT. At the unripe stages, *GLK2* provides a darker green fruits than *APRR2* but in all cases these were lighter than the dCEF phenotypes. Those phenotypic observations are presented in Figure 3, where a large range of variation is showed for the external color distribution in the different CEF lines. Differences between genotypes based on internal color development were lower, although they run always in parallel to the external surface values.

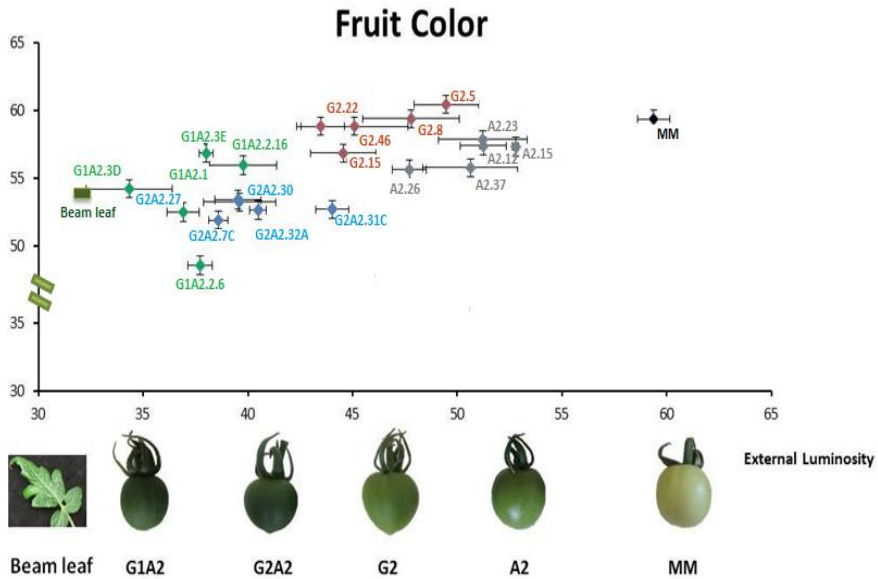


Figure 3: Surface color displays huge differences between single and double modified plants in external level. Internal and external color distribution based on luminosity parameter (higher values lighter green, smaller values darker green) at IMG stage (15 DAP) for dCEF (blue and green points), sCEF (gray and red points) and WT fruits (black point). In the abscissa axis (X) external luminosity (with Minolta color instrument, see M&M) and in ordinate axis (Y) internal fruit luminosity (Tomato Analyzer 3.0 software, see M&M)

To place our results in the context of the existing natural variability in green fruit color phenotype, we measured the variation in fruit “greenness” in a population composed by 560 different tomato accessions (traditional/ heirlooms: *S. pimpinellifolium*, *S. lycopersicum* *S. cersiforme* and *S. lycopersicoides*; and commercial varieties as used in Tieman et al., 2017) ¹⁵⁹. CEF presented a uniform color at MG, which is grouped within the higher values of the range of variability for this phenotype in the accessions diversity panel. In contrast to our engineering lines, most accessions showed a green color gradient along the longitudinal edge of the fruit, probably resulting from *GLK2* expression under the control of its own promoter (Supplementary Fig. 2).

Another observation related to the color distribution was an apparently negative correlation between green intensity in IMG fruits and weight in MG stages. Final weight at MG stage between sCEF and WT was slightly different; however, differences increased in dCEF parallel to the greener phenotype (Supplementary Fig. 3).

2.3 The largest increase in chlorophyll and carotenoid levels is observed in fruit at IMG stage from double engineered *GLKs-APRR2* plants

To assess whether the differences observed in fruit greenness between the engineered genotypes was associated with differences in the content of photosynthetic pigments (Chlorophyll a, b and carotenoids), a thorough analysis of the isoprenoid content in fruit was performed (Supplementary data on CD). Samples were taken at two stages: IMG (when TFM5 is highly active) and MG (when TFM5 promotor is less active), and analyzed for chlorophyll (Chl), carotenoids (Car) and total isoprenoid content. Figure 4a represents the fold change of the total isoprenoid content which is always higher in CEF at IMG. The levels of chlorophyll and carotenoid contents in immature fruit nicely correlated with the different intensities of green color previously presented in Fig. 3. However, those differences decrease in all cases at MG stage.

At the IMG stage, Chl content was, as expected, dramatically affected by TFs manipulation. Overexpression of either *GLK2* or *APRR2* increases total quantity of Chl around 80%, whereas dCEF foldchanges ranged from 250 to 300% (*GLK2-APRR2* and *GLK1-APRR2* respectively) respect to WT values (Fig. 4b). These results support at the molecular level the reinforced visual

phenotype observed in dCEF. Interestingly, in MG fruit the differences were less dramatic. Consistent with the visual phenotypes the Chl level is reduced as compared to IMG. At this stage sCEF are practically identical to WT, whereas dCEF still showed increased Chl content (1.4 fold of WT) (Fig. 4b). Chl a/b ratios also showed differences between CEF lines and control (Fig. 4c). This ratio is inversely proportional of photosynthetic capacity since Chl b is binding to LHCII proteins, therefore lower ratio values are indicative of more PSII proteins¹⁶⁰. CEF lines exhibit a decrease in Chl a/b ratio from IMG to MG, whereas MM does not present any change (ratio approx. 0.5). On the other hand, is necessary to take in consideration pheophytin values which are related to Chl degradation¹⁶¹. More pheophytin (a + b) could mean higher Chl b values before degradation. *GLKs-APRR2* fruits present the highest Chl b values of all samples at MG stage.

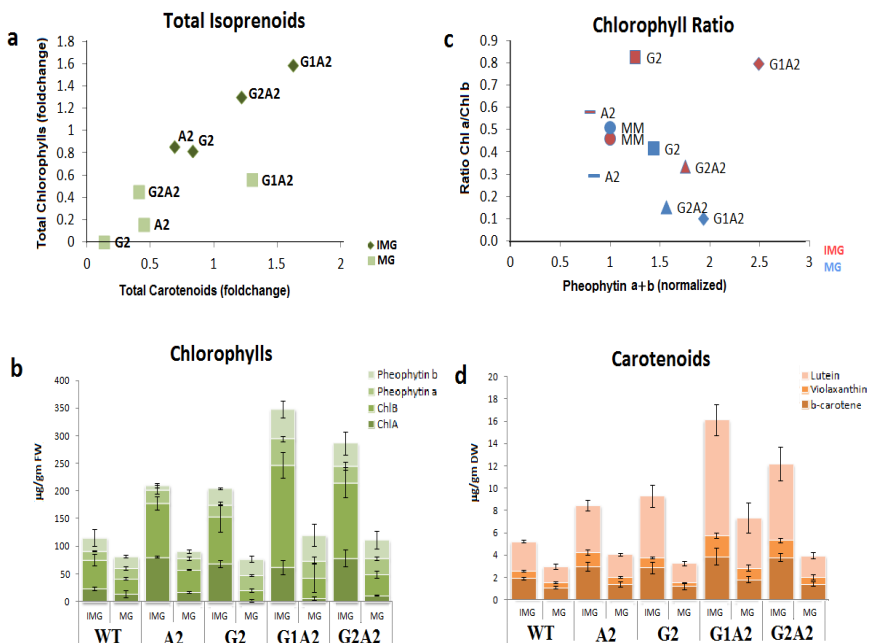


Figure 4: Combination of *GLK/APRR2* enhance major photosynthetic pigments of tomato fruit in green stages. (continued)

Figure 4 (continued): Isoprenoids composition was analyzed by HPLC (see M&M) at IMG and MG stage of development. **(a)** Total Isoprenoids, in IMG (diamond) and MG (square) fruit. Data showed was normalized and transformed by Log₂. **(b)** Chlorophylls content in µg/g FW (Fresh Weight) **(c)** Values for the ratio Chl a/b and total pheophytin (a + b) normalized to WT. Each genotype is represented by a symbol (G2: square, A2: line, G1A2: diamond, G2A2: triangle, MM: circle) **(d)** Carotenoids content in µg/g DW (dry weight). Genotypes represented in these graphics: *APRR2*, *GLK2*, *GLK2-APRR2* and *GLK1-APRR2* and WT: MM control.

Moreover, there is a strong relationship between chlorophyll and carotenoids as these pigments have interrelated functions¹⁶². Chl/Car ratio is used as an indicator of photooxidative damage¹⁶³ as carotenoids contribute to protect the photosynthetic apparatus against photooxidative stress¹⁶⁴. Chl/Car ratio was approximately 1.3 in CEF and WT, indicating that all chloroplast pigments increase in a coordinated manner.

During green stages, lutein is the most abundant carotenoid in fruit followed by to β -carotene and xanthophylls in both WT and CEF lines (Fig. 4d). Our results showed an increase of carotenoids in CEF at IMG stage that it is even most dramatic in dCEF lines; e.g. lutein level increase in sCEF over 100% related to WT, whereas levels reached 400% in dCEF. These data are consistent with the characterized fruit phenotype and again confirmed the additive visual effect mediated by simultaneously overexpression of *APRR2* and *GLKs* TFs observed in dCEF.

Higher levels in lutein, as well as β -carotene and violaxanthin are still obvious in MG stage for dCEF lines, whereas in sCEF the levels of these compounds are almost equal to WT. The differences in carotenoids between dCEF and WT decrease during fruit development, although at MG stage dCEF still remain over 50% higher than WT (Fig. 4b). In general, carotenoids profiles seem to be paralleled to chlorophylls and related metabolites, suggesting coordinated expression.

2.4 Primary metabolism also reflects the additive effect between *GLKs* and *APRR2* at IMG stage

Having determined the changes in photosynthetic-related pigment profiles during green stages of fruit development (IMG and MG), we next studied the effect of our engineering approach on fruit primary metabolism, using for that an established GC-MS method ¹⁶⁵ (see M&M).

A total of 70 primary metabolites were analyzed (Supplementary data on CD). Again, in general, dCEF present the most affected metabolite profiles. At IMG stage, the level of sugars was significantly higher in dCEF (Table 1), whereas no differences were detected in sCEF (except for the decrease of sucrose in *GLK2*). Glucose and fructose levels are significantly higher in *GLK1-APRR2* genotype, reaching up 56 and 50% of increase, respectively. Moreover, *GLK2-APRR2* showed significantly increase in sugar but due to sucrose (up to 42%); however, the differences in the accumulation of sugars were not significant when fruit reached MG stage. It is interesting to note that the higher sugar content observed at IMG stage was correlated with the more starch content presented in those fruit at MG stage (Supplementary Fig. 4).

Related to amino acids levels, significant changes were detected for both sCEF and dCEF lines. Amino acids such GABA increased only in dCEF (Table 1), others decrease such as asparagine, proline, aspartic acid or methionine (Supplementary data on CD) in all engineered genotypes. At MG stage, as a general rule, is observed that sCEF and dCEF showed an increase in aa levels (Alanine beta, Threonine, Proline, Asparagine, Glutamine, Glutamic acid, Phenylalanine, Arginine).

Group	Compound	IMG				MG			
		G1A2	G2A2	A2	G2	G1A2	G2A2	A2	G2
Sugars	Glucose	1.56	1.35	0.99	1.18	1.04	1.02	1.05	0.97
	Fructose	1.50	1.22	1.00	1.02	1.09	1.07	1.03	0.90
	Sucrose	1.16	1.42	0.96	0.68	0.99	1.04	1.11	0.71
Polyols	Putrescine	1.70	0.95	0.87	1.19	1.09	1.79	1.37	1.56
	Inositol, myo	1.58	1.31	0.93	1.05	1.74	1.17	1.11	1.16
	Inositol-1-phosphate	1.36	1.58	1.02	1.26	1.06	1.11	1.08	1.10
	Calystegine A3	1.08	2.68	3.10	0.35	1.36	2.26	1.98	0.64
	Erythrose	1.00	1.20	0.98	0.77	1.31	1.23	1.34	1.22
	Fructose-6-phosphate	0.90	0.73	0.69	1.10	1.05	1.06	0.76	0.85
Aminoacids	Arginine	1.28	0.74	0.96	0.82	1.59	1.13	1.00	1.14
	Glutamic acid	1.26	1.11	0.89	0.92	1.02	1.36	1.53	1.00
	GABA	1.22	1.34	1.11	0.92	0.89	1.01	1.15	1.12
	Cystine	1.18	1.59	1.02	1.22	1.04	1.20	1.12	1.13
	Alanine, beta	1.15	1.40	1.07	0.82	1.31	1.32	1.46	1.28
	Glycerol	1.11	1.31	1.13	1.59	1.40	1.00	1.05	0.97
	Cysteine, S-methyl	0.95	1.42	0.98	0.93	3.42	1.16	1.63	0.78
	Threonine	0.88	1.41	1.16	0.69	1.11	1.52	1.60	1.36
	Phenylalanine	0.75	1.13	1.25	0.98	1.09	1.70	1.67	1.31
	Valine	0.74	1.35	1.16	0.70	0.92	0.95	1.22	0.91
		Alanine	0.73	0.84	0.87	0.85	1.01	0.84	1.00
Organic acids	Galactonic acid-1, 4-lactone	4.64	4.19	0.56	11.35	1.87	1.48	1.25	0.91
	Quinic acid, 3-caffeoyl-trans	2.49	2.19	0.94	3.10	0.80	0.77	1.05	1.37
	Malic acid	1.52	1.21	1.12	1.07	1.12	2.38	1.94	1.61
	Succinic acid	1.50	1.21	1.03	1.57	1.20	1.19	1.19	0.96
	Gluronic acid	1.47	1.66	0.55	2.47	1.03	0.52	1.23	0.70
	Galactonic acid	1.36	1.43	0.81	2.57	1.35	1.01	1.14	0.72
	Quinic acid, 3-caffeoyl-cis	1.33	1.44	0.73	1.72	0.91	0.78	0.99	1.20
		Citric acid	1.14	1.11	1.00	1.05	1.18	1.80	1.51

Table 1: Main primary metabolites changes associated to CEF. Normalized metabolite contents (WT=1) for fruits at IMG and MG stage. Metabolites were grouped by compound classes: Sugars, Polyols, Aminoacids and Organic acids. Table shows some metabolite examples cited in the text (for the complete analysis see supplementary data on CD). A Student's t-test analysis was performed, statistical significance differences are in bold (pvalue<0.1).

Other products from Calvin cycle are also shown in table 1. Erythrose, a polyol compound displays significantly higher accumulation in *GLK1-APRR2* (30%) during MG stage. More polyols showed higher levels in MG samples of CEF lines such as Calystegine A3 (Polyhydroxy alkaloid which increases 126% in *GLK2-APRR2*) or Putrescine (diamine which presents levels of 79% in *GLK2-APRR2* and 37% in *APRR2*). Highlighted is myo-

inositol, the most abundant polyol, which is enriched in dCEF at IMG stage and reaches the highest values at MG stage in *GLK1-APRR2* genotype (75%).

We also studied levels of organic acids (such as quinic acids, succinic acid, gluronic acid, galactonic acid, malic acid) in CEF samples. In general, higher content was observed at IMG stage. For example, *GLK1-APRR2* lines significantly increase galactonic acid-1-4-lactone level, reaching the remarkable 364% related to WT. Others such as quinic acid (dCEF and *APRR2* genotypes) and succinic acid also significantly increase during IMG stage.

Related to malic acid, it is worthy to mention the enrichment in sCEF and dCEF at both stages of development, which was even greater in MG reaching 138% in *GLK2-APRR2*, 94% in *APRR2* and 61% in *GLK2* lines. This organic acid is highly interesting because could be directly related to an increase of fruit-photosynthesis²¹. In addition, differences in citric acid were observed in *GLK2-APRR2* genotype at MG stage (which level increases up to 80%).

2.5 *GLKs-APRR2* increases fruit plastid number, compartment size and grana stacking

In order to assess the effect of our chloroplast engineering module in the chloroplast structure of our plants, IMG fruit sections were analyzed using light and electron microscopes. First, we divided pericarp tissue in two areas of study (Area 1 and 2, Fig. 5a), from fruit epidermis (Fe) to the central part of the fruit pericarp (Pc) and from there to the locular area (La).

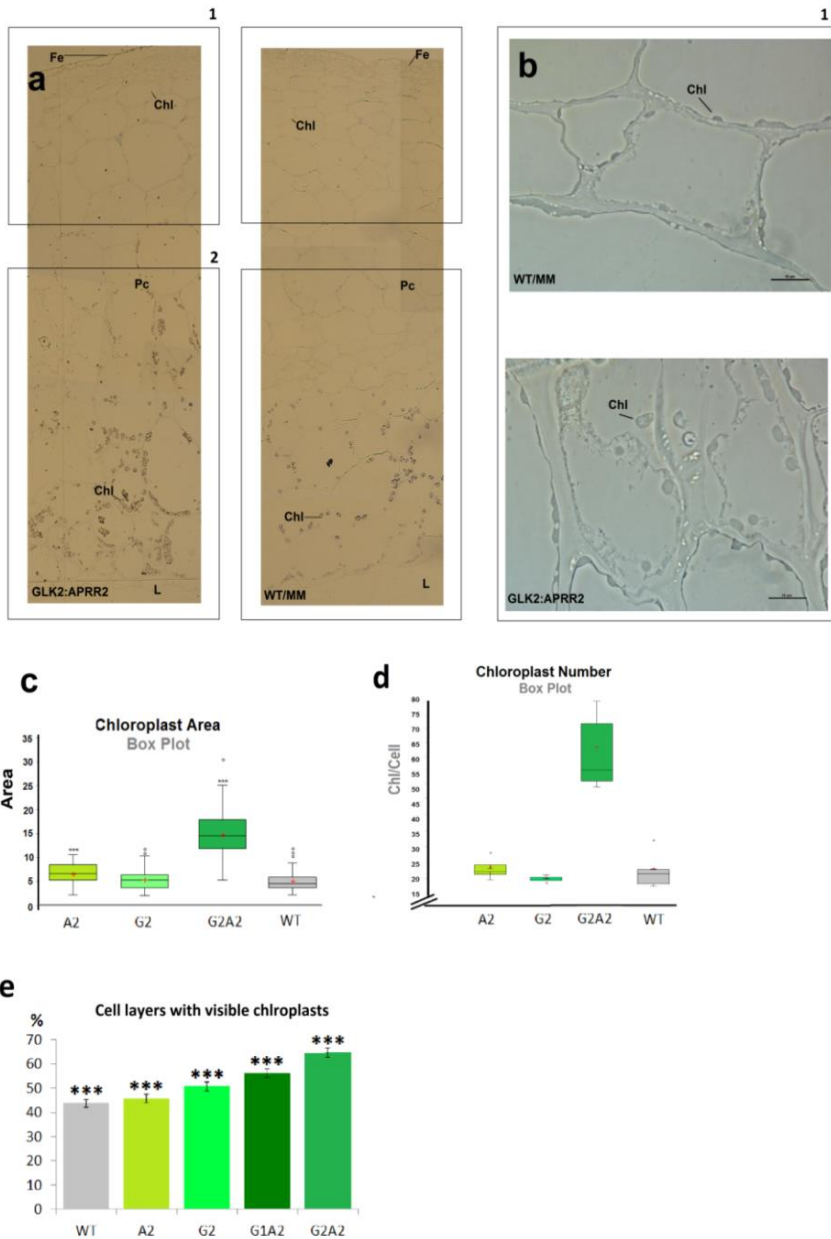


Figure 5: pTFM5::GLK2::TFM5::APRR2 expression increases plastid size in IMG stage of fruit development. (a) Light Microscope (10x) displays cells layers of fruit pericarp in detail from fruit epidermis (Fe) to the locular area (La) (Chl: chloroplast, Pc: center area) in GLK2-APRR2 (left) and WT (right) (continued)

Figure 5 (continued): (b) Light microscope magnification (100x) of a cell close to the fruit epidermis area where is possible to appreciate the higher size and number of the chloroplast in *GLK2-APRR2* genotype (down) compared to WT-MM (up). **(c)** Box plots show the data distribution for chloroplast area, **(d)** and chloroplast number/cell in the area between Fe and Pc, *** (ANOVA, $p < 0,0001$). **(e)** Percentage of cell layers with visible chloroplast at 10x in the area ranged from the Pc to L (*** ANOVA $p < 0,0001$).

In the area close to the Fe we could observed that cells in dCEF showed increases over 300%, in both chloroplast size (Fig 5b-c) and chloroplast number (Fig. 5d) ($p|t| < 0,0001$). Effects in sCEF were remarkably lower, suggesting again an additive effect mediated by simultaneous OE of *APRR2* and *GLKs* TFs in dCEF.

Cell layers spanning from Pc to the La did not show significance difference neither in chloroplast number nor size per cell. However, we could observe that the width of cell layers which present visible chloroplasts at 10X is bigger in CEF lines than WT (See area 2 in Fig. 5a). Whilst WT showed 40% of this area cells with visible chloroplast, dCEF increased up to 60% this trait. Differences were also observed in sCEF but to a lesser extent than dCEF (Fig. 5e).

Later, we studied in detail the area spanning from Pc to La, and we also observed changes at structural level. Under electron microscope we found that the number and size of grana stacking in dCEF samples was higher than WT. These observations were most significant in fruit of the *GLK1-APRR2* genotype (Fig. 6a). Figure 6b displays an image of the grana of *GLK1-APRR2* fruit chloroplast compared to WT.

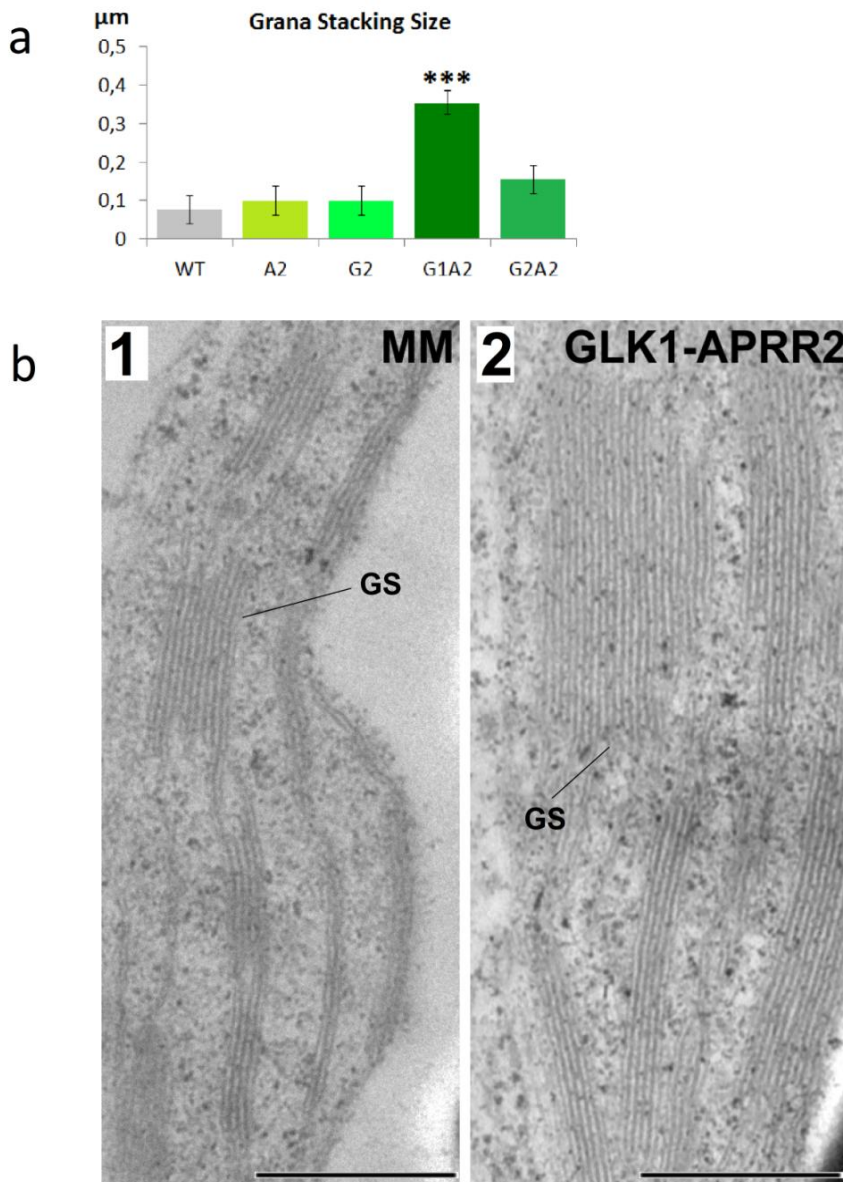


Figure 6: Expansion of the cell-enriched-chloroplast layer and effect on grana. Cells from central pericarp to the locular area were studied in *APRR2*, *GLK2*, *GLK1-APRR2* and *GLK2-APRR2* genotypes (see M&M). **(a)** Grana stacking size (μm) in CEF genotypes and WT (***) ANOVA $p < 0,0001$) **(b)** Electron Microscope detail of Grana Stacking (GS) in MM (left) and *GLK1-APRR2* fruit chloroplast (right), scale at 500nm.

2.6 Engineering *GLK* and *APRR2* early in ovary/ fruit development have consequences on fruit quality at the ripe stage

Our results so far indicated that OE of *GLK* and *APRR2* during early stages of ovary/ fruit produces changes in the structural and molecular levels in the IMG and also in MG stages, although the levels of expression of the engineered genes were very low at that MG stage. Next we analyzed whether these effects extend to fruit at the ripening stage of development. The following quality parameters were analyzed in Red Ripe (RR) fruit: TSS, carotenoids and volatiles.

2.6.1 dCEF lines accumulate higher levels of total soluble solids

The analysis of CEF indicated that the total content in TSS of red ripe RR fruit from engineered lines is higher than in the control MM background and correlates with the intensity of dark green phenotype observed in immature stages of development (Fig. 2).

sCEF genotypes showed slight differences in TSS when are compared to WT (Fig. 7) where *SIAPRR2* presents an increase of 3.1% in soluble solid levels, and *SIGLK2* 4.5%, barely significant. However, a significant change in Brix degree of fruit at the RR stage was observed when both TFs were co-expressed, achieving the 23% and 25% in *GLK1-APRR2* and *GLK2-APRR2* respectively.

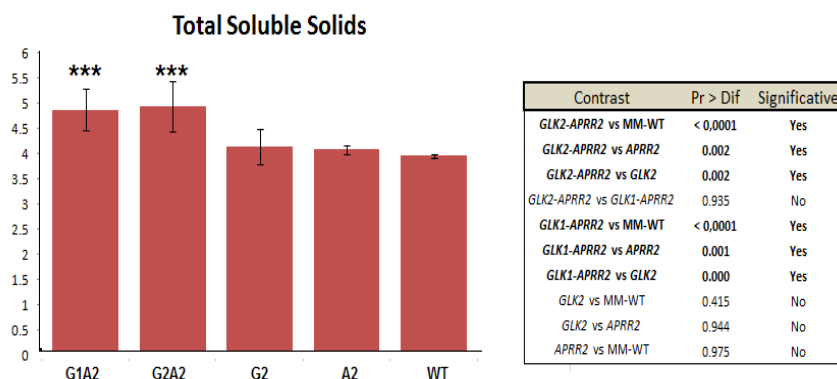


Figure 7: Total solids increases in *GLKs-APRR2* fruits. TSS were measured in RR fruits (left side) of *GLK1-APRR2*, *GLK2-APRR2*, *GLK2*, *APRR2* genotypes (see M&M). Significance difference between groups were detected between *GLKs-APRR2* and the rest of genotypes, whereas there was no difference between simple sCEF (*APRR2* and *GLK2*) and WT. ANOVA was carried out with Tukey (HSD), Fisher (LSD) and Bonferroni (99% confidence level). HSD contrast is represented in the table (right side of the chart).

2.6.2 dCEF lines show higher contents of nutritional and healthy compounds

HPLC analysis of medium and low polarity compounds was performed in RR fruit to define whether changes in nutritional and healthy compounds occurred. A first analysis examined changes in isoprenoid profile during ripening, using fruit from Br to RR (Fig. 8). Br samples are characterized by compounds such as lutein, β -carotene, α -carotene and violaxanthin, which display higher proportion in that stage of development. On the other hand, RR characterized for accumulating saturated carotenes such phytofluene and phytoene which are not detected in Br samples; and lycopene levels which increase in MM-control from 3 mg/g FW (Br) to 664 mg/g FW (RR).

When searching for the differences caused by our fruit engineering approach the largest differences between CEF and WT at Br stage were observed in Lutein levels which are notably

higher in engineered plants and mostly in dCEF. In addition, the carotenoid content of fruit from *APRR2* and *GLK1-APRR2* genotype displayed the highest amounts of carotenoids in general (200% compared to WT), mostly β -carotene at breaker and lycopene at RR. The β -carotene also increases in a significantly manner, (more than twofold) in fruit from *GLK2* and *GLK2-APRR2* lines.

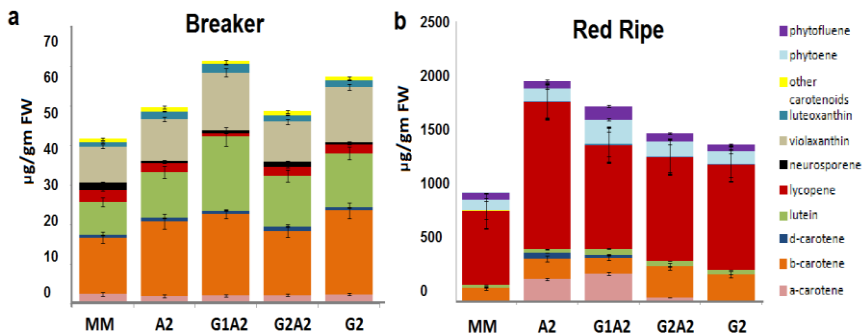


Figure 8: Nutritional and health quality related compounds increase in ripe fruit of all engineered genotypes. HPLC analysis (see M&M) performed in MM, *APRR2*, *GLK2*, *GLK1-APRR2* and *GLK2-APRR2* fruit samples at Breaker (left) and Red Ripe (right) stage of ripening development. Figure shows the carotenoid profiles in Br to RR fruits, as well as the contents in the different compounds for engineered genotype. Compounds references are in the right side of the chart, the color reference is keep for both graphics (purple phytofluene, light-blue phytoene, yellow other carotenoids, dark-green luteoxanthin, black neurosporene, red lycopene, light-green lutein, blue d-carotene, orange b-carotene, pink a-carotene).

2.6.3 Volatile profiles modified in *dCEF*

Double-engineered genotypes were selected to study the changes in volatile compounds at the ripe stage because they represent the strongest phenotypes achieved in our strategy. Five families were analyzed in the case of *GLK2-APRR2* and three representative families in the case of *GLK1-APRR2* based on the average values of green fruit color phenotype. A Principal

Component Analysis (PCA) was applied to assess the variability among genotypes, as well as to identify the main volatile compounds which contribute to those differences. PCA analysis explains approximately 42% (PC1 and PC2) of the total variance (Fig. 9a).

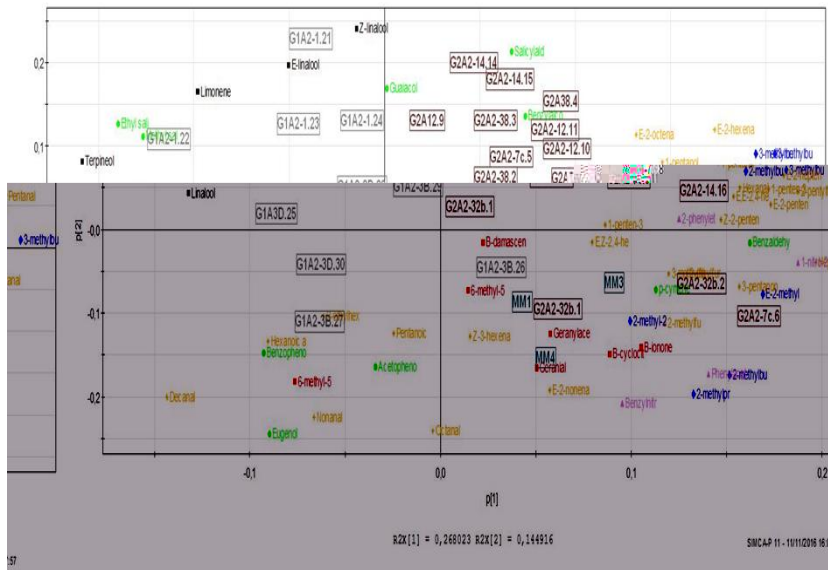


Figure 9: VOCs profile is modified in *GLKs::APRR2* fruits. Volatile compounds were analyzed in RR fruits of *GLK2-APRR2*, *GLK1-APRR2* and WT by GC (see M&M). Principal component (PC) analysis based on VOCs profile for each genotype. Loading plot is superimposed to the scores plot. VOCs: red Apocarotenoid, green Phenylalanine derived, blue Branched-chain amino acid derivatives, yellow Fatty acid, pink Benzenoid and black Terpenoid. Genotypes are highlighted with a square, in gray (*GLK1-APRR2*), brown (*GLK2-APRR2*) and blue (MM).

PC1 explains 27% of the variability with *GLK1-APRR2* located on the left, whereas MM and *GLK2-APRR2* on the right of the PCA. In general, fatty acid derivative compounds decrease in *GLKs-APRR2*, mainly in lines which express *GLK1* (Fig. 9). However, some derivative fatty acids show higher content in CEF, such as hexanal. Moreover, branched chain aa derivatives are highlighted because, as a general rule, increase levels in engineered fruits (2-methyl-2-butanol, 3-methylbutanol, E-2-

methyl-2-butenal), showing the largest amounts in *GLK2-APRR2* genotype (Table 2). On the other hand, the two ~~*GLK-APRR2*~~ genotypes are separated from WT along PC2 (explaining 15% of the variability). The most remarkable difference between CEF (besides higher branched aa ramified in *GLK2-APRR2*) are in the accumulation of terpenoids.

dCEF differ from WT plants as they produce ripe fruit with higher amount of a set of volatile compounds that may have an impact in flavor (pvalue \leq 0.05, table 2).

Terpenoid			
Family	Limonene	Linalool	Terpineol
G1A2-1	2.41	3.25	3.10
G1A2-3B	1.68	1.76	1.99
G1A2-3D	1.77	1.89	2.04
p value	0.02	0.03	0.03
G2A2-7	1.57	1.52	1.23
G2A2-38	1.87	1.56	1.24
G2A2-12	2.15	2.10	2.28
G2A2-14	1.54	1.85	1.61
G2A2-32	1.89	1.76	2.30
p value	0.02	0.03	0.03
Branched-chain amino acids derivatives			
Family	2-methyl-2-butanol	3-methylbutanol	E-2-methyl-2-butenal
G2A2-7	1.65	2.08	1.64
G2A2-38	2.12	2.37	1.59
G2A2-12	1.68	1.97	1.65
G2A2-14	1.55	2.19	1.26
G2A2-32	5.25	2.05	3.15
p value	0.01	0.01	0.04

Fatty acids derivative				
Family	1-penten-3-o	Heptanal	Decanal*	Octanal*
G2A2-7	0.60	0.93	0.80	0.86
G2A2-38	0.64	0.99	0.81	0.85
G2A2-12	0.78	0.90	0.75	0.77
G2A2-14	0.35	0.92	0.81	0.85
G2A2-32	0.32	0.96	1.00	0.88
p value	0.03	0.05	0.05	0.05
Family	1-penten-3-o	Heptanal	1-pentanone	1-penten-3-ol*
G1A2-1	0.89	0.71	0.72	0.59
G1A2-3B	0.86	0.75	0.66	0.57
G1A2-3D	0.75	0.40	0.65	0.29
p value	0.04	0.05	0.00	0.01
Family	E-2-nonenal*	2-penten-1o	Pentanal*	E-2-pentenal*
G1A2-1	0.83	0.55	0.73	0.89
G1A2-3B	0.76	0.75	0.54	0.85
G1A2-3D	0.81	0.26	0.39	0.57
p value	0.01	0.03	0.03	0.04

Table 2: Main family of VOC characteristic to each engineered genotype. Foldchange of terpenoids, branched-chain aa derivatives and fatty acids are represented in those tables (significance pValue (\leq 0.05), t-student). In color are highlighted the different families analyzed: G2A2 (red) and G1A2 (green).

3. DISCUSSION

There are many controversies about the idea of fruits as sink organs that exclusively rely on photosynthates import from leaves^{2,9,41,44}. The role of fruit photosynthesis in fruit quality has been much discussed^{5,166-168}; the contribution of photosynthetic activity in the fruit itself has been estimated as much as 20% of the total carbon⁹. It has been even associated the lower organoleptic quality in modern tomato varieties with selection against fruit chloroplasts⁵. During the last few years TFs implicated in fruit plastid biogenesis and metabolism have been identified^{8,78}. *GLKs* have been described as co-regulators of photosynthesis-related genes^{51,52}, and ectopic expression of *APRR2* has been shown to influence fruit plastid development⁸. Moreover, overexpression of those TF was shown to elevate organoleptic and nutritional quality in ripe fruits^{5,8}. However, the mechanisms underlying their function are not well understood. To date, those TFs have been only OE under a constitutive expression (35s promoter) and their co-expression has never been studied before.

Previous work suggested that a prerequisite to enhance chloroplast development is to act in early stages of fruit development. OE of *AtGLK1* or *AtGLK2* controlled under a late developmental promoter presented no change in fruit phenotype and no detection of mRNA⁵. In addition, tomato lines overexpressing those TFs under 35s promoter, in our hands, displayed negative effects on fruit development (low fruit set and small fruit, data not shown). For these reasons we chose to re-evaluate the relevance of enhancing fruit photosynthesis through the simultaneously OE of *GLKs* and *APRR2* controlled by a fruit specific promoter, as well as drive the expression from early stages of development. TFM5 promoter seemed to cover all the requirements as it was described to provide high levels of expression in IMG fruit and no activity in leaves², keeping a sustained expression until MG stages of fruit development¹⁵⁸.

As we anticipated, expressing *SIGLKs* or/and *APRR2* under the control of the TFM5 promoter generates darker green fruits at IMG stages; however, not only the expression but also the dark green phenotype disappeared sooner than expected. This pattern of expression driven by TFM5 promoter differs from what was expected based on the literature ¹⁵⁸ and Patent (number 5.608.150); where a more sustained expression was presented until mature green. However, the observed pattern is similar to that we obtained with TomExpress using the *TFM5* gene (Solyc12g014630.1.1) as a search string (<http://gbf.toulouse.inra.fr/tomexpress>). Those reports showing a more sustained expression could be due to the larger stability of the e.g. glucuronidase gene used as reporter ¹³⁷. In other cases where TFM5 was used for triggering RNAi silencing, early expression could have been sufficient for sustained silencing ². The loss of green color at MG stage could be also in part explained by the vacuolization taking place in cells during fruit expansion ¹⁶⁹.

In contrast to non-uniform tomatoes, which carry functional *GLK2* ⁵, CEF in MoneyMaker uniform background were homogeneously dark green. The ripening process in MoneyMaker, as in other varieties, is not synchronous; it starts in the stylar end and progresses later towards the basal end. Therefore, we believe that the “uniform” ripening phenotype is not a consequence of *GLK* expression and subsequently the name chosen for the “uniform” locus is a misnomer and should be better named green.

The observation of an additive effect obtained by co-expressing *GLKs* and *APRR2* in fruit is supported by all the analysis performed in this work. Color surface distribution plots of IMG fruit clearly clustered sCEF and dCEF lines separately. Chlorophyll accumulation followed a similar pattern and increase of 250 to 300% in double lines (e.g. *GLKs-APRR2*), compared to

the 80% in sCET lines. In the same line, carotenoids also showed higher content in *GLKs-APRR2* (fourfold over WT content) than either *APRR2* or *GLK2* line (twofold).

The dark color of CEF is not only associated to higher chlorophyll, but also to higher carotenoid levels which are in the line of previous studies⁶². At the same time, Chl/Car ratios are consistent with a normal coordinated regulation of carotenoid and chlorophyll synthesis¹⁷⁰, as a consequence of increasing both groups of compounds in the same proportion. Carotenoids increase parallel to chlorophyll, pointing out that *GLKs APRR2* strategy coordinates both pathways, which is essential for to protect the photosynthetic mechanism against excess light¹³⁹. In addition, chlorophyll a/b ratio was studied because is considered an indicator of the light adaptation which is taking place in photosynthesis machinery^{171,172}. This ratio showed differences between genotypes. Whilst in MM this ratio did not change from IMG to MG, CEF genotypes did, and this was more dramatic in *GLKs-APRR2* fruits which displayed the most reduced ratio. Thus, lower ratio is translated to higher amounts of Chl b and this pigment is associated to the antenna of PSII¹⁷³; a decrease in the Chl a/b ratio in engineered fruit could be interpreted as a greater development of the PSII antenna system.

Related to carotenoids, in general, lutein presents the larger increase followed by β -carotene and violaxanthin. Since these compounds are in different branch of the carotenoid pathway, it means that a global upregulation of carotenoid and chloroplast-bound isoprenoids is taking place. Under high light stress, plants reversibly convert lutein-epoxide and violaxanthin to lutein and zeaxanthin respectively, through two steps of de-epoxidation reaction. This mechanism reduces energy available for the light harvesting complex and is known as xanthophyll cycle¹⁷⁴⁻¹⁷⁶. As a response to more fruit chloroplast, CEF lines react by increasing

xanthophylls levels, possibly as a mechanism to prevent photooxidative damage in a richer chloroplast fruit context.

Primary metabolites were evaluated at pre-ripening stages. It is interesting to note that many aa increase at MG stage in all four CEF lines. Serine and glycine are inside this group; those aa have been described as important heme cofactor precursors, which chloroplast provides to mitochondria¹⁷⁷. Heme cofactor is a porphyrin derivative compound which plays an essential role in oxygen transport, respiration, signal transduction, and detoxification process¹⁷⁸. Taken together, isoprenoid and primary metabolite data would indicate that overexpression of *APRR2* and *GLKs* are not only limited to modify chloroplast/ photosynthesis but also transcends to other chloroplast processes and affect mitochondria organelle. Both organelles are considered the powerhouses of the cell because of the energy- transducing process which take place in photosynthesis and electron transport chain. Chloroplast-mitochondria development processes are coordinated by signals organelle-nucleus (retrograde), as well as nucleus-organelle (anterograde)^{177,179}.

Primary metabolite analysis also showed that dCEF, at IMG stage, present more sugar content and more starch, whereas no differences were detected in simple engineered lines. During IMG stage of development, a positive correlation between sucrose synthesis and ADP-glucose pyrophosphorylase activity on one side, and starch content on the other side has been demonstrated. The availability of glucose and fructose determines the final starch accumulation¹⁸⁰. It is interesting that previous study showed an increase of starch levels at IMG in response to *AtGLK1* or *AtGLK2* overexpression¹⁴¹.

To date, the mechanism of fruit photosynthesis has not been fully clarified. Given the specific fruit characteristics, fruit photosynthesis could be classified in the intermediate point between C3 and C4/CAM plants¹⁸¹. Tomato fruit have no

stomata, therefore CO₂ availability and fixation must be different from the process which takes place in leaves, and involves PEP carboxylase metabolism²¹. Some studies have shown a decreased Rubisco activity in fruits, accompanied by an increase in PEP carboxylase similar to that found in C₄ and CAM plants, where CO₂ is fixed by PEP and stored in vacuoles as malic acid, from where it is mobilized to provide a substrate for light dependent CO₂ fixation¹⁸². Indeed, it has been reported that, in fruit, mitochondrial respiration generate CO₂ and ATP which allow photosynthesis process, whilst chloroplast provides malate and O₂ in exchange¹⁷⁷. For example, in cucumber fruit, photosynthesis contributes to the 9.4% carbon fixation and PEPC activity has been identified as predominant in perivascular and placenta tissues. In addition, organic acids are higher in cucumber fruits than in leaves, presumably because they are synthesized by PEP¹⁸³. Here, our own results seem to support with this suggestion. We observed higher total amounts of organic acids in the four CEF genotypes, mainly malic acid content at both stages of green development (IMG and MG).

Previous studies⁵⁷ have demonstrated the importance of starch content at IMG stage for soluble solids and sugar content of the mature fruits. Similarly, overexpression of *GLKs* or *APRR2* was demonstrated to positively impact in red ripe fruit quality. *p35s::AtGLK* was reported to increase in fructose and glucose to reach 40%; and 21% of TSS in red fruit^{5,50}. Our results demonstrate that the early transgene expression has an effect on fruit at ripening stages of development. Again, the additive effect in dCET also works at this level. While simple lines increase over 3-4% soluble solids, ~~*GLKs*~~ *APRR2* reaches brix values around 23-25% higher than WT and reinforces the hypothesis that acting on chloroplast development at earlier stages, even IMG stage, could impact in red fruit quality. On the other hand, red ripe quality traits like aroma volatile profiles are affected in engineered lines. Previous work reported changes in volatile derivative compounds in chloroplast-rich *hp2* fruits⁶¹, with

increased values of (Z)-3-hexenal in mutants lines (1.6 times). Hexanal has been reported to affect tomato fresh flavor¹⁸⁴ and sweet taste¹⁸⁵. Here *GLK2-APRR2* fruits were found to produce more hexanal levels than WT or *GLK1-APRR2* fruits. Trans-2-hexenal has been described to be more important by mouth-feel properties than odor in itself^{186,187}. Fruit volatiles are classified into four main classes; derived from carotenoids, fatty acid volatiles, terpenoid volatiles and amino acid derived volatiles^{1,188}. Our results showed terpenoids as one of the important differential groups. *GLKs-APRR2* red fruits are enriched in volatile compounds derived from those compounds compared to MM. During ripening, plastidic terpenoids pathways increase but at the end tomato fruit, in general, contain low level of monoterpenes and sesquiterpenes¹⁸⁹. The increase in terpenoid VOCs in engineered fruit may be a direct consequence of the early enhancement of fruit chloroplasts.

Another interesting VOCs group is the phenylalanine derivative class that has also been associated to flavor/ aroma. This group of volatiles cluster *GLK2-APRR2* and *GLK1-APRR2* together and away from WT (pvalue<0.05). Interesting, guaiacol, which was found in lower levels for both genotypes is considered in general as undesirable compound for tomato (although the Asian consumers have a preference for it)¹⁹⁰ and is responsible for the 'smoky'^{191,192} or pharmaceutical fragrance and even associated to tissue damage^{24,193,194}. Interest in modulating their leaves have led to a number of specific genetic and biotech approaches to change the levels of phenyl-derivatives^{193,195} and our approach would lead inadvertently to a similar effect in ripe fruit.

The ability of a taste panel or consumer panel to feel the difference between those fruit will be the objective of future experiments.

In summary, we confirmed our fruit specific chloroplast-enhancing approach produced a consistent range of changes in fruit phenotype which ranged from an increase in chloroplast total surface (number, size and extend of cellular layer) and structure (higher develop of grana stacking) to metabolite changes occurring at green stages of development. We observed a dilution of the phenotype at MG stage, consequence of TFM5 pattern expression. However, the noticeable changes at IMG were enough to impact fruit quality at the red ripe stage, mostly as a consequence of a synergistic affect produced by co-expressing the two TFs. We conclude that *GLKs* and *APRR2* affects plastid metabolism and improve fruit quality individually, and that the combination of both has even a stronger effect. The molecular mechanism of this additive effect will be the subject of future studies to analyses at transcriptomic and proteomic levels.

4. MATERIALS AND METHODS

4.1 Plant materials and growth conditions

Tomato (*Solanum lycopersicum*) seeds from the cultivar cv. *MoneyMaker* (u/u) were provided by The Institute for Mediterranean and Subtropical Horticulture "La Mayora" (IHSM-UMA-CSIC) germoplasm bank. Transgenic plants for the different constructs were grown in P2 contention- level greenhouses during the spring-summer of 2015 at IHSM-UMA-CSIC (Algarrobo-Costa, Malaga). T3 generation lines were grown in spring-summer of 2016 at two different locations in Spain: The Institute for Plant Molecular and Cell Biology (IBMCP, Valencia) and IHSM-UMA-CSIC (Algarrobo-Costa, Malaga). Seedlings were transplanted to 16-L pots at the four true-leaf growth stage; they were watered daily and with soluble fertilizer once a week.

In some experiments, plants were cultivated in *in vitro* conditions provided by a growth chamber room set to a 16 h daylight cycle ($250 \mu\text{mol photons m}^{-2} \text{s}^{-1}$) and a relative humidity of 60-70% under 25°C (day/night temperatures). In this case seeds were sterilized with 50% of bleach, rinsed in sterile water, and sown in one-half strength Murashige and Skoog medium with 1.5% of sucrose, 1% agar, pH 5.8-6. After 3 days in darkness plants were grown under daylight cycle as indicated above.

All fruit samples for molecular characterization were collected around the same time of the day (9.00 to 11.00), pericarp excised, snap-frozen in liquid nitrogen and stored at -80°C until required. For each case a pool of 3 fruits for the defined developmental stage and coming from different plants representing different transgenic events for the same construct were mixed and stored as biological replicate. Each experimental data point is represented by 3-5 biological replicates.

To stage the fruits, flowers were tagged at anthesis and the fruit developmental stage was defined by the number of post-anthesis days (DPA). Immature green fruits (IMG) were defined as those fruits at 15-20 DPA and show as dull aspect. Mature green stage was defined by 35-40DPA and characterized for glossy green aspect of fruits that have reached the maximum size. Breaker stage (B) is when the fruit color starts to change from green to orange in the apical area. In the Light Red (LR) stage the fruit is completely orange and the Red Ripe (RR) was defined in days post-breaker, they were harvested between 7 and 10 days after breaker (depending on the season and the genotype).

4.2 Plasmid construction and generation of transgenic plants

To generate the genetic constructs used in this work the Golden Braid (GB) 2.0 cloning system was used¹⁹⁶.

DNA sequences corresponding to the *TFM5* (Solyc12g014630.1), *GLK1* (Solyc07g053630.2), *GLK2* (Solyc10g008160.2) and *APRR2* (Solyc08g077230.2) genes were fetched from Solgenomics and specific primers used to amplify the DNA of interest either from genomic (*TFM5* promoter) or cDNA (coding regions of *GLKs* and *APRR2*). All DNA fragments were confirmed by sequencing and introduced in the GB system. Promoter and terminator sequences were amplified from the *TFM5* gene and span 1200 bp upstream of the ATG of *TFM5* (SL2.50ch12:5631575..5630375) and 544pb downstream of *TFM5* gene (SL2.50ch12:5631496..5630952) respectively. A 934 bp PCR fragment of the tomato *GLK2-Like* gene from non-uniform background (*Ailsa Craig* U/U), a 1684 bp ORF corresponding to *APRR2-like* and a 1396 bp ORF of *GLK1-Like* were amplified from cDNA prepared from total RNA of fruit at breaker stage and leaf sample respectively. The ORF DNAs introduced in the sense orientation between *TFM5* promoter and *TFM5* terminator using GB cloning method. For double engineering purposes, transcriptional units corresponding to *GLKs* and *APRR2* genes were combined in a multipartite GB assembly to create two different final destination vectors with *GLK1-APRR2* and *GLK2-APRR2*; the kanamycin resistance gene (*nptII*) was also included and the final size of the assembled constructs were 7.907 pb and 7.283 pb respectively. Primers used for the cloning of the GB parts, domestication are listed in supplementary table 1.

Recombinant plasmids were checked at each stage of cloning with restriction digest analysis and by sequencing at the domestication step¹⁹⁶. Final constructs (pTFM5::*GLK2*,

pTFM5::APRR2, pTFM5::GLK2::pTFM5:APRR2, pTFM5::GLK1::pTFM5::APRR2) were independently introduced into plants by an *Agrobacterium tumefaciens* (strain LBA4404) mediated transformation cotyledon-explants protocol as described by Ellul et al. (2003) with minor modifications¹⁹⁷. Transgenic plants were identified by the ability to root on 100 mg/L kanamycin and genomic DNA extracted from young leaves¹⁹⁸, was confirmed for the presence of the T-DNA construct by PCR. Confirmed plants were transferred to the greenhouse. Gene-specific primers used in the genotyping are listed in supplementary table 2.

4.3 Transgenic plants evaluation

Plants were evaluated at T2 and T3 generation. Phenotypic analysis included external and internal color measurements, as well as, metabolites and organoleptic related traits. Objective color of the fruit surface (external color EC) was determined with a Minolta Chroma Meter model CR-400 (Konica Minolta, Inc., Tokyo, Japan). CIE Lab color space was used (+a= red, -a= green, +b= yellow and -b=green). The color space is three-dimensional and Luminosity parameter (L) defines if the sample is lighter (higher L) or darker (lower L), Chroma (C*) is a measure of color saturation and Hue-angle (H) is the color measurement taking in account "a" and "b" values. L values provided the largest differences between transgenic samples and WT. Each fruit color measurement is an average of 3 readings taken from the equatorial region of the tomato fruit. To determinate the color variation in populations of engineered fruits, the color essay was conducted on fruit from T2 generation plants. The measurements were obtained from 10 families for each genotype (APRR2, GLK2, GLK2-APRR2, GLK1-APRR2), 3 plants per family, 3 fruits per each stage of development.

In order to measure the internal color (IC), samples were harvested in IMG, MG, B and RR stages and cut along and across

of the longitudinal axis; after that were scanned with the Tomato Analyzer software¹⁹⁹

For TSS, between five and six fruits were grounded and the mixture was stored at -20°C previous to the analysis. Samples were thawed and centrifuged at 3500 rpm for 15 min and the supernatant was used for the analysis. TSS was performed using a digital refractometer (Atago CO LTD, Tokyo, Japan)

Fruit weight (in grams) was measured at MG stage on the second trusses of T3 plants. Approximately, 6-8 fruits were harvested from four-five plants per family using three-five families per CEF line.

Iodine staining was used as a qualitative test to see whether starch metabolism was affected in CEF. MG fruits samples from dCEF, sCEF and WT were harvested. Samples were cut transversally and let in iodine solution for two hours.

4.4 RNA extraction and qRT PCR expression analysis

Total RNA was extracted from ~200mg of tissue by using TRIzol RNA Isolation Kit (Thermo Fisher Scientific) according to manufacturer's instructions and it was treated with DNA-free™ DNA Removal Kit - Ambion following the manufacturer's recommendation. DNA-free RNA preps were reverse transcribed with SuperScript II (Takara). Single strand cDNA was then synthesized from 1 µg of RNA in 20 µl with oligo-dT and used for qRT PCR in 7500 fast Quantitative real-time PCR Applied Biosystem. PCR were run in triplicates using SYBR Green kit (Takara) following the manufacturer's instructions. RNA levels were expressed using the relative standard curve method (Applied Biosystems) and Actin 7 from tomato as an internal control Primers used in qRT PCR essay were designed using

Primer Express software (Applied Biosystem) and shown in supplementary table 2.

4.5 Carotenoid and chlorophyll analysis

LC-ESI(+/-)-MS (Liquid chromatography-electrospray ionization-tandem mass spectrometry) analysis of chlorophyll and carotenoids pigments in fruit samples were performance as previously described (Liu et al., 2014).

Fruit samples were harvested to represent 3 stages of fruit development as follows: IMG, MG and RR. Briefly, 20 (for IMG and MG samples) or 5 (RR samples) mg of freeze-dried powder were extracted with chloroform (spiked with 50 mg/l DL -alpha -tocopherol acetate as internal standard) and methanol (2:1 by volume); 1 volume of 50 mMTris buffer (pH 7.5, containing 1M NaCl) was then added and the samples were kept for 20 min on ice before a centrifugation step at 15,000 g for 10 min at 4°C. The hypophase was collected and the aqueous phase was re-extracted with the same amount of spiked chloroform; the two organic phases were combined and dried by speedvac and the resulting pellets were resuspended in 50 ul of ethyl acetate.

For each genotype, 4 fruits and at least two independent extractions were performed. LC-DAD analyses were carried out using an Accela U-HPLC system (Thermo Fisher Scientific, Waltham, MA, USA). LC separations were performed using a C30 reverse-phase column (100 x 3.0 mm) from YMC (YMC Europe GmbH, Schermbeck, Germany) with a mobile phases composed by methanol (A), water- methanol (20 : 80 by volume) containing 0.2% ammonium acetate (B) and tert -methyl butyl ether (C). The gradient was 95% A : 5% B for 1.3 min, followed by 80% A : 5% B : 15% C for 2.0 min and by a linear gradient to 30% A : 5% B : 65% C for 9.2 min. UV- visible detection was performed continuously from 220 to 700 nm with an online Accela Surveyor

photodiode array detector (PDA; Thermo Fisher Scientific). Mass ionization was performed with an atmospheric-pressure chemical ionization (APCI) probe, operating in both + and - voltage conditions. Nitrogen was utilized at 20 and 10 units as sheath and auxiliary gas, respectively. The vaporizer and capillary temperature were set at 300 °C and 250 °C, respectively. The discharge current was 5.5 μ A, while S-lens RF level was set at 50. A mass range of 110/1600 m/z was used both in positive and in negative voltage with the following parameters: resolution set at 70000; microscan, AGC target and maximum injection time equal to, respectively, 1, 1×10^6 and 50. All solvents used were HPLC-MS grade (Merck Millipore). All solvents used were LC-MS grade quality (CHROMASOLV® from Sigma-Aldrich).

Carotenoids and chlorophyll a and b were quantified on the basis of the internal standard amounts, obtained through comparison with peak areas of known amounts of external standard LC-MS runs and by extinction coefficient correction. Other isoprenoids (chlorophyll catabolites, tocochromanols, quinones) were identified on the basis of the accurate masses and by comparison with authentic standards when available.

4.6 Light and transmission electron microscopy

Samples of fruits from the different genotypes used in this work (*APRR2*, *GLK2*, *GLK2-APRR2* and *GLK1-APRR2*) were processed for light and transmission electron microscopy.

Two types of processing, using either High Pressure freezing and Freeze substitution (HPF-FS), or chemical fixation at room temperature, were performed. For HPF-FS, flat round, 3 mm-wide disks were taken with a dissector from the fruit pericarp. Pericarp disks were transferred to aluminium sample holders, cryoprotected with 150 mM sucrose, and high pressure frozen in a Leica HPM100 high-pressure freezer (Leica Microsystems,

Vienna, Austria). Then, samples were freeze substituted in 2% OsO₄ in anhydrous acetone at -80°C for 4 days, followed by slow warming to room temperature over a period of 24 h, according to Seguí-Simarro (2015). After rinsing in several acetone washes, they were removed from the holders and infiltrated with increasing concentrations of Spurr resin (Ted Pella, Redding, CA) in acetone according to the following schedule: 4 h in 2% and 5% resin, 12 h in 10%, 25%, 50% and 75% resin, and 40 h in 100% resin. In parallel, some pericarp samples, including the epidermis and mesocarp, were chemically fixed in Karnovsky's fixative (4% formaldehyde + 5% glutaraldehyde in 0.025M cacodylate buffer, pH 6.9,²⁰⁰ for 5 hours, postfixed in 2% OsO₄ in cacodylate buffer, dehydrated in increasing concentrations of ethanol during 30h and embedded with increasing concentrations of Spurr resin in ethanol as follows: 3 h in 30% resin, 15 h in 50% resin, 8 h in 70% resin and 40 h in 100% resin²⁰¹.

For both types of processing, all further steps were identical. Resin polymerization was performed at 70°C for 30 h. Thin sections (500 nm) from sample-containing polymerized blocks were obtained for light microscopy observation, and ultrathin sections (~80 nm) were obtained for electron microscopy, using in both cases a Leica UC6 ultramicrotome. Ultrathin sections were mounted on Formvar and carbon-coated, 150 mesh nickel grids, counterstained with uranyl acetate and lead citrate, and observed in a Jeol JEM 1010 transmission electron microscope.

Parameters measured in the images obtained from ultrathin sections were the followings: size of grana, starch granules and chloroplasts and number of grana and starch granules per chloroplast. From light microscopy images, information about number and distribution of chloroplast per area and per cell was obtained. All that data were evaluated by progressive mean test

²⁰² in order to know how many images were needed in each case, and then ANOVA analysis and LSD test were performed.

4.7 Volatile compounds profiling

Volatile compounds were analyzed in RR fruits as described in Rambla et al. 2017. Briefly, tomato pericarp was frozen in N₂ (l) and grounded. After this, 500 mg of frozen fruit powder was weighed for each sample and put in a closed vial (7 mL) at 37°C for 10 min of incubation. After that CaCl₂·2H₂O (1.1 g) and EDTA 100 mM pH 7.5 (0.5 mL) were added; the vials were shaken and sonicated for 5 min. 1 mL was vested into a 10 ml screw cap headspace vial and they were tempered at 50°C for 10 min and extracted by Headspace Solid Phase Microextraction (HS-SPME). Volatiles adsorbed in the fiber were desorbed at 250°C for 1 min in the injection port of a 6890N gas chromatograph (Agilent Technologies) in splitless mode. DB-5ms (60 m x 0.25 mm x 1 µm) column (J&W Scientific) was used with helium as carrier gas at a constant flow of 1.2 mL/min.

Compounds were identified by comparison of both mass spectrum (5975B, Agilent Technologies) and retention time to those of pure standards (SIGMA-Aldrich). For quantification, peak areas of selected specific ions were integrated for each compound and normalized by comparison with the peak area of the same compound in a reference sample injected regularly.

4.8 Primary metabolite profiling

Primary metabolite profiling was performed on methanol extracts after derivatization using gas chromatography mass spectrometry (GC-MS) coupled to a time-of-flight mass spectrometer (Leco Pegasus HT TOF-MS) as described in ²⁰³. Briefly samples were extracted in 100% methanol at 70°C for 15

min. After centrifugation, the resultant supernatant was dried under vacuum, and the residue was derivative for 120 min at 37°C (in 50 µl of 20 mg ml⁻¹ methoxyamine hydrochloride in pyridine) followed by a 30 min treatment at 37°C with 50 µl of MSTFA.

For the GCMS analysis, helium was used as carrier gas at a constant flow rate of 2 ml/s and gas chromatography was performed on a 30 m DB-35. Chromatograms and mass spectra were evaluated by using Chroma TOF 4.5 (Leco) and TagFinder 4.2 software^{203,204}.

4.9 Data analysis and visualization

MapMan software (version 3.6.0RC1) was used to visualize datasets on biochemical and regulatory pathways; DEG and DEP were mapped against the genome using ITAG 2.3.

The SIMCA current trial version software was used for PCA analysis (<http://mksdataanalytics.com/products/simca>).

The Sol Genomics Network web page (SGN, <http://solgenomics.net>) and tools in there was used to identify genes, search identification genes code, gen annotations (ITAG 2.3) and sequences.

Statistical analysis performance by excel program completed with the trial version of XLSTAT software. ANOVA analysis (comparison according to Tukey, Fisher and Bonferroni), Student's t test and histograms were performed with this tool.

CHAPTER 2: Molecular mechanisms underlying the fruit phenotypes driven by early expression of GLK and APRR2



1. INTRODUCTION

The importance of fruit photosynthesis and the light signal transduction machinery for tomato fruit quality have been recently reevaluated, opening new avenues for tomato breeding and biotech approaches to improve fruit quality^{59,61,70,205}. A number of transcription factors (TFs), which enhance chloroplast development were reported and associated not only to increase levels of fruit specialized metabolites, but also sugars in red ripe fruit¹⁴³.

Chloroplast biogenesis is a complex process where the role of light is indisputable, but its molecular genetic basis is not yet completely understood²⁰⁶. Expression of photosynthesis-related genes is also affected by developmental state or by organ-specific cues generating controversial results^{207,208}. The development of this organelle is a highly coordinated process and includes the coordination between two separate genomes: chloroplast and nucleus. In addition, occurs at multiple levels from transcription, translation, protein-import to RNA stability and protein-turnover¹⁷⁷. Approximately 3000 chloroplast proteins are nuclear-encoded, synthesized in cytoplasm and transported into the organelle often through the interaction of a peptide transit and chloroplast membrane (TOC-TIC, translocon at the outer/ inner envelope membrane of chloroplast)²⁰⁹. Many of these proteins participate in the main function of chloroplast: light absorption and conversion into chemical energy in thylakoid membrane by the action of complex machinery composed by two photosystems (PSI and PSII), a modular antenna, a reaction center (RC) and light harvested complexes proteins (LHCs). Most genes encoding photosystem reaction center subunits are localized in the plastid-genome, whereas LHC proteins (chlorophyll (Chl) and carotenoid-binding proteins) are nuclear-encoded²¹⁰. Photosystem I and II are mostly

associated to chlorophyll a, and LHC proteins require chlorophyll b for protein assembly ¹⁶⁰.

Chloroplast biogenesis is highly regulated by light ²⁰⁶. Some tomato mutants have been characterized to display an exaggerated light responsiveness such as *high pigment (hp) 1* and *2*; both mutants contain higher number of larger chloroplasts and increased carotenoid levels ²¹¹. *Hp1* and *hp2* are caused by mutations in different light regulatory genes ⁴³. Thus the *hp1* is defective in *UV-DAMAGED DNA BINDING PROTEIN 1 (DDB1)*, *hp2* mutant does in *Deetiolated1 (Det1)* ^{52,64}. Both genes are negative regulators of the photomorphogenesis and interact each other at genetic and protein levels in the well described CULLIN4-RING E3 ubiquitin ligase (CRL4s) complex ^{70,212}. DDB1 targets substrates for the proteasome complex and DET1 protein may act as a linker allowing substrate association ^{213,214}.

In contrast, other mutations have been described to affect specifically fruit chloroplast development. Tomato *uniform ripening (u)* is characterized by a pale green fruit phenotype including the loss of green shoulder in tomato fruits and shown to be a loss-of-function of *Golden 2-Like gene (SIGLK-2)*, caused an adenine insertion which generates an CDS shift, a premature stop codon and consequently a truncated protein ⁵. In contrast to *GLK2*, which is expressed in both fruit and vegetative tissues, the *GLK1* showed restricted expression to vegetative tissues ⁵². More interesting, *GLK2* protein has been recently identified as possible target for the proteasome CUL4-DET1-DDB1 complex ⁶⁸ affected in *hp* mutants. Another identified mutant in tomato, *uniform gray-green (ug)* displays the same phenotype as *u* but is due to *TKN4* gene (Knox) ⁴⁷. The only mutation so far described consists in a nucleotide substitution, which causes an aminoacidic change (Phe to Leu). *TKN4* has been shown to influence the latitudinal gradient of *SIGLK2* expression in tomato

fruit associated to the pale fruit color and lack of green shoulder phenotype⁸¹.

GLKs have been shown to be under the control of plant hormones and thus, auxin and brassinosteroids (BRs) have been described to influence fruit chloroplast through the control of *GLK* expression levels. The *BES1* (*bri1*-EMS-suppressor), a nuclear protein which regulates BRs biosynthesis, interacts with *GLK1* and *GLK2* repressing specific expression in dark conditions and consequently interrupting chloroplast development²¹⁵. *GLK2* has been demonstrated to act downstream of the light and auxin/ cytokinin signaling pathways, other effectors of chloroplasts and greening⁷¹. Down-regulation of *SIARF4* (auxin factor response) has been shown to upregulate *GLK2* levels in tomato fruit⁷.

It has been also described that the ectopic expression of *ARABIDOPSIS PSEUDO RESPONSE REGULATOR (APRR2)* enhances tomato fruit chloroplast; however, the participating mechanisms are not well understood⁸. *APRR* TFs roles have been implicated in circadian rhythm, control of flowering time and photomorphogenesis. It has been also described that for *APRR2* to function is required a calmodulin-like proteins binding (CMLs), which results in *CML9* releasing Ca^{2+} signals²¹⁶. Interestingly, *SIAPRR2* has sequence similarity to *GLK2*²¹⁷ and both TFs were identified as main contributors to the variation in fruit greenness in genetic studies in pepper^{8,154}.

A number of studies have used transcriptome analysis to elucidate the molecular basis of *GLKs* regulatory mechanism. Arabidopsis plants which overexpress (OE) ~~*GLK*~~ by 35s::*GLK1* and 35s::*GLK2* showed upregulation of a common set of genes involved in photosynthesis, being all the top genes related to Lhcb and chlorophyll biosynthesis. *GLK1* OE produced the most dramatic transcriptional upregulation⁵⁰. A more recent report on the transcriptome changes of IMG tomatoes, which OE *AtGLK1*

and *AtGLK2* under 35s promoter in *u/u* tomato background, resulted in around 60% of the unregulated genes being common for both types of transgenic lines ⁵. This and the previous ⁵⁰ study also provided indications on the importance of OE *GLKs* early in development to produce a fruit phenotype, since the use of late fruit developmental promoters with *GLKs* were ineffective in conferring phenotype. Moreover, it was reported that *GLKs* OE does not affect the timing of ripening and that *GLK2/TKN4* are responsible of the chlorophyll gradient normally produced from the styler to the basal end of the fruit ¹⁴¹.

We have shown, earlier in Chp I, that expression of *SIAPRR2* and *SIGLKs* genes, either singly or combined but exclusively expressed early in fruit development, provided different degrees of chloroplast-associated phenotypes. Here, we aim to understand the mechanism underlying this range of fruit phenotypes by performing detailed analyses at transcriptomic and proteomic levels, with special interest for the novel additive/synergistic phenotype achieved by the simultaneous ectopic expression of both TFs early in ovary/ fruit development.

2. RESULTS

2.1 Transcriptome data overview shows changes consistent with the metabolomic fruit phenotype

We reported previously the chloroplast structure and metabolite characterization of fruits from single (sCEF) and double (dCEF) *GLK/APRR2* engineered plants. Here, we first analyzed the changes in the transcriptome and tried to integrate that information with the metabolite and structural information obtained for the same fruit-set. Transcriptome analysis was performed at IMG and MG fruit stages of development (Genes

are listed in Supplementary tables 1-2 on CD). The RNAseq analysis was conducted on three biological replicates for each family corresponding to five types of genotypes, including single or combined engineered lines and controls (*APRR2*, *GLK2*, *GLK2-APRR2*, *GLK2-APRR2* and WT-MM). Libraries were generated with an average of 20000 and 25000 observed genes (Supplementary Fig. 5) and here we analyzed the foldchange values (log2) related to WT samples.

A first review of the transcriptome data was performed by studying functional categories and pathways affected by Differential Expression Genes (DEG). We chose for that the *GLK2-APRR2* expression dataset corresponding to the IMG stage of fruit development since, as we described previously, dCEF displayed the strongest phenotypes.

Gene Ontology (GO) term enrichment analysis revealed that the large portion of the 949 DEGs (P_{adj}<0.05) are related to photosynthesis (upregulated) and cell wall structure (downregulated) functional categories (Supplementary table 3 on CD). In order to visualize expression levels and the pathway responses in a compact manner, MAPMAN software²¹⁸ was used to map RNAseq reads. Figure 1 shows how the expression of *SIGLK2* and *SIAPRR2* affect the levels of transcripts involved in many biological processes. This figure and the enrichment analysis (table 3 on CD) reveal the upregulation of genes related to tetrapyrrole and chlorophyll biosynthesis processes.

Moreover transcripts involved in light photosynthesis and Calvin cycle are highly active in those fruit. Secondary metabolism is also enhanced as indicated by a group of representative genes. Genes involved in the synthesis of Geranylgeranyl phosphate (precursor of important metabolites as terpenoids, carotenoids, tocopherol and chlorophylls) are upregulated: farnesyl pyrophosphate synthase and geranylgeranyl pyrophosphate synthase (solyc12g015860.1.1,

solyc04g079960.1.1). Moreover, mitochondrial electron transporters and ubiquinone synthesis are also affected by upregulation of NADH dehydrogenase and Cytochrome c oxidase (solyc06g007160.2.1 and solyc03g043850.2.1).

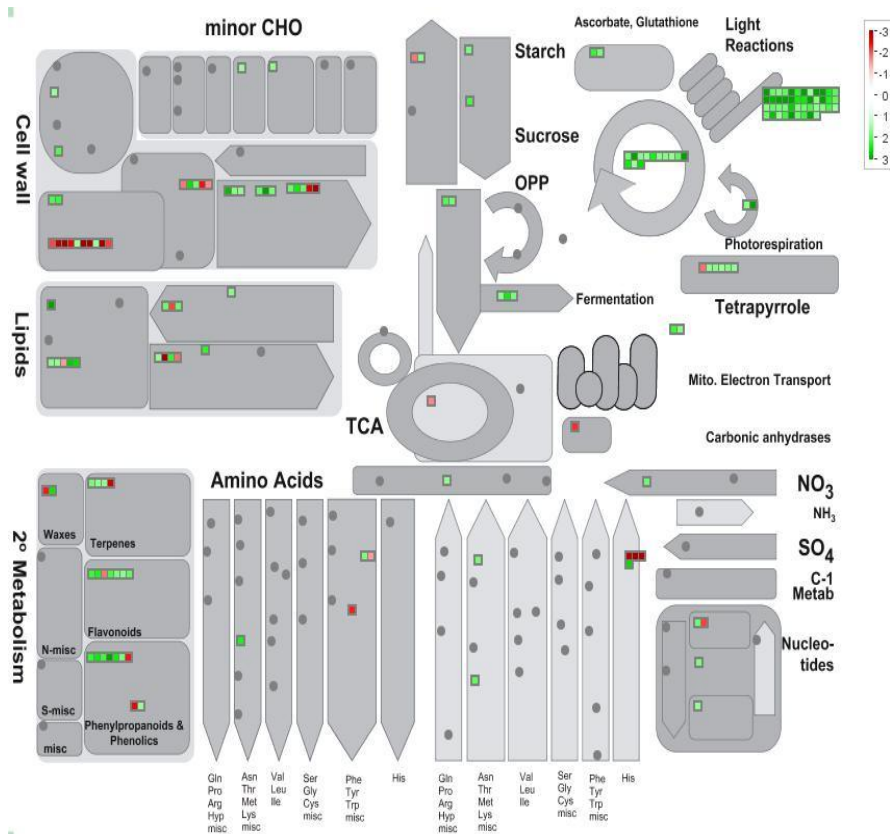


Figure 1: Transcriptome data of *GLK2-APRR2* genotype confirmed changes in primary metabolism and chloroplast composition. MapMan snapshot Metabolism overview display DEG in *GLK2-APRR2* compared to WT at IMG stage. Upregulated genes in green color (foldchange > 1.5) and downregulated in red (foldchange < -1.5) (PPadj < 0.05). 949 data point were mapped (table 3 on CD).

Minor carbohydrate metabolism-related genes (CHO) are upregulated indicating that pathways associated to sugar metabolism are affected. For instance, starch synthesis/degradation pathways showed higher levels of ADP-glucose

synthase and alpha-amylase (solyc01g079790.2.1/solyc04g078930.2.1) and sucrose degradation is probably altered by the upregulation of beta-fructofuranosidase (solyc08g079080.2.1).

In addition, there are indications of changes in the aminoacidic metabolism/ turnover as reflected by the upregulation of genes involved in synthesis-degradation of asparagine and aspartic acid (solyc04g078460.2.1, solyc10g075170.1.1, solyc09g007830.2.1, solyc03g098300.1.1).

2.2 Early activation of *GLK2* or *APRR2* under the control of a fruit specific promoter affects a common set of genes that are related to fruit photosynthesis and chloroplast development

Next we perform an analysis focused on investigate the structural-molecular-gene expression relationship that underlies the different aspects of the phenotype produce by the expression of *GLK* or *APRR2*, under an early fruit specific promoter. We examined either functions directly related to the fruit chloroplast, or in any other reprogramming unrelated to green/ chloroplasts development.

The total number of DEGs was higher at IMG stage in all the engineered lines involving *GLK2* or *GLK1*, whereas it was higher at MG stage for *APRR2* lines (Supplementary Fig. 6). This suggests *GLKs* affects the fruit transcriptome mainly at IMG stage, whereas *APRR2* effects remain high at later stage of MG fruit despite the lower levels of *APRR2* transgene.

The analysis of fruit transcriptome at IMG stage in sCEF was performance with a more restrictive cutoff criteria (PPadj < 0.01),

pursuing to find functional differences between TFs. A total of 47 common upregulated genes and 165 downregulated (Fig.2a) were detected in this analysis. Those common upregulated DEGs represent the 55% of *APRR2* upregulated genes, and 3% of *GLK2*. The proportion of common genes with a downregulated pattern is higher, representing 77% of *APRR2* and 17.5% of *GLK2* respectively.

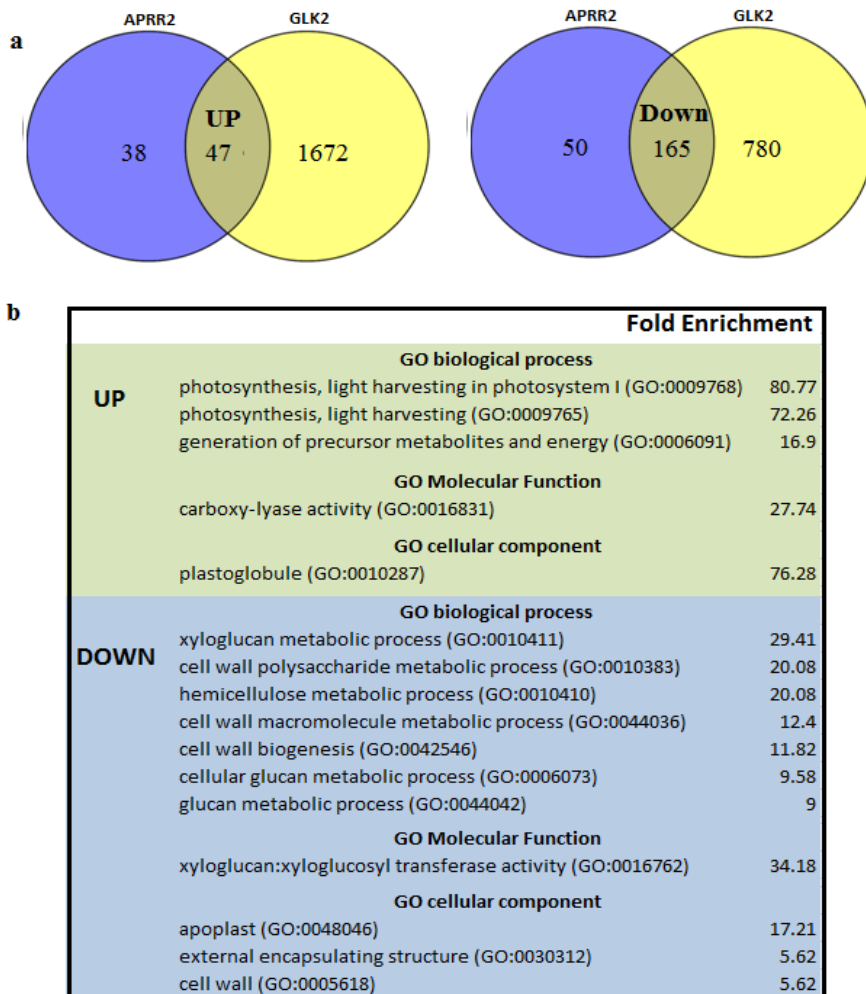


Figure 2: *APRR2* and *GLK2* upregulate photosynthesis function and show common pattern by downregulating cell wall-related genes (continued)

Figure 2 (continued): (a) On the left, Venn diagram with the upregulated list for pTFM5:*APRR2* and pTFM5:*GLK2* displays 47 genes shared for both genotypes. On the right, figure shows 165 overlapping genes from the downregulated lists (Supplementary table 4 on CD) **(b)** Fold enrichment of the 47 shared genes with upregulated pattern (green) and the 165 downregulated (blue). GO terms categorized for Biological Process, Molecular Function and Cellular Component (<http://geneontology.org/>) with P_{adjust}<0.01 (see M&M).

A GO term enrichment analysis applied to the set of common upregulated genes resulted in an obvious enhancement of transcripts related to photosynthesis functions (Fig. 2b). Genotype exclusive effects were also identified for the pTFM5::*APRR2* and pTFM5::*GLK2* lines which showed 38 upregulated genes and 1672 specifically activated genes respectively. Focusing on the set of genes specifically overexpressed in *APRR2* (38 genes), the largest majority of them fall in the unclassified class with the exception of *Pyruvate Decarboxylase* gene (solyc09g005110.2.1, foldchange over 2).

In contrast, when we investigate specifically the gene functions associated to genes overexpressed only in *GLK2*, the majority of them are clearly identified and include genes involved in the regulation of chloroplast development, organic nitrogen signaling and disease resistance. We could also visualize with Mapman that expressing *GLK2* early in ovary/ fruit development in a uniform background (MM) results in activation of those pathways, with photosynthesis being on top with 94 upregulated genes, 12 genes in tetrapyrrole synthesis class, 39 in redox, 49 in stress responses and 7 in Nitrogen metabolism (Supplementary Fig. 7).

It is interesting to note that in the downregulated list of genes, both sCEF genotypes display a remarkable alteration in cell wall remodeling related functions, identifying six xyloglucan-xyloglucosyl transferases genes with lower expression levels, from -3.8 to -2.6 of foldchange (Supplementary table 4 on CD).

2.3 Nuclear genes implicated in chloroplast retrograde signaling are DEG in CEF

Chloroplasts are implicated in many cellular roles and the regulation of its functions is coordinated by cross-talk between organelles, as well as nuclear signaling ²¹⁹. The mechanism through which nuclear genes are regulated by chloroplast signals is known as retrograde signaling ²²⁰ and 39 nuclear genes have been identified in retrograde-response regulation ^{221,222}.

Since it has been demonstrated that our CEF lines are enhanced in chloroplast functions, we analyzed expression levels of the tomato identified orthologues in this list of 39 nuclear genes ²²² implicated in retrograde signals at IMG stage (Supplementary table 5 on CD). Some of those were found significantly affected (P_{adj} <0.05) in our DEG list, implicating changes in Reactive Oxygen Species (ROS), Abscisic acid (ABA), sugar and auxins pathways.

A group of genes showed upregulation pattern (approx. two foldchange) as *Auxin responsive protein* (Solyc09g083280.2) *ABA-responsive protein-related-FIP1*, (Solyc04g008970.2), *Sugar/ inositol transporter* (Solyc02g079220.2) *Ubiquinol-Cytochrome-c* (Solyc10g048190.1), *Xylanase inhibitor* (Solyc01g080020.2.1), whereas others were displayed as downregulated: *MFS transporter glycerol-3-phosphate* (Solyc03g093140.2), *COSII-TF* implicated in initiation of pollen tube guidance (Solyc01g108910.2) and *UDP-glucose 4-epimerase* (Solyc08g080570.2).

Moreover, carotenoids have been recently proposed to be involved in a mechanism to regulate nuclear genes in response to oxidative stress. The singlet oxygen, which is an unstable ROS molecule, reacts oxidizing β -carotene and generating volatiles such β -cyclocitral ¹⁷⁹. Genes strongly induced by β -cyclocitral

have been previously studied and mainly characterized as 12 UDP-glycosyltransferases and 10 glutathione transferases (GST)²²³. We studied the expression of the tomato orthologues of those genes in our CEF (Supplementary table 6 on CD). Some of them are downregulated in our DEG lists as *UDP-sulfoquinovose* (Solyc10g085100.1), whereas we found others which showed higher expression, as in the case of 8 UDP-genes (PPadj<0.05). Levels of those genes presented similar expression in our *GLK2* lines than in the Lundquist (2013) study²²³.

Furthermore, GST genes were also found induced in sCEF and dCEF (Solyc06g009040.2) (PPadj<0.05, three foldchange), four previously described GST genes²²³ were listed in *GLK2* upregulated genes (Supplementary table 6 on CD).

2.4 Joint expression of *APRR2* and *GLK2* early in development results in further reinforcement of chlorophyll synthesis and photosynthesis machinery in the fruit

We described in chapter 1 that differences in “greenness” and chloroplast metabolite levels could be easily observed between single TFM5::*GLKs* and TFM5::*APRR2* lines on one side, and double TFM5::*GLK*::TFM5::*APRR2* genotypes on the other side; now we aimed to evaluate the molecular basis of the differences observed at the metabolite/ structural phenotypic level.

Subsets of statistically significant upregulated genes for *APRR2*, *GLK2*, *GLK1-APRR2* and *GLK2-APRR2* were compared and the largest numbers of DEGs (412) were clearly shared between *GLK2*, *GLK2-APRR2* and *GLK1-APRR2* lines. Only, 33 genes showed a common upregulated pattern also in *APRR2* (Fig. 3a). Common genes assessed by GO term enrichment analysis suggest that a substantial number of the common genes

shared by *GLK2* and *APRR2* are related to photosynthesis function, as well as generation of precursor metabolites and energy (Fig. 3b). Reviewing the annotation for each gene of the list is also highlighted the common upregulation in auxin responsive proteins, E3 ubiquitin-protein ligase complex, circadian clock associated protein and cell division (Supplementary table 7 on CD).

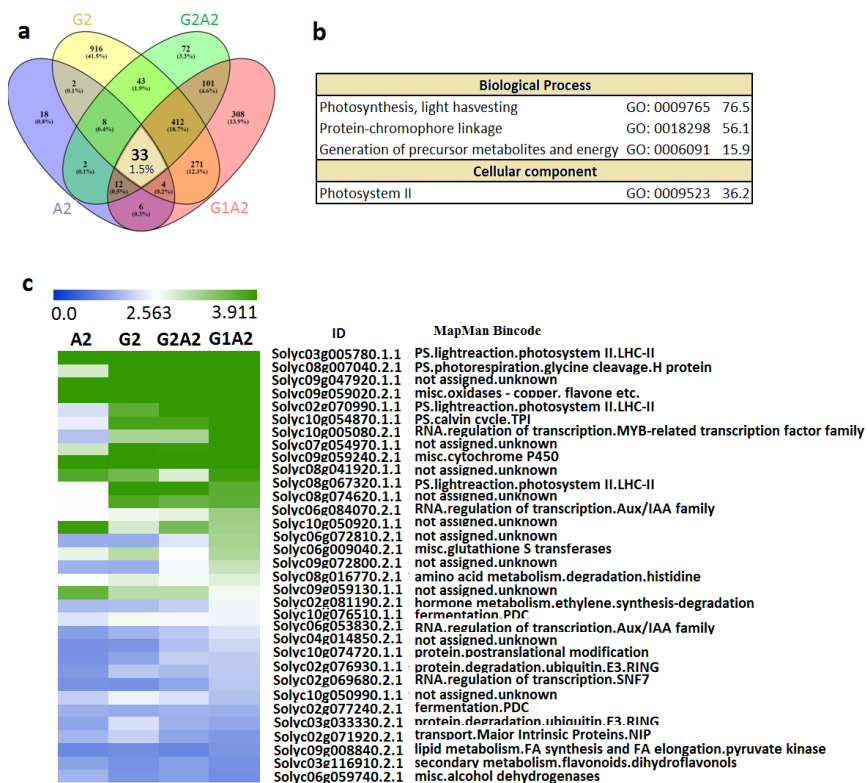


Figure 3: Common upregulated genes for all the sCEF genotypes display higher expression levels in GLKs::APRR2 plants. (a) Venn diagram with the upregulated list for *GLK2*, *APRR2*, *GLK2-APRR2* and *GLK1-APRR2* displays 33 genes with a common pattern among the four genotypes (PPadj < 0.01). (b) Fold enrichment of the 33 overlapping upregulated genes with GO terms categorized for Biological Process and Cellular Component (<http://geneontology.org/>) (c) Heat map with Foldchange (log2) data (from 1.2 to 5.9) for the 33 common genes in each construction (Supplementary table 7 on CD). Each gene ID is represented with the MapMan Bincode according to its function.

The Heat map in Fig 3c displays common upregulated genes (fold-change range from 1.2 to 5.9) and reveals that many genes showed higher expression levels in *GLKs-APRR2* lines than in the single ones. Gene expression levels suggest a boost of photosynthesis-related functions when both TFs are together.

Moreover, other genes increase in levels in dCEF more than they do in sCEF, such as in the case of auxin responsive proteins (Solyc06g053830.2 and Solyc06g084070.2) and one gene related to cell division (Solyc10g074720.1). These data emphasized a possible synergetic effect when both TFs are simultaneously expressing. Such hypothesis validates our previous fruit characterization through metabolomics analysis.

2.5 The boost in photosystem machinery in double engineered *GLK-APRR2* fruit is maximized when using *GLK1*

We previously reported (Chapter 1) that TFM5-driven expression of either *SIGLK1* or *SIGLK2* alone or in combination with *SIAPRR2* TF resulted in super dark green fruit, increasing chloroplast-related metabolites. Furthermore, the fruit color observed among the different CEF genotypes presents a wide range of green intensities. Darkest green fruit with higher chlorophyll and carotenoid levels at IMG belong to fruit from *GLK1-APRR2* families. To understand the molecular basis of this reinforced phenotype, as well as to identify additional molecular phenotypes and to analyze if they are functionally equivalent at the gene expression level, the *GLK1-APRR2* and *GLK2-APRR2* fruit transcriptome data at IMG stage was compared.

Most of *GLK2-APRR2* upregulated genes (81%) were shared with *GLK1-APRR2* (PPadj<0.05) (Fig. 4). However, this common list represents only a fraction of the differential genes in each case and only the 50% of those upregulated in *GLK1-APRR2*

genotype (1661 genes); using a more stringent cutoff criteria (P_{adj}<0.01) the percentage of shared genes between genotypes is similar (see upregulated genes table 8 on CD).

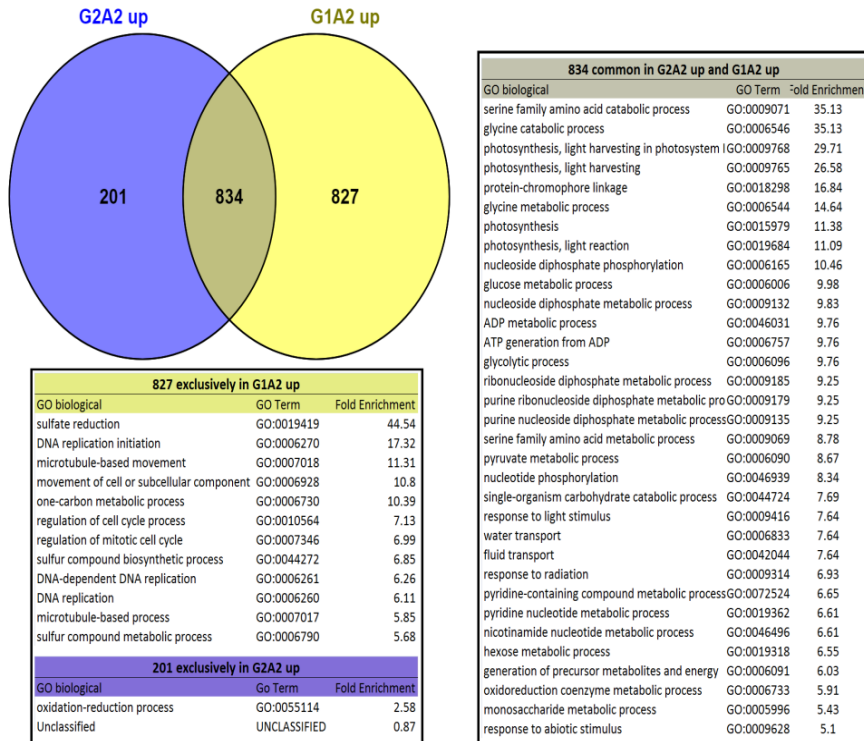


Figure 4: Upregulated genes in *GLK1-APRR2* and *GLK2-APRR2* are enriched in chloroplast related functions. Venn diagram (top-left) shows shared and unique upregulated genes for *GLK1-APRR2* and *GLK2-APRR2* genotypes, foldchange >1.5 (P_{adj}<0.05). Tables show GO term enrichment analysis for overlapping (brown) and exclusively upregulated genes: yellow (G1A2) and purple (G2A2). Terms are ranked by fold enrichment values, it was used a cutoff of 5% in yellow and brown tables.

Figure 4 shows an enrichment analysis of the common upregulated genes (834) in pTFM5::*GLK2*::pTFM5::*APRR2* and pTFM5::*GLK1*::pTFM5::*APRR2* lines. It is well known that chloroplasts provide to compounds involved in photorespiration such as serine and glycine to the mitochondria, and these

compounds are also exchanges with peroxisome in the characterized cross-talk between the tri-organelle-signalization¹⁷⁷. It is remarkable the effect of dCEF on serine and glycine metabolism, as well as photosynthesis function and chlorophyll biosynthesis.

Process involving glucose, ATP generation and glycolysis are also modified in both genotypes consistent with the chloroplast fortification described in our dCEF. *GLK1* and *GLK2* in combination with *APRR2* seem to affect chloroplast functions involved in metabolites biosynthesis and photosynthetic machinery.

Moreover, sulfur nutrient take part of the biological synthesis of many essential compounds such as cofactors or lipoic acids; e.g. the sulfolipid sulfolipid sulfolipid sulfolipid diacylglycerol (chloroplastid lipid). Sulfate reduction occurs in plastids since cysteine synthesis enzymes are found in plastids-mitochondria²²⁴. We examined whether *GLK1* combined with *APRR2* could enhance different functions or biological process than *GLK2*. In the list of exclusive genes for *GLK1-APRR2* genotype (Fig. 4) it is highlighted the enriched terms related to sulfur compound and sulfate reduction.

Focusing on genes upregulated only in *GLK2-APRR2* genotype (201 genes), oxidation-reduction process is upregulated compared with *GLK1-APRR2* fruits. Possibly the activation of chloroplast and mitochondria electron transport chains resulted in an enhancement of oxidation-reduction process (Supplementary table 8 on CD).

In summary, it is important to note that we described a higher number of upregulated genes when *APRR2* is combined with *GLK1* than when combined with *GLK2*; however, in both cases we observed shared changes related to chloroplast functions

(photosynthesis, chlorophyll, NO_3 , SO_4 , Calvin cycle, terpenoid and early steps in flavonoid biosynthesis among others).

2.6 Early expression of *GLK2* and *APRR2* still has a quantitatively large effect on fruit transcriptome at MG stage

According to TFM5 promoter pattern expression described in the previous chapter, green fruit stages showed the highest transgene expression level speaking at flower bud stage and decrease progressively until MG stage. Consistent with that, the total number of DEG in MG fruit decrease in all the modified tomato plants (*GLK1-APRR2*, *GLK2-APRR2* and *GLK2*). To our surprise this was not the case for the *APRR2* engineered lines, where the largest number of DEGs is found in MG (Supplementary Fig. 6).

In TFM5::*APRR2* fruit the ectopic and asynchronous expression of *APRR2* transgene occurs at a developmental stage when the endogenous *SIAPRR2* is still low expressed, and therefore our construct "extends" the pattern of expression anticipating it to stages before ripening. It is possible that this forward extension pattern of expression is responsible for producing the unexpected phenotype, *APRR2* could be recruiting genes normally activated by *GLK2* but not in MM (which lacks functional *GLK2*).

A total of 122 DEG were found in pTFM5-*APRR2* fruit at MG stage (listed in table 9 on CD). Figure 5 represents functions affected by the overexpression of *APRR2* in a MM *GLK2*-deficient background (*u/u*). Biological functions ranged from lipid, isoprenoid, major CHO metabolim, glycolysis, fermentation and tetrapyrrole metabolism genes.

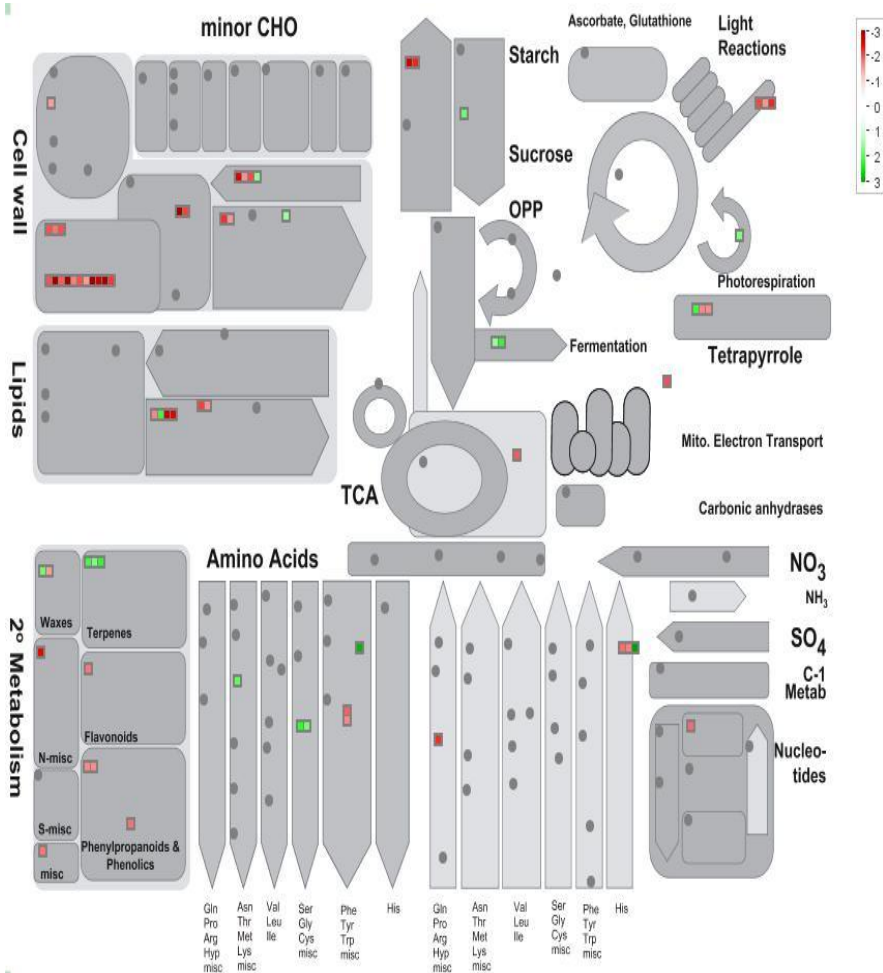


Figure 5: Ectopic expression of *APRR2* controlled by *TFM5* promoter affects diverse pathways at MG stage of fruit development. Snapshot of Metabolism Overview Panel was generated using MapMan software mapped with ITAG2.3 for *APRR2* DEG data at MG stage (See Table 9 on CD). Green means upregulated and red downregulated (PPadj < 0.05).

In addition, to assess the mechanism by which the early ectopic expression of *APRR2* exerts their primary influence in fruit at MG stage, we analyzed the expression pattern of those 122 DEG during normal fruit development. Data analysis suggests that *TFM5::APRR2* fruit are affected in the expression of genes with

ripening expression profile. Gene showed as upregulated by *APRR2* presented a more sustained expression between MG (almost all of them) and breaker stage under their own promoter (<http://gbf.toulouse.inra.fr/tomexpress>). There was an exception in this list, a phenylpropanoid gene (Solyc08g075210.1) which usually presents the highest levels at IMG stage.

Similarly the large majority of downregulated genes (76 genes) also showed an expression peak from MG to breaker (either up or down peak in wild type plants) suggesting that OE of *APRR2* at MG stage affects mainly genes which present a ripening expression pattern.

GLK2 lines display lower number of DEGs than *APRR2* at MG stage. However, several functions still remain selectively affected. Including the major CHO metabolism, where is shown the upregulation of genes associated to starch synthesis (foldchange 3.5) and sucrose metabolism (foldchange 2.73). In addition, the activation of photorespiration and aminoacid synthesis with foldchange values over 4 are highlighted.

Cell wall function is affected (as in IMG stage) by the downregulation of certain genes involved in the cell wall-modification, including for instance pectin-esterases and degradation. Furthermore, we found common changes when DEG of *GLK2-APRR2* and *GLK2* lines at MG stage were compared; with similar type of functions affected and ranging from starch synthesis and sucrose metabolism.

Affected functions In *GLK2-APRR2* at MG stage involve minor CHO pathway, which was also modified in *APRR2* lines (cell wall degradation and lipid metabolism) or upregulation of fermentation process (pyruvate decarboxylase), also higher in either pTFM5::*APRR2* or pTFM5::*GLK2* plants. Figure 6 shows changes in *GLK2-APRR2* lines at MG stage. The most remarkable differences between dCEF and sCEF at MG stage were

associated to secondary metabolism and photosynthesis (Fig. 6a-b). For instance, transcript levels for genes in PS light-reaction and Calvin cycle declined in MG fruit in contrast to IMG stage, where both pathways were extremely active. In summary, the largest group of upregulated genes at IMG stage declined at MG stage mainly in dCEF.

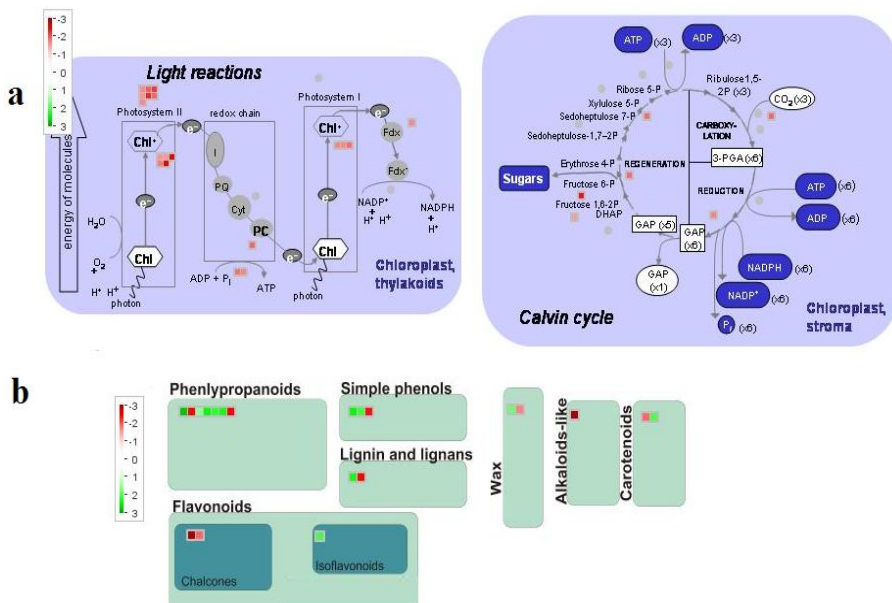


Figure 6: Light reaction, Calvin Cycle and Phenylpropanoids are the most affected pathways in *GLK2:APRR2* genotype at MG Stage. Snapshots were generated using MapMan software to map with ITAG2.3 for *GLK2:APRR2* DEG data at MG stage (see M&M). Green means higher expression and red lower expression (PPadj<0.05) (a) Represents Photosynthesis pathways (b) Secondary metabolism affected pathways.

2.7 Fruit chloroplast enhancement in *GLK1-APRR2* lines results in a higher total protein content

To study how much of the transcriptomic program, induced by providing high levels of active *GLK2* and *APRR2* at early

stages of development, ended up with observable modification at the protein level we carried out a proteomic analysis.

We first made an estimation of total protein content, using phenol extraction of fruit pericarp of each CEF genotype and WT, as well as of leaf samples as a reference. Figure 7 shows that *GLK2-APRR2* and *GLK2* fruit accumulate higher levels of total protein content (approx. 2.3 and 2 μg protein/mg-tissue respectively) and those are values closer to the leaf samples (4 μg protein/mg-tissue). Moreover, *GLK2-APRR2* and *GLK2* levels represent 400% the protein level present in *APRR2* and MM-WT fruit.

An ensuing proteomic analysis (see M&M) was able to detect more than 2500 proteins and further analysis were conducted to identify peptides and proteins differentially accumulating in the fruit pericarp of *GLK2*, *APRR2*, *GLK2-APRR2* and MM-WT.

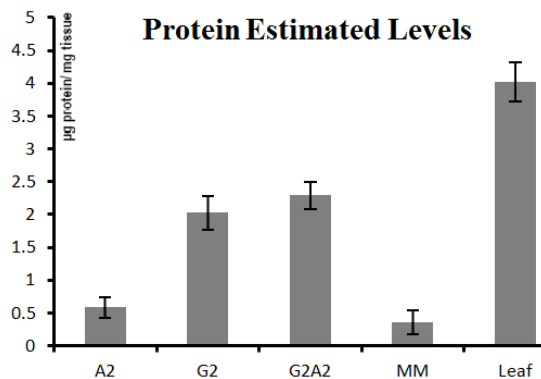


Figure 7: Fruits from *GLK2-APRR2* and *GLK2* lines contained increased protein content Quantification of protein content determined by modified Bradford protocol (see M&M) phenol extracts and expressed as μg total protein /mg fresh weight.

We initially focused on comparing the proteomic changes occurring in the fruit using MM leaves, as a control of typical photosynthetically active organ (Supplementary table 10 on CD).

Proteomic data was sorted in four classes: (1) proteins identified only in fruit samples, (2) upregulated in fruit, (3) identified only in leaves; and (4) upregulated in leaves (Fig. 8a).

GLK2-APRR2 presented the highest number of proteins that were exclusively found in fruit. This protein fraction is represented in Fig. 8b, where a cellular component enrichment analysis shows that the chloroplast is the most represented (Tylakoid lumen, 22.5%). This data suggests that most part of the exclusive proteins for *GLK2-APRR2* fruit (which are not present in leaves or MM fruits) belong to fruit chloroplasts protein isoforms.

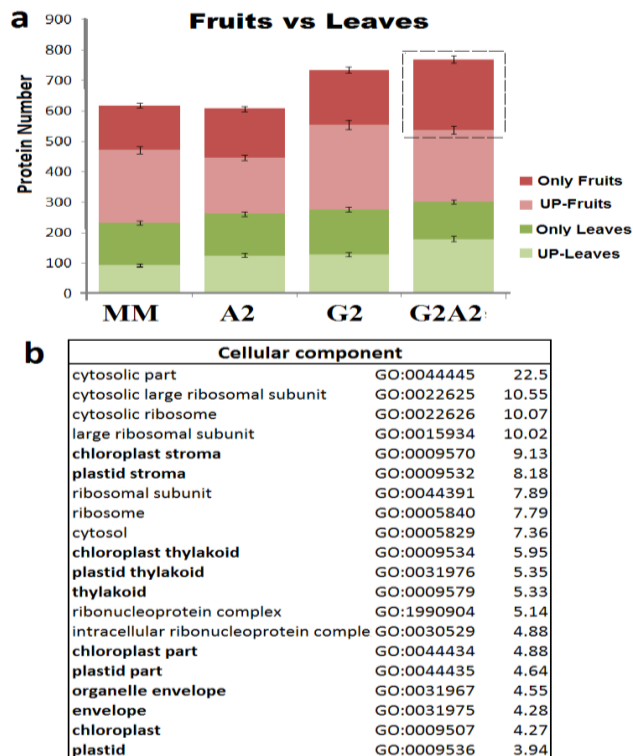


Figure 8: *GLK2::APRR2* fruits show higher number of identified proteins which are not found in leaves. (a) Fruit proteome was compared with that of leaf, subsequently several proteins were identified to be upregulated (red color, Up-fruits) or to be exclusive for each fruit genotype (continued)

Figure 8 (continued): (Pink color, only-Fruits) or in leaves (dark green color: Only-leaves and light green color: Up-Leaves). **(b)** Table represents the fold enrichment for cellular component category in the list of identified proteins exclusive for *GLK2-APRR2* fruits (protein fraction delimited by a box in the bar graphic). Chloroplast enriched parts are in bold (<http://geneontology.org/>).

Moreover, since photosynthetic genes are the most highly expressed in leaves, and their expression levels are drastically downregulated in non-photosynthetic tissues^{14,87}, we examined those differences at the GO level between fruit and leaves in MM and CEF genotypes. Fruit from MM-WT plant present higher differences compared to the leaf proteome than CEF samples. Figure 9 shows how those differences that we could name “typically photosynthetic proteome” decreased in a manner that is consistent with the strength of the fruit phenotype as described in chapter 1 (*APRR2* > *GLK2* > *GLK2-APRR2*).

A GO terms analysis revealed that the most downregulated functions in MM fruit compared with leaves are related to photosynthesis, monosaccharide metabolism, carbohydrates and protein-chromophore; differences which decrease in CEF. For example, light photosynthesis reaction proteins which present 32.6% of downregulation protein set in *MM*, is 29.8 in *APRR2*, 26.6 in *GLK2* and differences decrease to 15.3% in *GLK-APRR2*. In other categories is even more evident, e.g. LHC PS I ranged from the 94.1% of downregulation in MM fruits to the 35.3% presented in *GLK2-APRR2* fruit.

In summary, here we described that not only the upregulated proteins, but also the downregulated set (derived from the fruit vs leaves comparison), support the contribution of chloroplast fortification protein module to the increase observed in the total protein amount in *GLK2* and *GLK2::APRR2*.

Differences between Fruits and Leaves



Figure 9: *GLK2-APRR2* fruits display the most leaf-like protein profile. Fruits were compared with leaves and sets of downregulated proteins in fruit vs leaf were identified. Enrichment analysis of downregulated protein lists classified proteins in specific GO terms for MM, *APRR2* (A2), *GLK2* (G2) and *GLK2-APRR2* (G2A2) fruits at IMG. Fold enrichment values for the different GO terms (<http://geneontology.org/page/go-enrichment-analysis>) are represented in the heatmap. Green means higher enrichment values and red lower values (from 80% to 0%). Gray color means no identified class in the enrichment list.

2.8 Proteomics analysis of fortified fruit revealed that mitochondria and chloroplast are the most affected cellular components

A comparison between proteome profile of *APRR2*, *GLK2*, *GLK2-APRR2* at IMG stage with MM-fruit control resulted in two interesting sets of Differential Expression Proteins (DEP)

(Supplementary table 11 on CD). First, a set of upregulated proteins was identified in *GLK2* (131 proteins), *APRR2* (90) and *GLK2-APRR2* (51) fruit. Fifteen of those proteins are shared between *GLK2* and *GLK2-APRR2* but none overlaps with *APRR2* (Fig. 10a, left) indicating the existence of dramatic differences between the proteome in this type of fruit. Second, a set of proteins characteristic of the CEF genotypes, and no present in wild type MM, was defined. A Venn diagram shows the overlapping of exclusive proteins for CEF samples (highlighted in orange) (Fig. 10b, left). To learn more about these two set of proteins a functional analysis followed by gene ontology assessment was conducted.

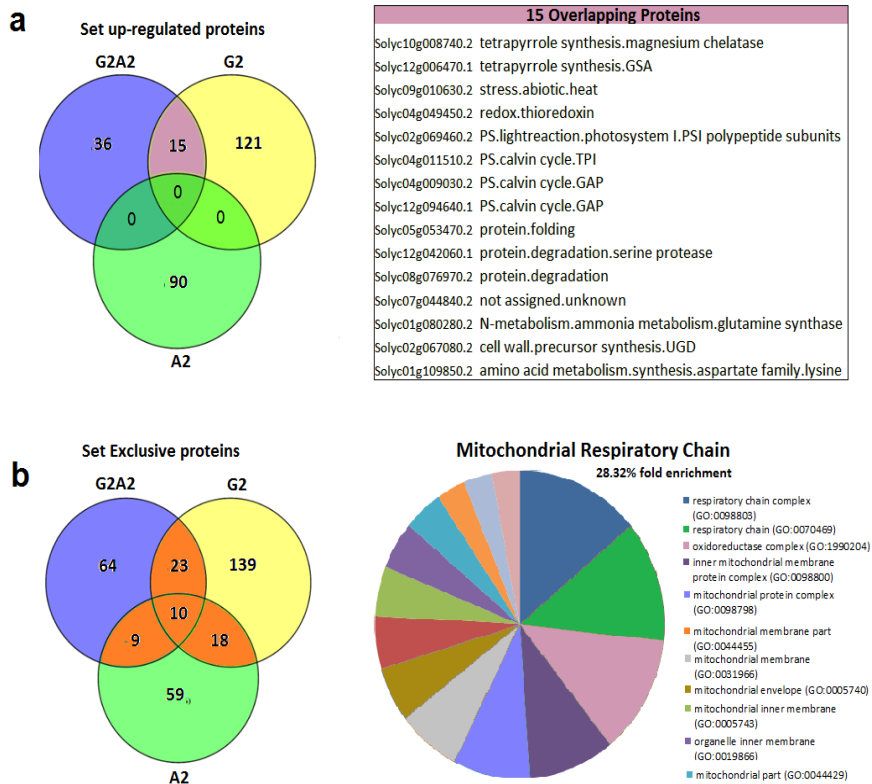


Figure 10: Exclusive or upregulated proteome in sCEF (*APRR2*, *GLK2* and *GLK2::APRR2*) when compared to MM fruits (continued).

Figure 10 (continued): (a) Venn diagram of upregulated proteins either unique to or shared between *GLK2*, *APRR2* and *SIGLK2::APRR2* compared with MM-WT at immature green fruit stage (left). Gene list of the 15 overlapping proteins found in *GLK2* and *GLK2::APRR2* (right). **(b)** Venn diagram of proteins identified in sCEF and no detected in MM-control (left) with overlapping proteins marked in orange. On the right categories affected in Mitochondrial Respiratory chain. GO enrichment analysis was performed in the list of exclusive proteins found in *GLK2* and 28.32% of the fold enrichment belongs to Mitochondrial Respiratory chain (Supplementary table 12 on CD).

The 15 overlapping upregulated proteins, according to MapMan Bins, are mainly related to chloroplast components (Fig. 10a, right). Moreover, we examined those lists by functional enrichment analysis (see table 12 on CD). GO terms show common changes in glucose (GO:0006006), hexose (GO:0019318) and monosaccharide metabolic process (GO:0005996) in the three CEF genotypes.

The set of upregulated proteins found only in *GLK2-APRR2* lines are involved in carbon-carbon lyase activity (GO:0016830) with an enrichment of chloroplast cellular components (GO:0009507). Proteins which characterized and differentiate *GLK2* fruit (121 proteins) affect several functions, which mainly compromise response to oxidative stress (GO:1900407), presenting an enrichment in mitochondrial (GO:0044429) and chloroplast (GO:0009534) organelles functions or composition. Meanwhile, proteins found in larger proportion in *APRR2* genotype relative to WT are associated to cellular amino acid metabolic process (GO:0006520), as well as there is an enrichment up to 100% in ATP citrate synthase activity (GO:0003878) (molecular function), cellular function is modified remarkable in citrate lyase complex (GO:0009346) and aminoacyl-tRNA synthase multienzyme complex (GO:0017101).

We later focused on the functional analysis of the set of exclusive proteins for each genotype. *GLK2-APRR2* exclusive proteins are related with cellular response to oxidative stress (GO:0034599), being overexpressed all the chloroplast

components (GO:0009570, GO:0009941, GO:0009526, GO:0009579, GO:0044445, GO:0009534). In the case of exclusive proteins found in *GLK2* lines, again pops up as over represented the oxidoreductase activity (GO:0016681, GO:0008121) and the class mitochondrial organelle with several subtypes specifying the compartments or functions of this organelle: e.g. mitochondrial respiratory chain (GO:0005746), inner mitochondrial membrane protein complex (GO:0098800), mitochondrial envelope (GO:0005740) and mitochondrial inner membrane (GO:0005743) (Fig. 10b, right).

All this support that chloroplasts, and to a lesser extent mitochondria, are the most affected cellular components in the cells of *GLK2* and *GLK2-APRR2* fruit.

2.9 Chloroplast biogenesis increases in CEF genotypes

We have just described sets of exclusive fruit proteins which are present in fruit from engineered lines but not in MM fruit. Here, we focus our attention on proteins which are detected in the fruit of at least two CEF engineered genotypes (Fig. 10b, orange area). Figure 11 shows a functional description, MapMan category and transit peptide for each overlapping protein (accumulating preferentially in at least two CEF). The clearest results are that the relative abundance of photosynthesis-chloroplast associated proteins increases in each case.

Gene ID	Annotation	Mapman Class
Solyc12g098150	Voltage-gated potassium channel beta subunit	hormone metabolism auxin induced-regulated-responsive-activated
Solyc12g043110	LETSW12 gene (Heat shock protein 4)	stress abiotic heat
Solyc12g009600	Thylakoid lumenal protein	not assigned no ontology
Solyc07g054210	Protochlorophyllide reductase like protein	tetrapyrrole synthesis protochlorophyllide reductase
Solyc07g005560	Eukaryotic translation initiation factor 5A	protein synthesis initiation
Solyc03g110960	Nucleoside diphosphate kinase	nucleotide metabolism phosphotransfer and pyrophosphatases nucleoside diphosphate kinase
Solyc10g086190	Adenosine kinase	nucleotide metabolism salvage nucleoside kinases adenosine kinase
Solyc09g075010	Prostaglandin E synthase 3	not assigned no ontology glycine rich proteins
Solyc09g008670	Threonine ammonia-lyase biosynthetic	aa metabolism synthesis branched chain group isoleucine specific threonine ammonia-lyase
Solyc07g055320	Protein FtsH homolog, ATP-dependent Zn protease cell division	protein degradation metalloprotease
Solyc06g083620	26S protease regulatory subunit 4	protein degradation ubiquitin proteasom
Solyc06g062840	Threonine synthase	aa metabolism synthesis aspartate family threonine threonine synthase
Solyc05g012270	Argininosuccinate synthase	aa metabolism synthesis glutamate family arginine argininosuccinate synthase
Solyc03g123610	Alanine aminotransferase	aa metabolism synthesis central aa metabolism alanine alanine aminotransferase
Solyc03g118910	T-complex protein 1 subunit delta	protein folding
Solyc01g006280	Formate-tetrahydrofolate ligase	C1-metabolism
Solyc04g071890	Peroxidase 4, Haem peroxidase, plant/fungal/bacterial	misc. peroxidases
Solyc07g006560	Hypersensitive response assisting protein	RNA processing ribonucleases
Solyc07g062130	Dtdp-glucose 4 6-dehydratase	cell wall precursor synthesis UDP-glucose 4,6-dehydratase
Solyc11g066410	S0S ribosomal protein L9	protein synthesis ribosomal protein prokaryotic chloroplast S0S subunit L9
Solyc09g082710	Histone H2A	DNA synthesis/chromatin structure histone
Solyc09g007940	Adenosine kinase	nucleotide metabolism salvage nucleoside kinases adenosine kinase
Solyc09g005740	Chloroplast lumen common family protein	not assigned no ontology
Solyc07g062870	30S ribosomal protein S20	protein synthesis ribosomal protein prokaryotic chloroplast 30S subunit S20
Solyc07g054290	Photosystem II family protein	PS lightreaction photosystem II PSII polypeptide subunits
Solyc07g041490	Stress responsive alpha-beta barrel domain protein	not assigned no ontology
Solyc06g050590	Acetyl-CoA carboxylase biotin carboxyl carrier protein	lipid metabolism FA synthesis and FA elongation Acetyl CoA Carboxylation
Solyc06g005710	cDNA clone 002-143-C11 full insert sequence	not assigned unknown
Solyc05g050600	Dynamine 2	misc. dynamin
Solyc05g018520	NADH-cytochrome b5 reductase-like protein	N-metabolism.nitrate metabolism.NR
Solyc04g074640	L-ascorbate peroxidase	redox ascorbate and glutathione ascorbate
Solyc02g091940	Fructokinase 3	major CHO metabolism degradation sucrose fructokinase
Solyc02g086830	Protease Do-like (S2 serine-type protease)	protein degradation serine protease
Solyc02g082000	30S ribosomal protein S19	protein synthesis ribosomal protein eukaryotic 40S subunit S15
Solyc01g108020	Thioredoxin m	redox thioredoxin
Solyc01g087040	Thylakoid lumenal 19 kDa protein, chloroplastic	PS.lightreaction photosystem II.PSII polypeptide subunits
Solyc03g098700	Cysteine Protease inhibitor 8	not assigned unknown
Solyc12g010060	Eukaryotic translation initiation factor 5A	protein synthesis initiation
Solyc07g042440	Alkyl hydroperoxide reductase/Thiol specific antioxidant family	redox peroxiredoxin
Solyc04g009540	S0S ribosomal protein L10	protein synthesis ribosomal protein prokaryotic chloroplast S0S subunit L10
Solyc07g005960	Serine carboxypeptidase K10B2.2	protein degradation serine protease

Figure 11: Common sets of photosynthesis-chloroplast associated proteins are differentially accumulated with respect to MM and shared between different CEF fruit. Functional information of overlapping proteins exclusive for sCEF (*APRR2*, *GLK2* and *GLK2::APRR2*) and not present in WT fruits (Figure 10b proteins empathized in orange). Genes implicated in PSII repair function are emphasized in bold. Three different color degrees characterized each kind of intersection; Light green means proteins shared between *GLK2* and *APRR2*, medium green represents *GLK2* and *GLK2::APRR2* and dark green means proteins shared by the three genotypes (*GLK2*, *APRR2*, *GLK2::APRR2*). Gene annotation used iTAG2.3 and Protein function is classified according to Mapman BIN code.

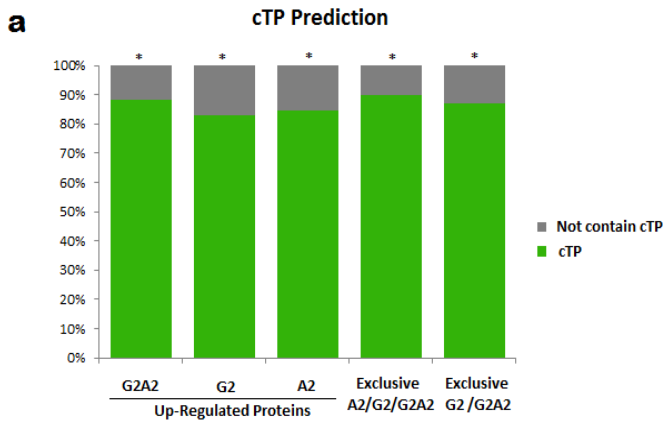
APRR2 and *GLK2* lines present 18 proteins that are specific for both fruit proteomes. One of them deserves special recognition because of its role in chloroplast, the ATP-dependent zinc

metalloprotease called *FTSH 2* homolog (Solyc07g055320.2.1) (Fig. 11). Furthermore, when protein exclusively present in *GLK2-APRR2* and *GLK2* fruit are analyzed is noteworthy the enrichment (100%) in photosystem II repair biological function (GO:0010206) related to PSII repair chloroplastic protein *psb27-h1* and chloroplastic protease do-like1.

Together, these data suggest a boost in chloroplast biogenesis in *APRR2*, *GLK2* and *GLK2-APRR2* lines. Not only because an increment in chloroplast parts and photosynthesis-chloroplast-associated proteins, but also for the presence of proteins implicated in the assembly of this highly complex PSII repair function and organelle division (*FTSH 2* homolog).

2.10 Largest differentially protein portion analyzed belong to chloroplast proteome

Plastidial targeting prediction software confirmed that the major proportion of DEP identified in CEF fruits belong to chloroplast proteome. The prediction was done with ChloroP program ([/www.cbs.dtu.dk/services/TargetP](http://www.cbs.dtu.dk/services/TargetP)), which it has been previously demonstrated to be reliable software ²²⁵. The chloroplast database reveals a range between 85-89% nucleus-encoded proteins, which present the pertinent transit peptide sequence (Fig.12a). Since 37 proteins did not show the predicted peptide signal, we analyzed those proteins to assess the subcellular localization by uniprot ID. As a result, 5 of them belong to cytoplasm, 4 to nucleus, 10 to cytosol, 2 to Endoplasmatic Reticulum, 1 is from glyoxisome and 5 unknown. The remaining 10 are also chloroplastic and mitochondrial proteins and are listed in Fig. 12b.



b

ID	Uniprot	Protein	Subcellular Localization
Solyc01g100360	IPR006258	Dihydropolipoamide dehydrogenase	mitochondrial
Solyc05g009530	IPR001078	2-oxoacid dehydrogenase acyltransferase, catalytic domain	mitochondrial
Solyc12g056230	IPR000889	Glutathione peroxidase	mitochondrial
Solyc12g056240	IPR000889	Glutathione peroxidase	mitochondrial
Solyc02g090030	IPR002628	Photosystem II manganese-stabilizing protein PsbO	chloroplast
Solyc03g112360	IPR001196	Ribosomal protein L15	chloroplast
Solyc06g071720	IPR001196	Ribosomal protein L15	chloroplast
Solyc07g065170	IPR001047	Ribosomal protein S8e	chloroplast
Solyc12g056830	IPR000711	ATPase, F1 complex, OSCP/delta subunit	chloroplast
Solyc10g006070	IPR001047	Ribosomal protein S8e	chloroplast

Figure 12: Most part of upregulated proteins in CEF have chloroplast transit peptides (cTP): (a) DEPs for CEF were screened by the ChloroP server (chloroplast transit peptides (cTP) predictor) (<http://www.cbs.dtu.dk/services/ChloroP/>). *ChloroP server demonstrated to has 88% of confidence ²²⁵. List of proteins comes from Venn Diagrams of Figure 10a-b: UP G2A2 (51 proteins upregulated in *GLK2-APRR2* genotype related to WT), UP G2 (38 proteins upregulated in *GLK2* related to WT), UP A2 (90 proteins upregulated in *APRR2* related to WT), Exclusive fortified fruits (10 identified proteins in *APRR2*, *GLK2*, *GLK2-APRR2* and not in WT) and Exclusive G2 and G2A2 (23 identified proteins in *GLK2* and *GLK2-APRR2* engineered fruits and not in WT). **(b)** Proteins without chloroplast signal peptide which present chloroplast or mitochondrial subcellular localization (<http://www.uniprot.org/>).

Next, we evaluated whether some of the identified DEP are plastid-encoded genes, founding that *GLK2-APRR2* fruit-proteome presents 50S ribosomal protein L14 chloroplast-encoded. This protein is upregulated in this genotype when is compared with MM fruit or leaves proteome. In addition, two

more chloroplast encoded photosynthesis light reaction proteins were identified: *atpB* detected exclusively in *GLK2-APRR2* fruit proteome, and *atpA* presented in higher levels in *GLK2* genotype.

2.11 Relationship between proteome and transcriptome

When up and downregulated protein levels, for different genotypes, are compared with transcriptomic data a low correlation is presented (data not shown). However, we could see across the CEF lines that changes in protein levels are generally higher than the corresponding changes in transcripts.

Figure 13 displays a histogram of the ratio protein/ transcript, where is possible to observe that the most common ratio ranges from 2.6 to 4.4 values, representing always higher protein levels than transcript.

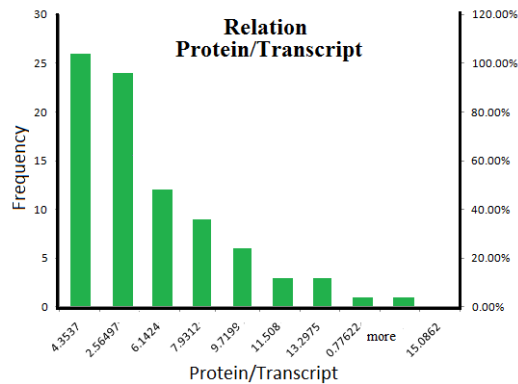


Figure 13: Higher Protein to transcript ratios for specific DEG in *GLK2* fruits. Transcript levels as from RNAseq data: PPadj<0.05 and protein levels from Proteomics analysis: p<0.01 (see M&M).

2.12 Over-represented motifs in common upregulated genes

Here, we wanted to examine whether the upstream regions of common DEG in CEF are sharing common *cis*-elements. A list of 84 upregulated genes ($P_{\text{Adj}} < 0.05$) found in the four engineered tomato lines (*APRR2*, *GLK2*, *GLK2-APRR2* and *GLK1-APRR2*) was analyzed (Fig. 14a). Two popular software programs, based in different consensus-method, to searching for conserved motifs were applied. MEME (Multiple Em for Motif Elicitation) and DREME (Discriminative Regular Expression Motif Elicitation) were used to identify overrepresented motifs in the 2000 pb 5'upstream sequences of these 84 genes. MEME evaluates the frequency of any motive in the input list while DREME compares motives which are over-expressed related to tomato genome (<http://meme-suite.org/>).

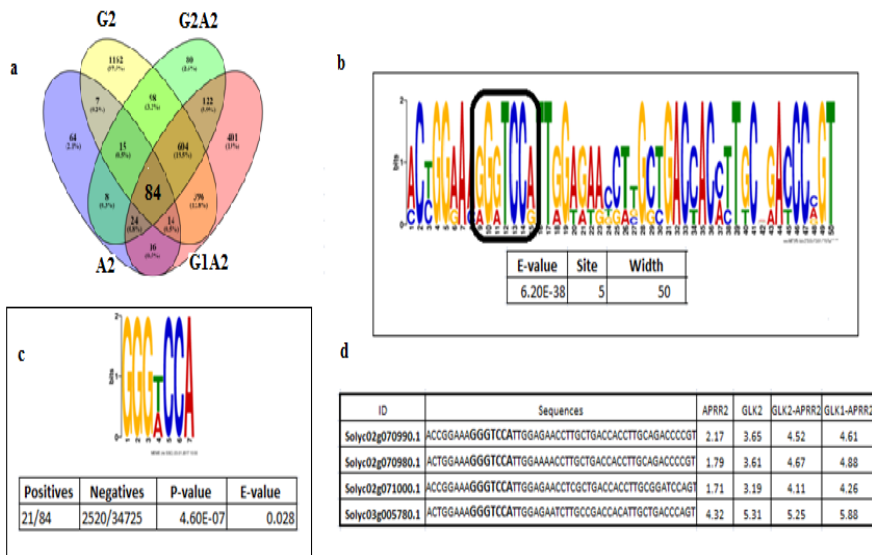


Figure 14: Over-represented motifs present in the 5'upstream regions of upregulated common set of genes. Analysis was done using MEME and DREME software and transcriptome data (continued).

Figure 14 (continued): (a) Venn Diagram with upregulated genes (PPadj<0.05) for each genotype (*APRR2*, *GLK2*, *GLK1-APRR2*, *GLK2-APRR2*) (b) DREME output of over-expressed motifs in 84 common UP regulated genes (c) One of the MEME outputs for the list of 84 genes using 2000pb upstream for the analysis. The E-value is an estimate of the expected number of motifs with the given log likelihood ratio. (d) RNAseq expression data corresponding to the four chlorophyll a/b genes found with GGGwCCA motive according to MEME.

DREME did report a significant well conserved 7pb motif of that was found in 21 of the 84 input genes (Fig.14b). It is highlighted that one of the significantly overrepresented MEME outputs indicates the same motif inside a sequence width of 50pb (Fig 14c). Five sites were identified by MEME, and all of them are chlorophyll a/b binding domain genes, being the 7pb motive presented in four of them without any offsets. Even more interesting is that the putative TF binding site is duplicated in the three genes *Cab* localized in Chr2. Therefore, in those genes both TFs could be interacting with the regulatory cis-elements at the same time elevating even more the gene expression by an additive or synergetic effect.

This hypothesis is supported by transcriptome data, which shows higher expression for both dCEF lines compared with sCEF (Fig. 14d).

3. DISCUSSION

Chloroplast development is a complex process that integrates developmental and environmental signals. Although we know a good deal about it^{177,206}, its molecular-genetic basis is still not completely understood. The tomato-plastome is composed of approximately 114 genes²²⁶ classified as photosynthesis-related genes. However, most of the genetic information (93-99%) to build and operate a chloroplast have been transferred to the

nucleus^{209, 227}. In addition to the environmental and developmental cues, organelle functions and composition are regulated by both retrograde (organelle to nucleus) and anterograde (nucleus to organelle) signals¹⁷⁷. Our CEF plants reveal that the multiple effects produced when *GLK* and *APRR2* factors expressed early in fruit development are derived from the complexity of the chloroplast development program in the fruit.

3.1 GLKs and APRR2 affecting an array of chloroplast biogenesis and chloroplast functions including retrograde signaling program

Among the number of nucleus-encoded transcriptional factors that have been identified to influence chloroplast development and chlorophyll content^{228,229}, only *GLK* and *APRR2* proteins are associated to a major quantitative trait locus (QTL). *CaGLK2* was identified as the gene governing green fruit phenotype *pc10*^{154,230}, whereas *APRR2* has been postulated as a candidate for *pc8.1* in pepper and later confirmed^{8,154}.

Overexpressing either *GLKs* or *APRR2* early in fruit development, resulted in immature tomato fruits with higher chlorophyll levels and higher chloroplast number^{5,8}. We have described in chapter 1 that our four CEF genotypes presented higher isoprenoid/ chlorophyll content during green stage of development, as well as an increased number and size of chloroplasts (more evident in dCEF). Consistently with those changes, in this chapter we showed the upregulation of genes involved in tetrapyrrole, chlorophyll biosynthesis processes, photosynthesis-related and Calvin cycle; as well as the increase levels of PEP protein in CEF samples.

Moreover, several reports revealed additional roles associated to those TFs besides those related to chloroplast/ photosynthesis functions. Since *GLKs* have been described to coordinate a retrograde regulation chloroplast-to-nucleus of process related to nitrogen signaling, disease resistance and redox state^{50,78,231-233}, we investigated genes activated in the fruit by overexpressing *GLK2* at IMG stage. In the line of those previous retrograde regulation studies we found 7 genes implicated in inorganic nitrogen signaling, 39 in redox function and 49 in stress responses (P_{adj}<0.01). On the other hand, *APRR2* has been described to affect genes related to ripening as *LeMADS-RIN*, *ACO*, and *PG*, being its main role associated to mature stage of fruit development⁸.

Moreover, both *GLKs* and *APRR2* have been described to be involved in circadian clocks that regulate several process such as flowering time, photosynthesis, carbon fixation or stress response^{50,217}. We also showed that one of the 33 common upregulated genes for the four CEF genotypes is a protein associate to circadian clock (*Solyc10g005080.2*). In addition, we described other function associated to the early OE of both TFs, which has not been previously described: Cell Wall Remodeling (CWR).

CWR is affected in CEF by the downregulation of 6 genes which encode xyloglucan-xyloglucosyl-transferases (changes ranging from -2.6 to -3.8 foldchange). Cellulose-xyloglucan matrix²³⁴ appears to be involved in cell expansion although the enzymes implicated and the regulation network of this matrix have not been characterized in detail²³⁵.

Our hypothesis that *APRR2* and *GLK2* affect CWR is supported by previous works because the retrograde signaling has been related to have effect on cell wall function. Arabidopsis mutant *ise (1-2)* is plastid defective and shows to be altered in cell wall synthesis and modification²³⁶. Indeed, Burch-Smith (2011)

showed 8 endotransglucosylase/ hydrolase (XTH) enzymes in the list of DEG. Almost all of them were upregulated as consequence of the chloroplast disruption²³⁶. In addition, within the 39 nuclear gene-set identified as responsible for retrograde signals²²² three of them are cell wall related (2 Pectinesterases and 1 xylanase enzyme). We also found that this xylanase, which is among the 39 gene-set, is in the DEG list of *GLK2* lines (PPadj< 0.01).

Moreover, it has been recently described that β -cyclocitral is related with oxidative stress in PSII and the fact that this molecule acts as signal from chloroplast to nucleus results in a readjustment and better PSII quantum²²³. More chloroplasts could be associated to induce β -cyclocitral signal pathway for keeping the oxidative stress balance in the photosystem. Previous work on Arabidopsis studied the *abc1k1* and *abc1k3* mutants, characterized by rapid chlorosis and higher β -cyclocitral accumulation upon high light stress²²³. Our CEF lines could be affected in CWR through the alteration in cross-talk regulation between chloroplast-nucleus as in the case of *abck* mutants.

More nuclear genes in the list of the 39 involved in retrograde signal²²² were found in our DEGs data and some are implicated in ROS, ABA, sugar and auxins pathway (PPadj<0.05). In summary, our analysis shows that not only *GLK* genes presented differential responses indicative of differences in plastid retrograde signals (as Waters et al., 2009) but also *APRR2* ectopic expression participate in this regulation process.

Chloroplasts produce a plethora of compounds associated to many functions as we described above, which are not only coordinate by nucleus response but also with other organelles. Redox signals are coordinated by the communication between chloroplast-mitochondria and peroxisomes²¹⁹. We demonstrated among chapter one that many metabolites associated to mitochondria organelle are also modified in CEF, such as serine and glycine which are considered heme cofactor

precursors (Woodson and Chory 2008). Compounds essential for the mitochondria to develop oxygen transport, respiration, signal transduction, and detoxification process (Azuma, Kabe et al. 2008). Among chapter two we demonstrated that mitochondria and chloroplast are the most affected cellular components in CEF cells with a fold enrichment of those GO terms.

3.2 Additive effect resulting from joint expression of *GLK* and *APRR2* early in fruit development and the possible underlying mechanism

Although the effect of constitutive expression of *GLKs* and *APRR2* has been reported, the use of an early specific fruit promoter to control de expression, as well as the effect to combine both TFs in the same construction has never been studied before. In chapter 1 we reported a novel additive effect when both *GLKs* and *APRR2* are expressed simultaneously. Hypothesis which was supported by the phenotype and metabolite profile in immature stage, as well as the positive quality impact found in red ripe *GLKs-APRR2* fruit. Here, we aimed to integrate the RNA and protein molecular information obtained in order to develop a comprehensive model that explains the additive effect previously described.

We have shown through transcriptome data at IMG stage that genes described as photosynthesis-related present, generally, higher expression levels in *GLKs-APRR2* lines compared to sCEF lines. Process such as chlorophyll metabolism, chlorophyll a/b binding proteins, PSII and Calvin cycle enzymes are even more enhanced in dCEF, indicating that the structural and small molecules data is supported by the protein and transcript program.

It has been suggested that *GLK* genes take part of a regulatory network which control the nuclear-photosynthetic-genes expression⁵⁰. *HY5* gene, which encodes a protein with bZIP motif, is identified as responsible for the regulation of fundamental developmental processes such as chloroplast development. This gene was described to modulate the signal transduction downstream in the pathways by auxin/ cytokinin signals²³⁷ and to acts with *GLKs*, coordinating the chloroplast biogenesis during root greening⁷¹. This network could be extended to fruits. The light signal transduction genes *LeCOP1LIKE*, *DET1* and *DDB1* were found to repress fruit pigmentation in tomato^{64,70,238}. Interestingly, the tomato mutants, *high pigment (hp)* are altered in chloroplast number and ultrastructure resulting in fruit with elevated levels of chlorophyll. Those genes, underlying the *hp* phenotype, are identified as *DDB1 (hp1)* and *DET1 (hp2)*^{64,70}.

Previous works demonstrated that the mechanism by which *DET1* and *DDB1* work is through their proteins take part of a proteasome complex²¹⁴, together with the E3 ligase (*CUL4*)²¹³. The complex regulates some factors relevant for plastid development, including chlorophyll accumulation. Recently the link between *DET1/DDB1 (hp)* and *GLK* has been established, where a physical interaction occurs between the *DET1/DDB1* ubiquitin-E3-ligase complex and the *GLK2* protein. This interaction requires two lysine residues present in *GLK2* protein that essential for the proteolytic activity of the complex⁶⁸. The *hp1-2* are, therefore, associated to mutations in elements implicated exclusively in the *CUL4*-proteasome complex, downregulation of *CUL4* also generates more plastid development⁶⁹. Although, the specific substrate for this degradation complex are still unknown, it has been proposed that *DDB1* targets *GLK2* for the degradation, and *DET1* protein may favor the association between *DDB1* and *GLK2*. The interaction of the last aa residues in N-terminal region of *DET1*

(26-87) with *GLK2* and the lack of this interaction in *hp-2* (*DET1* mutated) have further substantiated this model ⁶⁸.

This complex is probably involved in the degradation of other positive factors different to *GLK2* since Nguyen et al (2014) suggesting that *hp1* and *GLK2* act in independent pathways since overexpression of 35s:*GLK2* in *hp1* mutation background displays more chlorophyll levels than in *hp1* itself.

Even more interesting is the fact that *APRR2* gene was found in the list of DEG of *hp1* mutant plant. *hp1* fruit present higher expression levels of *APRR2* in 17 DPA than in breaker or red ripe fruit ⁷⁹.

It is possible that *APRR2* is also a substrate for CUL4-ligase complex and it would be worth to conduct experiments to confirm this contention. The joint targeting of both *GLK* and *APRR2* for ubiquitination and degradation by the proteasome could explain the synergistic effect observed when *SIAPRR2* and *SIGLK2* are co-expressed.

Interactions of the CUL4-DDB1 complex include *DDI1* (Solyc09g031610.2). The *DDI1* gene has been described to participate in salt, mannitol and UV-C responses through the ubiquitin-E3-ligase complex ²³⁹ and this gene is in the DEG list of our four CEF genotypes (PPadj<0.01).

Moreover, in this work to understand the mechanism underlying the described additive effect in chloroplast development, we studied over-represented motives in the promoter regions of common upregulated genes for sCEF and dCEF. A well conserved 7pb motif, GGGwCCA, was identified in tandem repetition in the 5' upstream regions of 21 of the 84 common upregulated genes. According to our result, both TFs could be functional on similar regulatory motives. The fact that the interaction could be additive, would be supported by the *Cab* genes that we found with duplicated binding domain sequences in the promoter region. Those genes showed higher expression when *GLKs* and *APRR2* were co-expressed. In

addition, there appears to be different affinity according to the TF, *APRR2* showed always lower expression levels than *GLK2* and both sCEF lower than dCEF. Future studies should be conducted to confirm those hypotheses, e.g. transactivation assays in *Cab* genes to demonstrate if both TFs bind to the identified sequence element.

Despite the fact that *SIAPRR2-like* has been associated to ripening it is also obvious its involvement in fruit plastid development from our results and others⁸. This is not restricted to Solanaceae, since a loss of function in *APRR2* has been found associated with the production of white fruit in cucumber⁸⁰, similarly to what was previously reported in pepper⁸. Fruit with *APRR2* loss of function lines showed fewer chloroplasts and lower chlorophyll accumulation at all stages of fruit development⁸⁰. The necessary presence of a functional *APRR2* gene in cucumber and pepper for correct chloroplast development, as well as the boost of fruit chloroplast development achieved in tomato when *SIAPRR2* is expressed early in fruit development suggest that *SIAPRR2* gene could have changed during tomato domestication. It will be interesting to study the pattern of expression of the *SIAPRR2* and check whether or not the *SIAPRR2* promoter could have lost an early fruit regulatory domain that may still be present in other species.

3.3 Increased in plastid development is well-correlated with the upregulation of genes and proteins process-associated

High pigment mutants are affected in many interrelated processes, ranging from light response and transduction signaling⁷⁰ to plastid development; increased fruit chloroplast size¹⁴⁵ and number⁶². Moreover, constitutive OE of *APRR2* or *GLK2* also contributed to chloroplast division and expansion^{5,8}.

The hypothesis by which silencing either, *DET1* or *DDB1* increases plastid number and size, was not clearly defined. It has been reported that *hp1* increases fivefold the DNA ratio plastome:genomic in hypocotyls⁶³, this fact was used as an explanation for the increase in plastid size¹⁴⁵. Another suggestion was the regulation of plastid development through cytokinin and photomorphogenic signal transduction^{64,145}. Here, in our sCEF effect on chloroplast number and size was described but lower than dCEF. Since dCEF plants are in the line with previous results, showing fruit with more dramatic changes and achieving over threefold increase in chloroplast number and size.

Upregulation of genes and proteins related to photosynthesis and chloroplast biogenesis correlate with the significantly increased in chloroplast development in IMG fruit of *GLKs-APRR2* lines, which could clarify the molecular mechanism underlying this change. In addition, we described among this chapter genes implicated in plastid division function affected in CEF.

FtsHs are a well-characterized family of membrane bound proteases important in chloroplast biogenesis and linked to the degradation of unassembled thylakoid proteins²⁴⁰. Meanwhile, those proteins play crucial role in PSII reaction center protein D1 during PSII repair^{241,242} some of those proteins were identified as DEG in our CEF lines (for example PSII repair Solyc07g054290.1 with 1.5 of foldchange).

Although previous studies have demonstrated that *AtGLKs* OE upregulate genes directly involved in photosynthesis, mainly Lhcb proteins^{50,141}, the proteomic changes occurring have not been characterized yet. Our proteomic analysis support that chloroplast biogenesis process is upregulated in our lines. We have shown an upregulation of ATP-dependent zinc metalloprotease *FTSH 2* homolog, as well as another two

chloroplastic proteins related to PSII repair (*psb27-h1* and chloroplastic protease *do-like1*). Those analyses also revealed that CEF have higher levels of total protein content than untransformed plants, reaching in *GLK2-APRR2* and *GLK2* fruit the highest amounts (about half of the total in developing leaves). More than 50% of total proteins in leaves are photosynthesis related, mainly represented by *RuBP* carboxylase (20-30% among species) and protein in the thylakoid membranes of the chloroplast (20-40%, being 60-85% pigment-protein) ²⁴³. Here, we confirmed through DEP and proteomic comparative analysis that the higher amounts of total proteins found in CEF are related to the positive impact in chloroplast development.

3.4 Fruit photosynthesis enhancement supported by upregulation of carbon fixation enzymes

As photosynthesis is the most conspicuous function of chloroplast, it was our interest to know the degree of change in photosynthesis related genes and proteins which is taking place in the plastid proteome of CEF. Early reports have provided novel and extensive information on the plastid proteome in plants ²⁴⁴ with strong emphasis on the chloroplast, where about 2000 proteins were identified in recent tomato papers. A range of 85-89% of DEPs present a chloroplast transit peptide and many of the remaining proteins are known to be chloroplast proteins. Not only proteins of light-harvesting complex, but also *RuBisCO* which is responsible for carbon fixation increased in CEF. However, whilst in leaves *RuBisCO* plays the most important role in photosynthesis through the assimilation of atmospheric CO_2 , in fruits re-fixation of CO_2 via malate by Phosphoenolpyruvate carboxylase (PEP) has been hypothesized as an important operating mechanism ²¹. *APRR2*, *GLK2* and *GLK2-APRR2* fruit presented higher PEP protein amounts than

leaves and MM fruit, showing dCEF the highest foldchange followed by *GLK2*.

The increase in levels of photosynthesis enzymes observed in proteomic analysis is in addition supported by transcriptomic data. Ribulose biphosphate carboxylase (Soly02g085950.2) and RuBisCO (Soly09g011080.2) gene expression was shown as upregulated in *GLK2*, *GLK2-APRR2* and *GLK1-APRR2* IMG fruit samples.

4. Materials and Methods

4.1 Plant materials and growth conditions

Tomato (*Solanum lycopersicum*) cultivar cv. MoneyMaker (u/u) used as WT plant was provided by The Institute for Mediterranean and Subtropical Horticulture "La Mayora" (IHSM-UMA-CSIC) germoplasm bank.

Transgenic tomato lines corresponding to *pTFM5:GLK2*, *pTFM5:APRR2*, *pTFM5::GLK2:pTFM5::APRR2*, *pTFM5:GLK1:pTFM5::APRR2* in MoneyMaker were the same as those used in the previous chapter.

T2 seeds were grown in greenhouses during the spring-summer of 2015 at IHSM-UMA-CSIC (Algarrobo-Costa, Málaga). Seedlings were transplanted to 16-L pots at the four true-leaf growth stage; they were watered daily and with soluble fertilizer once a week.

All fruit samples for RNA were taken at a similar time of the day, frozen in liquid nitrogen and stored at -80 until required. For each data point a pool of 3-4 fruits were collected and stored as a single biological replicate, each essay was performed with 3 biological replicates.

4.2 Protein analysis

Proteins were extracted from 1 gram of leaves (MM) and fruits (MM, *GLK2*, *APRR2* and *GLK2-APRR2*) using Tris-buffered phenol (pH 8) extraction protocol and precipitated as described Faurobert²⁴⁵. The protein pellet was dissolved in CHAPS and urea buffer.

Bradford assay²⁴⁶ was used to estimate the total protein concentration after acidification of the sample buffer²⁴⁷.

For proteomics analysis the method described in Kilambi (2016) was used²⁴⁸. Briefly, peptide fractionation was performed by in-gel digestion-SDS-PAGE according to Laemmli²⁴⁹. A total of 70 µg protein estimated as in Goldring JP et al., 2006, was used in each replicate. Protein were reduced with 10 mM DTT, alkylated with 55 mM IAA and subjected to trypsin digestion (1: 25, enzyme: protein) for 16 h at 37°C²⁵⁰. Peptides obtained from each extraction step were pooled and then concentrated using speed vacuum concentrator (Thermo Scientific), desalted using C18 stage tips and then subjected to LCMS analysis. For each sample, two replicates were run on SDS-PAGE, which were subsequently digested and processed for LCMS analysis. Separation of peptides was carried out on Easy nanoLC-II coupled with LTQ Velos Pro mass spectrometer (Thermo Scientific).

4.3 RNA extraction

Total RNA was extracted from ~200mg of tissue by using TRIzol RNA Isolation Kit (Thermo Fisher Scientific) according to manufacturer's instructions and it was treated with DNA-free™ DNA Removal Kit - Ambion following the manufacturer's recommendation. The DNA-free RNAs were reverse transcribed with SuperScript II (Takara).

4.4 Global transcriptomics profiling by RNAseq

Global expression of tomato genes was determined by replicated strand-specific Illumina RNAseq using RNA extracted from fruit at two green developmental stages: IMG and MG. Transgenic plants evaluated for RNAseq analysis were *APRR2*, *GLK2*, *GLK2-APRR2* and *GLK1-APRR2*. A pool between three and five fruits for each plant and stage was made. Three biological replicates from three independent plants corresponding to each transformation event were analyzed.

RNA was extracted and purified of genomic DNA as described above and quality checked by Agilent (2100 analyzer); samples with RNA Integrity Number (RIN) > 8.5 were used for Illumina sequencing. Six µg of the total RNA from pericarp tissues were used to construct strand-specific RNA libraries. Three biological replicates for each sample were sequenced using a HiSeq2500 sequencing system (Illumina) according to the manufacturer's procedure (hosted at the INRA Genomic Platform of Toulouse, France).

To process the RNA-seq data, the low quality reads from raw paired-end RNA-seq sequences in FASTQ format were removed with fastq quality filter using FASTX toolkit version 0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit/). The Phred quality score was set to 20, a probability of incorrect base call of 1 in 100. 3' adaptor sequences were trimmed from the Illumina reads. *S. lycopersicum* was the reference genome to map the sequence reads and the gene annotation version used was ITAG2.3 (Tomato_Genome_Consortium, 2012, <https://solgenomics.net/>) running tophat-2.0.14²⁵¹. HTSeq was used to perform differential gene analysis²⁵². All analyses have been performed with the R software. The normalization was done with DEseq2 package (version 1.6.3) with default settings from raw counts²⁵³. Raw p-value of multiple tests was adjusted as

padj using Benjamini and Hochberg methods²⁵⁴. Genes with a foldchange higher than 1.5 and lower than -1.5 and a PPadj<0.05 were defined as significantly differentially expressed genes (DEG). The cutoff criteria adopted changes according to the analysis, it was used PPadj <0.05 or in some cases a stringer criteria with a PPadj<0.01. In each analysis is specified which cutoff was applied.

4.5 RNA Data Analysis

The Sol Genomics Network web page (SGN, <http://solgenomics.net>) was used to identify genes, search identification genes code, gen annotations (ITAG 2.3) and sequences.

Enrichment analysis was performed in every DEG list created. GO terms were assigned and the ones significantly enriched evaluated in the three different GO categories: biological process, cellular components and molecular functions (<http://geneontology.org/page/go-enrichment-analysis>).

Clustering and dendograms graphics were generated with MeV (Multiple Experiment Viewer) free software, tool for genomic data analysis version 4.9 (<http://mev.tm4.org/#/welcome>).

Venn Diagram was obtained online with a free tool, based upon the venny tool first developed by J.C Oliveros (Oliveros, J.C (2007) VENNY (<http://bioinfogp.cnb.csic.es/tools/venny/index.html>))

MapMan software (version 3.6.0RC1) was used to supports the visualization of data sets in pathways; DEG and DEP were mapped by ITAG 2.3.

CHAPTER 3: Expression of tomato complex BHLH:MYB:WD40 under the light inducible promoter PLI results in high levels of anthocyanin accumulation in the fruit peel and Type VI glandular trichomes



1. INTRODUCTION

Tomato is one of the most cultivated vegetables worldwide (FAOSTAT) and an important source of antioxidants in a regular human diet. The potential health benefits rely on its composition in flavonoids, phenolic acids and lycopene^{255,256}. Tomato skin is particularly rich in healthy components and it has been reported that 98% of total fruit flavonols are accumulated in this tissue. In addition, stored 2.5 times higher lycopene levels than the pulp and also higher amounts of ascorbic acid.

Anthocyanin compounds are not naturally accumulated in domesticated tomato fruits, whereas wild tomato species such as *S. chilense* are able to produce this pigment in the fruit skin¹¹⁰. In addition to its role as attractive colors and healthy compounds, anthocyanins are involved in protecting plants against photodamage triggered by high light exposure^{164,257,258}. First, acting as a light screen, (absorption at 500-600 nm) and buffering an excess of photons that could overflow the photosynthetic machinery, reducing a toxic increase in the Reactive Oxygen Species (ROS) production. Second, by filtering UV light (absorption at 230-300nm) and third, acting as antioxidants reducing ROS toxicity^{257,259-261}. It has been hypothesized, that for this reason the leaves of many tropical trees (exposed to a higher UV levels) accumulate important levels of anthocyanins²⁶².

ROS has been reported in tomato under stress situations, such as salinity²⁶³, cold temperatures²⁶⁴ or drought²⁶⁵ and all of those hits conditions have been associated to anthocyanin accumulation in vegetative tissue. On the other hand, tomato engineered to accumulate anthocyanin in the fruit during ripening, showed increased ROS-scavenging capacity and to extended fruit shelf-life and resistance to grey mould¹⁰⁸.

Information about the genes responsible for anthocyanin accumulation in fruit of wild species in the tomato clade has been reported ²⁶⁶. Thus the phenotype of *S. chilense* was associated to a dominant allele in the *Anthocyanin fruit (Aft)* locus ¹¹⁷ and two homologous genes identified as possible candidates: *Anthocyanin 1 (SIANT1)* and *Anthocyanin 2 (SIAN2)* both mapping on *chromosome 10* of tomato ^{110,267}. Additional genes involved in anthocyanin production/ regulation were reported and mapped in more *Solanum lycopersicum* species. *Aubergine (Abg)* in chromosome 10 as responsible of accumulation of anthocyanins in the fruit peel of *S. lycopersicoides* ¹¹³ and *atviolaceum (atv)* locus on chromosome 7 from *S. cheesmaniae*, which in this case accounted for anthocyanin accumulation in the entire plant ¹¹⁵.

The dominant gene *Aft* from *S. Chilense* and the *atv* from *S. cheesmaniae* (L. Riley) were introgressed into domesticated tomato ¹¹¹, in separate lines or combined in the same line ^{115,119}. These genes introgressed from interspecific crosses are able to extend the flavonoid pathway, normally expressed in the tomato skin up to naringenin chalcone, downstream to anthocyanin. This is so because tomato fruit skin presents all necessary precursors and enzymes in the early pathway of flavonoid biosynthesis, but not down to anthocyanin whose specific genes are expressed below detection ^{121,123,124}. Cuticle that is synthesized from epidermal cells, is characterized to accumulate over 1% of its dry weight as Naringenin chalcona, and those cells express high levels of CHS, flavonol hydroxylase (F3H), and flavonol synthase (FLS) transcripts ²⁶⁸. In contrast, all the flavonoid pathway seems to be turned off in fruit flesh tissue ¹²⁴.

In all plants the anthocyanin pathway is regulated at transcriptional level that can be divided in three main steps: phenylpropanoid pathway (from phenylalanine to 4-coumaroyl CoA), early steps of the flavonoid pathway and late steps of the anthocyanin specific pathway (Late biosynthetic genes, LBGs)

^{125,126}. The transcriptional complex BHLH:MYB:WD40 (BMW) has been shown to regulate anthocyanin production by activating LBGs and this triad is required for regulating the synthesis of tannins and anthocyanins in many species ²⁶⁹⁻²⁷¹. BHLH:MYB (BM) complex was first described in maize where the R3 repeat of the MYB domain interacts with the N-terminal region of the BHLH factor ¹²⁷. Some interesting variations include the presence of two BHLH proteins, PhAN1 and PhJAF13, which interact with MYBs (*PhAN2*) in petunia ^{128,272}. The participation of WD40 (WDR; called *AN11* in petunia) was discovered later. This TF modulates the activity of the MYB or BHLH factors ¹²⁸⁻¹³⁰. Several studies found those regulators (MYB, BHLH and WDR) as key for the anthocyanin pigmentation in a number of fleshy fruits, such as grapevine, apples or tomato ^{125,131}.

Furthermore, a BMW type of regulatory complex has been described to control distinct cellular processes, including developmental and biochemical pathways ²⁷⁰. In the model plant *Arabidopsis*, initiation or inhibition of trichome development is dependent on BM regulation ^{273,274}. Interestingly, a MYB relevant for trichome formation called *Triptychon* (*TRY*) is able to interact with BHLH proteins described in anthocyanin pathway, such as JAF13 from petunia ²⁷⁵. It has been described that the specificity for this regulatory complex is provided by the MYB protein, whereas BHLH factor could participate in several biological processes. Therefore, it has been suggested that the MYB which control anthocyanin production in *Arabidopsis*, diverges during evolution to have a role in trichome formation. However, the *Arabidopsis* (Rosids clade) model differs from *Antirrhinum* and Solanaceous species (Asterids clade) ²⁷⁶. In Asterids have been described MIXTA (MYB) protein to be implicated in trichomes and conical cell formation on floral tissues; however, MIXTA does not has interaction with BHLH ²⁷⁷.

Environmental conditions are highly important as cues of anthocyanin synthesis regulation ²⁷⁸. The synthesis of

anthocyanin compounds is induced in most plants in response to abiotic stresses ²⁷⁹ and it reinforces the phenotype of anthocyanin mutants: *atv*, for example, displays a stronger anthocyanin pigmentation under cool conditions in stems, leaf veins, and green fruit ^{115,280}. Accumulation of anthocyanins is also stimulated by sugars ²⁸¹ and light ²⁸². The phenotypes of *Aft* and *Abg* require high light intensity to produce high levels of anthocyanins ¹¹⁹, but the combination of *Aft/Aft atv/atv* seems to reduce the strong light requirement for the anthocyanin production ¹¹⁸. Activated by high light and temperature treatments, *SAN2* is able to induce anthocyanin synthesis through the control of *SAN1* and *SJAF13* TFs expression. However, the *SAN11* and *SANT1* genes seem to be independent of *SAN2* levels and consequently not induced by environmental conditions ¹³¹.

Anthocyanins have been reported to inhibit polar auxin transport (PAT, occurs in cell-to-cell manner) ²⁸³⁻²⁸⁷, therefore, affecting developmental function from embryo to root, vascular tissue differentiation and general plant growth ²⁸⁸. In *Arabidopsis*, increased levels of acylated anthocyanin and flavonol glycosides have been correlated with a reduction in plant size and development ²⁸⁹. More precisely, in *Abg- atv/atv* and *Aft/Aft atv/atv* tomatoes were reported to show levels of anthocyanin in fruit that inversely correlated with the final weight of the mature fruit. Furthermore, fruit with higher content of pigment in flesh were always small, whereas it was possible to find normal fruit size and higher anthocyanin values in skin tissues ^{266,290}. Consequently, and to overcome the possible negative effects of anthocyanin accumulation in plant growth and development ²⁹¹, several engineering strategies have been followed to increase anthocyanin levels in fruit but not in vegetative tissues, as well as to restrict them as much as possible to skin tissues.

In addition, engineering strategies were developed although not always successful. Thus tomato plants overexpressing *PhDelila* gene (BHLH), under 35s promoter, resulted in increased the anthocyanins biosynthesis only in vegetative tissue but not in fruit. However, lines which overexpress *AtPAP1* and *SIANT1* (MYB) genes (also under the control of the 35s promoter) showed anthocyanin pigmentation in small fruit areas^{135,292}. Co-overexpression of TFs *Delila* and *Rosea1* (*Del*/BHLH and *Ros1*/MYB) from *Antirrhinum majus* under the ripening fruit-specific E8 promoter^{108,136} in MicroTom produced fruit with high levels of anthocyanins in both peel and flesh, with a notorious effect on both, color and health attributes^{293,294}. In a later report¹³⁸, the light-responsive PLI promoter¹³⁷ was chosen to drive the expression of *Ros1* (MYB) in transgenic tomato lines, and the 35s promoter was used to control *Del*, as a result in tomato fruit with lower levels of anthocyanin in flesh, as well as weak and variable accumulation of anthocyanins in the fruit peel¹³⁸ and other vegetative tissues (unpublished results).

Restricting anthocyanin accumulation to fruit peel has additional interests, since its UV-VIS filtering properties could protect the fruit against high light²⁹⁵⁻²⁹⁷. Therefore, this trait reduces ripening disorders^{90,199,298} or pathogens incidence¹⁰⁸ and consequently improving fruit quality and shelf-life.

In this chapter we report the engineering of two traditional tomato varieties to accumulate high levels of anthocyanins only in the fruit peel, whereas the inner fruit part exhibits its usual red color. Our engineering design is based on an all-tomato approach strategy, where tomato orthologues for BMW regulatory complex were driven by tomato specific promoters. The selection of PLI promoter to direct the main components of the complex allows, in addition to the specificity for the fruit peel, the possibility of modulating the accumulation of certain groups of flavonoids according to the light intensities. At the same time, we describe a novel production of anthocyanins in

glandular trichomes which sets the basis for further studies, on their potential use as biofactories and plant-pest interaction essays.

2. RESULTS

2.1 Generation of tomato plants overexpressing tomato BHLH:MYB:WD40 complex in fruit peel

Because *Petunia* is model specie for anthocyanins biosynthesis and their BMW complex components are the most studied in Solanaceae¹²⁸⁻¹³⁰ we chose *petunia* sequences as a query in a search against tomato genome to identify the orthologues genes. *PhAN11* (WD40), *PhAn2* (MYB) and *PhJAF13* (BHLH)²⁷⁰ were used to identify the corresponding orthologues in tomato (*S/AN11* XP_004245600.1, *S/ANT1* NP_001234417.1, *S/JAF13* NP_001333930.1) by BLASTP. Our protein output were transparent testa *Glabra-1-like* (89% identity), *MYB113-like* (59% identity) and TF *EGL-1* (80% identity). Those regulatory genes have been previously mapped in tomato¹¹² and recently identified by a phylogenetic approach as *AN11*, *ANT1* and *JAF13* respectively, based on their homology to the *petunia* genes¹³¹.

The expression profile of the genes (*S/AN11* Solyc03g097340.1, *S/ANT1* Solyc10g086260.1, *S/JAF13* Solyc08g081140.2) was analyzed (Supplementary Figure 8) (<http://gbf.toulouse.inra.fr/tomexpress17b>. *JAF13* (BHLH) displayed a peak of expression pattern from breaker to RR fruit, whereas *ANT1* (MYB) was almost non-expressed in all the analyzed tissues. Despite endogenous *S/AN11* (WD40) expression showed steady high levels in fruits from IMG to RR stages, and therefore could be enough for our tomato engineering needs, *S/AN11* was also included in the multigene

strategy of this work, to ensure efficiency of the anthocyanin pathway and ectopic activation in fruit peel during development/ripening.

In order to confer, a ripening specific tissue-stage of anthocyanin accumulation, to our fruit peel expression strategy phytoene desaturase promoter (PDS)²⁹⁹ was used to drive *AN11* expression. Furthermore *PDS* displayed a similar expression pattern for both peel and flesh tissues (Supplementary Fig. 9a). To direct anthocyanin accumulation to the outer fruit layers, we choose Plastidial Light Inducible promoter (PLI) to control the expression of *ANT1* and *JAF13*. Previous studies, described the role of PLI protein in chloroplast to chromoplast transitions³⁰⁰, with a characteristic peak of abundance in outer pericarp of fruit, expressing from immature stage with highest expression at breaker^{137,300}. Our analysis using TomExpress database also confirmed a fruit peel specific expression (Supplementary Fig. 9b).

Anthocyanin Engineered Plants (AEP) were generated harboring the designed anthocyanin regulatory complex (Fig. 1), and the presence of each of the transcriptional units in those plants was verified by PCR as described in M&M.

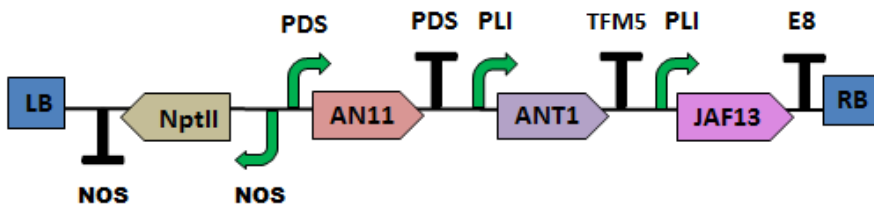


Figure 1: Anthocyanin module used in our anthocyanin engineering peel (AEP) approach. The Anthocyanin construct (12Kb) consisted in four transcriptional units assembled as indicated in M&M in the following order from left border to right border: *NptII* (KmR), *AN11* (WDR), *ANT1* (MYB like) and *JAF13* (BHLB). The promoter (arrow) and terminator (T) sequence used in each case is indicated by the gene from where they derive: Phytoene desaturase (PDS), light inducible (PLI), E8 and nopaline synthase (NOS).

A total of fifteen plants of tomato cv. Muchamiel and nineteen for tomato cv. Pera were selected based on Kanamycin resistance (KmR) and on the integration of the engineered TUs in the genomic DNA. Three lines which displayed the representative phenotype (see later) were selected for each variety (Muchamiel and Pera) and T2 homozygous plants used for further evaluation.

2.2 Engineered fruit display intense purple pigment accumulation in fruit peel

Fruit from all AEP lines, in both genetic backgrounds Pera and Muchamiel, showed a deep purple coloration of the peel which commenced to be visible from early mature green and breaker stages to the ripe stage, whereas no purple pigmentation was visible in the inner pericarp, locular tissue or columnella, which remained at the corresponding green/ pink or red stages according to their stage of ripening (Fig. 2). These results are consistent with the expected expression pattern of the PLI promoter³⁰⁰ (Supplementary Fig. 9b).



Figure 2: Plants engineered with anthocyanin module produced tomatoes that were purple on the surface. (a) Fruit from engineered plants (upper part) and the corresponding wild type untransformed controls (bottom) at three representative stages: MG, Breaker and Red Ripe. Two different tomato cultivars were engineered with the anthocyanin module cv. Pera (left) and cv. Muchamiel (right). **(b)** Figure shows Muchamiel fruit (left) and longitudinal section at Red Ripe stage with a close-up of peel-pericarp border (right) showing the purple pigmentation constrained to fruit peel.

To get a deeper insight into the tissue localization of the accumulated pigments, peel samples from engineered fruit at breaker stage were examined by light microscopy. As shown in Fig. 3, the purple color/ pigment accumulated in the sub-epidermal collenchyma cells, whereas no significant pigmentation was observed in the epidermis or parenchyma layers. These observations were confirmed and made more precise by confocal microscopy, showing that the accumulation of pigments with the characteristic anthocyanin excitation and emission spectra (530 and 675 nm respectively) occurred in these cell layers³⁰¹. Furthermore it was possible to localize the

anthocyanins inside the vacuole of collenchyma cells (Supplementary Fig. 10).

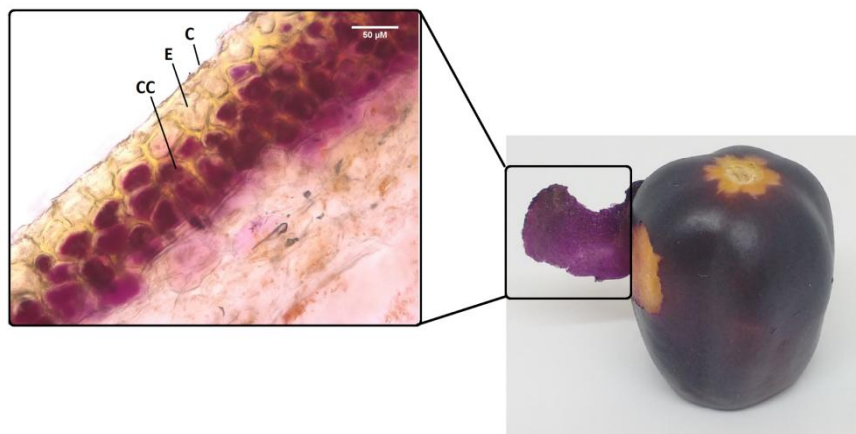


Figure 3: Anthocyanin pigment is accumulated in collenchyma sub-epidermal cell layers. C: Cuticle, E: Epidermis, CC: Collenchymas Cells. Peel fruit at breaker stage was analyzed with light microscope (20X). Bar represent a scale of 50 μ m.

It is expected a similar expression level in the transgenes driven by the same promoter. Here, to confirm the spatiotemporal expression profile of the transgenes driven by PLI, *ANT1* gene expression was analyzed by qRT-PCR. As we previously described, *ANT1* (MYB) gene is almost non-expressed in all the analyzed tomato plant tissues, therefore there won't be expected any interferences in the quantification assessment. As it is shown in Figure 4, *ANT1* transgene expression was substantially higher in fruit at breaker stage, showing the highest level of expression in peel tissue. This result confirmed that the activation of the anthocyanin pathway in the fruit is triggered by the PLI driven expression.

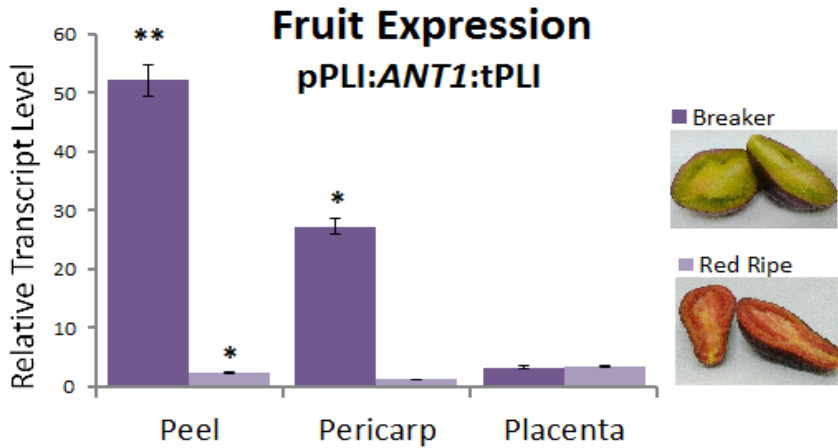


Figure 4: Expression level of *ANT1* driven by PLI promoter displayed the maximum in peel tissue at breaker stage. Expression level of *ANT1* in tomato peel, pericarp and placenta were measured at Breaker and Red Ripe stages. The $2^{-\Delta\Delta CT}$ method was used to determine the relative mRNA abundance (see M&M). The assay was repeated two times (\pm sd) with three biological replicates each. Gene expression is relativized against the condition which presents the lowest expression and statistical significance differences (t-test) are represented with an asterisk (** $p < 0.01$, * $p < 0.05$).

2.3 Biochemical analysis revealed the detailed composition of anthocyanin-enriched fruits

High-Performance Liquid Chromatography (HPLC) was used to define the flavonoid composition of AEP and to find out if the expression of anthocyanin module has any effect on the endogenous peel flavonoid pathway. Analysis of Flavonoid/anthocyanin composition of the peel from AEP fruit harvested at two ripening stages (Br and RR) confirmed the accumulation of anthocyanin compounds. Figure 5a shows the range of identified compounds in AEP samples; the major anthocyanins identified in our samples were Petudins (Petunidin-3-coumaroyl-rutinoside-5-glucoside and Petunidin-3-feruloyl-rutinoside-5-glucoside) followed by Malvidin (Malvidin-3-coumaroyl-rutinoside-5-

glucoside and Malvidin-3-feruloyl-rutinoside-5-glucoside) (Supplementary table 1 on CD). Fruit at both, Br and RR stages accumulated the same set of anthocyanin compounds but in different amounts (Fig. 5b). RR presents the highest levels, for Malvidin-3-feruloyl-rutinoside-5-glucoside which was 227 times higher than in Br samples. In addition, Petunidin-3-feruloyl-rutinoside-5-glucoside, Petunidin-3-caffeoyl-rutinoside-glucoside and Malvidin-3-coumaroyl-rutinoside-5-glucoside increase with ripening 78x, 60x and 51x respectively. Anthocyanin accumulation in the fruit peel thus lagged behind BMW transcript levels during fruit ripening.

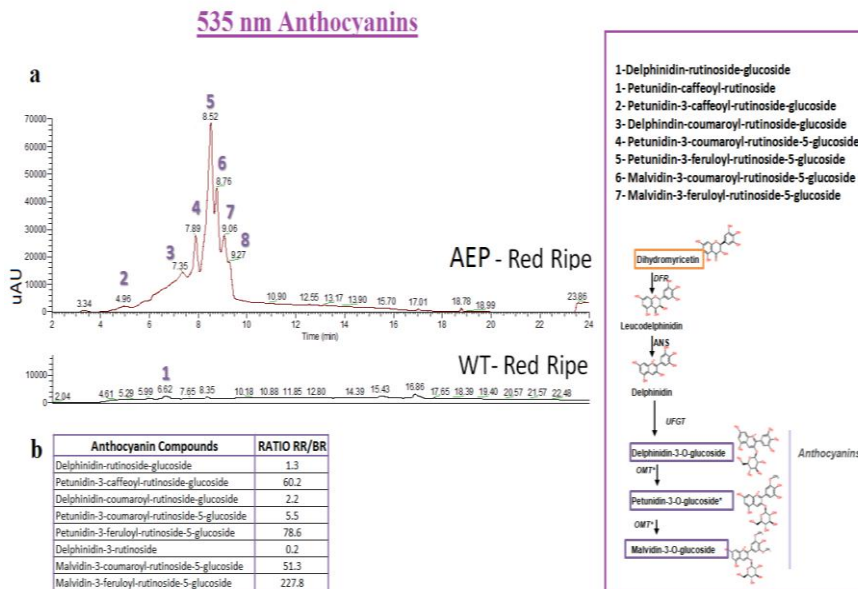


Figure 5: Purple Pigmentation in AEP fruits is mostly due to a small number of highly abundant anthocyanin compounds. (a) HPLC chromatogram of methanol extracts from AEP and WT peel tissues. HPLC chromatogram recorded at 535 nm with compounds for each peak given in the right side **(b)** Ratios of anthocyanin accumulation in the peel of fruits between RR and Br stage showing the increase accumulation at the latest stages of ripening/development.

In contrast with the unusual high levels of anthocyanin accumulation in the peel, other flavonoids like Chalcones, Flavanones, Flavonols and Naringenin derivatives decreased their levels in AEP relative to WT samples. This depletion was observed at both stages of ripening (B and RR) with some exception e.g. dihydronaringenin chalcona, which showed higher levels in AEP fruit (Fig. 6).

In spite of this general pattern, some differences in flavonoids associated to our AEP were dependent on the stage of fruit development. Many phenolic acids increased only in Br and sometimes even decrease in the peel RR fruit related to WT. This is the case for Tricaffeoyl quinic acid (4.5 vs -2.3 foldchange), Dicaffeoylquinic acid (2.6 vs 1.2 foldchange) and Feruloyl quinic acid (1.4 vs -1.6 foldchange).

All this data reflect the dynamism/ plasticity of the anthocyanin pathway and open new possibilities to generate different combinations of specialized metabolites whose effects can be analyzed by functional assays (health bioassays and effect of those changes in fruit biology/ interaction with the environment).

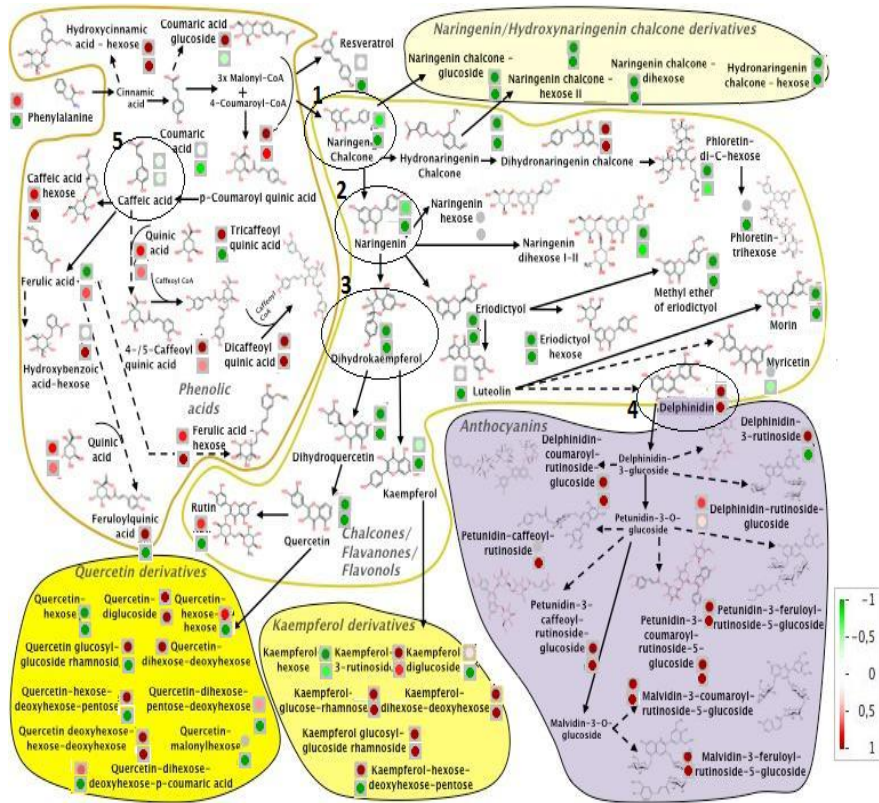


Figure 6: Flavonoid pathway remodeling in AEP fruit peel at two ripening stages. Flavonoid data was acquired by HPLC_MS analysis for AEP Pera cv. at two ripening stages: Breaker (up color) and Red Ripe (down color) (see supplementary table 2 on CD), normalized to WT values and transformed by Log2. For data visualization Mapman was used and shades of red indicating higher foldchange and shade of green lower levels of metabolite amounts. Groups of compounds are numbered in the graphic and circled with a specific color or colorless according to their pigmentation. Colorless: (1) chalcones (2) flavonones, (3) Dihydroflavonols, (5) Phenolyc acids, Purple: (5) anthocyaninis. Dark-yellow: quercetins and Light-yellow: kaempferols.

2.4 Unexpected anthocyanin accumulation in type VI trichomes

Ectopic expression of BMW complex in the peel of AEP was accompanied by an unexpected phenotype. In addition to the

fruit peel anthocyanin accumulation described so far, all AEP lines showed purple pigments in trichomes. Those compounds were easily released upon touching the plant and stained everything being readily soluble in water (Supplementary Fig. 11).

Fig. 7 shows *ANT1* transgene expression levels (*pPLI::ANT1*) on different AEP tissues: isolated trichomes, peel, leaves, and stem without trichomes. Real time PCR data correlated quite well with the purple pigmentation: fruits peel exhibited the highest levels of *ANT1* transcripts followed by trichomes, whereas *ANT1* was practically inexistent in leaf and stem tissues without trichomes.

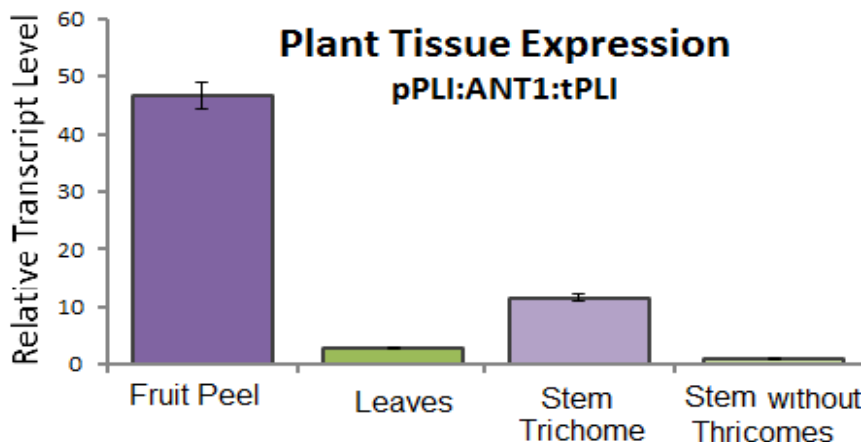


Figure 7: Trichomes and fruit peel of AEP plants showed high expression levels of *ANT1* transgene. *ANT1* levels were analyzed with RT-PCR in different plant tissues: peel fruit (at breaker), trichomes, leaves, stem without trichomes (see M&M). *ANT1* values in relation to those of the sample with the least expression (stem without trichomes). Experiments were repeated twice with similar results. Data are shown as mean ± SD of three technical replicates.

2.5 Trichome-type Specificity

To assess the organ specificity of the unexpected expression of the anthocyanin pathway in trichomes different organs/ tissues were observed in detail. Figure 8 shows a magnification of trichome cells in different tissues: flowers, leaves, stems and fruits. Out of the eight types of tomato trichomes (4 glandular and 4 non-glandular) described in the tomato clade³⁰², three types were identified in our study: I, V and VI (Fig. 8e).

Interesting, the anthocyanin activation was observed only in type VI trichomes, whereas no accumulation was shown in type I (glandular) or V (non-glandular). To investigate in more detail the localization of the purple compounds in the glandular head and perform subsequent analysis, trichome heads were isolated by glass bead abrasion and filtering with nylon meshes (M&M) (SF. 10c). Figure 8f displays the morphology of isolated type VI heads (AEP and WT). Confocal studies confirmed that anthocyanin pigment is accumulated inside the four secretory cells according to the emission spectra (Fig. 8f, SF. 11) and not in the stalk cell.

To further analyze the specificity of our approach for type VI trichomes, the anthocyanin genetic module (Fig. 1) was introduced by crossing in other genetic backgrounds presenting additional types of glandular trichomes, as it is the case of type IV trichomes (check supplementary table 5 to see crosses and different type of trichomes evaluated). Type IV is glandular and belongs to the capitate class as type I (assayed earlier). They are shorter than type I and exclusively presented in wild species, such as *S. pennelli*, *S. pimpinellifolium* and *S. habrochaites*. Moreover, type IV trichomes present most of the genes for specialized metabolites in glandular trichomes, including acyl sugar, terpenoid, or flavonoid pathways, as type I and VI³⁰³.

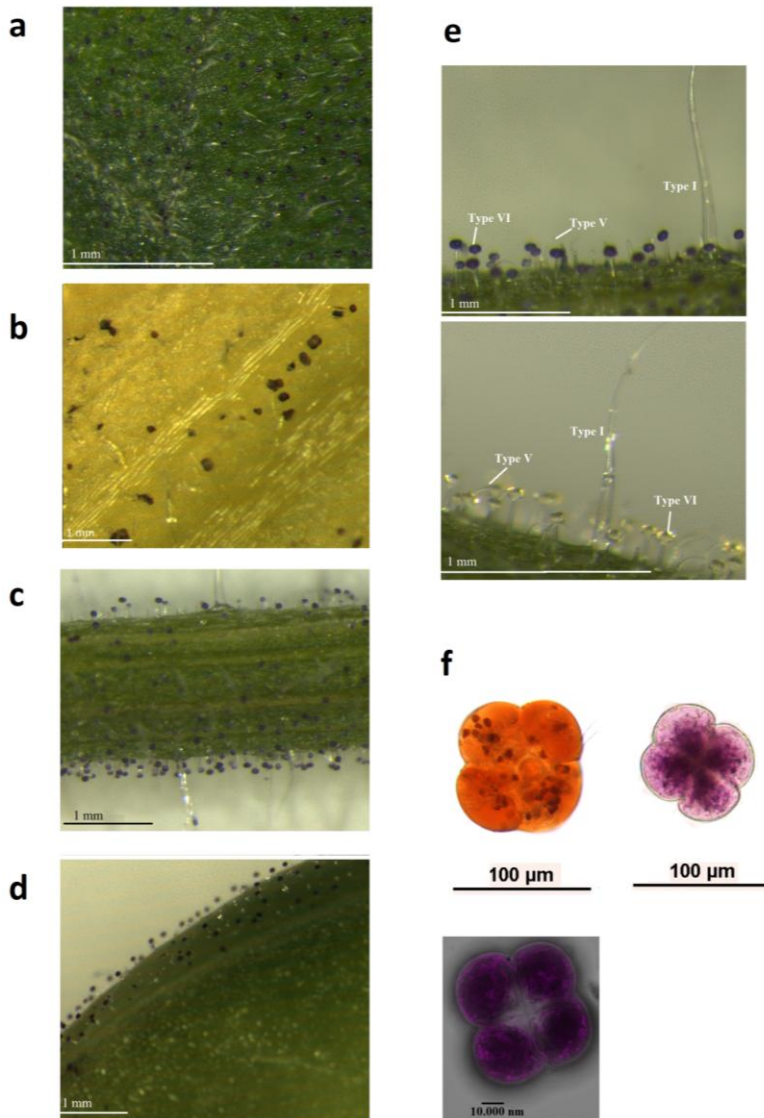


Figure 8: Anthocyanin accumulates only in type VI trichomes of AEP engineered plants. (a-d) *S. lycopersicum* showing trichomes on the surface of (a) Leaves, (b) Petal (c) Stem (d) Fruit samples (e) Detail of type I, V, VI trichomes of AEP engineered plants (upper) and WT (bottom) where the accumulation of anthocyanins is detected only in type VI trichomes. (f) Isolated type VI trichome heads from stems WT (upper left) and AEP plants (bottom right). Accumulation of anthocyanins is detected by the purple color under the light microscope (upper right) and the fluorescence emitted at 675 nm (bottom).

Here, we used BC₅ lines developed by Fernandez-Muñoz et al., 2016 (unpublished data) from an initial cross between *S. pimpinellifolium* (TO-937) and MM background. Those plants presenting highly bushy hairy organs, especially rich in type IV trichomes, were crossed with our homozygous AEP (Supplementary table 5). F1 hybrids, obtained with either TO-937 or the introgressed MM, showed anthocyanin accumulation only in type VI trichomes (Fig. 9).

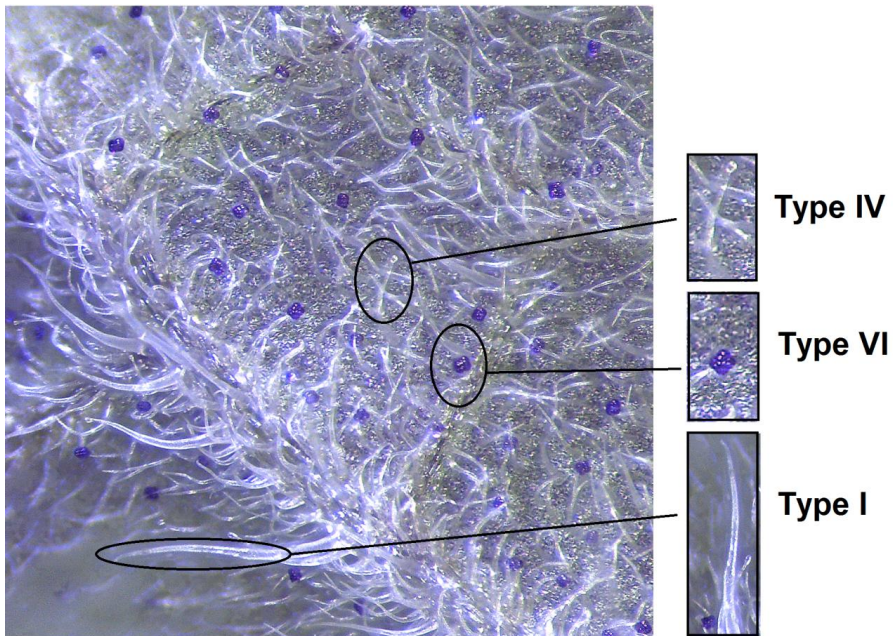


Figure 9: Anthocyanins accumulated in Type VI but not in type IV glandular trichomes. Evaluation of trichome-rich plants resulting from the cross of *S. pimpinellifolium* introgression lines (MM as a recurrent parent) and AEP genotypes. Steam tissue magnification allows us to identify several trichome types: Type I, IV and VI.

Interestingly, a preliminary experiment suggested that the F1 hybrid obtained by crossing our AEP lines with the *S. habrochaites f. glabratum* accession showed blue coloration, not only in type VI trichomes but also in the multi-headed glandular type VII trichomes. However, the expression of the trait seems to

be not consistent in this F1, and significant numbers of type VI and type VII trichomes showed only faint or null coloration. New crosses between our engineered lines and a number of wild species accessions with type IV trichomes, e.g. *S. galapagense*, *S. habrochaites f. typicum*, and *S. pennellii* are on the way to complete the picture.

Since a similar gene regulatory network involving BMW complex controls the trichome development in *Arabidopsis*³⁰⁴, we examined trichome density in leaves and stems of AEP plants and compared to WT (table 3 on CD). Not significant change in trichome density or number was observed in any case, and for instance, density of trichomes in stem remaining mostly constant between genotypes (about 23-25 trichomes/mm²) ruling out the possibility that the anthocyanin module could be affecting trichome density.

2.6 Anthocyanin trichome expression appears to require BM transcriptional complex to be driven the PLI promoter

MicroTom transgenic plants (PDR) overexpressing *AmDelila* (BHLH) by PLI promoter and *Rosea1* (MYB-like) under CaMV35s¹³⁸, did not report any anthocyanin expression in trichomes. To test whether the reason was the different environmental conditions, we cultivated PDR lines¹³⁸ in our conditions and then, confirmed the absence of anthocyanin in trichomes (Fig. 10a, down). PDR plants used CaMV35s promoter instead of PLI to drive the expression of BHLH. Therefore we used MM tomato lines, which overexpress *dsRED* reporter under 35s promoter, to check whether 35s is able to drive expression of the reporter in trichomes. Stem samples of 35s::*DsRed* plants showed the characteristic high levels of red fluorescence, whereas no significant fluorescence was detected in trichomes (Fig. 10b),

indicating that the 35s could not drive high levels of expression to trichomes tissues.

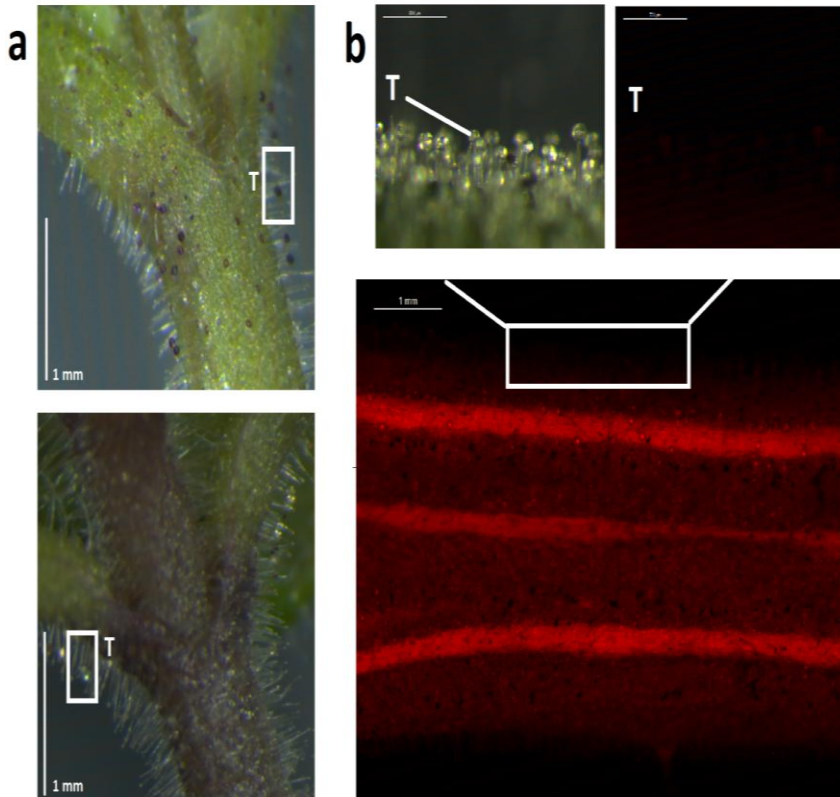


Figure 10: PLI promoter but no 35s results in anthocyanin accumulation in Type VI trichome. T: type VI trichomes (a) Trichome from plants carrying PL1BMW (pPLI:*SIANT1*:pPLI:*SIJAF13*, upper) and PDR lines (pPLI:*PhRosea*:p35s:*PhDelila*, lower) reveal that accumulation of anthocyanins occurs only when PLI promoter controls both type of transcriptional factors. (b) Fluorescence emitted by p35s:*DsRED*:tNos in stem (bottom) but no fluorescent was reported in trichomes (upper). Upper part of the figure shows a detail of trichome in visible field on the left and DsRED on the right.

2.7 Specificities of the alteration of the flavonoid pathway in type VI glandular trichomes driven by the PLI::BMW

To study the effect of PLI::BMW on the flavonoid pathway, anthocyanin and non-anthocyanin flavonoids were analyzed by HPLC-MS in trichome samples from AEP and WT control plants. Three main anthocyanin compounds were identified in stem trichome samples as the major anthocyanins accumulating at high levels in PLI::BMW lines and they were below detection levels in trichome preparations from wild type stems (Fig. 11).

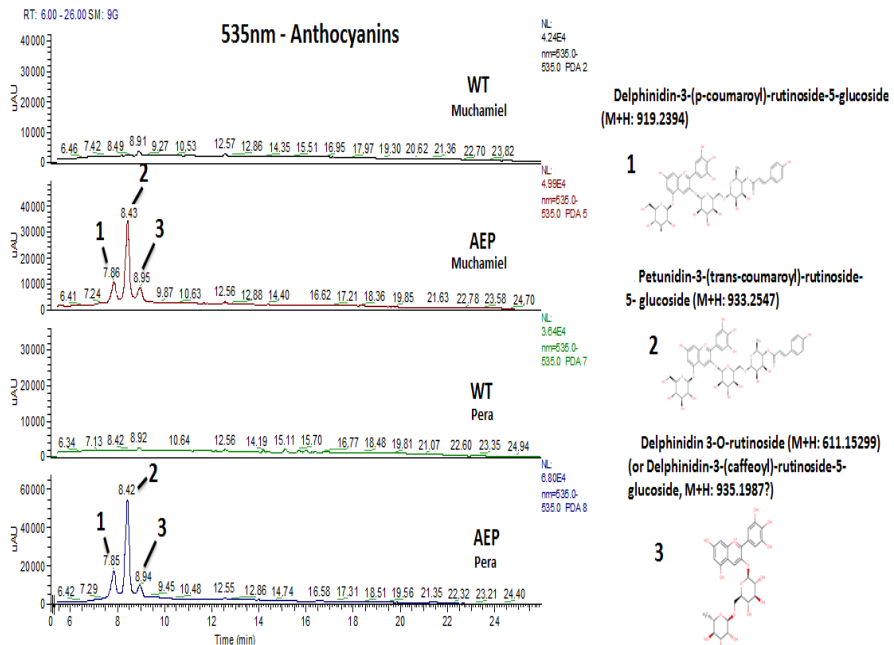


Figure 11: AEP trichomes display three main identified anthocyanin pigments not present in WT samples. Methanol extracts for stem samples of two AEP varieties (Pera and Muchamiel) and their respective controls (WT) were analyzed by HPLC-PDA-MS (see M&M). Anthocyanin profile was determined at 535 nm wavelength. Peaks numbered in the chromatogram correspond to the names and chemical structure indicated on the right of the figure.

Similarly to what happens to the metabolite profile of fruit peels from AEP lines, the most abundant anthocyanin compound was identified as Petunidin-3-coumaroyl-rutinoside-5-glucoside. However, the second most abundant anthocyanin was Delphinidin-coumaroyl-rutinoside-glucoside, a compound which is not highly present in the peel of AEP fruit, indicating that there are specific differences in the activation of the anthocyanin pathway between trichome and sub-epidermal cells.

To better understand the effect that activating anthocyanin accumulation in the trichome may have on the upstream flavonoid pathway, total flavonoids were estimated by measuring at 254 and 325 nm.

Globally, both Pera and Muchamiel AEP trichome samples, presented lower content in total flavonoids than WT (Fig. 12a). However, some specific non-anthocyanin flavonoid compounds increased in AEP trichomes, similarly to what happens in the fruit peel (Figure 12b). When AEP samples are compared to WT, derivatives from caffeoyl acids seem to present higher levels; e.g. Feruoyl quinic acid in Muchamiel. On the other hand, flavonones, which are precursors of dihydroflavonols showed an increase in Pera trichomes (Naringenin).

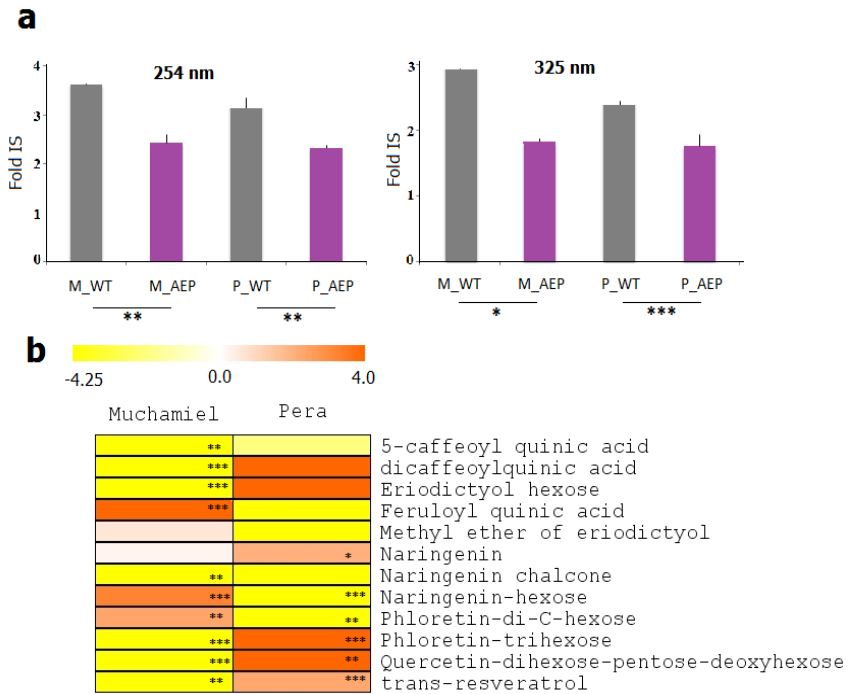


Figure 12: Trichomes of AEP showed lower levels of the other flavonoids compounds. HPLC analysis of phenylpropanoid compounds in earlier steps of the anthocyanin biosynthetic pathway. **(a)** Total flavonoids content recorded at 254 nm (left) or 325 nm (right) of trichome extracts from AEP (purple bar) and WT (gray bar) (M: Muchamiel and P: Pera) **(b)** Heatmap of other identified flavonoids upstream in anthocyanin pathway (table 4 on CD). Values are normalized to WT for both AEP varieties, Pera and Muchamiel. Student's t-test was performance, * = $p\text{Value} < 0.05$; ** = $0.01 < p\text{Value} < 0.05$; *** = $p\text{Value} < 0.01$.

2.8 AEP phenotype is dependent on light

The anthocyanin accumulation in the fruit peel of AEP displays a light-dependent phenotype. We observed that anthocyanin accumulation in fruit of plants growing in the greenhouse decreased, when fruit developed under the canopy, in opposite side from light of fruit exposed to lateral illumination, or in fruit region underneath sepals (Supplementary Fig. 12). As the *PLI* gene has been reported to has photo-protective roles under

light stress conditions, with increasing *PLI* levels according to light intensity³⁰⁵, we determined the levels of expression of the engineered genes driven by *PLI* on fruit samples exposed to different environmental light conditions.

In a first experiment, some fruit were covered with aluminum foil for 2 week while others fruit were exposed to increasing amounts of light provided by artificial lamps. The results showed that we can produce a range of fruit phenotypes with increasing "purpleness" in the fruit by simply modulating the amount of light in the greenhouse (Fig. 13a) opening the way for the environmental control of purple coloring in the peel in these lines. Furthermore, the expression levels of transgenes driven by *PLI* (*ANT1* and *JAF13*) were analyzed in fruit exposed to 0 W/m², 230 W/m² and 270W/m² (constant temperature, 21°C) (Fig. 13b).

Expression levels of *ANT1* and *JAF13* were estimated by quantitative real-time PCR and found highly correlated with the levels of pigments accumulated in the fruit skin under the three different tested conditions.

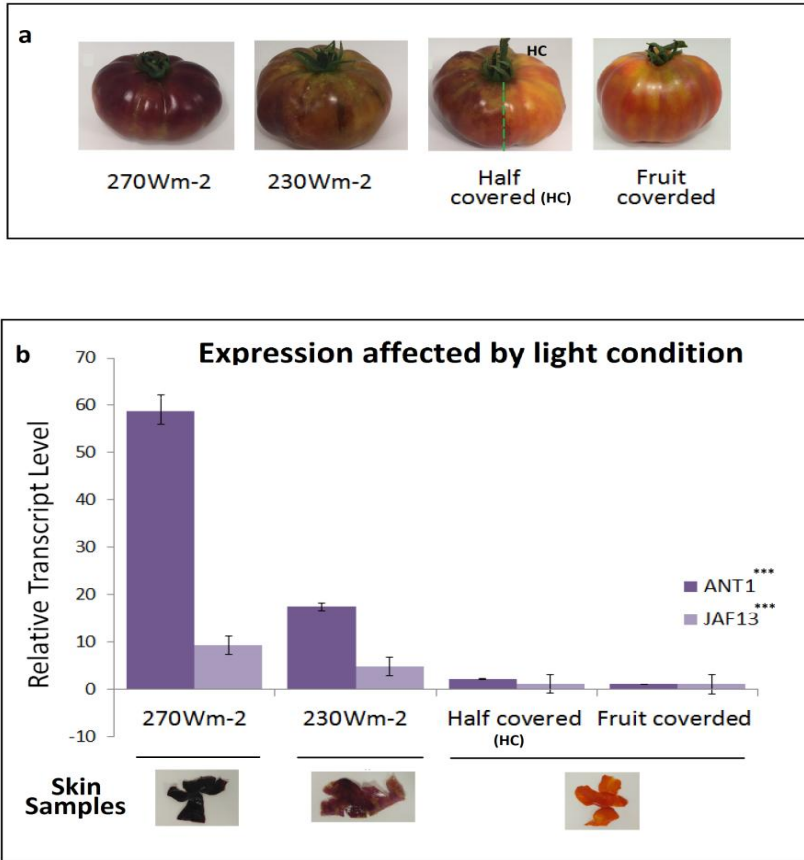


Figure 13: The degree of fruit purple color and transgene expression correlated directly with the increase in light irradiance. (a) Fruit phenotype of AEP is highly dependent on light intensity. Four different conditions were tested by exposing fruit for 2 weeks to: 270W/m² of light intensity, 230 W/m², half fruit covered (0 and 230W/m²) or fruits totally covered (0 W/m²) (b) Relative transgene expression levels driven by PLI promoter (*ANT1* and *JAF13*) analyzed by RT-PCR. Gene expression relativized against the condition which presents the lowest expression (fruit covered). Standard deviations (SD) are the result of three independent replicate of skin samples, statistically significant difference between sample groups (***)ANOVA $p < 0.01$.

To analyze the effect of light on anthocyanin accumulation in AEP trichomes, a similar experiment to that of fruit peel was conducted. Seeds from T3 generation of AEP and PDR

transgenic lines ¹³⁸ were germinated and grown in vitro (see M&M) under different light conditions (Fig. 14), ranging from 13 $\mu\text{mol m}^{-2} \text{s}^{-1}$ to 130 $\mu\text{mol m}^{-2} \text{s}^{-1}$ and under standard light conditions (in vitro chamber, 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$). After three weeks the phenotype of the seedlings was observed under a Leica binocular magnifier.

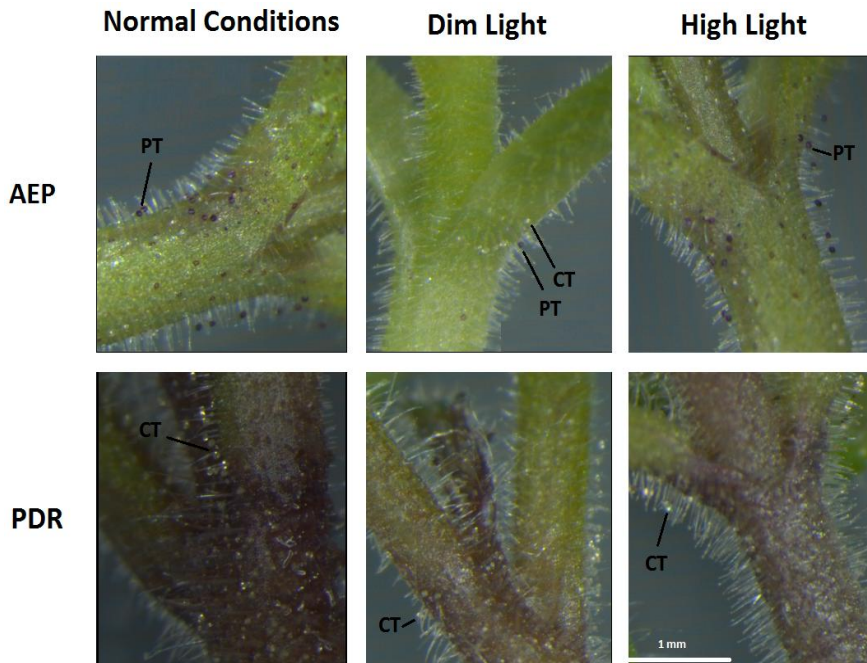


Figure 14: The number of purple trichomes decreases in dim light conditions. AEP Muchamiel (upper) and PDR transgenic plants (bottom) were grown for three weeks in different light conditions: 130 $\mu\text{mol m}^{-2} \text{s}^{-1}$ (High Light), 13 $\mu\text{mol m}^{-2} \text{s}^{-1}$ (Dim Light) and Normal Conditions (Phytotron, 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$). Examples of Purple Trichomes (PT) and Colorless Trichomes (CT) are marked in the pictures.

Under light conditions reduced to 13 $\mu\text{mol m}^{-2} \text{s}^{-1}$ (Dim light), Type VI trichomes of our AEP were still accumulating anthocyanins. However, lower number of purple trichomes were detected (Fig. 14, up). In addition, there was no difference between normal and high light conditions. In contrast, when the

Bassolino MicroTom plants (PDR)¹³⁸, which harbor the pPLI::*Ros1::35s::Del1*, were cultivated under the same conditions no anthocyanins were observed in trichomes. Those results were independently of the light intensities (Fig. 14, down), even after high light treatment.

As far as plant phenotype is concerned, our engineered plants display no other obvious changes except that roots of plantlets accumulated obvious amounts of anthocyanins growing under high light conditions (Fig. 15a), probably in response to the presence of sucrose in the media³⁰⁶. No phenotype was observed in roots when plants were cultivated in soil.

In contrast, PDR plants displayed a phenotype not restricted to the fruit peel¹³⁸ showing accumulation of anthocyanins in different parts of the plant, even in normal light conditions (Fig. 15b). Those observations indicate that the phenotype is leaky and not fruit peel specific.

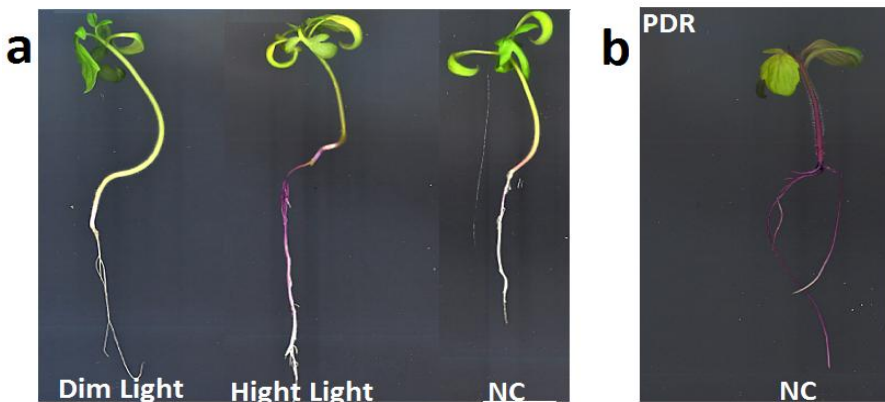


Figure 15: AEP Roots show anthocyanin accumulation under high light conditions, whereas PDR plant accumulates anthocyanins in vegetative tissues under normal conditions. (a) AEP Muchamiel (left) grown for three weeks in Dim Light ($13 \mu\text{mol m}^{-2} \text{s}^{-1}$), High Light ($130 \mu\text{mol m}^{-2} \text{s}^{-1}$) and Normal Conditions ($100 \mu\text{mol m}^{-2} \text{s}^{-1}$). **(b)** PDR transgenic plants show anthocyanin accumulation in vegetative tissues under light normal conditions. Seedlings were grown in MS with sucrose (See M&M).

2.9 Differences in light intensity produces dramatic changes in the flavonoid profile of the fruit peel

Our results so far suggested that the anthocyanin/ flavonoids levels in our AEP lines were much dependent on the amount of light, what made us to wonder if the flavonoid profile was also dependent on light conditions. To get a better insight, on how the flavonoid profile changes with light, we first analyzed the anthocyanin profile of fruit from Weak Phenotypes (WP, Dim Light conditions) and Strong Phenotypes (SP, High Light conditions) samples (Fig. 16). Compared to WP fruit, SP fruit presented higher levels of anthocyanins and this was true for all the identified compounds (most of them around 30 times higher). Malvidin compounds however, were the most dramatically affected (Malvidin-3-coumaroyl-rutinoside-5-glucoside and Malvidin-3-feruloyl-rutinoside-5-glucoside) reaching values over 150 times higher than WP samples (ratio strong/weak phenotype).

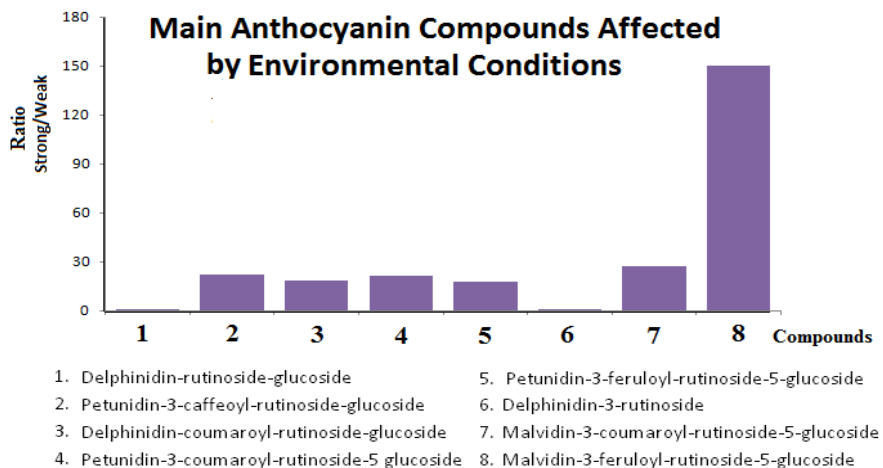


Figure 16: High irradiance light conditions promoted specific anthocyanin accumulation in the peel of AEP fruits (continued).

Figure 16 (continued): HPLC chromatogram at 535 nm of methanol extracts from peel of AEP fruit displaying strong and weak phenotypes. Ratios of strong/weak phenotypes due to differences in lighting/shading are represented in the bar graphic

Our results also revealed that levels of some flavonoid compounds, upstream from anthocyanin pathway (precursors or in other branches), increase up to double in WP samples, whereas the peel of SP fruit halves WT total levels. Our data suggest that the environmental conditions affect the regulation of phenylpropanoid metabolism in the peel of our engineered fruit and this occurs at different levels including early steps on the pathway.

Mapman comparison between WP and SP flavonoid profiles (Fig. 17) allow us to see in an organized and detailed way the differential effect through the phenylpropanoid-flavonoid-anthocyanin pathway. The peel samples from WP fruit presented higher levels of several precursors of anthocyanin compounds such as Naringenin chalcona, Naringenin, Dihydrokaempferol, Eriodictyol, Ferulic acid, p-coumaric acid-glucoside, Rutin, among others, reaching values even eightfold higher relative to WT. In contrast, these compounds showed a slight decrease in SP peel samples (see supplementary table 5 on CD).

The situation is different in other branches of the pathway, and thus dim light conditions (WP fruit) stimulate a more dramatic accumulation of quercetin and Kaempferol derivative compounds. Although the increasing light exposure also has a moderate positive effect on quercetin and Kaempferol derivative compounds levels, most of the flavonoid pathway seem to be depleted as it is probably diverted to the synthesis of anthocyanins (Fig. 17).

These results suggest that strong activation of anthocyanin synthesis leads to a partial depletion of other upstream flavonoids components. Whilst low irradiance environmental

conditions provide visually weak fruit phenotypes, with less anthocyanin accumulation, but increased in other flavonoids. In summary, our AEP can be used to produce fruits accumulating preferentially different subsets of flavonoids by adjusting the light conditions.

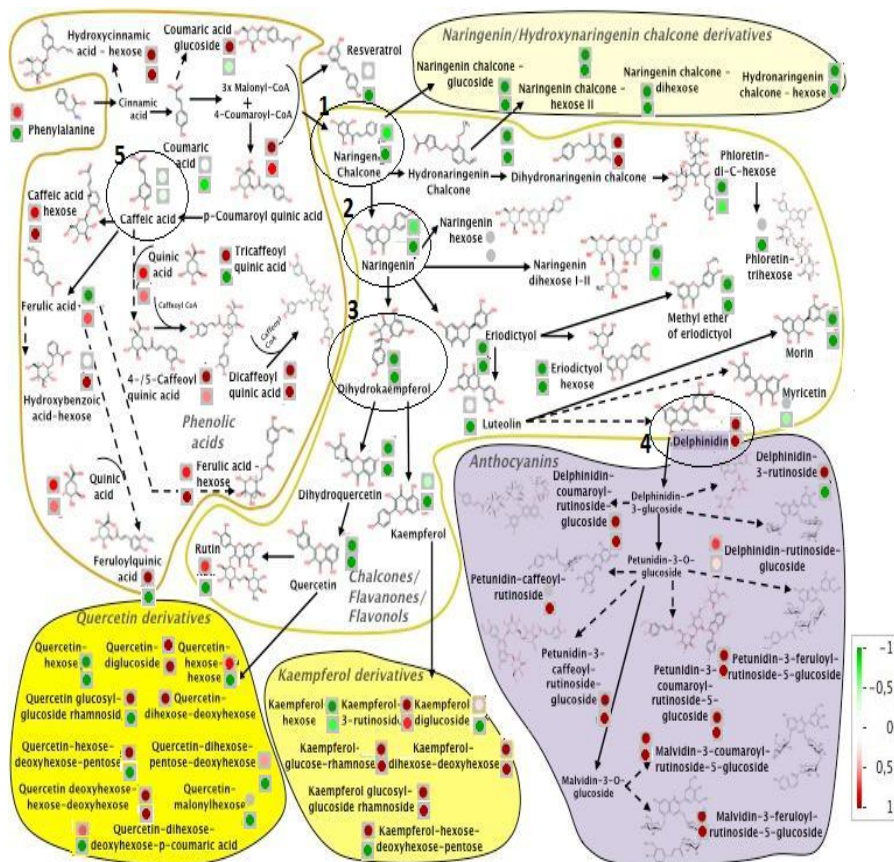


Figure 17: The peel of engineered fruits accumulate differential flavonoids up and downstream in the pathway (a) Mapman of flavonoids extracted from peel of fruit under low light / weak phenotype (WP upper dots) and high light/ strong phenotype (SP, lower dots). Data represents flavonoids content in AEP previously normalized to WT values and transformed with Log base 2. Three representative plants from Pera transformants and WT were used under each light condition. Groups of compounds are numbered in the graphic and circled with a specific color or colorless according to their pigmentation. Colorless: (1) chalcones (2) flavanones, (3) Dihydroflavonols, (5) Phenolic acids, Purple: (5) anthocyaninis. Dark-yellow: quercetins and Light-yellow: kaempferols.

3. DISCUSSION

3.1 Purple pigment in our AEP is restricted to peel similarly as it happens in fruit of wild species

Although the distribution of anthocyanin pigment could include the inner part of the fruits/ tuber ³⁰⁷, in specific *Solanaceae* members, most natural variation identified for this trait is restricted to the peel and this is especially true in the tomato and pepper clade. Furthermore several genes have been described to control the expression of this peel trait ³⁰⁸ in the corresponding species such as in tomato (*Aft*, *Abg*), potato (*D*, *R*, *P* loci in the tuber) or pepper (*A* and *Fc* loci) ^{112,307,309} and are responsible for the peel specific phenotype.

Wild species (*S. chilense*, *S. peruvianum*) in the tomato clade, and some others less closely-related (*S. lycopersicoides*), accumulate relatively high amounts of anthocyanins in sub-epidermal tissue ¹¹⁹. However, this characteristic trait is not presented in the red species such as *S. lycopersicum*, *S. lycopersicum. Cerasiforme* or *S. pimpinellifolium*. Therefore, we engineered this trait in traditional tomato varieties and our plants displayed uniform anthocyanin distribution across the whole collenchyma cell layer (2-4 sub-epidermal), whereas no purple pigmentation was detected in the epidermis or in the fruit flesh. Except for the unexpected expression in the glandular trichomes, this is the first engineered tomato to accumulate anthocyanins specifically in the fruit surface/ peel; since PDR lines ¹³⁸ showed (in our hands) ectopic expression in many other parts of the plant (see Fig. 15b and unpublished results).

PDR lines ¹³⁸, showed a low anthocyanin accumulation in tomato flesh, previous to our work and as in Butelli et al., ¹³⁶ used snapdragon (*AmMYB* and *AmBHLH Antirrhinum majus*) genes only that in this latter case was to induce anthocyanin

accumulation in both, peel and flesh, by expressing *the two TF* under E8 promoter in cv. MicroTom.

Furthermore, we have reported for the first time that the combined expression of an all-Tomato BMW complex leads to anthocyanin accumulation in fruit. However, the main reason underlying the peel restricted phenotype relies on promoter selection. *PLI* was described to drive the expression on the outer pericarp tissues of fruit from immature stage with a peak at breaker stage of development^{137,300}. This expression pattern is consistent with the endogenous *PLI*, as can be obtained from TomExpress analysis. The specificity expression pattern was; however, not a guarantee that the levels of expression achieved would be enough to cause accumulation of anthocyanins (Mathews 2003, Zuluaga, Gonzali et al. 2008). Our success is consequence of the enough triad-BMW expression levels specifically expressed in the target tissue.

3.2 Unexpected novel localization of anthocyanin in tomato trichomes

All vegetative and reproductive tissues of our engineered lines present type VI trichomes with high levels of anthocyanin accumulation. Trichomes that accumulate anthocyanins are very unusual in nature and when they appear they are mostly restricted to floral organs³¹⁰. Only a few cases of anthocyanic trichomes in vegetative tissues have been reported and most of them within the genus Gesneriaceae^{311,312}.

Here, the novel anthocyanin trichome phenotype is provided by the expression of *ANT1* and *JAF13* cisgenes driven by the *PLI* promoter. The Trichome database (http://bioinfo.bch.msu.edu/trichome_est) and literature³⁰² revealed a remarkable high expression of *PLI* gene in trichomes.

Thus, Schillmiller shows that within the 25 most overrepresented reads/ contigs in trichomes *PLI* takes the 18th position³⁰². In addition, it is in 4th position in LA1777 accession of *S. habrochaites* (Trichome database). Therefore, from this information it makes sense that our anthocyanin engineered constructs based on *PLI*, are expressed also in glandular trichomes. In contrast, the lack of anthocyanin trichome phenotype in the PDR plants¹³⁸ could be due to the relatively low level of expression driven by 35s promoter in glandular trichomes, supported by our experiments with 35s::*DsRed* tomato lines.

Although many studies used 35s promoter to drive the transgen expression to trichomes^{17,313}, attempts to characterize trichome specific promoters have been made to avoid pleiotropic effects and also to increase expression levels in trichomes³¹⁴⁻³¹⁷. A success example is the recombinant promoter, MSgt-FSgt (hybrid sub-genomic transcript promoter) which showed 25 times more expression in stem trichomes of *Arabidopsis* than the CaMV35s promoter.

We have shown that promoter, used in our construct, drives specific TFs expression to trichome to high enough levels as to produce a strong phenotype. However, there could be additional reasons for the high levels of anthocyanin accumulation in these cells. Type VI trichomes in tomato are specialized cells with endogenous high flavonoid and terpenoid metabolic activity^{318,319}, which means that phenylpropanoid pathway is already active in those cells. Indeed, rutin has been described as the major flavonoid accumulated in type VI trichomes³²⁰.

In addition to that, a clearly relationship between trichome regulation and anthocyanin biosynthetic pathway has been described in *Arabidopsis*, that may also explain our phenotype³²¹. For example, a trichome MYB called Triptychon (TRY) is able

to interact with BHLH TF described in anthocyanin pathway such as *JAF13*²⁷⁵. Silencing *AtTTG1*, (WD40 TF) implicated in initiation of trichomes process not only affects trichome but also repress the plant ability to accumulate anthocyanins²⁷¹. Although glandular trichome biogenesis in tomato maybe different from that of *Arabidopsis* (Serna and Martin, 2003) a certain connection between anthocyanin and trichomes regulation could be emerging in tomato, as in the following examples and manifested in our ectopic overexpression experiments. Thus, *Anthocyanin free (af)* mutant^{112,322} lacks the ability of accumulating anthocyanins in any tomato tissue, but it also has lower trichome type VI density and shows high susceptibility to flea indicative of an effect on secondary metabolism in the trichomes³²³. Furthermore the *Af* has been shown to encode the CHI enzyme which is involved in the biosynthesis of naringenin chalcone the metabolic precursor of other flavonoids including anthocyanins³²⁴.

Here, no differences were detected in the trichome density of our AEP lines; however, trichome flavonoid profiles were changed. AEP trichomes showed in general lower levels of flavonoids compared to WT samples. In addition, it was observed that some compound decreases but derivative of it increases, e.g. Phloretin-dihexose and Phloretin-Trihexose. In the case of rutin, the major metabolite produced in type VI³²⁰, is not detected in our analysis (even in WT samples). However, precursors of those compounds are upregulated in AEP trichomes (cv Pera); e.g. quercetin or eryodictyol. Most important, anthocyanin compounds, which are not naturally present in WT, were detected in trichomes samples of AEP (two Delphinidin and one Petunidin).

Since both, trichome density and anthocyanins have been described as photoprotective mechanisms developed by plants to mitigate an excess of light^{296,325,326}, the novel trichome phenotype that we have generated can be used in future

studies; aimed to define the possible use of anthocyanins as a UV protective filter or to increase plant photoprotection.

3.3 Engineering trichomes as biofactories

Our findings clearly support that it is possible to modify trichome metabolism to accumulate high levels of anthocyanins, even in species that do not usually do it.

Plant trichomes appear to be suitable not only as biofactories for the synthesis of natural products, but also to provide plants with better protection against several kind of biotic stresses³²⁷. In contrast with *Arabidopsis*, that only presents non-glandular trichomes³²⁷, tomato produces several types of trichomes including the glandular types where several compound such as terpenes, acyl sugars and phenylpropanoid-derived metabolites are synthesized^{302,328}. In cultivated tomato, two glandular types, I and VI have been described. In this work, we also used Moneymaker ILs which incorporates type IV glandular trichomes introgressed from TO-937 (*S. pimpinellifolium*). Type I is considered a capitate type integrated by multi-cellular stalk with a single small gland cell, whereas type VI contains a unicellular stalk with four gland cells³⁰². Type IV is described as capitate and similar to type I but shorter and restricted to wild species such *S. pimpinellifolium*, *S. pennellii* and *S. habrochaites*³²⁹. Here, not only AEP resulted to accumulate anthocyanin restricted to type VI, but also plants from our crosses which are endowed of type IV keep type VI specificity.

Since glandular trichomes have been described to produce different set of metabolites³²⁹, our results make us search for the differences in metabolite profiles between type I, IV and VI in tomato cultivated species. We would like to evaluate to what extent the ability to accumulate anthocyanins of our engineered

PLI::BMW lines will depend on the active specialized pathways operating in each trichome type. Type I, IV and IV trichomes are considered the main site of synthesis of a variety of acyl glucoses³⁰³. However, type VI is known as the major terpenoid producer in cultivated species³⁰². Indeed, a comparative genomic approach³²⁹ showed no qualitative but quantitative differences in the expression of genes involved in acyl sugars and terpenoid pathways in those trichome types. The flavonoid rutin has been described as the major specialized metabolite accumulating in type VI trichomes³²⁰, suggesting that the flavonoid pathway is highly active at least for the EBGs, and therefore with all precursors for anthocyanin engineering in place.

Our preliminary experiment, crossing AEP and *S. habrochaites f. glabratum*, resulted in a F1 generation with weak/null blue coloration in type VI and VII trichomes. Since acyl sugar, phenylpropanoid, flavonoid, terpenoid, and specific alkaloid pathways are expressed in *S. habrochaites* trichomes, new crosses in the rest of wild species having type IV, VI and VII trichomes, e.g. *S. galapagense*, *S. habrochaites f. typicum*, and *S. pennellii* are on the way to explain the trichome specificity of our construct.

3.4 Enriched fruit flavonoid profile

We have reported here that BMW tomato complex has an effect on several branches of the general phenylpropanoid pathway acting probably even upstream in the biosynthetic route. Butelli et al. (2008) reported that AmBHLH (*Delila*) and AmMYB (*Rosea1*) regulate both, EBGs and the LBGs of flavonoid biosynthetic pathway. Here, weak phenotypes of our AEP fruits were characterized by an increase up to double in total amounts of other flavonoid compounds including upstream flavonoids. In contrast, peel of fruit with strong purple pigmentation showed a

decrease in the total levels of flavonoids to half the levels of WT, probably because they are used as precursors of the highly activated anthocyanin biosynthesis and in other branches of the pathway. Consequently, fruit with strong phenotypes show higher levels of anthocyanin compounds than mild-expressor. For example, Malvidin compounds ranged from 30 to 150 times higher in SP than WP, in part depleting precursors of early flavonoids normally accumulating.

In AEP fruits we identified the nature of the anthocyanin compounds and found they all derive from dihydromyricetin, malvidin, petunidin and delphinidin (Fig. 17). These anthocyanin compounds are relatively common in tomato tissues under stress conditions in other *Solanaceae* plants, and derivative from some of the six most common which are cyanidin, peonidin, pelargonidin, malvidin, petunidin, and delphinidin³³⁰.

The anthocyanin profile of *Aft* lines indicates that petunidin is the main anthocyanin accumulated in *Aft* fruits followed by malvidin and delphinidin¹¹⁷. Transgenic tomato which carries 35s:ANT1 reported the same anthocyanin profile in fruits¹¹⁰. The major anthocyanin detected in Microtom transgenic lines which overexpress AmMYB (*Rosea1*) and AmBHLH (*Delila*) under E8 promoter, was Delphinidin 3-(trans-coumaroyl)-rutinoside-5-glucoside followed by Petunidin 3-(trans-coumaroyl)-rutinoside-5-glucoside¹³⁶. Presence of Malvidin in those samples was reported although in a lower proportion³³¹. However, in our case and conditions, AEP fruits display some specific differences in anthocyanin composition (Supplementary table 13): the late compounds derived from dihydromyricetin (ie. Malvidin derivatives) represented a much higher relative proportion of the anthocyanin pool that in *Ros-Del* fruit³³¹. This indicates that the downstream late transcripts/ activities may be further activated in our lines.

3.5 Light dependent phenotype

The anthocyanin accumulation phenotype of AEP is strongly dependent upon light. Under high light conditions, fruits can accumulate anthocyanin compounds over 30-150 fold higher than they do under dim light conditions. In contrast, fruits grown under lower light display enrichment in other flavonoids (up twofold WT). This environmental control of the phenotype allows us to adjust the purple accumulation and flavonoid composition of the fruit peel.

Understanding the underlying basis of those results includes analyzing the light regulation of the PLI promotor and also the effect on the phenylpropanoid /Flavonols/ Antocyanin pathway. The regulation by light of PLI protein has been described previously^{137,300,305}. Light regulation of the *PLI* gene is not always required such as in the fruit during the chloroplast-to-chromoplast, where the PLI protein is expressed even in the dark³³². However, our report showed that the lowest levels of transgene expression driven by the PLI promoter and the weakest phenotype occur in the absence of light. On the other hand, we have to consider that anthocyanin production often starts in tomato upon exposure to stress conditions, i.e. Cold or high light, whereas dim light could involve a negative regulation mechanism¹¹⁹. In addition, tomato plant homozygotes for *Aft/Aft/atv/atv* present a synergistic anthocyanin phenotype with reduced high light requirements¹¹⁸.

Interesting, previous works also proved that is possible to stimulate flavonoid synthesis in low irradiance (kaempferol and quercetin glycosides) when other stresses are present such as lack of nitrogen or salinity¹⁰⁶. It would be interesting in the near future to study the effect of other stresses in the AEP lines phenotype.

3.6 Photo-oxidative Stress Protection

Among the biological functions of anthocyanins, is noteworthy the role as protective components against stressful events leading to ROS, such as high irradiance or pathogen attack ²⁵⁷. The mechanism of action appears to be associated to counteracting the negative effect of ROS, in addition to protect against UV-B damage and to avoid the photoinhibition ^{261,333}. These pigments are considered scavengers of ROS because of their antioxidant properties ³³⁴, and also due to its ability to filter the light in the yellow-green waveband ³³⁵.

It has been reported that anthocyanins decrease the excess of photons absorbed by Chl ²⁹⁵⁻²⁹⁷. Other effective plant strategies such as epoxidation of xanthophylls have been documented ^{164,175} but it showed to be more effective in short term exposure, whereas anthocyanins can compensate the plant for a longer period ³³⁶.

Their role against photoinhibition is still debated, since anthocyanin is not able to absorb in red and blue wavebands which is the maximum wavelengths of Chl ³³⁷; however, anthocyanin compounds are accumulated in internal mesophyll which is an optimal site for UV interception ³³⁸.

Moreover anthocyanin accumulation is an interesting trait to be incorporated in tomatoes since this could increase fruit healthy composition ³³⁹, decrease susceptibility to grey mould ¹⁰⁸, extend the shelf-life ¹³⁸ and decrease detrimental excess of light as in the Yellow Shoulder Disorder (YSD) ^{82,83}.

The two traditional varieties, Pera and Muchamiel, were selected for this study because of their susceptibility inherent to the non-uniform ripening type (green shoulder phenotype). Since, excessive sunlight absorbed by light-harvesting complex

in the shoulder of tomatoes could cause oxidative stress and fruit damage associated to the so called YSD^{90,199,298}. The effect that anthocyanin accumulation in the peel of the engineered varieties could have on UV-shielding, photoinhibition or YSD is currently being investigated in the lab. In addition, improving fruit shelf-life is one of the most important goals for breeders and increase of anthocyanin content in the fruit has been shown to be an effective measure (10). Susceptibility to *B. cinerea* (responsible of tomato gray mold) was dramatically reduced in tomatoes enriched in anthocyanins¹⁰⁸. That could increase even more the commercial and research interest of EAP.

3.7 Conclusion and future perspectives

Our findings show that the simultaneous expression of *SIANT1*, *SIJAF13* and *SIAN11* activate anthocyanin pathway in tomato with no signs of silencing after several generations. To date, ectopic expression of foreign BMW complex genes provided tomato fruits with flesh anthocyanin accumulation. Fruit with restricted purple pigmentation to peel opens new avenues for breeders because it avoids possible negative effect in fruit growth (polar auxin transport), decrease fruit photosynthesis efficiency (photoinhibition) and could even increase consumer acceptance due to maintaining red color in the inner parts.

The tomatoes we have engineered could show an extended postharvest shelf-life, as well as they present benefits in human health due to anthocyanins enrichment. Moreover, the surface photoprotection could allow the use of genetic resources with non-uniform ripening trait, which is associated to better fruit quality but also to increase susceptibility to YSD and fruit-cracking. On the other hand, no previous cases of trichome anthocyanin accumulation in tomato have been reported in the literature. Future uses of these lines include studying the effect of

high levels of anthocyanins in the tomato trichomes, focusing on the high irradiation response or in plant-pest interaction. In addition, it will be interesting to elucidate the molecular basis of the light anthocyanin biosynthesis control; in peel and trichome samples of our AEP.

4. MATERIALS AND METHODS

4.1 Plant materials and growth conditions

Traditional Tomato (*Solanum lycopersicum*) cultivars cv. Pera (variety 7) and Muchamiel (variety 18) were provided by Dr Juan José Ruiz of the Genetic Department of the University Miguel Hernandez, Orihuela, Spain. Genetically Engineered T2 plants were grown in greenhouse located in IHSM-UMA-CSIC (Algarrobo-Costa, Málaga) and The Institute for Plant Molecular and Cell Biology (IBMCP, Valencia) during the spring-summer of 2016. Seedlings were transplanted to 16-L pots at the four true-leaf growth stage; they were watered daily and fed with soluble fertilizer once a week.

The in vitro experiments involved sterilization of seeds under 50% of bleach, rinsed in sterile water, and sowing in one-half strength Murashige and Skoog medium in 1.5% of sucrose, 1% agar, pH 5.8-6. After 3 days pre-germination in darkness plants were cultivated in a growth chamber with 16 h daylight (250 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$) and a relative humidity of 60-70% under 25°C day / 25°C night temperature.

Fruit samples at green stages of development were harvested at the following stages defined by the number of days after anthesis (DPA) 15-20 DPA (Immature green fruits, IMG), 35-40DPA (Mature green, MG). For ripening stages fruits were sampled at Breaker (B), defined by change from green to orange

in the apical area and at Red Ripe (RR) between 7 and 10 days after breaker (depending on the season and the genotype).

4.2 Anthocyanin module construct and plant transformation

The sequences of *Petunia* proteins: *PhAn11* (Genbank accession number AAC18914.1), *PhAn2* (Genbank accession number BAO51604.1), and *PhJAF13* (Genbank accession number AAC39455.1) were used with BLASTP (NCBI) to search homologous tomato proteins. Sequences with the highest homology were considered tomato homologues after confirmation of gene expression (TomExpress). Sequences *SIAN11* (Solyc03g097340.1), *SIANT1* (Solyc10g086260.1), *SJAF13* (Solyc08g081140.2) were amplified from cDNA prepared from total RNA of fruit at breaker stage (*AN11* and *JAF13*) or from genomic DNA (*ANT1*).

To generate the genetic constructs used in this work the Golden Braid (GB) 2.0 cloning system was used¹⁹⁶. The full-length sequence of *SIAN11* (1029 pb), *SIANT1* (1013 pb) and *SJAF13* (1882 pb) had been already domesticated and available in GB-elements Collection (GB 2.0)¹⁹⁶ the identification part numbers are GB0076, GB0077 and GB0078 respectively.

CDSs were introduced in the sense orientation between promoters: PDS (*AN11*) and PLI promoter (*ANT1* and *JAF13*) and terminators (PDS, TFM5, E8) as indicated in Fig. 1. PDS promoter was defined as 1637 pb upstream (Solyc03g123760.2) whereas the size of PLI was 1293 pb (Solyc09g082690.2). E8 promoter and terminator were taken from GB collection with the respective GB numbers GB0914 and GB0144. TFM5 promoter and terminator were domesticated as indicated in M&M of

chapter 1. Primers used for GB parts cloning-domestication are listed in supplementary table 3.

The three anthocyanin TUs and kanamycin resistance TU (*nptII*) were combined in a multipartite GB assembly to create a final construct (12 Kb).

Plasmids were checked at each stage of cloning with restriction digest analysis and by sequencing in the case of the domestication step¹⁹⁶. Anthocyanin module was introduced into plants by *Agrobacterium tumefaciens* (strain LBA4404) mediated transformation cotyledon-explants protocol³⁴⁰. Transformed plants were rooted on 100 mg/L kanamycin and genomic DNA was extracted from young leaves (Doyle and Doyle, 1990) that was used for genotyping by PCR using gene-specific primers (ST. 2); once confirmed to carry the engineering constructs, plants were cultivated in the greenhouse.

4.3 Transgene expression determination by quantitative real time PCR

Total RNA was extracted from ~200 mg of tissue by using TRIzol RNA Isolation Kit (Thermo Fisher Scientific) according to manufacturer's instructions and treated with DNA-free™ DNA Removal Kit - Ambion following the manufacturer's recommendations. The DNA-free RNAs were reverse transcribed with SuperScript II (Takara). Single strand cDNA was then synthesized from 1 µg of RNA in 20 µl with oligo-dT and used for qRT PCR performance in 7500 Fast Quantitative real-time PCR Applied Biosystem. PCR products were amplified in triplicates using SYBR Green kit (Takara) following the manufacturer's instructions. Gene expression was calculated using the relative standard curve method (Applied Biosystems) where actin gene was used as a reference for normalization³⁴¹. The $2^{-\Delta\Delta C_t}$ method

was used to calculate relative changes in gene expression ³⁴². Primers for qRT PCR were designed with Primer Express software (Applied Biosystem) sequences used are displayed in ST. 4. Each sample was analyzed using three biological replicates.

4.4 Sampling method

Samples for RNA extraction were taken at the same time each day, frozen in liquid nitrogen and stored at -80 until required. For each case a pool of 3 fruits at the specified developmental stage were mixed and stored as a biological replicate, each assay was performed with 3 biological replicates

In the case of Trichome heads were isolated by glass bead abrasion and filtering with nylon meshes according to Schillmiller 2010 ³⁰². This resulted in leaf trichome preparations which contained around 80% type VI glandular head cells and 20% of other trichome cells including other glandular and stalks. Similar results have been reported ³²⁰.

Phenylpropanoid and isoprenoid analysis was performed in trichomes and fruits by using different sampling procedures. For metabolite extraction Trichome were collected from liqN2 frozen stems (2cm from the same plant area) by scrapping using the flat-end of a liqN2 frozen plastic spatula. Peel samples were obtained by peeling the fruit by making incisions on the surface and peeling away the tissue with the help of a knife.

4.5 Microscopy characterization

Trichomes: Several organs from both WT and AEP engineered plants were sampled from the greenhouse: flower, stem, branch, leaf and fruit. Tissues were observed under

Magnifier MacroFluo MZZ16F (Leica) with digital camera DFC300 FX (Leica). Images were processed with Fiji software (Schindelin et al., 2012) and used for counting trichomes and identification of the different types of trichomes.

A Confocal Microscope (ZEISS LSM 780 AxioObserver) was used to characterize in more detail the heads of Type VI trichomes after isolation as indicated above. Anthocyanins were recorded by using their characteristic emission at 675 nm when an excitation wavelength of 530 nm was used. Images were taken with digital camera DFC550 (Leica).

Typical DsRed Fluorescence of p35s::*DsRed* tomato lines was detected with Magnifier MacroFluo MZZ16F (Leica) using DsRed (violet/blue excitation) filter with excitation 546/10nm and emission 590nm.

Fruit Peel: For anthocyanin pigment localization in the engineered fruit peel tissues, tomatoes sections were cut longitudinally (aprox. 500 micron thick), placed on slide under a cover slip and observed using a white field

4.6 LC-ESI(+/-)-MS analysis of phenylpropanoids and anthocyanins in tomato fruits and trichomes

Phenylpropanoid extraction from trichome samples was carried out as previously described³⁴³. Briefly, ten mg of ground freeze-dried powder corresponding to the peel of ripe fruit were extracted with 0.75 mL cold 75% (v/v) methanol, 0.1% (v/v) formic acid and spiked with 10 µg/ml formononetin. Samples were vortexed for 30 sec, and shaken at 15 Hz for 15 min using a Mixer Mill 300 (Qiagen) and kept at RT for 5 min (repeat twice). After centrifugation for 15 min at 20,000 g at 4°C, 0.6mL of supernatant was removed and transfer to HPLC tubes. For each

genotype, 4 fruits and at least two independent extractions were performed. Chromatographic separation was carried out using an Ultimate 3000 HPLC coupled to a Q-EXACTIVE mass spectrometer (ThermoFisher) equipped with a C18 Luna reverse-phase column (150 x 2.0mm, 3 m; Phenomenex, Macclesfield, UK) and a gradient system as follows: 95%A:5%B for one minute, followed by a linear gradient to 25%A:75%B over 40 minutes. LC conditions were kept for 2 more minutes, before going back to the initial LC conditions in 18 minutes. Ten μl of each sample were injected and a flow of 0.2 mL was carried out during the whole LC runs.

Detection was performed continuously from 230 to 800 nm with an online Ultimate 3000 photodiode array detector (PDA, Thermo Fischer Scientific, Waltham, MA). All solvents used were LC-MS grade quality (CHROMASOLV[®] from Sigma-Aldrich). The Exactive Plus Orbitrap mass spectrometer was equipped with a heated electrospray probe (H-ESI) operating in both + and - ion modes. ESI and MS parameters were as follows: spray voltage -5.0 kV, sheath gas and auxiliary nitrogen pressures 30 and 10 arbitrary units, respectively; capillary and heater temperatures were set at, respectively 250 and 150 °C, while tube lens voltage was 50 V.

Data were acquired in profile mode. Identification was performed through comparison of chromatographic and spectral properties of authentic standards and reference spectra, and on the basis of the m/z accurate masses, as reported on Pubchem database (<http://pubchem.ncbi.nlm.nih.gov/>) for monoisotopic masses identification, or on Metabolomics Fiehn Lab Mass Spectrometry Adduct Calculator (<http://fiehnlab.ucdavis.edu/staff/kind/Metabolomics/MS-Adduct-Calculator/>) in case of adduct ion detection. Metabolites were quantified relatively on the basis of the internal standard amounts. ANOVA was applied to the absolute metabolite quantifications, and pair-wise comparisons were used in Tukey's

HSD test running on the Past 3.x software
(<http://folk.uio.no/ohammer/past/>).

GENERAL DISCUSSION



In this thesis, we designed new strategies to improve tomato fruit quality based on molecular breeding and the new insights in tomato genetic research. First, we implemented a fruit plastid/photosynthesis strategy to improve fruit metabolite composition, adopting a re-evaluation of the importance of fruit plastid/photosynthesis in tomato red-fruit quality (CEF plants). Second, we provided a strategy to overcome a possible increase in fruit photo-oxidative stress associated to chloroplast/photosynthesis enhancement (AEP), with the additional interest for the healthy composition of non-uniform type of tomatoes. The general phenotype and metabolomics' profile of CEF was discussed in chapter one, while the -omics approaches were discussed in chapter 2. Characterization of AEP was developed and discussed in chapter three. This general discussion is aimed to examine our engineered fruit lines from a common framework.

1. Fruit Photosynthesis

Many structural features indicate that photosynthesis in fruit is special. Fruits lack the typical palisade and spongy cell layers where photosynthesis normally occurs in leaves, they also present none or a reduced number of stomata and the epidermis is covered by a thick cuticle, all this indicating that CO₂ diffusion may find certain difficulties. To overcome these limitations fruit cells have developed a special photosynthesis mechanism, where RUBISCO is not as important for assimilation of atmospheric CO₂ as it is in leaves and it is rather the PEP activity, responsible for CO₂ fixation under fruit environmental conditions²¹. Therefore, fruit photosynthesis could be classified in the intermediate point between C3 and C4/CAM plants. In addition, the importance of mitochondria and chloroplast coordination is of paramount importance in this process¹⁷⁷ (Fig. 1).

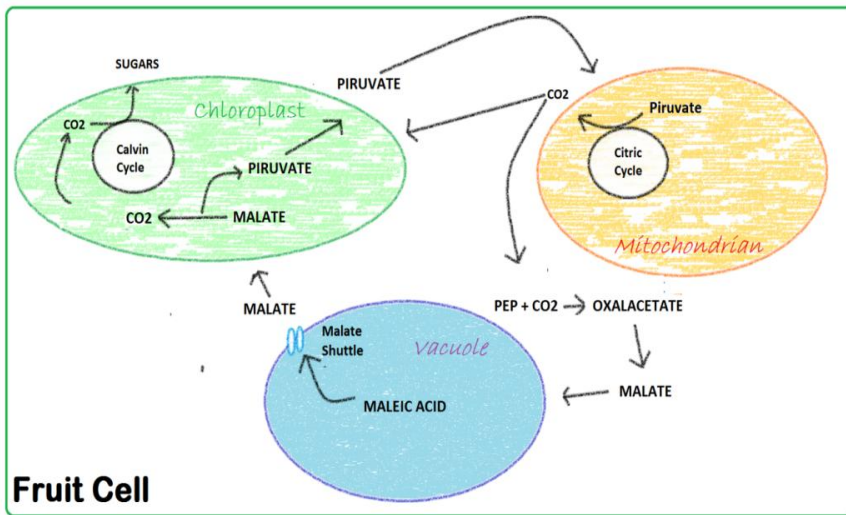


Figure 1: CO₂ re-fixation as it occurs in a fruit pericarp cell. First, PEP captures CO₂ from internal fruit cavities generating oxaloacetate. This is converted into malate and could be stored in vacuoles as organic acids. Second, CO₂ from respiration would be use directly in the chloroplast.

The role of fruit photosynthesis in fruit quality has been much discussed ^{9,166,183,344}. The recently identified tomato *u/u* mutants, exhibit fruits with lower chloroplast number and thylakoid grana and show a decrease of 10-15% in sugar accumulation in red ripe stage, compared to *U/U* genotype. The detriment in fruit quality has been associated to a reduction in fruit photosynthesis capacity ⁵. However, the higher chloroplast development in *U/U* fruits makes them more prone to develop physiological disorders associated to oxidative stresses, mainly after exposure to high irradiance/temperatures conditions. An example for that are 1) the yellow shoulder disorder ⁹⁰, which affects tomatoes carrying the *u* mutation to a lesser extent than wild type *U* varieties ⁹¹ and, 2) cuticle cracking ⁹²⁻⁹⁴, which is more frequent in darker green non-uniform tomatoes ⁸³. Therefore, in modern varieties of tomato the enhanced fruit chloroplast trait has been avoided by breeders to protect fruit from the deleterious consequences of the photooxidative stress.

Interestingly, enhancing photoprotection has been proposed and demonstrated as a useful strategy to improve photosynthesis⁸⁴. Anthocyanin compounds not only act as UV-b filter against high energetic irradiance (absorbing blue light and reflecting red wavelengths), but also are considered a sunscreen protector of PSII^{9,104,105}. In addition of the antioxidant protection, it has been reported the capacity of flavonoids/anthocyanins to extend the shelf life of the fruits¹⁰⁸. Those characteristics make them interesting for fruit breeding with impact in different stakeholder interests¹⁰⁷⁻¹⁰⁹.

2. Early-fruit Specific Overexpression of *GLKs* and *APRR2* TFs: Chloroplast and Photosynthesis Boost with Additive Effect

During the last few years a number of TFs implicated in fruit plastid biogenesis have been identified³⁴⁵. In tomato, *GLKs* have been recently described as co-regulators of photosynthesis-related genes^{5,141}. Likewise, ectopic expression of *APRR2* has been reported to influence fruit plastid development⁸. Constitutive expression of each of those TFs resulted in greener fruits with higher chloroplast development; however, the mechanisms underlying their function are not well understood. In addition, tomato lines overexpressing *GLKs* under 35s promoter, displayed negative effects on fruit development (low fruit set and small fruit, unpublished results).

Aiming to overcome this drawback and include additional novelty steps, we used a spatio-temporal specific promoter to drive the expression of tomato *GLK* and *APRR2*, either separately or by co-expression of both TFs.

2.1 A Multi-level Analysis Confirmed the Chloroplast Boost

In chapter 1 and 2, we reported the characteristic chloroplast enhancement that is taking place in CEF. In previous reports, larger and more numerous plastids, and consequently, higher chlorophyll levels in immature fruits were observed in *APRR2* and *GLK2* constitutively overexpressed (under 35S promoter) in tomato plants ^{5,8}. In general, sCEF lines showed weaker fruit phenotypes than those driven by the 35s, probably due to differences in the strength of the promoter used in each case and in the temporal expression pattern provided by TFM5. Nevertheless, sCEF were always greener than WT and fruit phenotype correlated with an increase of chlorophyll, carotenoids and several primary metabolites concomitant with the upregulation of photosynthesis-related genes, as detected both at the transcript and protein levels. Joint expression of both TF *GLK* and *APRR2*, in dCEF lines, did achieve a fruit chloroplast enhancement similar to those reported in previous works ^{8,50,141}.

Those results, described in chapter 1 and 2, emphasize the reinforcing effect of both TFs co-expression. IMG fruit cells of *GLKs-APRR2* lines presented larger and more abundant chloroplasts (over threefold change) close to the fruit epidermis. In addition, showed to be more structured (more grana stacking in thylakoids) in the area close to the fruit locule.

This together with the wider cell layer with enhanced chloroplast observed from the center inner pericarp to the fruit locule, indicates that *GLKs-APRR2* acts at multiple levels of chloroplast-cell development. Interestingly, this phenotype reminds to *hp1* mutants, which also use a multilevel strategy to achieve a similar enlarged chloroplast area in the fruit pericarp ¹⁴⁵. The increase in plastid number and size was no clearly defined; however, it has been reported that *hp1* increase fivefold

the DNA ratio plastome:genomic in hypocotyls⁶³. This ratio was used as an explanation for the increase in plastid size¹⁴⁵; another suggestion was the regulation of plastid development through cytokinin and photomorphogenic. Previous reports which constitutively overexpress *APRR2* or *GLK2* also observed a modification in chloroplast division and expansion^{5,8}. Our observations in chloroplast development at the cell and structural levels were also supported by isoprenoid analysis (chapter 1) that showed, as in previous works^{5,8}, an increase of chlorophyll/ carotenoid levels in immature CEF.

Additionally, omics- approaches were described in chapter 2, and the results were consistent with the chloroplast boost and the additive effect observed in chapter 1. As in the 35S::*GLK* lines⁵⁰, our transcriptomic data showed that *GLKs* target chlorophyll related genes and Lhc proteins. Processes affected involved, among others, the tetrapyrrole synthesis, photosynthesis-related, and Calvin cycle genes. In addition to that, we confirmed that *SIGLKs* is involved in plastid retrograde signals involved in responses to ROS, ABA, sugar and auxin pathways^{50,222}. Moreover, we found those genes also affected in pTFM5:*APRR2* plants, which lead us to suggest that similarly to *GLKs*, *SIAPRR2* could also be implicated in retrograde signals responses.

Genes related to plastid division were studied but we did not find expression changes in the well-characterized *FtsZ1-1* and *MinD* plastid division-related genes³⁴⁶, neither in sCEF nor dCEF lines similarly to the 35s plants⁵⁰. We found, however, other components of the thylakoid biogenesis machinery in DEG and DEP lists e.g. *FtsHs* proteases, which are also important in chloroplast biogenesis and linked to the degradation of unassembled thylakoid proteins (Ostersetzer and Adam 1997) and more precisely, *FTSH 2* and chloroplastic proteins related to PSII repair were upregulated in our lines.

2.2 Beyond Chloroplasts- Mitochondria and other Related Functions

It is well known that chloroplasts provide mitochondria with compounds involved in photorespiration such as serine and glycine. These compounds are also exchanged with peroxisomes as part of a tri-organelle-signalization cross-talk that seems to be required for efficient use of photosynthesis¹⁷⁷. Chapter 1 showed that the accumulation of metabolites associated to mitochondria organelle, such the higher levels of serine and glycine are affected in s/dCEF, and in chapter 2 an enhancement of mitochondria functions was revealed by the DEP enrichment analysis. Indeed, both mitochondria and chloroplast were the cellular component GO terms-enrichment class for all CEF lines.

Moreover, in chapter 2 we found novel functions affected by joint OE of *GLKs* and *APRR2* TFs that suggest *GLKs* and *APRR2* could target genes beyond chlorophyll and photosynthesis-related functions. Cell Wall Remodeling genes (CWR), were found as affected in our s/dCEF, and more precisely genes encoding xyloglucan-endotransglucosylase/hydrolases (XTH, specifically type 7 and 9). Cell elongation process, typically associated to fruit growth, is characterized by high expression levels of XTH²³⁵. XTH enzymes have been described to act by two different activities: transglucosylase (XET) and hydrolases (XEH) which present different effects in cell wall integrity³⁴⁷. Although, cell enlargement is associated to a cell wall loosening³⁴⁸, it has been speculated that XET activity is able to keep the cell wall integrity participating in cell wall organization process³³². In chapter 2, six genes which encodes XHT were found downregulated (PPadj<0.001) and xyloglucan metabolic process was the biological process affected for all CEF lines. Highlighted is the pleiotropic effect in size observed in 35s::*GLK2* fruits (data not shown) and the slightly decrease of

weight, parallel to the genotype, observed in our sCEF lines, which was higher in dCEF. Whether these XTH expression changes results in actual modifications of cell elasticity / stiffness with consequences in final cell size, and this is responsible for the decrease in final fruit size observed in parallel with the fruit greening phenotype is currently being investigated in our lab.

2.3 Molecular Basis of the *GLK::APRR2* additive Effect

Putative light signal transduction genes were identified to be antagonist in chloroplast regulation, including *DDB1* (*hp1*) and *DET1* (*hp2*) that together with the E3 ligase (*CUL4*)²¹³ take part of the well described 26s proteasome during photomorphogenesis. Higher chloroplast number and chlorophyll content in tomato *hp* mutant is associated to the lack of function of this complex, and it seems to be the connection between the photomorphogenesis mutation effects on fruit quality traits and *GLK2*. Indeed, *GLK2* has been proposed to be substrate of the DET/DDB1/CUL4 E3 ubiquitin ligase^{68,349}. Even more interesting is that *APRR2* gene is upregulated in *hp1* mutants at immature stages⁷⁹. In chapter 1 and 2, we described an additive effect on different fruit phenotypes when *GLKs* and *APRR2* are co-expressed. If both *GLK* and *APRR2* are targeted for ubiquitination and degradation by the same proteasome-system, a joint coexpression under a fruit specific promotor beyond that provided by the endogenous counterparts would explain the more chloroplast development achieved in dCEF. Interestingly in this direction, the DEG lists in CEF included genes related to the CUL4 E3 ubiquitin ligase complex (*DDI1*).

An additive effect was reported between *GLK2* OE and *DDB1* repression. This could suggest that different routes coexist to

regulate chloroplast biogenesis via photomorphogenesis and *GLKs* function¹⁴¹.

Furthermore, to try to understand the mechanism underlying the additive effect described here, we performed a preliminary search for putative binding domains in the promoters of DEG in CEF lines. Previous works described the interaction of *GLKs* with two G-box binding factors³⁵⁰; however, the binding to conserved motifs in the case of *APRR2* remained unknown. We described in chapter 2, a well conserved 7pb motif which was found even duplicated in three *CAB* genes. Interestingly those genes showed higher expression in dCEF than sCEF. Future studies to elucidate this hypothesis will include transactivation assays using *GLK*, *APRR2* or a combination of the two on *CAB* promoter luc/renilla expression system.

2.4 Fruit Photosynthesis Boost Demonstrated through Metabolites Associated

The enhancement of fruit photosynthesis-related functions was supported at different levels in our CEF lines. Chapter 1 showed higher accumulation of photosynthesis-associated primary metabolites. As examples of that we found: Calvin cycle intermediates such as Erythrose-4-P or fructose-6-P, several organic acids and sugars. Previous works associated malic acid levels directly to the particular type of photosynthesis process taking place in the fruit²¹. It has been suggested that PEP is responsible for malic acid synthesis and its activity is predominant in perivascular and placenta fruit tissues¹⁸³. Our four CEF genotypes showed increase in malic acid levels, which is therefore in the line of the photosynthesis boost. Moreover, the impulse in fruit photosynthesis is confirmed by our proteomic profile (PEP, Rubisco protein levels) and transcriptomic data (photosynthesis related genes).

Although previous studies have demonstrated that *AtGLKs* overexpression upregulates genes directly involved in photosynthesis, mainly Lhcb proteins^{50,141}, no proteomic characterization has been performed so far neither in *GLKs* nor in *APRR2* overexpression plants. In chapter 2 we confirmed and analyzed the upregulation of several photosynthesis-related genes, which represent the top of the DEG lists in s/dCEF.

Additionally, we examined the fruit protein profile at IMG stage. Almost all of the DEP in CEF are synthesized as precursors with a chloroplast transit peptide and most of the remaining proteins are characterized by chloroplast or mitochondria sub-localization. The list of DEP revealed many proteins related to photosynthesis function such as light-harvesting complex, PEP and RuBisCO, where ~~*GLKs*~~ *APRR2* fruits presented the highest levels of PEP protein followed by *GLK2*. All this confirms that *GLK/APRR2* work jointly in enhancing chloroplast functions by affecting both the light and dark stages of photosynthesis.

2.5 An Early Chloroplast Boost Impacts on Red Fruit Quality

Chapter 1 analyzed red fruit quality in our CEF. Previous works suggested that a fruit chloroplast/photosynthesis enhancement by using 35s promoters is associated with an increase in organoleptic and nutritional quality^{5,8,43,61}. For example, p35s::*AtGLK* was reported to increase fructose and glucose reaching 21% more TSS in red fruits, as well as to increase lycopene from 10 to 60%^{5,50}. Here, we achieved a similar increase in TSS of the red fruits when expression was driven with an early fruit promoter, but only in the case of *GLKs-APRR2* (23-25%). sCEF lines did not show significant changes in sugar levels at ripening, again indicating that it was necessary the additive effect in dCEF to have a significant effect in sugar accumulation.

However, lycopene levels increase up to double either in sCEF or dCEF lines, indicating that the effect on plastid metabolism is more sensitive to our engineering strategy.

To date, how enhancement of *GLK/APRR2* TFs function early in fruit development can affect fruit traits at the red ripe stage is not well understood. Previous studies⁵⁷ have demonstrated the importance of starch content at IMG/MG stage for soluble solids and sugar content of the mature fruits. In chapter 1, our primary metabolite results supported this hypothesis, as dCEF at IMG present more sugar content and more starch. This data reinforces the hypothesis that is possible to improve fruit quality by just acting on chloroplast development, at earlier stages of fruit development, even at IMG stage.

Additionally, it is important to mention that we observed a decrease in fruit weight (MG) parallel to the green intensity (IMG) of our CEF lines and therefore, also correlated with the level of sugar in red ripe fruits (darker genotypes showed more Brix). Previous reports, which study tomato genetic diversity, found that high fruit weight is correlated to low TSS and carotenoids³⁵¹.

3. A Possible Solution for Tomato Fruit Photooxidative Stress

In chapter 1 and 2, we followed an all-tomato intragenesis approach to improve chloroplast in tomato fruits aimed to increase fruit photosynthesis and red fruit quality. However, since an excessive sunlight absorbed by light-harvesting complex in tomato fruits could cause oxidative stress and fruit damage associated to the so called YSD^{90,199,298} or cuticle cracking⁹²⁻⁹⁴, our next goal was reinforce the fruit with a protective mechanism against oxidative stress, thereby to avoid possible fruit impairments (chapter 3).

Chapter 3 described the generation of AEP lines, our designed strategy to engineer fruits to accumulate anthocyanins in peel. Anthocyanins compounds are also considered an interesting trait for tomatoes, since this could increase fruit in healthy metabolite composition ³³⁹, decrease susceptibility to grey mould ¹⁰⁸, extend the shelf-life ¹³⁸ and decrease detrimental effect of excess of light as in the Yellow Shoulder Disorder (YSD) ^{82,83}.

AEP showed a fruit light dependent anthocyanin accumulation that could be interesting for basic and applied research studies since the "light protection" is regulated by the same stressor (light). Although we don't know in detail the basis of the light regulation mechanism, it provides a way to adjust the anthocyanin accumulation ³³² and modify flavonoid composition in the fruit peel almost at will, e.g. with lower light it is possible to have red tomatoes with higher flavonoid levels in the peel, whereas at higher light intensity fruits can accumulate anthocyanin compounds, over 30-150 fold higher than they do under dim light conditions. Additionally, fruits from AEP grown under lower light display enrichment in other flavonoids.

In this thesis, we generated engineered tomato lines which overexpress a Tomato BMV complex in the background of traditional varieties (U/U background). Previous to this work, transgenic approaches involving the overexpression of Snapdragon genes (MYB and bHLH) in cv. Microtom accumulated anthocyanins in both peel and flesh ^{136,138}. Here, we achieved tomatoes enriched in anthocyanin with a peel restricted accumulation, as well as engineered plants with only tomato DNA.

The success of our approach relies in the choice of the promoter to control the expression of MYB and bHLH. We confirmed in this work that PLI drives transgene expression to the outer pericarp tissues ^{137,300}.

The restricted fruit peel accumulation, described in chapter 3, presents some advantages for breeders. For example, the regular red color in the inner fruit parts could increase the consumer acceptance. Moreover, evaluation of shelf life trait in *Aft/atv* genotype confirmed that accumulation of anthocyanins in the skin is sufficient to delay the post-harvest over-ripening and to extend the life span¹³⁸. In addition to the demonstrated effect on photooxidative UV_VIS stress³³⁷.

Among the biological functions of anthocyanins, is noteworthy the role as protective components against stressful events leading to ROS; such as high irradiance or pathogen attack²⁵⁷. It has been reported that anthocyanins decrease the excess of photons absorbed by Chl²⁹⁵⁻²⁹⁷. The effect of anthocyanin accumulation in the fruit peel of AEP against UV-shielding, photoinhibition or YSD is currently being investigated in the lab.

3.1 Novel Anthocyanin Trichome Phenotype

To date, no previous cases of anthocyanin accumulation in tomato trichomes have been reported in the literature. We have shown that PLI promoter drives specific TFs expression to trichome to high enough levels to produce a strong phenotype (which was not the case of 35s promoter). We also described expression specificity in type VI trichome for our anthocyanin module in tomato cultivated background (type I, V and VI). Subsequently, analysis by crossing with other genetic backgrounds, which present additional types of glandular trichomes (type IV and VII trichomes), also showed expression in the case type VII trichomes.

Anthocyanin in trichomes could be used as an extra protection to high irradiance, combining the efficiency of trichomes in this issue with the sunscreen and antioxidant capacities of anthocyanins. In addition, the importance of this novel protective trait relies in the chemical secretory capacity of plant trichomes which make them suitable as biofactories for the synthesis of natural products and also in its ability to provide protection against pests.

4. Future perspectives for CEF and AEP

Our CEF re-evaluates the importance of fruit photosynthesis, supported by the contribution of this process to the fruit carbon economy and, therefore, to the tomato red fruit quality. Although previous works have been reported that use similar chloroplast enhancement strategy, CEF are the first designed based on an intragenesis approach, and therefore with higher consumer acceptance.

Most of the tomato cultivated varieties that fill our groceries carry the *u* mutation, and therefore have defective fruit chloroplasts, with the associated effects of lower sugars and lycopene levels. Breeders introduced this mutation in 1953 as *u* fruits are less prone to develop certain photooxidative injuries. In this work, AEP strategy opens new avenues for breeding tomato for higher content of sugars in fruit, as the CEF lines could benefit by combining it with the ability to strongly accumulate anthocyanins in peel of our AEP lines. Anthocyanins have been demonstrated to act, not only as UV-b filter against high energy radiation, or as a photooxidative stress to protect PSII in high irradiance, but also to provide extended postharvest shelf-life in tomato fruit ¹⁰⁸. Our *GLK2-APRR2* CEF line has been already combined with our anthocyanin module and is currently being investigated in our lab.

The restricted anthocyanin accumulation to the fruit peel has interest for consumer and others stakeholder in the tomato food chain. That could prevent possible negative effect in fruit growth (polar auxin transport) or in the decrease of fruit photosynthesis efficiency (by photoinhibition). Additionally to the previously described regular red color in the inner parts, the benefit of anthocyanins in human health and the use of only tomato DNA in our engineering strategy.

In addition to that, our CEF and AEP lines generated new knowledge that will be the basis of future studies. We show here some interesting examples, such as the additive effect that we found with *APRR2* and *GLK2* co-expression, which could be the start point to deeply elucidate the network signaling which control chloroplast biogenesis and development. CWR function affected in CEF could help to assess the activity of XTH in tomato, to date, poorly studied. In addition, the novel trichome phenotype we have generated can be used in future studies aimed to define the possible use of anthocyanins as a UV protective filter or to increase plant photoprotection, as well as studies which involve plant-pest interaction.

Tomato is one of the most important cultivated vegetables worldwide and a source of antioxidants in a regular human diet. By combining CEF and AEP strategies in the future we aim to improve fruit quality while endowing the fruit with protective mechanisms to improve its performance in the field and post-harvest life.

CONCLUSION



C1. Tomato plants (*S. Lycopersicum* var. MoneyMaker) engineered to express *SIGLK1*, *SIGLK2* and *SIAPRR2* transcription factors (TFs), either separately or in combination, during early fruit development, showed a range of chloroplast enhancement phenotypes in their fruit. As a result, Chloroplast Engineered Fruit (CEF) accumulated more sugars, carotenoids and certain specific volatiles than their wild type counterparts, affecting the final ripe fruit quality. Moreover, those lines expressing combinations of *SIAPRR2* and *SIGLK*, showed an additive effect in the accumulation of the same compounds.

C2. Chloroplast number, size and total plastid surface are increased in CEF lines. Modifications in fruit photosynthesis-related gene expression, protein profile and metabolite accumulation, support a general activation of fruit photosynthesis in all engineered lines, which affects both light and dark phases.

C3. Two traditional varieties of tomato (Muchamiel and Pera) were engineered with the tomato BMW multigene construct regulating anthocyanin biosynthesis, which comprises the SIJAF13 Basic Helix-Loop-Helix (BHLH) and the SIANT1 (MYB) TFs, both under the control of the fruit epidermis-specific PLI promoter, next to the SIAN11 (WD40) TF regulated by the fruit-ripening PDS promoter. The resulting Anthocyanin Engineered Plants (AEP) accumulated high levels of flavonoids in the fruit peel. The flavonoid composition changes with light intensity: fruit exposed to high light conditions are enriched mainly in anthocyanins, whereas dim light regimes induced the accumulation of other upstream components of the flavonoids biosynthetic pathway.

C4. Plants carrying the BMW multigene construct accumulate high levels of anthocyanins in the glandular head cells of type VI trichomes, as these cells divert their endogenously active

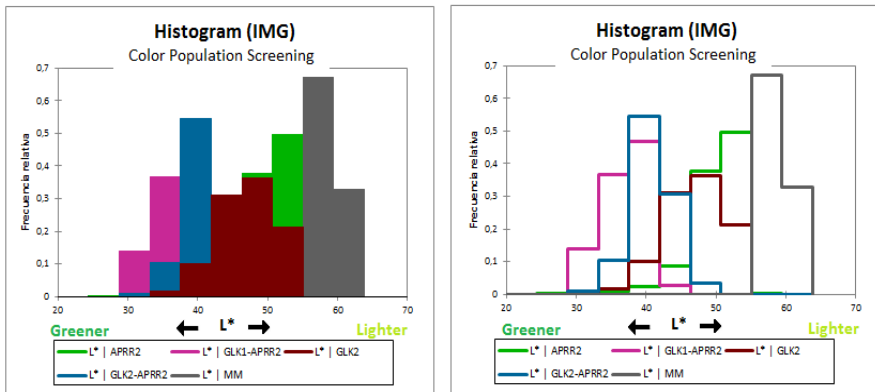
Conclusions

phenylpropanoid metabolism towards the synthesis of anthocyanins. AEP lines are a useful tool to assess the effect of anthocyanin accumulation on plant growth under UV/ high light stresses, as well as for the study of plant-pest interactions.

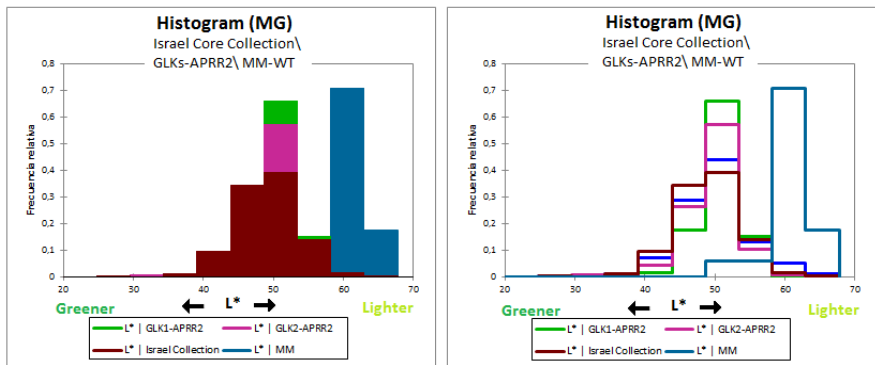
SUPPLEMENTARY FIGURES



Supplementary figures

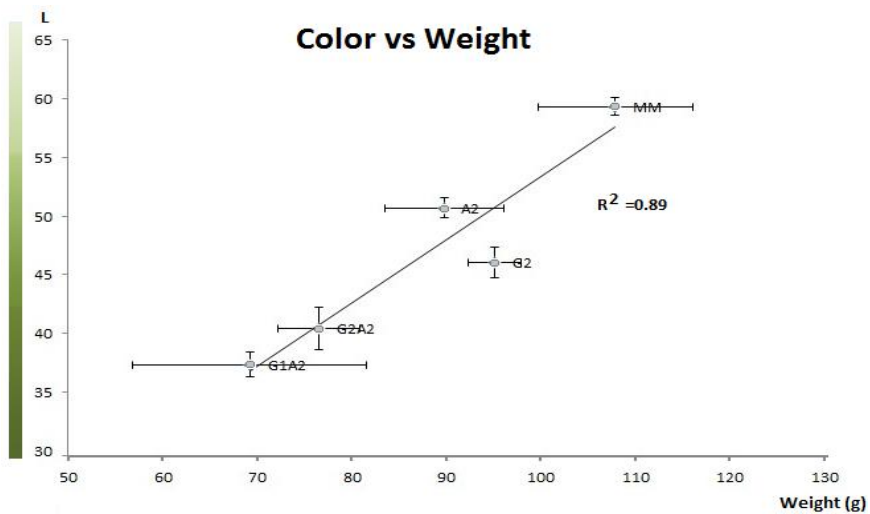


Supplementary Figure 1: Color range is overlapping between populations showing the similarities in phenotypes according to the genotype. Histogram (XLSTAT, 2015) represents the color pattern in each genotype at IMG stage, 1053 observations are represented, the lighter color is 63 (luminosity) and the darker 27; global average is 47 and standard deviation 7.29. On the left image shows overlapping population, on the right is possible to appreciate the range for each genotype: *GLK1-APRR2* (pink), *GLK2-APRR2* (blue), *GLK2* (red), *APRR2* (green) and *MM* (gray). Around 500 plants were used for this screening; color was taken on the surface of the fruits as is described in M&M.



Supplementary Figure 2: Color variability in a natural population reveals GLKs-APRR2 still being greener at MG stage. Histogram (XLSTAT, 2015) represents the color variability in a natural population of 560 accessions (Israel Core Collection), *GLKs-APRR2* and *MM* at MG stage. 881 observations are represented in the graphic, 560 comes from the average of each line (where in total 10.000 measurements were taken), the rest of them from *GLKs-APRR2* and *MM* (continued)

Supplementary Figure 2 (continued): The lighter color is 67 (luminosity) and the darker 27; global average is 50 and standard deviation 4.8. On the left image shows overlapping population, on the right is possible to appreciate the range for each genotype: *GLK1-APRR2* (green), *GLK2-APRR2* (pink), MM (blue) and Israel Core Collection (red). Color was taken on the surface of the fruits as is described in M&M.



Supplementary Figure 3: Relationship between green intensity in IMG fruit and final weight at MG stage according to the CEF line. Weight was measured on the second trusses of T3 plants (M&M). Weight represented in grams (g) and color by Luminosity (L) (higher values lighter green, smaller values darker green). The color scale bar is an approximation of the real color. In the graph are represented the averages (\pm SD) for CEF genotypes (G1A2, G2A2, A2, G2) and MM (WT).

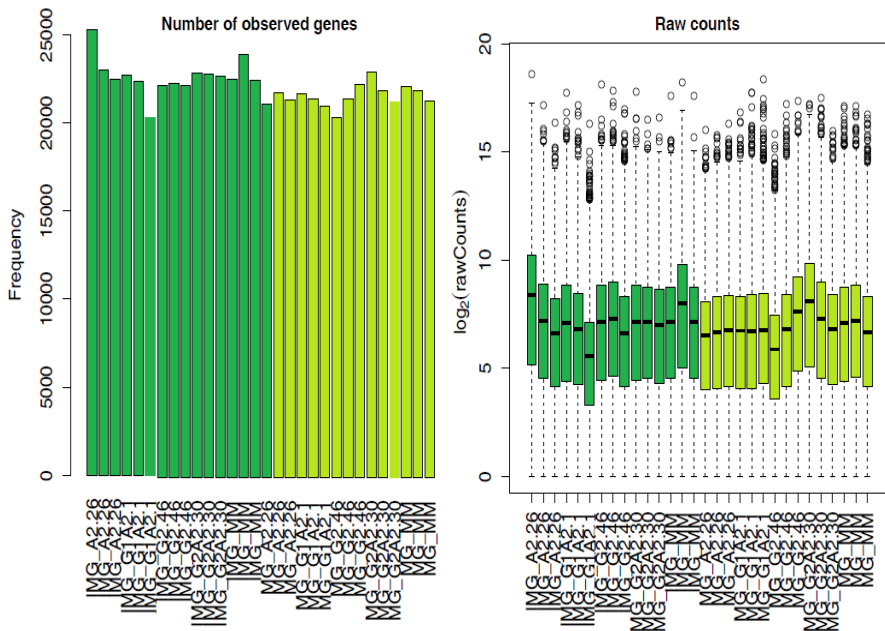
MM



G2A2

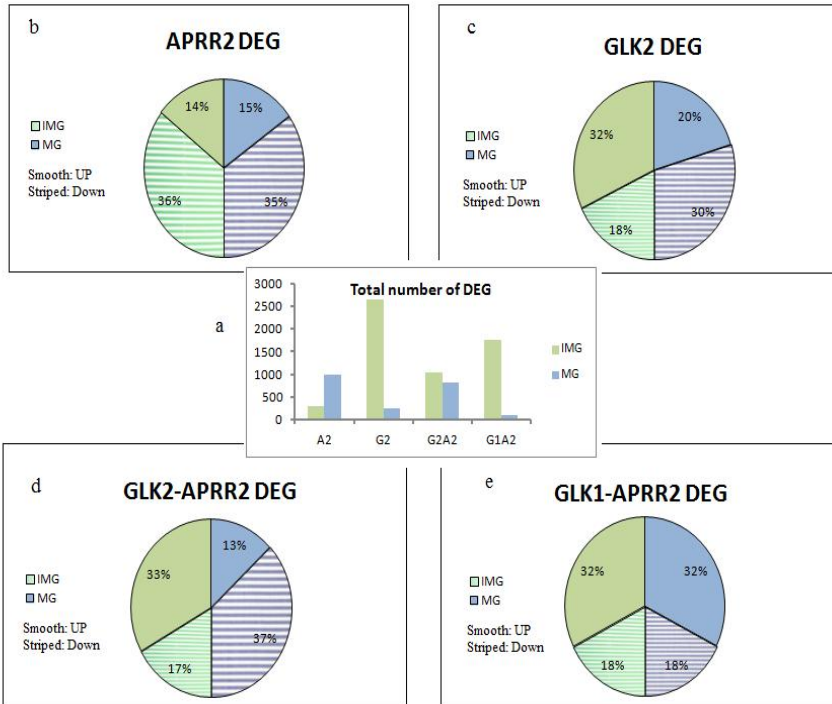


Supplementary Figure 4: Higher amounts of starch storage in G2A2 fruits at MG stage. Starch iodine staining pattern used to estimate in a qualitative manner the starch content in MM-WT fruits and G2A2 fruits at MG stage of development.



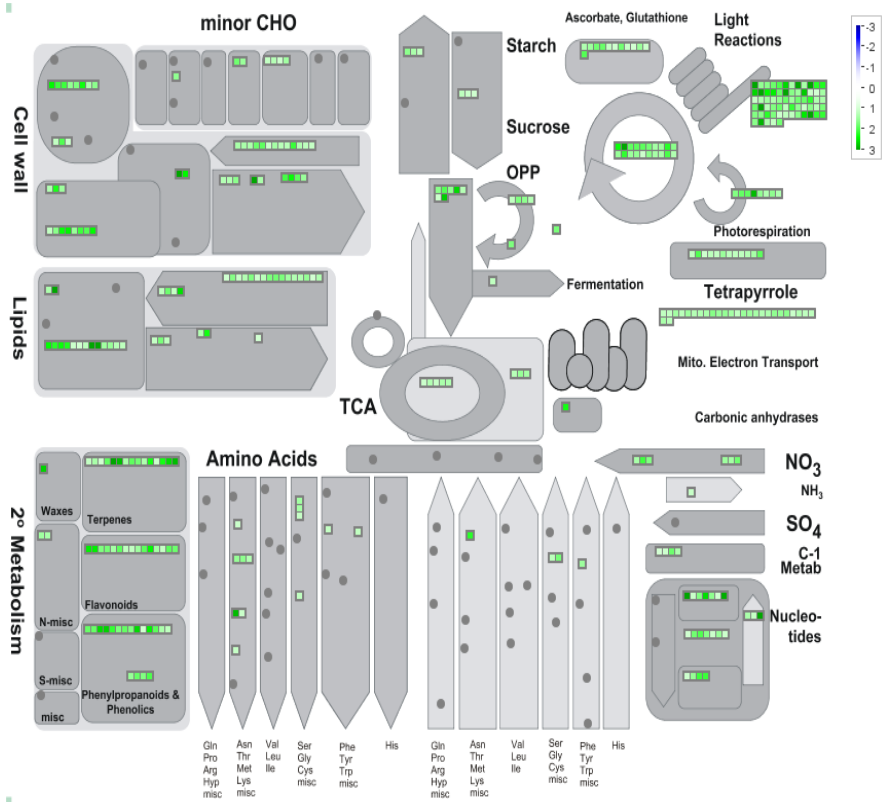
Supplementary Figure 5: Similar number of observed genes and raw counts in the 5 analyzed genotypes (continued).

Supplementary Figure 5 (continued): Code for each sample is made by the stage of development (Immature green "IMG" or Mature green "MG") and genotype (*APRR2* "A2", *GLK2* "G2", *GLK2-APRR2* "G2A2" and *GLK1-APRR2* "G1A2"). On the left, number of observed genes (genes with at least one read); all samples have between 20000 and 25000 non-zero genes. On the right, represents the log₂ of raw counts normalized by the library sizes for each biological replicate.

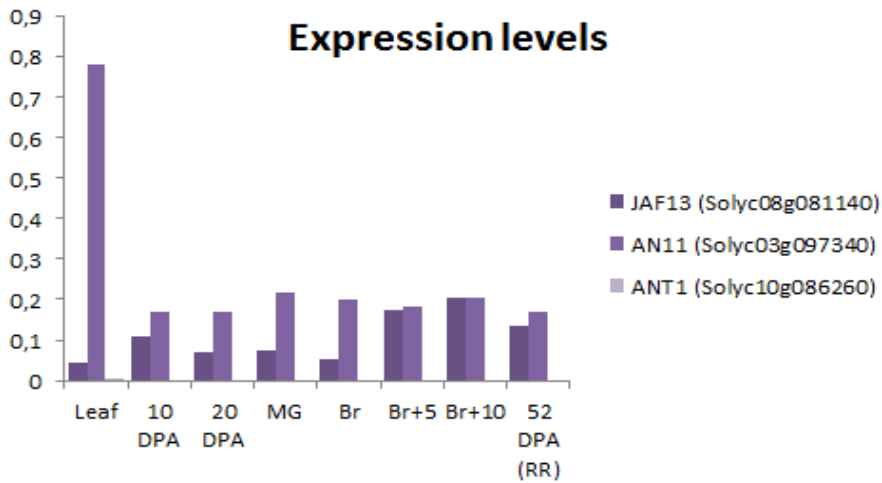


Supplementary Figure 6: GLKs influence mainly at IMG stage whereas APRR2 exert its primary influence at IMG stage manifesting the effect in MG stage. (a) Numbers of differentially expressed genes (DEGs) identified in tomato transgenic lines. The DEGs were identified using $P_{adj} \leq 0.01$ and a Log₂ fold-change >1 and <-1 . Gene expressions analysis based on RNASeq data. **(b)** *APRR2* DEG profile (up and down) at IMG and MG stages. **(c)** *GLK2* DEG profile (up and down) at IMG and MG stages. **(d)** *GLK2-APRR2* DEG profile (up and down) at IMG and MG stages. **(e)** *GLK1-APRR2* DEG profile (up and down) at IMG and MG stages.

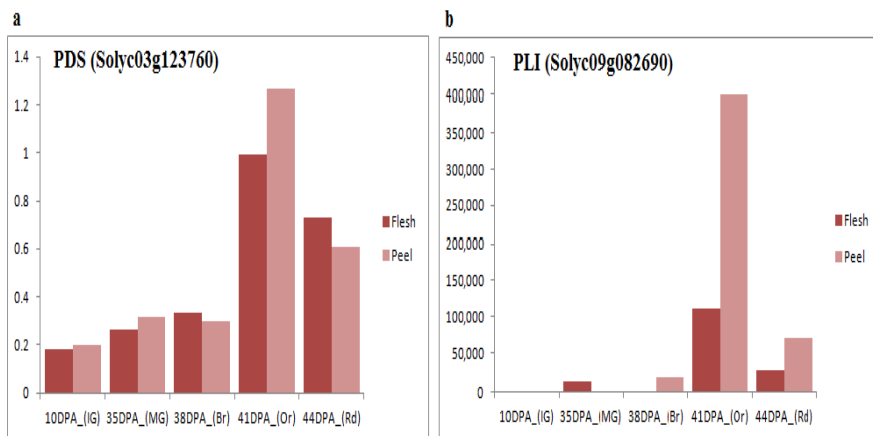
Supplementary figures



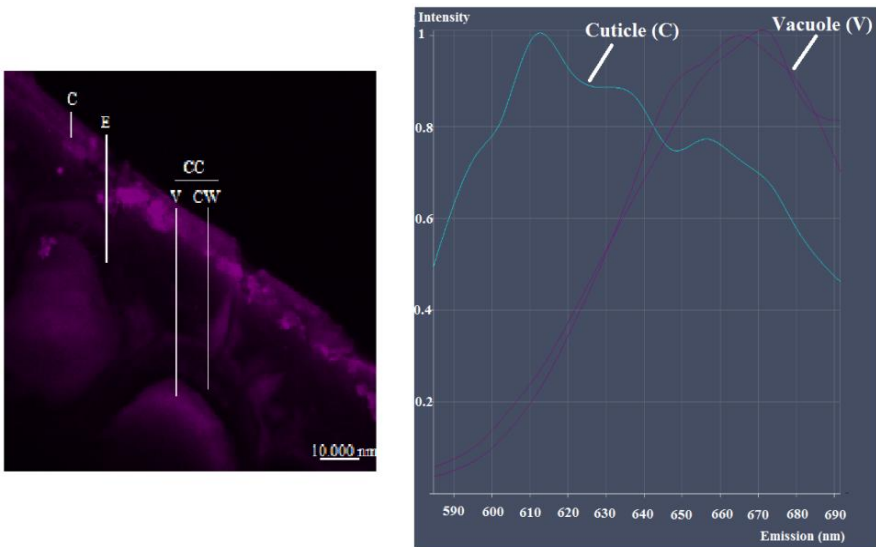
Supplementary Figure 7: Photosynthesis, Nitrogen-metabolism and stress are some of the modified pathway as a consequence of GLK2 expression. Mapping the list of 1672 exclusively DEG for *GLK2* at IMG stage of fruit development against tomato metabolism. This snapshot was generated using MapMan software (metabolism overview), mapped with ITAG2.3. green means higher expression.



Supplementary Figure 8: *AN11* and *JAF13* displays significant levels of expression during fruit development, whereas, *ANT1* is practically inexistent. TomExpress database has been used to determined expression patterns in fruit tissues at different development stages and leaves of tomato orthologues: *JAF13*, *AN11* and *ANT1*. Data comes from middle section of the fruit and Ailsa Craig cultivar (<http://gbf.toulouse.inra.fr/tomexpress17b>).



Supplementary Figure 9: *PDS* and *PLI* are ripening fruit promoters with expression in peel fruit tissue. TomExpress database has been used to determined expression patterns in fruit tissues at different development stages of *PDS* and *PLI* genes. Data comes from peel and flesh of *Microtom* cultivar (<http://gbf.toulouse.inra.fr/tomexpress17b>).



Supplementary Figure 10: Anthocyanin seems to be accumulated in vacuoles of collenchyma cells. C: Cuticle, E: Epidermis, CC: collenchymas Cells, V: Vacuole, CW: Cell Wall. Confocal image of EAP peel sample (left) excitation at 530 nm. Emission spectrum (right) of soluble anthocyanin compounds in vacuoles at 670nm (purple curve) and autofluorescence in cuticle at 610nm (light-blue curve). Unique vacuoles from collenchyma cells emitted at anthocyanin described spectrum (670nm).

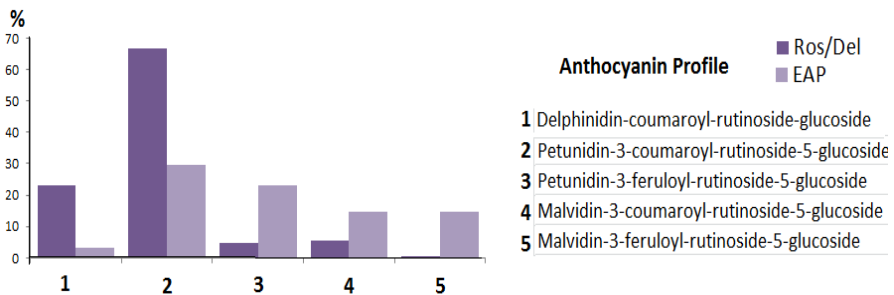


Supplementary Figure 11: Unexpected plant phenotype in trichomes. (continued)

Supplementary Figure 11 (continued): (a) Trichomes on tomato engineered plants contain purple pigment that is released when the plant is touched **(b)** Pigment shows hydrosoluble properties correlated to anthocyanins compounds removing with water (Eppendorf) **(C)** Magnification of trichomes which accumulates purple pigments. Type VI trichome heads were isolated from EAP steams by glass bead abrasion and filtering with nylon meshes (M&M)



Supplementary Figure 12: Fruit of AEP displays a light-dependent phenotype. Fruit anthocyanin accumulation decreases in areas exposed to lower light (non-uniform anthocyanin accumulation on the left). Different intensities of purple pigmentation were observed in the same plant (on the right).



Supplementary Figure 13: AEP fruits display higher levels of Malvidin than Rosea/ Delila lines. (continued)

Supplementary Figure 13 (continued): Comparison between AEP and Rosea/Delila (Ros/Del) fruits anthocyanin profile. Each component is represented as a percentage of the total and numbered in the graphic bar. Data from Ros/Del plants comes from Xiaoyu-Su et al. (2012).

SUPPLEMENTARY TABLES



Supplementary tables

Patch	Forward Primer	Reverse Primer
pTFM5 Patch1	GCGCCG TC <u>TCG</u> CTCG GGAGTACTTGCTTCCATTAGCTTAG ATAAAGTATTAAGTTATTTCGAACATAATAGGTTCTTGAGGC	GCGCCG TC <u>TCG</u> CTCG CATTGGAAGCTAAAGCACACAAATCGTTGCA
tTFM5 Patch1	GCGCCG TC <u>TCG</u> CTCG CTTGGCATTCTGCTCTTCTCTATATAT	GCGCCG TC <u>TCG</u> CTCG AGCGTTTGAGGCTTCATTGGTTTACT
GLK2 Patch1	GCGCCG TC <u>TCG</u> CTCG AATGCTTGCTCTATCTTCATCATTGA	GCGCCG TC <u>TCG</u> CTCG AGGCATGAGTTGCCATAAGCTCCAAAATT
GLK2 Patch2	GCGCCG TC <u>TCG</u> CTCG CTACTAGACATAACATTGCTAGTCATCTTC	GCGCCG TC <u>TCG</u> CTCG AAGCTCAAGTTGGGGTATTTTGGTAAT
GLK1 Patch1	GCGCCG TC <u>TCG</u> CTCG AATGCTAGTTGTGCACCTTTTAGTAACACAACG	GCGCCG TC <u>TCG</u> CTCG AATGGCGAAGTGCTAGATGT
GLK1 Patch2	GCGCCG TC <u>TCG</u> CTCG ATCTCCACAACATCCATCACCAGCAT	GCGCCG TC <u>TCG</u> CTCG AAGCCTATGCACAAGTTGGTATTTTAGGT
APRR2 Patch1	GCGCCG TC <u>TCG</u> CTCG AATGATTGCTTGGAGAAATATTGGGT	GCGCCG TC <u>TCG</u> CTCG AAGCTCATCTCCGACAACGGGAGT

Supplementary Table 1: Primers for GB 2.0 domestication process. NTI Designed primers (NTI10 Software) for promoter: pTFM5, terminator: tTFM5 and CDSs: SIGLK2, SIGLK1 and SIAPRR2, BsmBI recognition site is represented in bold, whereas cleavage is underlined.

Essay	Sequences	Forward Primer	Reverse Primer
Genotyping - PCR	TFM5:APRR2	GCTTTAGCTTCCAATGATTTGCA	GCGCAATACATTTTCATCATGGTACT
	TFM5:GLK1	AGCTTCCAATGCTAGTTGTGCACC	GAAGCTGAAACCGAGCCTGTT
	TFM5:GLK2	TTTAGCTTCCAATGCTTGCTCTATC	GGGACTGCCTTATCAACACCTAATT
qRT-PCR	Actin7	CCTCAGCACATTCCAGCAG	CCACCAAACCTTCTCCATCCC
	TFM5:APRR2	TCAACGGAGAGCGTTCTTGA	GAGCAGAATGCCAAGCTCATC
	TFM5:GLK2	CCCCTCAAATGAGAGCATAGATG	GCAGAATGCCAAGCTCAAGTT
	TFM5:GLK1	ACCACTTCCCCTAGGACTCAAAC	CAGAATGCCAAGCCTATGCA

Supplementary Table 2: QRT and PCR primers. Primers used for PCR (dark green) and qRT-PCR (light green)

Supplementary tables

Patch	Forward Primer	Reverse Primer
pPDS Patch1	GCGCC <i>GTCTC</i> <u>GCTCG</u> GCTTCTTTGTGGTTATTATTAGCT	GCGCC <i>GTCTC</i> <u>GCTCG</u> AGCGTAAAGTTACCCATCCCCA
tPDS Patch1	GCGCC <i>GTCTC</i> <u>GCTCG</u> GCTTCTTTGTGGTTATTATTAGCT	GCGCC <i>GTCTC</i> <u>GCTCG</u> AGCGTAAAGTTACCCATCCCCA
pPLI Patch1	GCGCC <i>GTCTC</i> <u>GCTCG</u> GGAGTTCATCTGCAATACTTCGGAC	GCGCC <i>GTCTC</i> GAGCGCCACATAGTCATGG
pPLI Patch2	GCGCC <i>GTCTC</i> GGCTGAGTTACGCAGATACGTG	GCGCC <i>GTCTC</i> GATGGCGTTTTATAGGGCCC
pPLI Patch3	GCGCC <i>GTCTC</i> GCCATGCGAATTAACCGGAGC	GCGCC <i>GTCTC</i> <u>GCTCG</u> ATGGTGAAGAAGTACTAGTATTGTATAGATCGCG

Supplementary Table 3: Primers for GB 2.0 domestication process. NTI Designed primers (NTI10 Software) for promoter: pPLI and pPDS terminator: tPDS, BsmBI recognition site is represented in bold, whereas cleavage is underlined and overhand in italics.

Essay	Sequences	Forward Primer	Reverse Primer
Genotyping - PCR	ANT1	GCCACATCACCTTGTACCT	CTATTCAGACCTGCACGTAC
	JAF13	GGATTGCCATACAGTTCAGTCTTC	CGGAAATTGAGTTGAAGGCAC
	AN11	GGAGTTGCCAACTGTTGCAG	GTAAAGACTGGAAGACACGGTTCA
qRT-PCR	Actin7	CCTCAGCACATTCCAGCAG	CCACCAAACCTTCTCCATCCC
	ANT1	GATTTTTCTGCTGAAATTGACTTATGG	CAACTGGTTGACGAATCTGCTTA
	JAF13	GGGATTCTATCCATTGCTATCGA	TCAAGCTCAAGATTTCCATACTACTCTC
	AN11	CTGCACAGCGTGATTGGATT	CCACAAGAAGTCCATACTTTAAGCA

Supplementary Table 4: QRT and PCR primers. Primers used for PCR (dark green) and qRT-PCR (light green)

CROSSES	Type I	Type IV	Type V	Type VI	Type VII
AEP cv Pera	+	-	++++	++	+
AEP cv Muchamiel	+	-	++++	++	+
BC5	+	++	+++	+++	++
<i>S. pimpinellifolium</i> TO-937	++	++++	-	+	+
<i>S. habrochaites</i> f. <i>glabratum</i> PI 134417	+	++++	-	++++	+
AEP cv Pera x TO-937 (F1)	+	+	+++	++	++
AEP cv Muchamiel x TO-937 (F1)	+	+	++++	++	++
AEP cv Pera x BC5 (F1)	+	+	++++	++	++
AEP cv Pera x PI 134417 (F1)	+	++	+++	++ +	+ +

Supplementary Table 5: Type VI and VII trichome showed anthocyanin accumulation. Relative abundance of glandular type I, IV, VI and VII and non-glandular V trichomes is represented in the table (+). Purple-highlighted indicates anthocyanin accumulation in the glandular head of the trichomes. Genotypes evaluated: AEP (Pera and Muchamiel), BC5 (MM x *S. pimpinellifolium*), *S. habrochaites* (PI 134417) and respective crosses.

REFERENCES



- 1 Klee, H. J. & Giovannoni, J. J. Genetics and control of tomato fruit ripening and quality attributes. *Annu Rev Genet* **45**, 41-59, doi:10.1146/annurev-genet-110410-132507 (2011).
- 2 Lytovchenko, A. *et al.* Tomato fruit photosynthesis is seemingly unimportant in primary metabolism and ripening but plays a considerable role in seed development. *Plant Physiol* **157**, 1650-1663, doi:10.1104/pp.111.186874 (2011).
- 3 Causse, M. *et al.* Consumer preferences for fresh tomato at the European scale: a common segmentation on taste and firmness. *J Food Sci* **75**, S531-541, doi:10.1111/j.1750-3841.2010.01841.x (2010).
- 4 Gillaspay, G., Ben-David, H. & Gruissem, W. Fruits: A Developmental Perspective. *Plant Cell* **5**, 1439-1451, doi:10.1105/tpc.5.10.1439 (1993).
- 5 Powell, A. L. *et al.* Uniform ripening encodes a Golden 2-like transcription factor regulating tomato fruit chloroplast development. *Science* **336**, 1711-1715, doi:10.1126/science.1222218 (2012).
- 6 Seymour, G. B., Chapman, N. H., Chew, B. L. & Rose, J. K. Regulation of ripening and opportunities for control in tomato and other fruits. *Plant Biotechnol J* **11**, 269-278, doi:10.1111/j.1467-7652.2012.00738.x (2013).
- 7 Sagar, M. *et al.* SlARF4, an auxin response factor involved in the control of sugar metabolism during tomato fruit development. *Plant Physiol* **161**, 1362-1374, doi:10.1104/pp.113.213843 (2013).
- 8 Pan, Y. *et al.* Network inference analysis identifies an APRR2-like gene linked to pigment accumulation in tomato and pepper fruits. *Plant Physiol* **161**, 1476-1485, doi:10.1104/pp.112.212654 (2013).
- 9 Hetherington, A. M. The control of specificity in guard cell signal transduction. *Philosophical Transactions of the Royal Society B: Biological Sciences* **353**, 1489-1494, doi:10.1098/rstb.1998.0304 (1998).

- 10 Carrara, S., Pardossi, A., Soldatini, G., Tognoni, F. & Guidi, L. Photosynthetic activity of ripening tomato fruit. *Photosynthetica* **39**, 75-78 (2001).
- 11 Manzara, T., Carrasco, P. & Gruissem, W. Developmental and organ-specific changes in DNA-protein interactions in the tomato *rbcS1*, *rbcS2* and *rbcS3A* promoter regions. *Plant Mol Biol* **21**, 69-88 (1993).
- 12 Sugita, M. & Gruissem, W. Developmental, organ-specific, and light-dependent expression of the tomato ribulose-1,5-bisphosphate carboxylase small subunit gene family. *Proc Natl Acad Sci U S A* **84**, 7104-7108 (1987).
- 13 Wanner, L. A. & Gruissem, W. Expression dynamics of the tomato *rbcS* gene family during development. *Plant Cell* **3**, 1289-1303, doi:10.1105/tpc.3.12.1289 (1991).
- 14 Piechulla, B., Glick, R. E., Bahl, H., Melis, A. & Gruissem, W. Changes in Photosynthetic Capacity and Photosynthetic Protein Pattern during Tomato Fruit Ripening. *Plant Physiol* **84**, 911-917 (1987).
- 15 Piechulla, B. & Gruissem, W. Diurnal mRNA fluctuations of nuclear and plastid genes in developing tomato fruits. *EMBO J* **6**, 3593-3599 (1987).
- 16 Nielsen, M. T., Akers, C. P., Järlfors, U. E., Wagner, G. J. & Berger, S. Comparative Ultrastructural Features of Secreting and Nonsecreting Glandular Trichomes of Two Genotypes of *Nicotiana tabacum* L. *Botanical Gazette* **152**, 13-22, doi:10.1086/337858 (1991).
- 17 Pyke, K. A. & Howells, C. A. Plastid and stromule morphogenesis in tomato. *Annals of Botany* **90**, 559-566 (2002).
- 18 Carrari, F. *et al.* Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. *Plant Physiol* **142**, 1380-1396, doi:10.1104/pp.106.088534 (2006).
- 19 Steinhauser, M. C. *et al.* Enzyme activity profiles during fruit development in tomato cultivars and *Solanum pennellii*. *Plant Physiol* **153**, 80-98, doi:10.1104/pp.110.154336 (2010).
- 20 Schaffer, A. A. & Petreikov, M. Sucrose-to-starch metabolism in tomato fruit undergoing transient starch accumulation. *Plant Physiology* **113**, 739-746 (1997).

- 21 Blanke, M. M. & Lenz, F. Fruit photosynthesis. *Plant, Cell & Environment* **12**, 31-46, doi:10.1111/j.1365-3040.1989.tb01914.x (1989).
- 22 Lemaire-Chamley, M. *et al.* Changes in transcriptional profiles are associated with early fruit tissue specialization in tomato. *Plant Physiol* **139**, 750-769, doi:10.1104/pp.105.063719 (2005).
- 23 Laval-Martin, D. Light versus Dark Carbon Metabolism in Cherry Tomato Fruits: I. Occurrence of Photosynthesis. Study of the Intermediates. *Plant Physiol* **60**, 872-876 (1977).
- 24 Zantor, M. I. *et al.* Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents. *J Exp Bot* **60**, 2139-2154, doi:10.1093/jxb/erp086 (2009).
- 25 Do, P. T., Prudent, M., Sulpice, R., Causse, M. & Fernie, A. R. The influence of fruit load on the tomato pericarp metabolome in a *Solanum chmielewskii* introgression line population. *Plant Physiol* **154**, 1128-1142, doi:10.1104/pp.110.163030 (2010).
- 26 Schauer, N. *et al.* Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat Biotechnol* **24**, 447-454, doi:10.1038/nbt1192 (2006).
- 27 Hackel, A. *et al.* Sucrose transporter LeSUT1 and LeSUT2 inhibition affects tomato fruit development in different ways. *The Plant Journal* **45**, 180-192 (2006).
- 28 Baldet, P. *et al.* The expression of cell proliferation-related genes in early developing flowers is affected by a fruit load reduction in tomato plants. *J Exp Bot* **57**, 961-970, doi:10.1093/jxb/erj082 (2006).
- 29 Burstin, J. *et al.* Developmental genes have pleiotropic effects on plant morphology and source capacity, eventually impacting on seed protein content and productivity in pea. *Plant physiology* **144**, 768-781 (2007).
- 30 Fridman, E., Carrari, F., Liu, Y. S., Fernie, A. R. & Zamir, D. Zooming in on a quantitative trait for tomato yield using interspecific introgressions. *Science* **305**, 1786-1789, doi:10.1126/science.1101666 (2004).
- 31 Fridman, E., Pleban, T. & Zamir, D. A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar

- content to 484 bp within an invertase gene. *Proc Natl Acad Sci U S A* **97**, 4718-4723 (2000).
- 32 Frary, A. *et al.* fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. *Science* **289**, 85-88 (2000).
- 33 Cong, B., Barrero, L. S. & Tanksley, S. D. Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. *Nat Genet* **40**, 800-804, doi:10.1038/ng.144 (2008).
- 34 Araujo, W. L. *et al.* Antisense inhibition of the iron-sulphur subunit of succinate dehydrogenase enhances photosynthesis and growth in tomato via an organic acid-mediated effect on stomatal aperture. *Plant Cell* **23**, 600-627, doi:10.1105/tpc.110.081224 (2011).
- 35 Nunes-Nesi, A., Araujo, W. L. & Fernie, A. R. Targeting mitochondrial metabolism and machinery as a means to enhance photosynthesis. *Plant Physiol* **155**, 101-107, doi:10.1104/pp.110.163816 (2011).
- 36 Bertin, N., Borel, C., Brunel, B., Cheniclet, C. & Causse, M. Do genetic make-up and growth manipulation affect tomato fruit size by cell number, or cell size and DNA endoreduplication? *Ann Bot* **92**, 415-424, doi:10.1093/aob/mcg146 (2003).
- 37 Heuvelink, E. Effect of fruit load on dry matter partitioning in tomato. *Scientia Horticulturae* **69**, 51-59 (1997).
- 38 Gautier, H., Guichard, S. & Tchamitchian, M. Modulation of competition between fruits and leaves by flower pruning and water fogging, and consequences on tomato leaf and fruit growth. *Annals of Botany* **88**, 645-652 (2001).
- 39 Yin, Y. G. *et al.* Salinity induces carbohydrate accumulation and sugar-regulated starch biosynthetic genes in tomato (*Solanum lycopersicum* L. cv. 'Micro-Tom') fruits in an ABA- and osmotic stress-independent manner. *J Exp Bot* **61**, 563-574, doi:10.1093/jxb/erp333 (2010).
- 40 Davies, J. & Cocking, E. Changes in carbohydrates, proteins and nucleic acids during cellular development in tomato fruit locule tissue. *Planta* **67**, 242-253 (1965).
- 41 Tanaka, A., Fujita, K. & Kikuchi, K. Nutrio-physiological studies on the tomato plant III. Photosynthetic rate of individual leaves in relation to the dry matter production of plants. *Soil Science and Plant Nutrition* **20**, 173-183, doi:10.1080/00380768.1974.10433240 (1974).

- 42 Giliberto, L. *et al.* Manipulation of the blue light photoreceptor cryptochrome 2 in tomato affects vegetative development, flowering time, and fruit antioxidant content. *Plant Physiol* **137**, 199-208, doi:10.1104/pp.104.051987 (2005).
- 43 Azari, R. *et al.* Overexpression of UV-DAMAGED DNA BINDING PROTEIN 1 links plant development and phytonutrient accumulation in high pigment-1 tomato. *J Exp Bot* **61**, 3627-3637, doi:10.1093/jxb/erq176 (2010).
- 44 Obiadalla-Ali, H., Fernie, A. R., Lytovchenko, A., Kossmann, J. & Lloyd, J. R. Inhibition of chloroplastic fructose 1,6-bisphosphatase in tomato fruits leads to decreased fruit size, but only small changes in carbohydrate metabolism. *Planta* **219**, 533-540, doi:10.1007/s00425-004-1257-y (2004).
- 45 Wang, H. *et al.* Regulatory features underlying pollination-dependent and -independent tomato fruit set revealed by transcript and primary metabolite profiling. *Plant Cell* **21**, 1428-1452, doi:10.1105/tpc.108.060830 (2009).
- 46 Schwender, J., Goffman, F., Ohlrogge, J. B. & Shachar-Hill, Y. Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature* **432**, 779-782, doi:10.1038/nature03145 (2004).
- 47 Bohn, G. W. & Scott, D. H. A 2nd gene for uniform unripe fruit color in the tomato. *Journal Hered* **36**, 169-172. (1945).
- 48 AF., Y. The uniform fruit color gene in the tomato. *Proceedings American Society Horticultural Science* **33**, 512 (1935).
- 49 Kemp, G. & Nonnecke, I. Differences in intensity of unripe fruit colour in the tomato. *Canadian Journal of Plant Science* **40**, 306-309 (1960).
- 50 Waters, M. T. *et al.* GLK transcription factors coordinate expression of the photosynthetic apparatus in Arabidopsis. *Plant Cell* **21**, 1109-1128, doi:10.1105/tpc.108.065250 (2009).
- 51 Waters, M. T. & Langdale, J. A. The making of a chloroplast. *EMBO J* **28**, 2861-2873, doi:10.1038/emboj.2009.264 (2009).
- 52 Waters, M. T., Moylan, E. C. & Langdale, J. A. GLK transcription factors regulate chloroplast development in a cell-autonomous manner. *Plant J* **56**, 432-444, doi:10.1111/j.1365-313X.2008.03616.x (2008).
- 53 Bükler, M., Schünemann, D. & Borchert, S. Enzymic properties and capacities of developing tomato (*Lycopersicon esculentum*

- L.) fruit plastids. *Journal of experimental botany* **49**, 681-691 (1998).
- 54 Kahlau, S. & Bock, R. Plastid transcriptomics and translaticomics of tomato fruit development and chloroplast-to-chromoplast differentiation: chromoplast gene expression largely serves the production of a single protein. *Plant Cell* **20**, 856-874, doi:10.1105/tpc.107.055202 (2008).
- 55 Barry, C. S., McQuinn, R. P., Chung, M. Y., Besuden, A. & Giovannoni, J. J. Amino acid substitutions in homologs of the STAY-GREEN protein are responsible for the green-flesh and chlorophyll retainer mutations of tomato and pepper. *Plant Physiol* **147**, 179-187, doi:10.1104/pp.108.118430 (2008).
- 56 Egea, I. *et al.* Chromoplast differentiation: current status and perspectives. *Plant Cell Physiol* **51**, 1601-1611, doi:10.1093/pcp/pcq136 (2010).
- 57 DINAR M, M. S. The relationship between starch accumulation and soluble solids content of tomato fruits. *Journal of the American Society for Horticultural Science*, 415-418 (1981).
- 58 Isaacson, T., Ronen, G., Zamir, D. & Hirschberg, J. Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of beta-carotene and xanthophylls in plants. *Plant Cell* **14**, 333-342 (2002).
- 59 Galpaz, N., Wang, Q., Menda, N., Zamir, D. & Hirschberg, J. Abscisic acid deficiency in the tomato mutant high-pigment 3 leading to increased plastid number and higher fruit lycopene content. *Plant J* **53**, 717-730, doi:10.1111/j.1365-313X.2007.03362.x (2008).
- 60 Nashilevitz, S. *et al.* An orange ripening mutant links plastid NAD(P)H dehydrogenase complex activity to central and specialized metabolism during tomato fruit maturation. *Plant Cell* **22**, 1977-1997, doi:10.1105/tpc.110.074716 (2010).
- 61 Bino, R. J. *et al.* The light-hyperresponsive high pigment-2dg mutation of tomato: alterations in the fruit metabolome. *New Phytol* **166**, 427-438, doi:10.1111/j.1469-8137.2005.01362.x (2005).
- 62 Kolotilin, I. *et al.* Transcriptional profiling of high pigment-2dg tomato mutant links early fruit plastid biogenesis with its overproduction of phytonutrients. *Plant Physiol* **145**, 389-401, doi:10.1104/pp.107.102962 (2007).

- 63 Yen, H. *et al.* The tomato high-pigment (hp) locus maps to chromosome 2 and influences plastome copy number and fruit quality. *TAG Theoretical and Applied Genetics* **95**, 1069-1079 (1997).
- 64 Mustilli, A. C., Fenzi, F., Ciliento, R., Alfano, F. & Bowler, C. Phenotype of the tomato high pigment-2 mutant is caused by a mutation in the tomato homolog of DEETIOLATED1. *Plant Cell* **11**, 145-157 (1999).
- 65 Davuluri, G. R. *et al.* Manipulation of DET1 expression in tomato results in photomorphogenic phenotypes caused by post-transcriptional gene silencing. *Plant J* **40**, 344-354, doi:10.1111/j.1365-313X.2004.02218.x (2004).
- 66 Davuluri, G. R. *et al.* Fruit-specific RNAi-mediated suppression of DET1 enhances carotenoid and flavonoid content in tomatoes. *Nat Biotechnol* **23**, 890-895, doi:10.1038/nbt1108 (2005).
- 67 Enfissi, E. M. *et al.* Integrative transcript and metabolite analysis of nutritionally enhanced DE-ETIOLATED1 downregulated tomato fruit. *Plant Cell* **22**, 1190-1215, doi:10.1105/tpc.110.073866 (2010).
- 68 Tang, X. *et al.* Ubiquitin-conjugated degradation of golden 2-like transcription factor is mediated by CUL4-DDB1-based E3 ligase complex in tomato. *New Phytol* **209**, 1028-1039, doi:10.1111/nph.13635 (2016).
- 69 Wang, S. *et al.* Altered plastid levels and potential for improved fruit nutrient content by downregulation of the tomato DDB1-interacting protein CUL4. *The Plant Journal* **55**, 89-103, doi:10.1111/j.1365-313X.2008.03489.x (2008).
- 70 Liu, Y. *et al.* Manipulation of light signal transduction as a means of modifying fruit nutritional quality in tomato. *Proc Natl Acad Sci U S A* **101**, 9897-9902, doi:10.1073/pnas.0400935101 (2004).
- 71 Kobayashi, K. *et al.* Regulation of root greening by light and auxin/cytokinin signaling in Arabidopsis. *Plant Cell* **24**, 1081-1095, doi:10.1105/tpc.111.092254 (2012).
- 72 Jones, B. *et al.* Down-regulation of DR12, an auxin-response-factor homolog, in the tomato results in a pleiotropic phenotype including dark green and blotchy ripening fruit. *Plant J* **32**, 603-613 (2002).
- 73 Liu, L. *et al.* Ectopic expression of a BZR1-1D transcription factor in brassinosteroid signalling enhances carotenoid accumulation

- and fruit quality attributes in tomato. *Plant Biotechnology Journal* **12**, 105-115, doi:10.1111/pbi.12121 (2014).
- 74 Martineau, B., Houck, C. M., Sheehy, R. E. & Hiatt, W. R. Fruit-specific expression of the *A. tumefaciens* isopentenyl transferase gene in tomato: effects on fruit ripening and defense-related gene expression in leaves. *The Plant Journal* **5**, 11-19 (1994).
- 75 Borovsky, Y. & Paran, I. Chlorophyll breakdown during pepper fruit ripening in the chlorophyll retainer mutation is impaired at the homolog of the senescence-inducible stay-green gene. *Theoretical and Applied Genetics* **117**, 235-240 (2008).
- 76 Barry, C. S. *et al.* Altered chloroplast development and delayed fruit ripening caused by mutations in a zinc metalloprotease at the lutescent2 locus of tomato. *Plant Physiol* **159**, 1086-1098, doi:10.1104/pp.112.197483 (2012).
- 77 Chen, G., Bi, Y. R. & Li, N. EGY1 encodes a membrane-associated and ATP-independent metalloprotease that is required for chloroplast development. *Plant J* **41**, 364-375, doi:10.1111/j.1365-313X.2004.02308.x (2005).
- 78 Fitter, D. W., Martin, D. J., Copley, M. J., Scotland, R. W. & Langdale, J. A. GLK gene pairs regulate chloroplast development in diverse plant species. *Plant J* **31**, 713-727 (2002).
- 79 Rohrmann, J. *et al.* Combined transcription factor profiling, microarray analysis and metabolite profiling reveals the transcriptional control of metabolic shifts occurring during tomato fruit development. *Plant J* **68**, 999-1013, doi:10.1111/j.1365-313X.2011.04750.x (2011).
- 80 Liu, H. *et al.* Map-based cloning, identification and characterization of the *w* gene controlling white immature fruit color in cucumber (*Cucumis sativus* L.). *Theoretical and Applied Genetics* **129**, 1247-1256, doi:10.1007/s00122-016-2700-8 (2016).
- 81 Nadakuduti, S. S., Holdsworth, W. L., Klein, C. L. & Barry, C. S. KNOX genes influence a gradient of fruit chloroplast development through regulation of GOLDEN2-LIKE expression in tomato. *Plant J* **78**, 1022-1033, doi:10.1111/tpj.12529 (2014).
- 82 Hobson, G., Davies A review of blotchy ripening and allied disorders of the tomato. *Glasshouse Crops Research Institute Annual Repor*, 139-147 (1976).

- 83 Peet, M. Fruit cracking in tomato. *HortTechnology* **2**, 216-223 (1992).
- 84 Murchie, E. H. & Niyogi, K. K. Manipulation of photoprotection to improve plant photosynthesis. *Plant Physiol* **155**, 86-92, doi:10.1104/pp.110.168831 (2011).
- 85 Burgess, D. G. & Taylor, W. C. Chloroplast photooxidation affects the accumulation of cytosolic mRNAs encoding chloroplast proteins in maize. *Planta* **170**, 520-527, doi:10.1007/BF00402986 (1987).
- 86 Kawata, E. E. & Cheung, A. Y. Molecular analysis of an aurea photosynthetic mutant (Su/Su) in tobacco: LHCP depletion leads to pleiotropic mutant phenotypes. *EMBO J* **9**, 4197-4203 (1990).
- 87 Piechulla, B., Pichersky, E., Cashmore, A. R. & Gruissem, W. Expression of nuclear and plastid genes for photosynthesis-specific proteins during tomato fruit development and ripening. *Plant Mol Biol* **7**, 367-376, doi:10.1007/BF00032566 (1986).
- 88 Livne, A. & Gepstein, S. Abundance of the Major Chloroplast Polypeptides during Development and Ripening of Tomato Fruits: An Immunological Study. *Plant Physiol* **87**, 239-243 (1988).
- 89 Pfannschmidt, T., Nilsson, A., Tullberg, A., Link, G. & Allen, J. F. Direct transcriptional control of the chloroplast genes *psbA* and *psaAB* adjusts photosynthesis to light energy distribution in plants. *IUBMB Life* **48**, 271-276, doi:10.1080/713803507 (1999).
- 90 Francis, D. M., Barringer, S. A. & Whitmoyer, R. E. Ultrastructural characterization of yellow shoulder disorder in a uniform ripening tomato genotype. *HortScience* **35**, 1114-1117 (2000).
- 91 Picha, D. H. Physiological factors associated with yellow shoulder expression in tomato fruit. *Journal of the American Society for Horticultural Science* **112**, 798-801 (1987).
- 92 LW., H. The inheritance of resistance to fruit cracking in the tomato (*Lycopersicon esculentum* L.). *MSc thesis, Oregon State College, Corvallis, USA.*, 85 (1956).
- 93 PA, Y. & JW, M. Horticultural characters of tomatoes. *College Station, Texas agricultural and mechanical college* **698**, 61 (1947).
- 94 PA, Y. Cuticle cracking in green tomato fruits. *Phytopathology* **36**, 413 (1946).
- 95 Jimenez, A. *et al.* Changes in oxidative processes and components of the antioxidant system during tomato fruit

- ripening. *Planta* **214**, 751-758, doi:10.1007/s004250100667 (2002).
- 96 Mondal, K., Sharma, N., Malhotra, S. P., Dhawan, K. & Singh, R. Antioxidant systems in ripening tomato fruits. *Biologia Plantarum* **48**, 49-53 (2004).
- 97 DL, S., JA, A. & KC, G. Down-regulation of tomato beta-galactosidase IV results in decreased fruit softening. *Plant Physiology* **129**, 1755-1762 (2002).
- 98 Mehta, R. A. *et al.* Engineered polyamine accumulation in tomato enhances phytonutrient content, juice quality, and vine life. *Nature biotechnology* **20**, 613 (2002).
- 99 Petrusa, E. *et al.* Plant flavonoids--biosynthesis, transport and involvement in stress responses. *Int J Mol Sci* **14**, 14950-14973, doi:10.3390/ijms140714950 (2013).
- 100 Chalker-Scott, L. Environmental significance of anthocyanins in plant stress responses. *Photochemistry and photobiology* **70**, 1-9 (1999).
- 101 Saure, M. C. External control of anthocyanin formation in apple. *Scientia Horticulturae* **42**, 181-218 (1990).
- 102 Kerckhoffs, L. H. J. *et al.* Physiological characterization of exaggerated-photoresponse mutants of tomato. *Journal of Plant Physiology* **150**, 578-587, doi:10.1016/s0176-1617(97)80322-7 (1997).
- 103 Kaltenbach, M., Schroder, G., Schmelzer, E., Lutz, V. & Schroder, J. Flavonoid hydroxylase from *Catharanthus roseus*: cDNA, heterologous expression, enzyme properties and cell-type specific expression in plants. *Plant J* **19**, 183-193 (1999).
- 104 Rozema, J., van de Staaij, J., Bjorn, L. O. & Caldwell, M. UV-B as an environmental factor in plant life: stress and regulation. *Trends in ecology & evolution* **12**, 22-28 (1997).
- 105 Pietrini, F. & Massacci, A. Leaf anthocyanin content changes in *Zea mays* L. grown at low temperature: Significance for the relationship between the quantum yield of PS II and the apparent quantum yield of CO₂ assimilation. *Photosynthesis Research* **58**, 213-219, doi:10.1023/A:1006152610137 (1998).
- 106 Agati, G., Azzarello, E., Pollastri, S. & Tattini, M. Flavonoids as antioxidants in plants: location and functional significance. *Plant Sci* **196**, 67-76, doi:10.1016/j.plantsci.2012.07.014 (2012).
- 107 Cassidy, A. *et al.* Higher dietary anthocyanin and flavonol intakes are associated with anti-inflammatory effects in a population of

- US adults. *Am J Clin Nutr* **102**, 172-181, doi:10.3945/ajcn.115.108555 (2015).
- 108 Zhang, Y. *et al.* Anthocyanins double the shelf life of tomatoes by delaying overripening and reducing susceptibility to gray mold. *Curr Biol* **23**, 1094-1100, doi:10.1016/j.cub.2013.04.072 (2013).
- 109 Rein, M. (University of Helsinki Helsinki, 2005).
- 110 Schreiber, G. *et al.* ANTHOCYANIN1 from *Solanum chilense* is more efficient in accumulating anthocyanin metabolites than its *Solanum lycopersicum* counterpart in association with the ANTHOCYANIN FRUIT phenotype of tomato. *Theor Appl Genet* **124**, 295-307, doi:10.1007/s00122-011-1705-6 (2012).
- 111 Jones, C. M. Evaluation of Carotenoids and Anthocyanins in High Pigment, Processing, Heirloom and Anthocyanin Fruit Tomatoes. *Degree of Master of Science in Horticulture* (2000).
- 112 De Jong, W. S., Eannetta, N. T., De Jong, D. M. & Bodis, M. Candidate gene analysis of anthocyanin pigmentation loci in the Solanaceae. *Theor Appl Genet* **108**, 423-432, doi:10.1007/s00122-003-1455-1 (2004).
- 113 CM, R., P, C., RT, C. & JW, D. Abg, a gene on chromosome 10 for purple fruit derived from *S. lycopersicoides*. *Rep. Tomato Genet. Coop.* **44**, 29-30 (1994).
- 114 Canady, M. A., Ji, Y. & Chetelat, R. T. Homeologous recombination in *Solanum lycopersicoides* introgression lines of cultivated tomato. *Genetics* **174**, 1775-1788, doi:10.1534/genetics.106.065144 (2006).
- 115 Rick, C., Reeves, A. & Zobel, R. Inheritance and linkage relations of four new mutants. *Tomato Genet Coop Rep* **18**, 34-35 (1968).
- 116 Ibrahim, H., Shannon, S. & Robinson, R. Influence of the high pigment gene on anthocyanin content. *Tomato Genet. Coop. Rpt* **18**, 21 (1968).
- 117 Jones, C. M. Characterization and Inheritance of the Anthocyanin fruit (Aft) Tomato. *Journal of Heredity* **94**, 449-456, doi:10.1093/jhered/esg093 (2003).
- 118 Povero, G., Gonzali, S., Bassolino, L., Mazzucato, A. & Perata, P. Transcriptional analysis in high-anthocyanin tomatoes reveals synergistic effect of Aft and atv genes. *J Plant Physiol* **168**, 270-279, doi:10.1016/j.jplph.2010.07.022 (2011).

- 119 Mes, P. J., Boches, P., Myers, J. R. & Durst, R. Characterization of tomatoes expressing anthocyanin in the fruit. *Journal of the American Society for Horticultural Science* **133**, 262-269 (2008).
- 120 Luo, J. *et al.* AtMYB12 regulates caffeoyl quinic acid and flavonol synthesis in tomato: expression in fruit results in very high levels of both types of polyphenol. *The Plant Journal* **56**, 316-326 (2008).
- 121 Verhoeven, M. E. *et al.* Increasing antioxidant levels in tomatoes through modification of the flavonoid biosynthetic pathway. *Journal of Experimental Botany* **53**, 2099-2106, doi:10.1093/jxb/erf044 (2002).
- 122 Hutzler, P. *et al.* Tissue localization of phenolic compounds in plants by confocal laser scanning microscopy. *Journal of Experimental Botany* **49**, 953-965 (1998).
- 123 Allan, A. C., Hellens, R. P. & Laing, W. A. MYB transcription factors that colour our fruit. *Trends Plant Sci* **13**, 99-102, doi:10.1016/j.tplants.2007.11.012 (2008).
- 124 Bovy, A. High-Flavonol Tomatoes Resulting from the Heterologous Expression of the Maize Transcription Factor Genes LC and C1. *The Plant Cell Online* **14**, 2509-2526, doi:10.1105/tpc.004218 (2002).
- 125 Petroni, K. & Tonelli, C. Recent advances on the regulation of anthocyanin synthesis in reproductive organs. *Plant Sci* **181**, 219-229, doi:10.1016/j.plantsci.2011.05.009 (2011).
- 126 Shi, M. Z. & Xie, D. Y. Biosynthesis and metabolic engineering of anthocyanins in *Arabidopsis thaliana*. *Recent patents on biotechnology* **8**, 47-60 (2014).
- 127 Goff, S. A., Cone, K. C. & Chandler, V. L. Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. *Genes & development* **6**, 864-875 (1992).
- 128 Spelt, C., Quattrocchio, F., Mol, J. N. & Koes, R. anthocyanin1 of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. *Plant Cell* **12**, 1619-1632 (2000).
- 129 Quattrocchio, F. *et al.* Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color. *Plant Cell* **11**, 1433-1444 (1999).
- 130 Nick de Vetten, F. Q., Joseph Mol, and Ronald Koes. <AN11 PETUNIA.pdf>. (1997).

- 131 Kiferle, C. *et al.* Tomato R2R3-MYB Proteins SIAN1 and SIAN2: Same Protein Activity, Different Roles. *PLoS One* **10**, e0136365, doi:10.1371/journal.pone.0136365 (2015).
- 132 Li, S. Transcriptional control of flavonoid biosynthesis. *Plant Signaling & Behavior* **9**, e27522, doi:10.4161/psb.27522 (2014).
- 133 Georgiev, C. Anthocyanin fruit (Af). *Tomato Genet. Coop. Rpt* **22** (1972).
- 134 Goodrich, C. A. Ureilites: A critical review. *Meteoritics* **27**, 327-352, doi:10.1111/j.1945-5100.1992.tb00215.x (1992).
- 135 Zuluaga, D. L. *et al.* Arabidopsis thaliana MYB75/PAP1 transcription factor induces anthocyanin production in transgenic tomato plants. *Functional Plant Biology* **35**, 606, doi:10.1071/fp08021 (2008).
- 136 Butelli, E. *et al.* Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat Biotechnol* **26**, 1301-1308, doi:10.1038/nbt.1506 (2008).
- 137 Estornell, L. H. *et al.* A multisite gateway-based toolkit for targeted gene expression and hairpin RNA silencing in tomato fruits. *Plant Biotechnology Journal* **7**, 298-309, doi:10.1111/j.1467-7652.2009.00402.x (2009).
- 138 Bassolino, L. *et al.* Accumulation of anthocyanins in tomato skin extends shelf life. *New Phytol* **200**, 650-655, doi:10.1111/nph.12524 (2013).
- 139 Fraser, P. D., Truesdale, M. R., Bird, C. R., Schuch, W. & Bramley, P. M. Carotenoid Biosynthesis during Tomato Fruit Development (Evidence for Tissue-Specific Gene Expression). *Plant Physiol* **105**, 405-413 (1994).
- 140 Egea, I. *et al.* Chloroplast to chromoplast transition in tomato fruit: spectral confocal microscopy analyses of carotenoids and chlorophylls in isolated plastids and time-lapse recording on intact live tissue. *Ann Bot* **108**, 291-297, doi:10.1093/aob/mcr140 (2011).
- 141 Nguyen, C. V. *et al.* Tomato GOLDEN2-LIKE transcription factors reveal molecular gradients that function during fruit development and ripening. *Plant Cell* **26**, 585-601, doi:10.1105/tpc.113.118794 (2014).
- 142 Pesaresi, P., Mizzotti, C., Colombo, M. & Masiero, S. Genetic regulation and structural changes during tomato fruit

- development and ripening. *Front Plant Sci* **5**, 124, doi:10.3389/fpls.2014.00124 (2014).
- 143 Cocaliadis, M. F., Fernandez-Munoz, R., Pons, C., Orzaez, D. & Granell, A. Increasing tomato fruit quality by enhancing fruit chloroplast function. A double-edged sword? *J Exp Bot* **65**, 4589-4598, doi:10.1093/jxb/eru165 (2014).
- 144 Biswal, B., Krupinska, K. & Biswal, U. C. *Plastid development in leaves during growth and senescence*. (Springer, 2013).
- 145 Cookson, P. J. *et al.* Increases in cell elongation, plastid compartment size and phytoene synthase activity underlie the phenotype of the high pigment-1 mutant of tomato. *Planta* **217**, 896-903, doi:10.1007/s00425-003-1065-9 (2003).
- 146 Liu, J., Van Eck, J., Cong, B. & Tanksley, S. D. A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proc Natl Acad Sci U S A* **99**, 13302-13306, doi:10.1073/pnas.162485999 (2002).
- 147 Tanksley, S. D. The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. *Plant Cell* **16 Suppl**, S181-189, doi:10.1105/tpc.018119 (2004).
- 148 Azzi, L. *et al.* Fruit growth-related genes in tomato. *J Exp Bot* **66**, 1075-1086, doi:10.1093/jxb/eru527 (2015).
- 149 Czerednik, A., Busscher, M., Angenent, G. C. & de Maagd, R. A. The cell size distribution of tomato fruit can be changed by overexpression of CDKA1. *Plant Biotechnol J* **13**, 259-268, doi:10.1111/pbi.12268 (2015).
- 150 Robinson, N. L., Hewitt, J. D. & Bennett, A. B. Sink metabolism in tomato fruit : I. Developmental changes in carbohydrate metabolizing enzymes. *Plant Physiol* **87**, 727-730 (1988).
- 151 Petreikov, M. *et al.* Carbohydrate balance and accumulation during development of near-isogenic tomato lines differing in the AGPase-L1 allele. *Journal of the American Society for Horticultural Science* **134**, 134-140 (2009).
- 152 Llorente, B. *et al.* Tomato fruit carotenoid biosynthesis is adjusted to actual ripening progression by a light-dependent mechanism. *The Plant Journal* **85**, 107-119 (2016).
- 153 Toledo-Ortiz, G., Huq, E. & Rodríguez-Concepción, M. Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. *Proceedings of the National Academy of Sciences* **107**, 11626-11631 (2010).

- 154 Brand, A. *et al.* CaGLK2 regulates natural variation of chlorophyll content and fruit color in pepper fruit. *Theor Appl Genet* **127**, 2139-2148, doi:10.1007/s00122-014-2367-y (2014).
- 155 Rossini, L., Cribb, L., Martin, D. J. & Langdale, J. A. The maize golden2 gene defines a novel class of transcriptional regulators in plants. *Plant Cell* **13**, 1231-1244 (2001).
- 156 Wang, P. *et al.* Evolution of GOLDEN2-LIKE gene function in C(3) and C (4) plants. *Planta* **237**, 481-495, doi:10.1007/s00425-012-1754-3 (2013).
- 157 Chen, M. *et al.* GOLDEN 2-LIKE Transcription Factors of Plants. *Frontiers in Plant Science* **7**, 1509, doi:10.3389/fpls.2016.01509 (2016).
- 158 Santino, C. G., Stanford, G. L. & Conner, T. W. Developmental and transgenic analysis of two tomato fruit enhanced genes. *Plant Mol Biol* **33**, 405-416 (1997).
- 159 Tieman, D. *et al.* A chemical genetic roadmap to improved tomato flavor. *Science* **355**, 391 (2017).
- 160 Green, B. R. & Durnford, D. G. The Chlorophyll-Carotenoid Proteins of Oxygenic Photosynthesis. *Annu Rev Plant Physiol Plant Mol Biol* **47**, 685-714, doi:10.1146/annurev.arplant.47.1.685 (1996).
- 161 Eckardt, N. A. A New Chlorophyll Degradation Pathway. *The Plant Cell* **21**, 700-700, doi:10.1105/tpc.109.210313 (2009).
- 162 Ramirez, D. A. & Tomes, M. L. Relationship between Chlorophyll and Carotenoid Biosynthesis in Dirty-Red (Green-Flesh) Mutant in Tomato. *Botanical Gazette* **125**, 221-226 (1964).
- 163 Hendry, G. Oxygen, free radical processes and seed longevity. *Seed Science Research* **3**, 141-153, doi:doi:10.1017/S0960258500001720 (1993).
- 164 Liakopoulos, G. *et al.* The photoprotective role of epidermal anthocyanins and surface pubescence in young leaves of grapevine (*Vitis vinifera*). *Ann Bot* **98**, 257-265, doi:10.1093/aob/mcl097 (2006).
- 165 Fernie, A. R., Trethewey, R. N., Krotzky, A. J. & Willmitzer, L. Metabolite profiling: from diagnostics to systems biology. *Nat Rev Mol Cell Biol* **5**, 763-769, doi:10.1038/nrm1451 (2004).
- 166 Hiratsuka, S., Yokoyama, Y., Nishimura, H., Miyazaki, T. & Nada, K. Fruit photosynthesis and phosphoenolpyruvate carboxylase activity as affected by lightproof fruit bagging in Satsuma

- mandarin. *Journal of the American Society for Horticultural Science* **137**, 215-220 (2012).
- 167 Blanke, M. M. in *II Workshop on Pome Fruit* 466. 19-22.
- 168 Lopez, Y., Riano, N., Mosquera, P., Cadavid, A. & Arcila, J. Activities of phosphoenolpyruvate carboxylase and ribulose-1, 5-bisphosphate carboxylase/oxygenase in leaves and fruit pericarp tissue of different coffee (*Coffea* sp.) genotypes. *Photosynthetica* **38**, 215-220 (2000).
- 169 Amemiya, T., Kanayama, Y., Yamaki, S., Yamada, K. & Shiratake, K. Fruit-specific V-ATPase suppression in antisense-transgenic tomato reduces fruit growth and seed formation. *Planta* **223**, 1272, doi:10.1007/s00425-005-0176-x (2005).
- 170 Stange, C. & Flores, C. in *Advances in photosynthesis-fundamental aspects* 77–96 (InTech, 2012).
- 171 Lichtenthaler, H. K. Biosynthesis, accumulation and emission of carotenoids, α -tocopherol, plastoquinone, and isoprene in leaves under high photosynthetic irradiance. *Photosynthesis Research* **92**, 163-179, doi:10.1007/s11120-007-9204-y (2007).
- 172 Lichtenthaler, H. K. *et al.* Photosynthetic activity, chloroplast ultrastructure, and leaf characteristics of high-light and low-light plants and of sun and shade leaves. *Photosynth Res* **2**, 115-141, doi:10.1007/bf00028752 (1981).
- 173 Xu, H., Vavilin, D. & Vermaas, W. Chlorophyll b can serve as the major pigment in functional photosystem II complexes of cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 14168-14173, doi:10.1073/pnas.251530298 (2001).
- 174 Jahns, P., Latowski, D. & Strzalka, K. Mechanism and regulation of the violaxanthin cycle: The role of antenna proteins and membrane lipids. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1787**, 3-14, doi:<https://doi.org/10.1016/j.bbabi.2008.09.013> (2009).
- 175 Jahns, P. & Holzwarth, A. R. The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1817**, 182-193, doi:<https://doi.org/10.1016/j.bbabi.2011.04.012> (2012).
- 176 Liakopoulos, G. *et al.* The photoprotective role of epidermal anthocyanins and surface pubescence in young leaves of grapevine (*Vitis vinifera*). *Annals of botany* **98**, 257-265 (2006).

- 177 Woodson, J. D. & Chory, J. Coordination of gene expression between organellar and nuclear genomes. *Nature reviews. Genetics* **9**, 383-395, doi:10.1038/nrg2348 (2008).
- 178 Azuma, M. *et al.* Adenine Nucleotide Translocator Transports Haem Precursors into Mitochondria. *PLoS ONE* **3**, e3070, doi:10.1371/journal.pone.0003070 (2008).
- 179 Bobik, K. & Burch-Smith, T. M. Chloroplast signaling within, between and beyond cells. *Front Plant Sci* **6**, 781, doi:10.3389/fpls.2015.00781 (2015).
- 180 Wang, F., Sanz, A., Brenner, M. L. & Smith, A. Sucrose synthase, starch accumulation, and tomato fruit sink strength. *Plant Physiology* **101**, 321-327 (1993).
- 181 Visser, A. Photosynthesis: Photoreactions to Plant Productivity. Edited by Yash Pal Arbol, Prasanna Mohanty and Govindjee. *Plant Ecology* **146**, 118-119 (2000).
- 182 Kringstad, R., Kenyon, W. H. & Black, C. C. The rapid isolation of vacuoles from leaves of crassulacean Acid metabolism plants. *Plant Physiol* **66**, 379-382 (1980).
- 183 Sui, X. *et al.* The complex character of photosynthesis in cucumber fruit. *J Exp Bot*, doi:10.1093/jxb/erx034 (2017).
- 184 Kazeniak, S. J. & Hall, R. M. FLAVOR CHEMISTRY OF TOMATO VOLATILES. *Journal of Food Science* **35**, 519-530, doi:10.1111/j.1365-2621.1970.tb04799.x (1970).
- 185 Krumbein, A., Peters, P. & Brückner, B. Flavour compounds and a quantitative descriptive analysis of tomatoes (*Lycopersicon esculentum* Mill.) of different cultivars in short-term storage. *Postharvest Biology and Technology* **32**, 15-28, doi:<https://doi.org/10.1016/j.postharvbio.2003.10.004> (2004).
- 186 Petro-Turza, M. Flavor of tomato and tomato products. *Food Reviews International* **2**, 309-351, doi:10.1080/87559128609540802 (1986).
- 187 Yilmaz, E. Oxylipin pathway in the biosynthesis of fresh tomato volatiles. *Turkish Journal of Biology* **25**, 351-360 (2001).
- 188 Vogel, J. T. *et al.* Carotenoid content impacts flavor acceptability in tomato (*Solanum lycopersicum*). *Journal of the Science of Food and Agriculture* **90**, 2233-2240, doi:10.1002/jsfa.4076 (2010).
- 189 Buttery, P. J. & Rowsell, E. V. Enzymic assays for amoonia and L-glutamine in tissue extracts. *Anal Biochem* **39**, 297-310 (1971).

- 190 Ballester, A. R. *et al.* Biochemical and molecular analysis of pink tomatoes: deregulated expression of the gene encoding transcription factor SIMYB12 leads to pink tomato fruit color. *Plant Physiol* **152**, 71-84, doi:10.1104/pp.109.147322 (2010).
- 191 Kennison, K. R., Wilkinson, K. L., Williams, H. G., Smith, J. H. & Gibberd, M. R. Smoke-derived taint in wine: effect of postharvest smoke exposure of grapes on the chemical composition and sensory characteristics of wine. *J Agric Food Chem* **55**, 10897-10901, doi:10.1021/jf072509k (2007).
- 192 Arumugam, M. *et al.* Isolation of an unusual metabolite 2-allyloxyphenol from a marine actinobacterium, its biological activities and applications. *Applied microbiology and biotechnology* **86**, 109-117 (2010).
- 193 Tikunov, Y. M. *et al.* Non-smoky glycosyltransferase1 prevents the release of smoky aroma from tomato fruit. *Plant Cell* **25**, 3067-3078, doi:10.1105/tpc.113.114231 (2013).
- 194 Zierler, B., Siegmund, B. & Pfannhauser, W. Determination of off-flavour compounds in apple juice caused by microorganisms using headspace solid phase microextraction–gas chromatography–mass spectrometry. *Analytica Chimica Acta* **520**, 3-11, doi:<http://doi.org/10.1016/j.aca.2004.03.084> (2004).
- 195 Mageroy, M. H., Tieman, D. M., Floystad, A., Taylor, M. G. & Klee, H. J. A *Solanum lycopersicum* catechol-O-methyltransferase involved in synthesis of the flavor molecule guaiacol. *Plant J* **69**, 1043-1051, doi:10.1111/j.1365-313X.2011.04854.x (2012).
- 196 Sarrion-Perdigones, A. *et al.* GoldenBraid 2.0: a comprehensive DNA assembly framework for plant synthetic biology. *Plant Physiol* **162**, 1618-1631, doi:10.1104/pp.113.217661 (2013).
- 197 Ellul, P. *et al.* The ploidy level of transgenic plants in *Agrobacterium*-mediated transformation of tomato cotyledons (*Lycopersicon esculentum* L. Mill.) is genotype and procedure dependent. *Theoretical and Applied Genetics* **106**, 231-238 (2003).
- 198 Doyle, J. J., Doyle, J. L., Brown, A. H. & Grace, J. P. Multiple origins of polyploids in the *Glycine tabacina* complex inferred from chloroplast DNA polymorphism. *Proc Natl Acad Sci U S A* **87**, 714-717 (1990).
- 199 Darrigues, A., Schwartz, S. J. & Francis, D. M. Optimizing sampling of tomato fruit for carotenoid content with application

- to assessing the impact of ripening disorders. *J Agric Food Chem* **56**, 483-487, doi:10.1021/jf071896v (2008).
- 200 Karnovsky, M. J. A formaldehyde-glutaraldehyde fixative of high osmolality for use in electron microscopy. *J. Cell Biol.* **27**, 137A–138A (1965).
- 201 Seguí-Simarro, J. M. & Nuez, F. Embryogenesis induction, callogenesis, and plant regeneration by *in vitro* culture of tomato isolated microspores and whole anthers. *Journal of Experimental Botany* **58**, 1119-1132 (2007).
- 202 Williams, M. in *Practical methods in electron microscopy* (ed A.M. Glauert) (North Holland/American Elsevier, 1977).
- 203 Schauer, N. *et al.* GC-MS libraries for the rapid identification of metabolites in complex biological samples. *FEBS Lett* **579**, 1332-1337, doi:10.1016/j.febslet.2005.01.029 (2005).
- 204 Roessner, U. *et al.* Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *Plant Cell* **13**, 11-29 (2001).
- 205 Levin, I. *et al.* High pigment tomato mutants—more than just lycopene (a review). *Israel Journal of Plant Sciences* **54**, 179-190, doi:10.1560/ijps_54_3_179 (2006).
- 206 Pogson, B. J., Ganguly, D. & Albrecht-Borth, V. Insights into chloroplast biogenesis and development. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847**, 1017-1024, doi:<http://doi.org/10.1016/j.bbabi.2015.02.003> (2015).
- 207 Chory, J. A genetic model for light-regulated seedling Arabidopsis. *Development* **115**, 337-354 (1992).
- 208 Kapoor, S., Maheshwari, S. C. & Tyagi, A. K. Organ-specific expression of plastid-encoded genes in rice involves both quantitative and qualitative changes in mRNAs. *Plant and cell physiology* **34**, 943-947 (1993).
- 209 Jung, H.-S. & Chory, J. Signaling between Chloroplasts and the Nucleus: Can a Systems Biology Approach Bring Clarity to a Complex and Highly Regulated Pathway? *Plant Physiology* **152**, 453-459, doi:10.1104/pp.109.149070 (2010).
- 210 Dall'Osto, L., Unlu, C., Cazzaniga, S. & van Amerongen, H. Disturbed excitation energy transfer in Arabidopsis thaliana mutants lacking minor antenna complexes of photosystem II. *Biochim Biophys Acta* **1837**, 1981-1988, doi:10.1016/j.bbabi.2014.09.011 (2014).

- 211 Kendrick, R. E., Kerckhoffs, L. H. J., Van Tuinen, A. & Koornneef, M. Photomorphogenic mutants of tomato. *Plant, Cell & Environment* **20**, 746-751, doi:10.1046/j.1365-3040.1997.d01-109.x (1997).
- 212 Schroeder, D. F. *et al.* De-etiolated 1 and damaged DNA binding protein 1 interact to regulate Arabidopsis photomorphogenesis. *Curr Biol* **12**, 1462-1472 (2002).
- 213 Zhang, Y. *et al.* Arabidopsis DDB1-CUL4 ASSOCIATED FACTOR1 forms a nuclear E3 ubiquitin ligase with DDB1 and CUL4 that is involved in multiple plant developmental processes. *The Plant Cell* **20**, 1437-1455 (2008).
- 214 Chen, H. *et al.* Arabidopsis CULLIN4-damaged DNA binding protein 1 interacts with CONSTITUTIVELY PHOTOMORPHOGENIC1-SUPPRESSOR OF PHYA complexes to regulate photomorphogenesis and flowering time. *The Plant Cell* **22**, 108-123 (2010).
- 215 Yu, X. *et al.* A brassinosteroid transcriptional network revealed by genome-wide identification of BES1 target genes in Arabidopsis thaliana. *The Plant Journal* **65**, 634-646, doi:10.1111/j.1365-313X.2010.04449.x (2011).
- 216 Perochon, A. *et al.* Interaction of a plant pseudo-response regulator with a calmodulin-like protein. *Biochem Biophys Res Commun* **398**, 747-751, doi:10.1016/j.bbrc.2010.07.016 (2010).
- 217 Fujimori, T., Yamashino, T., Kato, T. & Mizuno, T. Circadian-controlled basic/helix-loop-helix factor, PIL6, implicated in light-signal transduction in Arabidopsis thaliana. *Plant Cell Physiol* **45**, 1078-1086, doi:10.1093/pcp/pch124 (2004).
- 218 Thimm, O. *et al.* MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J* **37**, 914-939 (2004).
- 219 Foyer, C. H. & Noctor, G. Redox sensing and signalling associated with reactive oxygen in chloroplasts, peroxisomes and mitochondria. *Physiologia plantarum* **119**, 355-364 (2003).
- 220 Fey, V. *et al.* Retrograde plastid redox signals in the expression of nuclear genes for chloroplast proteins of Arabidopsis thaliana. *The Journal of biological chemistry* **280**, 5318-5328, doi:10.1074/jbc.M406358200 (2005).
- 221 Gläßer, C. *et al.* Meta-analysis of retrograde signaling in Arabidopsis thaliana reveals a core module of genes embedded

- in complex cellular signaling networks. *Molecular plant* **7**, 1167-1190 (2014).
- 222 Ramel, F. *et al.* Carotenoid oxidation products are stress signals that mediate gene responses to singlet oxygen in plants. *Proceedings of the National Academy of Sciences* **109**, 5535-5540 (2012).
- 223 Lundquist, P. K. *et al.* Loss of Plastoglobule Kinases ABC1K1 and ABC1K3 Causes Conditional Degreening, Modified Prenyl-Lipids, and Recruitment of the Jasmonic Acid Pathway. *The Plant Cell* **25**, 1818-1839, doi:10.1105/tpc.113.111120 (2013).
- 224 Leustek, T. Sulfate Metabolism. *The Arabidopsis Book / American Society of Plant Biologists* **1**, e0017, doi:10.1199/tab.0017 (2002).
- 225 Emanuelsson, O., Nielsen, H. & von Heijne, G. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci* **8**, 978-984, doi:10.1110/ps.8.5.978 (1999).
- 226 Kahlau, S., Aspinall, S., Gray, J. C. & Bock, R. Sequence of the tomato chloroplast DNA and evolutionary comparison of solanaceous plastid genomes. *J Mol Evol* **63**, 194-207, doi:10.1007/s00239-005-0254-5 (2006).
- 227 Shimada, H. & Sugiura, M. Fine structural features of the chloroplast genome: comparison of the sequenced chloroplast genomes. *Nucleic acids research* **19**, 983-995 (1991).
- 228 Terry, M. J. & Smith, A. G. A model for tetrapyrrole synthesis as the primary mechanism for plastid-to-nucleus signaling during chloroplast biogenesis. *Frontiers in Plant Science* **4**, 14, doi:10.3389/fpls.2013.00014 (2013).
- 229 Ruckle, M. E., DeMarco, S. M. & Larkin, R. M. Plastid signals remodel light signaling networks and are essential for efficient chloroplast biogenesis in Arabidopsis. *Plant Cell* **19**, 3944-3960, doi:10.1105/tpc.107.054312 (2007).
- 230 Brand, A. *et al.* pc8.1, a major QTL for pigment content in pepper fruit, is associated with variation in plastid compartment size. *Planta* **235**, 579-588, doi:10.1007/s00425-011-1530-9 (2012).
- 231 Gutierrez, R. A. *et al.* Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene CCA1. *Proc Natl Acad Sci U S A* **105**, 4939-4944, doi:10.1073/pnas.0800211105 (2008).

- 232 Savitch, L. V., Subramaniam, R., Allard, G. C. & Singh, J. The GLK1 'regulon' encodes disease defense related proteins and confers resistance to *Fusarium graminearum* in *Arabidopsis*. *Biochem Biophys Res Commun* **359**, 234-238, doi:10.1016/j.bbrc.2007.05.084 (2007).
- 233 Kakizaki, T. *et al.* Coordination of Plastid Protein Import and Nuclear Gene Expression by Plastid-to-Nucleus Retrograde Signaling. *Plant Physiology* **151**, 1339-1353, doi:10.1104/pp.109.145987 (2009).
- 234 Cavalier, D. M. *et al.* Disrupting Two *Arabidopsis thaliana* Xylosyltransferase Genes Results in Plants Deficient in Xyloglucan, a Major Primary Cell Wall Component. *The Plant Cell* **20**, 1519-1537, doi:10.1105/tpc.108.059873 (2008).
- 235 Rose, J. K. C. & Bennett, A. B. Cooperative disassembly of the cellulose-xyloglucan network of plant cell walls: parallels between cell expansion and fruit ripening. *Trends in Plant Science* **4**, 176-183, doi:[https://doi.org/10.1016/S1360-1385\(99\)01405-3](https://doi.org/10.1016/S1360-1385(99)01405-3) (1999).
- 236 Burch-Smith, T. M., Brunkard, J. O., Choi, Y. G. & Zambryski, P. C. Organelle-nucleus cross-talk regulates plant intercellular communication via plasmodesmata. *Proc Natl Acad Sci U S A* **108**, E1451-1460, doi:10.1073/pnas.1117226108 (2011).
- 237 Oyama, T., Shimura, Y. & Okada, K. The *Arabidopsis* HY5 gene encodes a bZIP protein that regulates stimulus-induced development of root and hypocotyl. *Genes & development* **11**, 2983-2995 (1997).
- 238 Levin, I., Frankel, P., Gilboa, N., Tanny, S. & Lalazar, A. The tomato dark green mutation is a novel allele of the tomato homolog of the DEETIOLATED1 gene. *Theor Appl Genet* **106**, 454-460, doi:10.1007/s00122-002-1080-4 (2003).
- 239 Miao, M. *et al.* The tomato DWD motif-containing protein DDI1 interacts with the CUL4-DDB1-based ubiquitin ligase and plays a pivotal role in abiotic stress responses. *Biochem Biophys Res Commun* **450**, 1439-1445, doi:10.1016/j.bbrc.2014.07.011 (2014).
- 240 Ostersetzer, O. & Adam, Z. Light-stimulated degradation of an unassembled Rieske FeS protein by a thylakoid-bound protease: the possible role of the FtsH protease. *Plant Cell* **9**, 957-965, doi:10.1105/tpc.9.6.957 (1997).

- 241 Sakamoto, W., Tamura, T., Hanba-Tomita, Y. & Murata, M. The VAR1 locus of Arabidopsis encodes a chloroplastic FtsH and is responsible for leaf variegation in the mutant alleles. *Genes to cells : devoted to molecular & cellular mechanisms* **7**, 769-780 (2002).
- 242 Zhang, D. *et al.* The FtsH protease heterocomplex in Arabidopsis: dispensability of type-B protease activity for proper chloroplast development. *Plant Cell* **22**, 3710-3725, doi:10.1105/tpc.110.079202 (2010).
- 243 Evans, J. R. Photosynthesis and nitrogen relationships in leaves of C 3 plants. *Oecologia* **78**, 9-19 (1989).
- 244 van Wijk, K. J. & Baginsky, S. Plastid proteomics in higher plants: current state and future goals. *Plant Physiol* **155**, 1578-1588, doi:10.1104/pp.111.172932 (2011).
- 245 Faurobert, M., Pelpoir, E. & Chaib, J. Phenol extraction of proteins for proteomic studies of recalcitrant plant tissues. *Methods Mol Biol* **355**, 9-14, doi:10.1385/1-59745-227-0:9 (2007).
- 246 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254 (1976).
- 247 Ramagli, L. S., Capetillo, S., Becker, F. F. & Rodriguez, L. V. Alterations in nonhistone chromatin proteins during hepatocarcinogenesis induced by diverse acting carcinogens. *Carcinogenesis* **6**, 367-375 (1985).
- 248 Kilambi, H. V. *et al.* Shotgun Proteomics of Tomato Fruits: Evaluation, Optimization and Validation of Sample Preparation Methods and Mass Spectrometric Parameters. *Frontiers in Plant Science* **7**, 969, doi:10.3389/fpls.2016.00969 (2016).
- 249 Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685 (1970).
- 250 Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature protocols* **1**, 2856-2860, doi:10.1038/nprot.2006.468 (2006).
- 251 Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-1111, doi:10.1093/bioinformatics/btp120 (2009).

- 252 Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169, doi:10.1093/bioinformatics/btu638 (2015).
- 253 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
- 254 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300 (1995).
- 255 Stewart, A. J. *et al.* Occurrence of flavonols in tomatoes and tomato-based products. *J Agric Food Chem* **48**, 2663-2669 (2000).
- 256 George, B., Kaur, C., Khurdiya, D. S. & Kapoor, H. C. Antioxidants in tomato (*Lycopersium esculentum*) as a function of genotype. *Food Chemistry* **84**, 45-51, doi:[http://doi.org/10.1016/S0308-8146\(03\)00165-1](http://doi.org/10.1016/S0308-8146(03)00165-1) (2004).
- 257 Gould, K. S. Nature's Swiss Army Knife: The Diverse Protective Roles of Anthocyanins in Leaves. *Journal of biomedicine & biotechnology* **2004**, 314-320, doi:10.1155/s1110724304406147 (2004).
- 258 Raiola, A., Rigano, M. M., Calafiore, R., Frusciante, L. & Barone, A. Enhancing the Health-Promoting Effects of Tomato Fruit for Biofortified Food. *Mediators of Inflammation* **2014**, 139873, doi:10.1155/2014/139873 (2014).
- 259 Mehlenbacher, S. A. & Thompson, M. M. Inheritance of a Chlorophyll Deficiency in Hazelnut. *Journal HORTSCIENCE* **26**, 1414-1416 (1991).
- 260 Winkel-Shirley, B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol* **126**, 485-493 (2001).
- 261 Hughes, N. M., Neufeld, H. S. & Burkey, K. O. Functional role of anthocyanins in high-light winter leaves of the evergreen herb *Galax urceolata*. *New Phytol* **168**, 575-587, doi:10.1111/j.1469-8137.2005.01546.x (2005).
- 262 Dominy, N. J. *et al.* Why Are Young Leaves Red? *Oikos* **98**, 163-176 (2002).
- 263 Murshed, R., Lopez-Lauri, F. & Sallanon, H. Effect of salt stress on tomato fruit antioxidant systems depends on fruit

- development stage. *Physiology and Molecular Biology of Plants* **20**, 15-29, doi:10.1007/s12298-013-0209-z (2014).
- 264 Subramanian, P. *et al.* Cold Stress Tolerance in Psychrotolerant Soil Bacteria and Their Conferred Chilling Resistance in Tomato (*Solanum lycopersicum* Mill.) under Low Temperatures. *PLOS ONE* **11**, e0161592, doi:10.1371/journal.pone.0161592 (2016).
- 265 Florido Bacallao, M. & Bao Fundora, L. Tolerancia a estrés por déficit hídrico en tomate (*Solanum lycopersicum* L.). *Cultivos Tropicales* **35**, 70-88 (2014).
- 266 Mes, P. J., Boches, P., Myers, J. R. & Durst, R. Characterization of tomatoes expressing anthocyanin in the fruit. *Journal of the American Society for Horticultural Science* **133**, 262-269 (2008).
- 267 Boches, P. & Myers, J. in *HortScience*. 856-856 (AMER SOC HORTICULTURAL SCIENCE 113 S WEST ST, STE 200, ALEXANDRIA, VA 22314-2851 USA).
- 268 Adato, A. *et al.* Fruit-surface flavonoid accumulation in tomato is controlled by a SIMYB12-regulated transcriptional network. *PLoS Genet* **5**, e1000777, doi:10.1371/journal.pgen.1000777 (2009).
- 269 Borevitz, J. O., Xia, Y., Blount, J., Dixon, R. A. & Lamb, C. Activation Tagging Identifies a Conserved MYB Regulator of Phenylpropanoid Biosynthesis. *The Plant Cell* **12**, 2383-2393 (2000).
- 270 Feller, A., Machemer, K., Braun, E. L. & Grotewold, E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *The Plant Journal* **66**, 94-116, doi:10.1111/j.1365-313X.2010.04459.x (2011).
- 271 Walker, A. R. *et al.* The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. *Plant Cell* **11**, 1337-1350 (1999).
- 272 Quattrocchio, F. <AN2 PETUNIA.pdf>. (1999).
- 273 Ramsay, N. A. & Glover, B. J. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci* **10**, 63-70, doi:10.1016/j.tplants.2004.12.011 (2005).
- 274 Schellmann, S. *et al.* TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in Arabidopsis. *Embo j* **21**, 5036-5046 (2002).
- 275 Quattrocchio, F. *et al.* PH4 of Petunia is an R2R3 MYB protein that activates vacuolar acidification through interactions with basic-helix-loop-helix transcription factors of the anthocyanin

- pathway. *Plant Cell* **18**, 1274-1291, doi:10.1105/tpc.105.034041 (2006).
- 276 Serna, L. & Martin, C. Trichomes: different regulatory networks lead to convergent structures. *Trends Plant Sci* **11**, 274-280, doi:10.1016/j.tplants.2006.04.008 (2006).
- 277 Glover, B. J., Perez-Rodriguez, M. & Martin, C. Development of several epidermal cell types can be specified by the same MYB-related plant transcription factor. *Development* **125**, 3497-3508 (1998).
- 278 Gould, K., Davies, K. M. & Winefield, C. *Anthocyanins: biosynthesis, functions, and applications*. (Springer Science & Business Media, 2008).
- 279 Kovinich, N., Kayanja, G., Chanoca, A., Otegui, M. S. & Grotewold, E. Abiotic stresses induce different localizations of anthocyanins in Arabidopsis. *Plant Signaling & Behavior* **10**, e1027850, doi:10.1080/15592324.2015.1027850 (2015).
- 280 Von Wettstein-Knowles, P. MUTATIONS AFFECTING ANTHOCYANIN SYNTHESIS IN THE TOMATO. *Hereditas* **60**, 317-346, doi:10.1111/j.1601-5223.1968.tb02212.x (1968).
- 281 Weiss, D. Regulation of flower pigmentation and growth: Multiple signaling pathways control anthocyanin synthesis in expanding petals. *Physiologia plantarum* **110**, 152-157, doi:10.1034/j.1399-3054.2000.110202.x (2000).
- 282 Cominelli, E. *et al.* Expression analysis of anthocyanin regulatory genes in response to different light qualities in Arabidopsis thaliana. *J Plant Physiol* **165**, 886-894, doi:10.1016/j.jplph.2007.06.010 (2008).
- 283 Michniewicz, M., Brewer, P. B. & Friml, J. Polar Auxin Transport and Asymmetric Auxin Distribution. *The Arabidopsis Book / American Society of Plant Biologists* **5**, e0108, doi:10.1199/tab.0108 (2007).
- 284 Geisler, M. *et al.* Cellular efflux of auxin catalyzed by the Arabidopsis MDR/PGP transporter AtPGP1. *The Plant Journal* **44**, 179-194, doi:10.1111/j.1365-313X.2005.02519.x (2005).
- 285 Brown, D. E. *et al.* Flavonoids act as negative regulators of auxin transport in vivo in Arabidopsis. *Plant Physiol* **126**, 524-535 (2001).
- 286 Peer, W. A. & Murphy, A. S. Flavonoids and auxin transport: modulators or regulators? *Trends Plant Sci* **12**, 556-563 (2007).

- 287 Kant, S., Bi, Y.-M., Zhu, T. & Rothstein, S. J. SAUR39, a Small Auxin-Up RNA Gene, Acts as a Negative Regulator of Auxin Synthesis and Transport in Rice. *Plant Physiology* **151**, 691-701, doi:10.1104/pp.109.143875 (2009).
- 288 Tanaka, H., Dhonukshe, P., Brewer, P. B. & Friml, J. Spatiotemporal asymmetric auxin distribution: a means to coordinate plant development. *Cellular and Molecular Life Sciences CMLS* **63**, 2738-2754, doi:10.1007/s00018-006-6116-5 (2006).
- 289 Besseau, S. *et al.* Flavonoid accumulation in Arabidopsis repressed in lignin synthesis affects auxin transport and plant growth. *The Plant Cell* **19**, 148-162 (2007).
- 290 Mes, P. J. *Breeding tomatoes for improved antioxidant activity*, (2004).
- 291 Taylor, L. P. & Grotewold, E. Flavonoids as developmental regulators. *Curr Opin Plant Biol* **8**, 317-323, doi:10.1016/j.pbi.2005.03.005 (2005).
- 292 Mathews, H. Activation Tagging in Tomato Identifies a Transcriptional Regulator of Anthocyanin Biosynthesis, Modification, and Transport. *The Plant Cell Online* **15**, 1689-1703, doi:10.1105/tpc.012963 (2003).
- 293 Rocha, M. d. C., Deliza, R., Corrêa, F. M., Carmo, M. G. F. d. & Abboud, A. C. S. A study to guide breeding of new cultivars of organic cherry tomato following a consumer-driven approach. *Food Research International* **51**, 265-273, doi:10.1016/j.foodres.2012.12.019 (2013).
- 294 LOPEZ CAMELO, A. F. & GOMEZ, P. A. Comparison of color indexes for tomato ripening. *Hortic. Bras.* **22**, 534-537 (2004).
- 295 Tuohy, J. & Choinski, J. Comparative photosynthesis in developing leaves of *Brachystegia spiciformis* Benth. *Journal of Experimental Botany* **41**, 919-923 (1990).
- 296 Gould, K. S., Vogelmann, T. C., Han, T. & Clearwater, M. J. Profiles of photosynthesis within red and green leaves of *Quintinia serrata*. *Physiologia plantarum* **116**, 127-133, doi:10.1034/j.1399-3054.2002.1160116.x (2002).
- 297 Manetas, Y., Petropoulou, Y., Psaras, G. K. & Drinia, A. Exposed red (anthocyanic) leaves of *Quercus coccifera* display shade characteristics. *Functional Plant Biology* **30**, 265-270 (2003).

- 298 Darrigues, A. *Dissecting variation in tomato fruit color quality through digital phenotyping and genetic mapping*, The Ohio State University, (2007).
- 299 Corona, V. *et al.* Regulation of a carotenoid biosynthesis gene promoter during plant development. *Plant J* **9**, 505-512 (1996).
- 300 Bruno, A. K. & Wetzels, C. M. The early light-inducible protein (ELIP) gene is expressed during the chloroplast-to-chromoplast transition in ripening tomato fruit. *J Exp Bot* **55**, 2541-2548, doi:10.1093/jxb/erh273 (2004).
- 301 Miyazawa, T. & Nakagawa, K. Structure-related emission spectrometric analysis of the chemiluminescence of catechins, theaflavins and anthocyanins. *Bioscience, biotechnology, and biochemistry* **62**, 829-832 (1998).
- 302 Schillmiller, A. L. *et al.* Studies of a biochemical factory: tomato trichome deep expressed sequence tag sequencing and proteomics. *Plant Physiol* **153**, 1212-1223, doi:10.1104/pp.110.157214 (2010).
- 303 Glas, J. J. *et al.* Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *Int J Mol Sci* **13**, 17077-17103, doi:10.3390/ijms131217077 (2012).
- 304 Pattanaik, S., Patra, B., Singh, S. K. & Yuan, L. An overview of the gene regulatory network controlling trichome development in the model plant, Arabidopsis. *Frontiers in Plant Science* **5**, 259, doi:10.3389/fpls.2014.00259 (2014).
- 305 Adamska, I. <1071167.pdf>. (1995).
- 306 Solfanelli, C., Poggi, A., Loreti, E., Alpi, A. & Perata, P. Sucrose-specific induction of the anthocyanin biosynthetic pathway in Arabidopsis. *Plant Physiol* **140**, 637-646, doi:10.1104/pp.105.072579 (2006).
- 307 De Jong, H. Inheritance of pigmented tuber flesh in cultivated diploid potatoes. *American Potato Journal* **64**, 337-343, doi:10.1007/BF02853595 (1987).
- 308 Livingstone, K. D., Lackney, V. K., Blauth, J. R., van Wijk, R. & Jahn, M. K. Genome mapping in capsicum and the evolution of genome structure in the solanaceae. *Genetics* **152**, 1183-1202 (1999).
- 309 Aza-González, C., Núñez-Palenius, H. G. & Ochoa-Alejo, N. Molecular biology of chili pepper anthocyanin biosynthesis. *Journal of the Mexican Chemical Society* **56**, 93-98 (2012).

- 310 Davies, K. M., Albert, N. W. & Schwinn, K. E. From landing lights to mimicry: the molecular regulation of flower colouration and mechanisms for pigmentation patterning. *Functional Plant Biology* **39**, 619-638 (2012).
- 311 Zhang, H., Wang, M., Huang, X. & Roth, H. Application of adaptive genetic algorithm for the parameter selection of support vector regression. *International Journal of Modelling, Identification and Control* **21**, 29-37 (2014).
- 312 Winefield, C. S. *et al.* Investigation of the biosynthesis of 3-deoxyanthocyanins in *Sinningia cardinalis*. *Physiologia plantarum* **124**, 419-430 (2005).
- 313 Tissier, A. *Trichome Specific Expression: Promoters and Their Applications*. (INTECH Open Access Publisher, 2012).
- 314 Gutierrez-Alcala, G. *et al.* A versatile promoter for the expression of proteins in glandular and non-glandular trichomes from a variety of plants. *J Exp Bot* **56**, 2487-2494, doi:10.1093/jxb/eri241 (2005).
- 315 Ennajdaoui, H. *et al.* Trichome specific expression of the tobacco (*Nicotiana glauca*) cembratrien-ol synthase genes is controlled by both activating and repressing cis-regions. *Plant Molecular Biology* **73**, 673-685, doi:10.1007/s11103-010-9648-x (2010).
- 316 Wang, S. *et al.* Control of Plant Trichome Development by a Cotton Fiber MYB Gene. *The Plant Cell* **16**, 2323-2334, doi:10.1105/tpc.104.024844 (2004).
- 317 Szymanski, D. B., Jilk, R. A., Pollock, S. M. & Marks, M. D. Control of GL2 expression in *Arabidopsis* leaves and trichomes. *Development* **125**, 1161-1171 (1998).
- 318 Kennedy, G. G. Tomato, pests, parasitoids, and predators: tritrophic interactions involving the genus *Lycopersicon*. *Annual Review of Entomology* **48**, 51-72 (2003).
- 319 Van Schie, C. C. N. *et al.* Geranyl diphosphate synthase is required for biosynthesis of gibberellins. *The Plant Journal* **52**, 752-762, doi:10.1111/j.1365-313X.2007.03273.x (2007).
- 320 Balcke, G. *et al.* Multiomics of tomato glandular trichomes reveals distinct features of central carbon metabolism supporting high productivity of specialized metabolites. *Plant Cell*, doi:10.1105/tpc.17.00060 (2017).
- 321 Gonzalez, A., Zhao, M., Leavitt, J. M. & Lloyd, A. M. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb

- transcriptional complex in Arabidopsis seedlings. *The Plant Journal* **53**, 814-827, doi:10.1111/j.1365-313X.2007.03373.x (2008).
- 322 Tanksley, S. *et al.* High density molecular linkage maps of the
tomato and potato genomes. *Genetics* **132**, 1141-1160 (1992).
- 323 CM, R., CF, Q., WH, L. & MA, S. Monogenic control of resistance
in tomato to tobacco flea beetle: probable repellence by foliar
volatiles. *Euphytica* **25**, 521–530 (1976).
- 324 Kang, J. H. *et al.* The flavonoid biosynthetic enzyme chalcone
isomerase modulates terpenoid production in glandular
trichomes of tomato. *Plant Physiol* **164**, 1161-1174,
doi:10.1104/pp.113.233395 (2014).
- 325 Karabourniotis, G. & Bornman, J. F. Penetration of UV-A, UV-B
and blue light through the leaf trichome layers of two
xeromorphic plants, olive and oak, measured by optical fibre
microprobes. *Physiologia plantarum* **105**, 655-661,
doi:10.1034/j.1399-3054.1999.105409.x (1999).
- 326 Gould, K. S., Kuhn, D. N., Lee, D. W. & Oberbauer, S. F. Why
leaves are sometimes red. *Nature* **378**, 241-242 (1995).
- 327 Wagner, G. J., Wang, E. & Shepherd, R. W. New approaches for
studying and exploiting an old protuberance, the plant trichome.
Ann Bot **93**, 3-11, doi:10.1093/aob/mch011 (2004).
- 328 Luckwill, L. C. The genus *Lycopersicon*: an historical, biological,
and taxonomic survey of the wild and cultivated tomatoes.
Aberdeen University Press **44** (1943).
- 329 McDowell, E. T. *et al.* Comparative functional genomic analysis
of *Solanum* glandular trichome types. *Plant Physiol* **155**, 524-
539, doi:10.1104/pp.110.167114 (2011).
- 330 Iwashina, T. The structure and distribution of the flavonoids in
plants. *Journal of Plant Research* **113**, 287-299 (2000).
- 331 Su, X. *et al.* Identification and quantification of anthocyanins in
transgenic purple tomato. *Food Chem* **202**, 184-188,
doi:10.1016/j.foodchem.2016.01.128 (2016).
- 332 Alba, R., Cordonnier-Pratt, M.-M. & Pratt, L. H. Fruit-localized
phytochromes regulate lycopene accumulation independently of
ethylene production in tomato. *Plant Physiology* **123**, 363-370
(2000).
- 333 Feild, T. S., Lee, D. W. & Holbrook, N. M. Why Leaves Turn Red in
Autumn. The Role of Anthocyanins in Senescing Leaves of Red-
Osier Dogwood. *Plant Physiology* **127**, 566-574 (2001).

- 334 Neill, S. O. & Gould, K. S. Anthocyanins in leaves: light attenuators or antioxidants? *Functional Plant Biology* **30**, 865-873 (2003).
- 335 Neill, S. & Gould, K. S. Optical properties of leaves in relation to anthocyanin concentration and distribution. *Canadian Journal of Botany* **77**, 1777-1782 (2000).
- 336 Harvaux, M. & Kloppstech, K. The protective functions of carotenoid and flavonoid pigments against excess visible radiation at chilling temperature investigated in *Arabidopsis* npq and tt mutants. *Planta* **213**, 953-966 (2001).
- 337 Burger, J. & Edwards, G. E. Photosynthetic efficiency, and photodamage by UV and visible radiation, in red versus green leaf coleus varieties. *Plant and Cell Physiology* **37**, 395-399 (1996).
- 338 Lee, D. W. & Collins, T. M. Phylogenetic and ontogenetic influences on the distribution of anthocyanins and betacyanins in leaves of tropical plants. *International Journal of Plant Sciences* **162**, 1141-1153 (2001).
- 339 Gonzali, S., Mazzucato, A. & Perata, P. Purple as a tomato: towards high anthocyanin tomatoes. *Trends Plant Sci* **14**, 237-241, doi:10.1016/j.tplants.2009.02.001 (2009).
- 340 Marti, C. *et al.* Silencing of DELLA induces facultative parthenocarpy in tomato fruits. *Plant J* **52**, 865-876, doi:10.1111/j.1365-313X.2007.03282.x (2007).
- 341 Phillips, M. A., D'Auria, J. C., Luck, K. & Gershenzon, J. Evaluation of Candidate Reference Genes for Real-Time Quantitative PCR of Plant Samples Using Purified cDNA as Template. *Plant Molecular Biology Reporter / Isymb* **27**, 407-416, doi:10.1007/s11105-008-0072-1 (2009).
- 342 Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods (San Diego, Calif.)* **25**, 402-408, doi:10.1006/meth.2001.1262 (2001).
- 343 Fasano, C. *et al.* Transcriptome and metabolome of synthetic *Solanum* autotetraploids reveal key genomic stress events following polyploidization. *New Phytol* **210**, 1382-1394, doi:10.1111/nph.13878 (2016).
- 344 Atkins, C. A., Kuo, J. & Pate, J. S. Photosynthetic Pod Wall of Pea (*Pisum sativum* L.): Distribution of Carbon Dioxide-fixing

- Enzymes in Relation to Pod Structure. *Plant Physiol* **60**, 779-786 (1977).
- 345 Listiawan, D. A., Tanoue, R., Kobayashi, K. & Masuda, T. Expression Analysis of Transcription Factors Involved in Chloroplast Differentiation. *Procedia Chemistry* **14**, 146-151, doi:<http://dx.doi.org/10.1016/j.proche.2015.03.021> (2015).
- 346 Chikkala, V. R., Nugent, G. D., Stalker, D. M., Mouradov, A. & Stevenson, T. W. Expression of Brassica oleracea FtsZ1-1 and MinD alters chloroplast division in Nicotiana tabacum generating macro- and mini-chloroplasts. *Plant cell reports* **31**, 917-928, doi:10.1007/s00299-011-1212-x (2012).
- 347 Rose, J. K., Braam, J., Fry, S. C. & Nishitani, K. The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature. *Plant Cell Physiol* **43**, 1421-1435 (2002).
- 348 Cosgrove, D. J. Growth of the plant cell wall. *Nat Rev Mol Cell Biol* **6**, 850-861, doi:http://www.nature.com/nrm/journal/v6/n11/supinfo/nrm1746_S1.html (2005).
- 349 Tang, Y., Fu, X., Shen, Q. & Tang, K. Roles of MPBQ-MT in Promoting α/γ -Tocopherol Production and Photosynthesis under High Light in Lettuce. *PLoS ONE* **11**, e0148490, doi:10.1371/journal.pone.0148490 (2016).
- 350 Tamai, H., Iwabuchi, M. & Meshi, T. Arabidopsis GARP transcriptional activators interact with the Pro-rich activation domain shared by G-box-binding bZIP factors. *Plant Cell Physiol* **43**, 99-107 (2002).
- 351 Mohan, V. *et al.* Tomato Fruits Show Wide Phenomic Diversity but Fruit Developmental Genes Show Low Genomic Diversity. *PLoS ONE* **11**, e0152907, doi:10.1371/journal.pone.0152907 (2016).