Molecular and Biochemical Mechanisms Underlying Carotenoid Biosynthesis in *Capsicum annuum*

Rebecca Mary Nohl

This thesis was submitted for the degree of Doctor of Philosophy at Royal Holloway, University of London, August 2017

Declaration of Authorship

I, Rebecca Nohl, hereby declare that the work presented in this thesis is the original work of the author unless otherwise stated. Original material used in the production of this thesis has not been previously submitted either in part or whole for a degree of any description from any institution.

Signed:_____

Date:_____

Abstract

Colour is a commercially important quality trait for many fruit crops, including the sweet bell pepper (*Capsicum annuum*). Due to the significant financial value and nutritional benefits conferred on crop products by colour, it is essential that we understand the mechanisms which govern its development. Mature fruit colour in Capsicum is controlled by three loci: *c1, c2*, and *y*. The *c2* locus has been associated with the gene encoding the fruit specific enzyme catalysing the first step of carotenoid biosynthesis, phytoene synthase (*Psy-1*), the y locus with the gene encoding the enzyme capsanthin-capsorubin synthase (*Ccs*), and the identity of the *c1* locus remains undetermined. The present work aims to investigate fruit colour from two different perspectives: orange and yellow fruit phenotypes in relation to the ripening-fruit copy of the enzyme catalysing the first step in carotenoid biosynthesis (*Psy-1*), and yellow and white fruit phenotypes in relation to the *c1* locus.

Firstly, a discovery panel of sixteen *Capscium annuum* accessions ranging in colour from red to white was used to further investigate the role of the enzyme PSY-1 in determining pepper mature fruit colour. Reverse-phased chromatography and quantitative real-time PCR (qRT-PCR) were used to investigate the molecular mechanisms governing colour in this panel of lines. It was demonstrated that *Psy-1* expression correlated with colour, and that expression levels of most other genes in the carotenoid biosynthesis pathway were unaffected by a lack of *Psy-1* expression; validating that this is the most important regulatory gene in the pathway. The structure of the plastids in which carotenoids accumulate and the specialisations of their sub-compartments were investigated with electron microscopy, sub-chromoplast fractionation and enzyme activity assays. These studies showed that PSY-1 protein was absent from accessions lacking Psy-1 transcript; that the enzyme is located in the same plastid sub-structure as most of the carotenoids it synthesises; and that this is also the region of the plastid where the enzyme is active. Gas Chromatography-Mass Spectrometry (GC-MS) was also employed to elucidate the impact of a lack of *Psy-1* on the wider metabolism of the fruit, revealing that the largest metabolic changes were in fatty acids capable of esterifying carotenoids. The multiple lines of investigation validate the contribution of *Psy-1* to fruit colour.

A second diversity panel consisted of twelve lines, none of which possessed a functional copy of either the *y* or *c2* loci. Six were positive for the *c1* locus and produced yellow ripe fruit, and six were negative for all three loci and produced cream coloured or white ripe fruit. Carotenoids were profiled, and broader fruit metabolism was investigated with GC-MS. This demonstrated that carotenoids were only detectable at trace levels in ripe fruit of lines lacking the c1 locus, and that these lines possessed elevated levels of Very Long Chain Fatty Acids (VLCFAs), which are too long to esterify carotenoids. This could have implications for their storage during the chloroplast to chromoplast transition. Chloroplast and chromoplast structures were studied with electron microscopy, revealing that accessions lacking functional copies of the c1 locus possessed aberrant chloroplast structures at the mature green stage, which could alter their ability to retain carotenoids throughout ripening. Furthermore, the transcriptomes of *c1* positive and negative accessions were generated with RNA-seq, and interrogated for potential candidate genes for association with the *c1* locus. Candidate genes for further analysis were generated; these include a NAC transcription factor, an acyl transferase, and a plastid-specific kinase, among others.

Acknowledgements

I have thoroughly enjoyed my time working on this project in Professor Paul Fraser's research group. I am so grateful for Professor Fraser's inspiring supervision and all the wonderful opportunities I have received: learning a huge range of techniques, meeting the international academic community through attending meetings and conferences, and the freedom to follow fascinating research avenues.

I would also like to thank my industrial partners at Syngenta; I have appreciated the support and feedback of Dr. Daniel Rickett and Dr. Brandon Hurr. Plant material provided by Syngenta was essential to the project, as was access to their sequence databases. I would also like to thank Syngenta for the generation RNA-seq data at their Research Triangle Park facility in North Carolina.

I am thankful to the DISCO EU consortium for the opportunity to learn about the processing of RNA-seq data at Istituto di Genomica Applicata (IGA) in Udine, Italy. I gained so much from the experience of working in another high-quality institute; the informatics skills I was taught were extremely useful for the project.

Completing my PhD in the Fraser Lab has been an absolute pleasure; throughout the last four years I have thoroughly enjoyed working with such an interesting and enthusiastic group of people. I would like to thank everyone for their constructive feedback and useful discussions. I would particularly like to thank Dr. Genny Enfissi, Dr. Marilise Nogueira, Dr. Harriet Berry, Dr. Margit Drapal and Dr. Elliott Price for their consistent help and advice. Within the Fraser Lab, it has been a joy to be part of 'Team Pepper' which will hopefully continue to grow.

This project would have been much more difficult without my friends, both nearby and far away, and my boyfriend Matthew. Thank you all for your support! Partticular thanks to Dan and Rosa for living with me throughout the project and listening to a lot of pepper talk!

Finally, I would like to thank my family; my parents, Jenny and Francis, and my siblings Miri and James, for their constant love and encouragement.

Contents

Table of Figures

Table of Tables

List of Abbreviations

ABA	Abscisic acid
ABCK	Activity of bc_1 complex kinase
AMD	Age-related macular degeneration
B	Breaker ripening stage
BCH-1	Beta-carotene hydroxylase-1
BCH-2	Beta-carotene hydroxylase-2
BCO	Beta-carotene oxygenase
C	Carbon
Catf1	Capsicum acyl transferase-1
Catf2	Capsicum acyl transferase-2
CCD	Carotenoid Cleavage Dioxygenase
CCS	Capsanthin-capsorubin synthase
Chl	Chlorophyll
CL	Chlorophyll retainer protein
CoA	CoEnzyme A
COMPASS	Collaborative Optical Macular Pigment Assessment Study
CRISPR	Clustered Regularly Interspersed Short Palindromic Repeats
CRTI	Bacterial phytoene desaturase
CRTISO	Carotene isomerase
CVD	Cardiovascular disease
CYC-B	Fruit specific lycopene β-cyclase
DGDG	Digalactosyl-diacylglycerol
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
DSB	Double stranded break
DW	Dry weight
DXP	1-deoxy-D-xylulose 5-phosphate
DXR	1-deoxy-D-xylulose 5-phosphate reductase
DXS	1-deoxy-D-xylulose 5-phosphate synthase
ECR	2,3-enoyl-CoA reductase
EDTA	Ethylenediaminetetraacetic acid
ELO	Elongase
ER	Endoplasmic reticulum
FAE	Fatty acid elongase
FAME	Fatty acid methyl ester
FAS	Fatty acid synthase complex
FIB	FIBRILLIN protein
FPKM	Frequency per kilobase million
FW	Fresh weight
GC-MS	Gas chromatography-mass spectrometry
GDL62	Glycine-aspartic acid-serine-leucine esterase/lipase
GGPP	Geranylgeranyl pyrophosphate
GGPS	Geranylgeranyl pyrophosphate synthase
H	Hydrogen
HCD	3-hydroxy-acyl-CoA dehydratase
HLEC	Human lens epithelial cells

HPLC	High Performance Liquid Chromatography
ip	Isoprene unit
IPP	Isopentenyl pyrophosphate
IPPI	Isopentenyl pyrophosphate isomerase
KCR	3-keto-acyl-CoA reductase
KCS	3-keto-acyl-CoA synthase
kDa	Kilodaltons
LC-MS	Liquid chromatography-mass spectrometry
LDL	Low density lipoprotein
LEDC	Less economically developed country
LHC	Light harvesting complex
εLCY	Lycopene ε-cyclase
MAS	Marker assisted selection
MEP	2-C-methylerythritol 4-phosphate
MG	Mature green ripening stage
MGDG	Monogalactosyl-diacylglycerol
MTA	Material Transfer Agreement
MUFA	Monounsaturated fatty acid
NAC	No-apical-meristem (NAM), ATAF-1 and 2, Cup shaped cotyledon (CUC2) transcription factor
NAC37	No-apical-meristem (NAM), ATAF-1 and 2, Cup shaped cotyledon (CUC2) 37 transcription factor
NaCl	Sodium chloride
NCBI	National Center for Biotechnology Information
NIH	National Institutes of Health
NOR	Nonripening
NPQ	Non-photocemical quenching
NXS	Neoxanthin synthase
O	Oxygen
OPR1	Oxophytodienoate reductase
OR	ORANGE chromoplast development protein
PAGE	Poly-acrylimide gel electrophoresis
PAS2	PASTICCINO2 protein
PC	Phosphatidylcholine
PCA	Principal Components Analysis
PDA	Photo Diode Array
PDS	Phytoene desaturase
PE	Phosphatidylethanolamine
PG	Plastoglobule OR Phosphatidylglycerol
PLS-DA	Partial Least Squares Discriminant Analysis
PS	Phosphatidylserine
PSI	Photosystem I
PSII	Photosystem II
PSY-1	Phytoene synthase-1; fruit specific
PSY-2	Phytoene synthase-2; leaf specific
PSY-3	Phytoene-synthase-3; root specific
PUFA	Poly-unsaturated fatty acid
Pun1	Pungency locus encoding acyl transferase AT3
PVDF	Polyvinylidene fluoride

qPCR	Quantitative polymerase chain reaction
QTL	Quantitative trait loci
R	Ripe fruit stage
RAE	Retinol Activity Equivalents
RbcL	Rubisco large subunit
RC	Reaction centre
RFLP	Restriction fragment length polymorphism
RIN	Ripening inhibitor
RNA	Ribonucleic acid
ROS	Reactive Oxygen Species
RT-PCR	Reverse transcription polymerase chain reaction
RuBisCO	Ribulose 1,5-bisphosphate carboxylase/oxygenase
SD	Standard deviation
SDN	Site-directed nuclease
SDS	Sodium dodecyl sulphate
SEM	Standard error of the mean
SGR	STAY GREEN protein
T	Turning ripening stage
TAG	Triacylglycerol
TB	Transfer buffer
TEM	Transmission Electron Microscopy
Tris	Tris-hydroxymethyl-aminomethane
UPLC	Ultra-high Performance Liquid Chromatography
UQ	Ubiquinone
V	Volt
v	Volume
VDE	Violaxanthin de-epoxidase
VLCFA	Very long chain fatty acid
w	Weight
ZDS	Zeta-carotene desaturase
ZEP	Zeaxanthin epoxidase
Z-ISO	Zeta-carotene isomerase

1 Introduction

1.1 Capsicum annuum

1.1.1 The *Capsicum* Genus

The genus *Capsicum* is a member of the Solanaceae family. This group of dicotyledonous angiosperms contains some of the world's most economically important crops, including tomato, tobacco, potato, and aubergine as well as pepper. Over the last few years, the publication of complete genomes for several members of the family has provided a platform from which researchers in phylogenetics can build an evolutionary framework (Hirakawa et al., 2014; Kim et al., 2014; The Potato Genome Consortium, 2011; The Tomato Genome Consortium, 2012). In 2013, a phylogeny generated from DNA sequences of seven loci from 1075 Solanaceae species placed *Capsicum* as a sister group to the Lycianthes (blue potato bush) (Figure 1.1). Together, they form a sister clade to the *Physalis* genus, with the split from the *Solanum* clade occurring approximately 19 million years ago (Särkinen, Bohs, Olmstead, & Knapp, 2013) (Figure 1.1).

The *Capsicum* genus contains 20-27 species, five of which have been domesticated: *Capsicum annuum, Capsicum baccatum, Capsicum chinense, Capsicum fructescens* and *Capsicum pubescens* (Heiser & Pickersgill, 1969; Walsh & Hoot, 2012). Fruit morphologies are diverse in shape, size and colour (Figure 1.2) (Sarpras et al., 2016; Zhigila, Abdulrahaman, Kolawole, & Oladele, 2014).

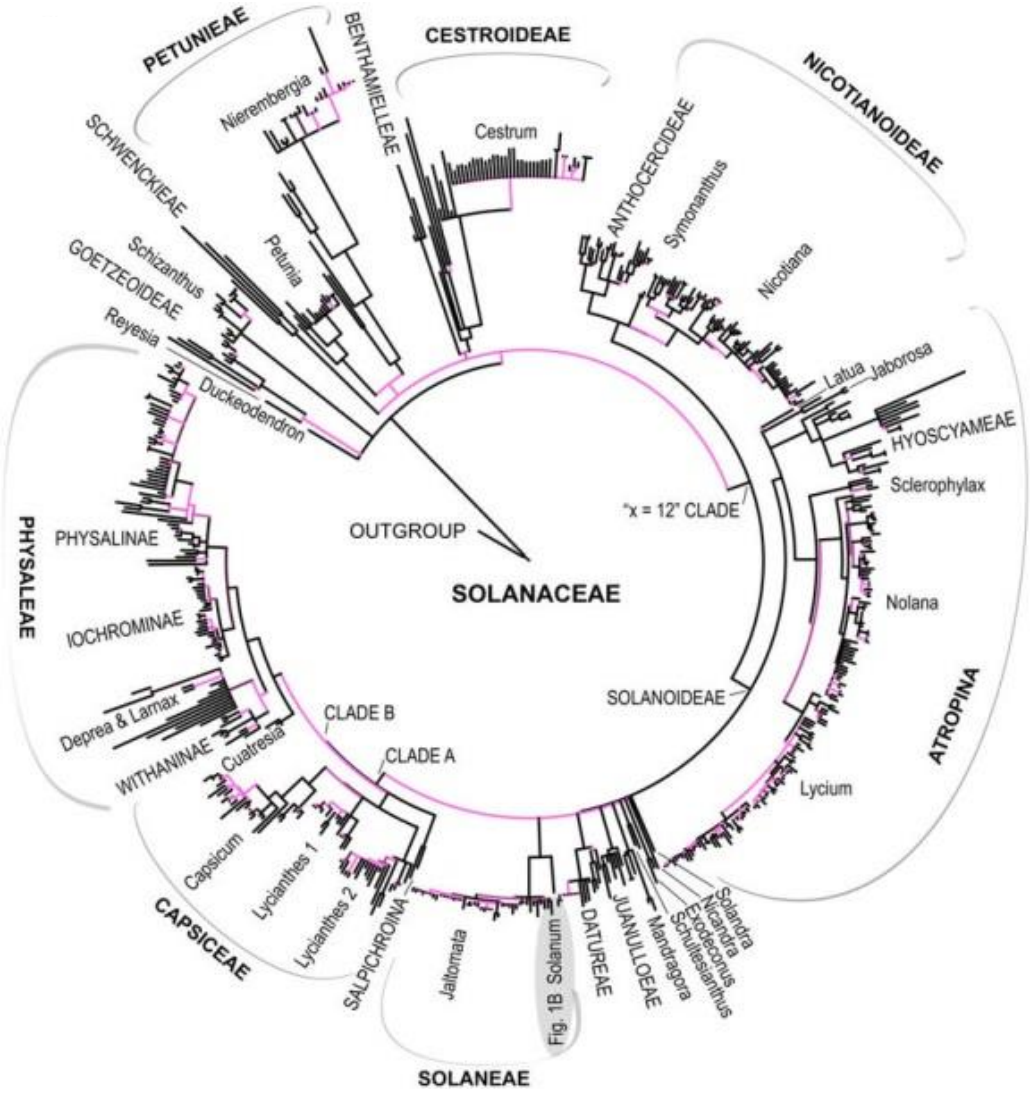


Figure 1.1. Phylogeny of the Solanaceae family (from Sarkinen et al. 2013)

Figure 1.2. Fruit morphologies within the *Capsicum* genus (Van der Knapp Laboratory, 2012)

1.1.2 Origins

The *Capsicum* genus originated in South America, in modern day Bolivia and southern Brazil (Nunn & Qian, 2010). Domesticated *Capsicum* peppers have their origins in Mexico (Pickersgill, 1971). Evidence from several lines of investigation supports this hypothesis. A 2014 study combined species distribution modelling and paleobiolinguistics with genetic microsatellite and archaeobotanical data. Genetic data supports an origin of domestication for *Capsicum annuum* in north eastern Mexico (Figure 1.3; Kraft et al., 2014); however, combining evidence from all four sources of information supports a central eastern location (Figure 1.4) (Kraft et al., 2014).

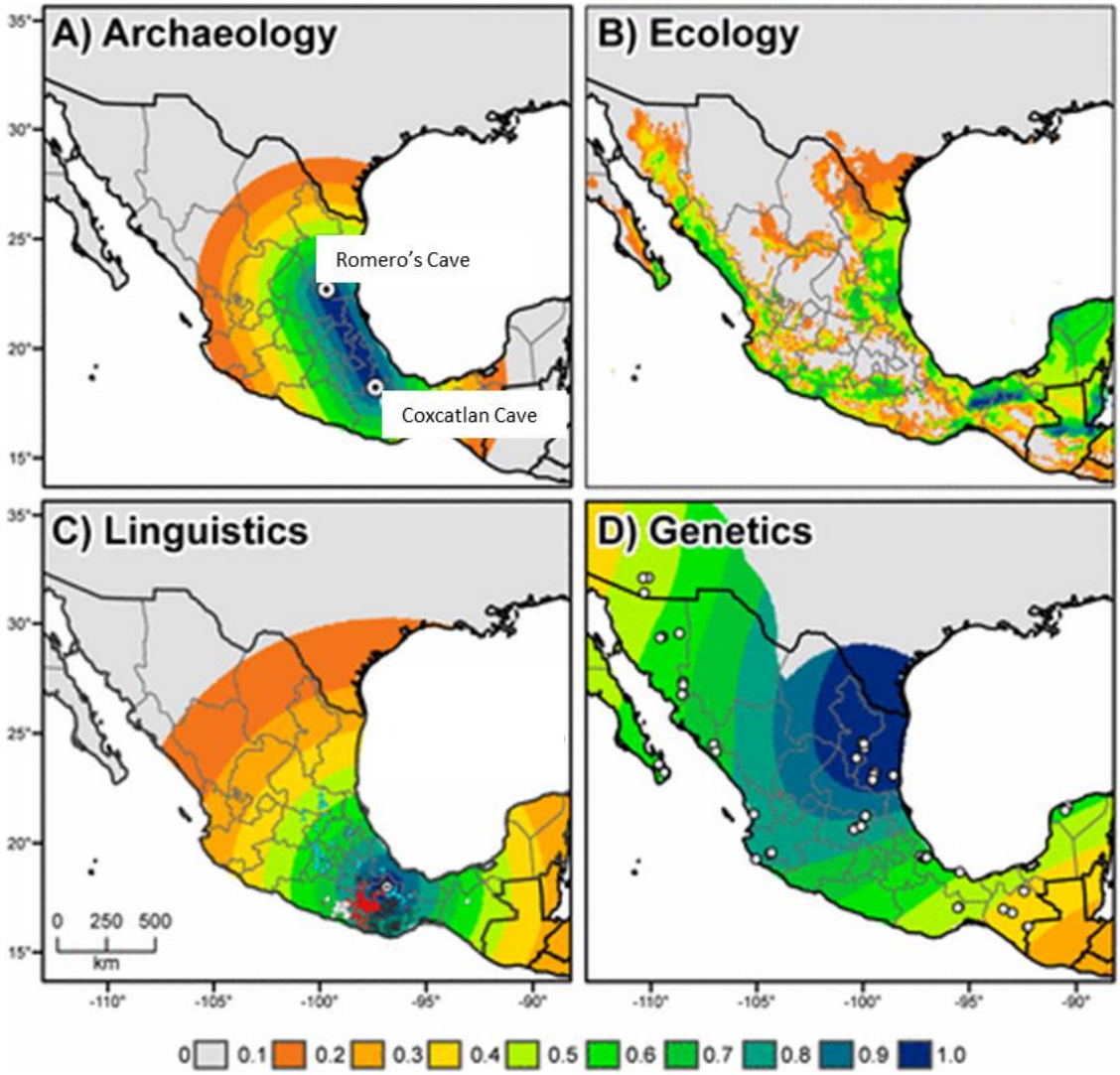


Figure 1.3 Separate origins of *Capsicum annuum* domestication predicted by four lines of evidence: archaeobotany, ecology, paleo-linguistics, and microsatellite genetics (Kraft et al., 2014)

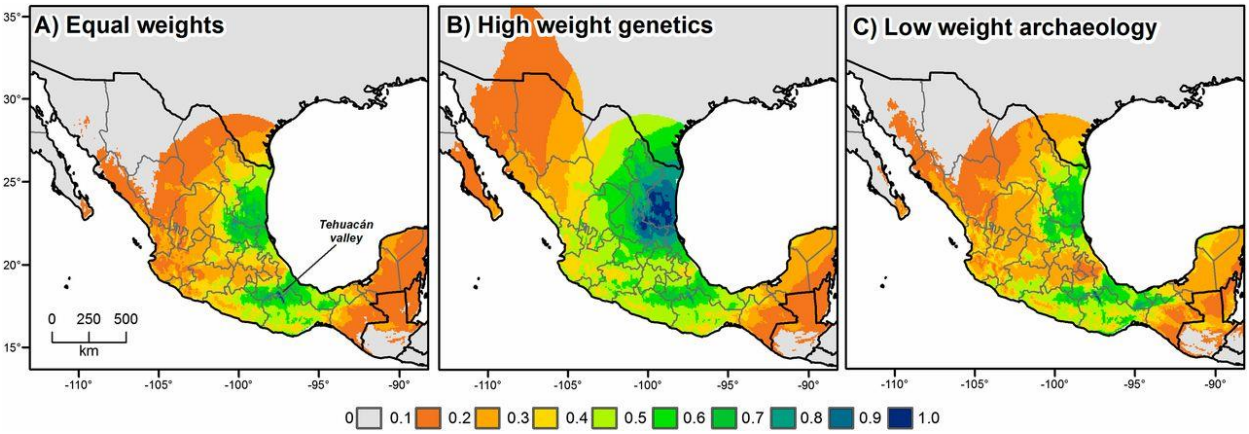


Figure 1.4. Models of the origin of *Capsicum* domestication, produced by combining separate lines of investigation. A) Equal weights given to all four types of evidence. B) Genetic evidence is given a greater weight than the other three sources of information (genetics 1/2, all others 1/6). C) Archaeology is given a lower weight then the other three lines of investigation (archaeology 1/10, all others 3/10).

The oldest remains unambiguously identified as *Capsicum* pepper were found in preceramic strata from dry caves at two sites in Mexico: Puebla (Tehuacán Valley) and Tamaulipas (Ocampo caves). Dating these with other crop remains (those of maize and squash) found at the same locations give the chilli remains a rough date of 6000-5000 years Before Present (B. P.); accelerator mass spectrometry (AMS) ^{14}C radiocarbon dating was carried out (Long, Benz, Donahue, & Jull, 1989; B. D. Smith, 1997). Its wild ancestor was probably the bird pepper (or Chiltepin) (Eshbough, 1993). This plant was markedly different to the domestic peppers we are familiar with today: its fruit were extremely small (measuring only 1cm in length), they grew upward, fell from the plant when ripe (deciduous), were exclusively red in colour, soft-fleshed, and pungent. Their pungency in particular selectively recruited birds as effective seed dispersal agents: the compound Capsaicin was shown to attract birds, while repelling less effective seed dispersers such as small mammals (Tewksbury & Nabhan, 2001).

Following the Columbian Exchange in the 15th and 16th centuries, *Capsicum* peppers were transported from the Americas to the Old World, along with other crops and technologies such as potatoes and tomatoes. On the 15th January 1493, Columbus discussed chilli peppers in South America his diary, stating: *"…everybody does not eat without it, which is very healthy"* (Smith, 2014). They then introduced to parts of Asia via European travellers, where they became integral components of everyday cuisine (Nunn & Qian, 2010). By 1493, *Capsicum* peppers had reached Spain and Africa; they had reached the East Indies by 1540, and India by 1542 (Andrews, 1993). For the last few hundred years, peppers have continued to be used all over the world on a daily basis, and have as a result become extremely economically valuable.

1.1.3 Pepper and Tomato

The most widespread model fleshy fruit is *Capsicum*'s close relative, the tomato (*Solanum lycopersicum*). During the course of their domestication, similar genes enhancing taste, texture, size and pericarp thickness were selected (Paran & Knaap, 2007). The high level of synteny between the two genomes means that the tomato can be used as a model for development of certain traits in pepper; the fact that genetic transformation techniques are much more advanced in tomato than in pepper means this is occasionally necessary (Livingstone, Lackney, Blauth, Wijk, & Jahn, 1999; Rinaldi et al., 2016).

1.2 Colour

1.2.1 Carotenoids

The major class of compounds conferring colour on ripe *Capsicum* fruit are the carotenoids, a class of C_{40} terpenoids consisting of eight five-carbon isoprene (*ip*) units arranged in a head-to-tail manner

(Britton, Liaaen-Jensen, & Pfander, 2004; Fraser & Bramley, 2004). Synthesised by all photosynthetic organisms, these secondary metabolites are derived from end-products of the plastid-specific Methylerythritol 4-Phosphate (MEP) pathway, whose precursors are pyruvate and 3-phospho-glycerate synthesised by the plant's primary carbon metabolism (Figure 1.5) (Rohmer et al., 1999). Characteristic colours are conferred on the molecules by the system of conjugated double bonds which comprises their chromophore. Since the elucidation of the first carotenoid structure (β-carotene) in the late 1920s (by Richard Kuhn and Paul Karrer separately), over 700 have been identified; at the steepest rate of discovery over 20 new structures a year were reported (Maoka, 2009). As they are not usually synthesised by non-photosynthetic organisms, animals must obtain them through the diet; therefore carotenoid containing crops are extremely important both nutritionally and economically. However, in rare instances, carotenoids are synthesised by non-photosynthetic organisms: an example of this is canthaxanthin synthesis in the bacterial species *Rhodococcus erythropolis* (Tao & Cheng, 2004).The market value of carotenoids was $1.5 billion in 2014, and is expected to reach $1.8 billion in 2019 (Marz, 2015).

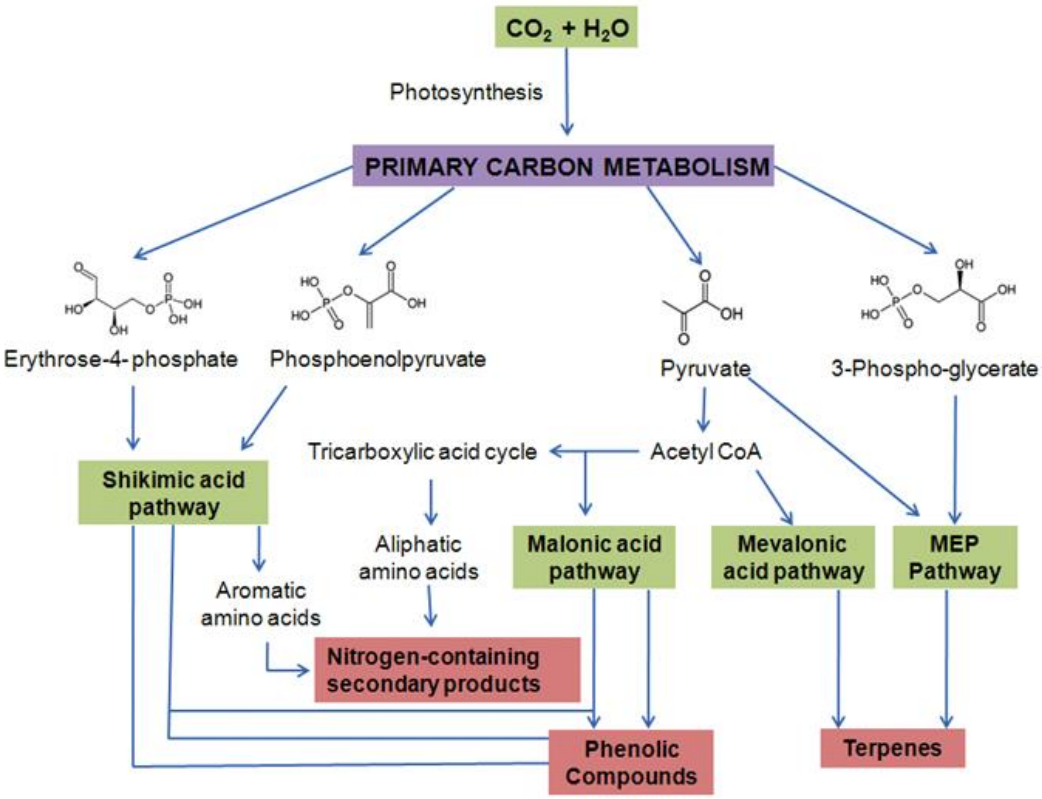


Figure 1.5. Broad overview of plant metabolism. Primary metabolism: purple box; secondary metabolism: red box. Adapted from Plant Biotechnology (www.plantbiotech.org)

1.2.1.1 Structure and function

Structures of the carotenoids are extremely diverse. Though they all share a common C_{40} backbone, the end-groups can vary significantly. In the C_{40} backbone, the eight isoprene units are joined

together in such a way that the double bond order is reversed at the centre of the molecule, giving the methyl groups at 20 and 20' a 1,6 relationship, whereas other methyl groups have a 1,5 positional relationship (Figure 1.6) (Fraser & Bramley, 2004).

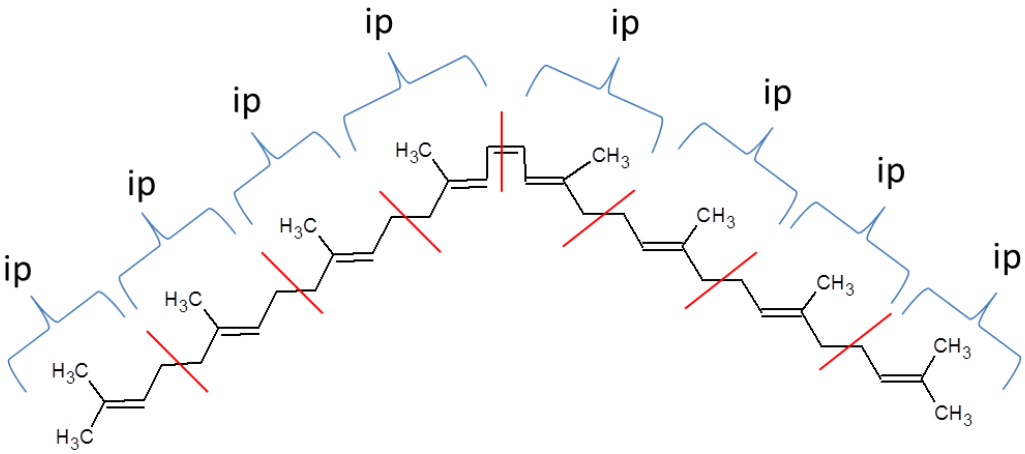


Figure 1.6. Isoprene units making up a C40 carotene backbone, with the double bond order reversed at the centre of the molecule.

There are seven possible end groups for a carotenoid: β, γ, ϕ, χ, ψ, ε, and κ. Only four are found in higher plant carotenoids: β, ψ, ε, and κ (Figure 1.7) (Britton et al., 2004).

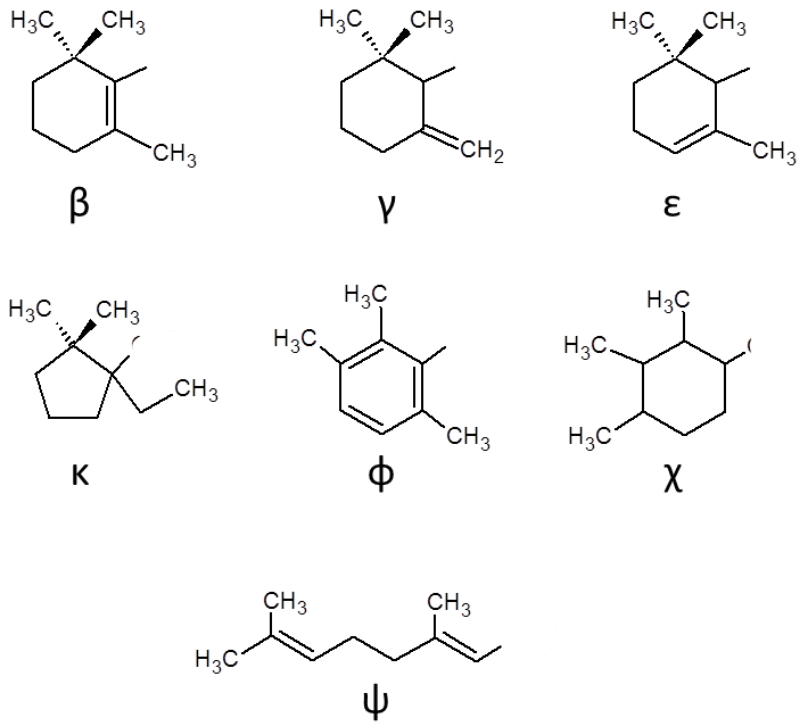


Figure 1.7. The seven possible carotenoid end groups. (Britton et al., 2004)

End rings can be modified by a range of functional groups, resulting in hydroxycarotenoids, ketocarotenoids, epoxycarotenoids, aldehydes and carboxylic acids. Based on their structural

components carotenoids are divided into two groups: carotenes, made up of only carbon and hydrogen, and xanthophylls, which contain at least one oxygen atom. The length of the hydrocarbon backbone can also be increased or decreased, giving norcarotenoids (in which one or more carbons have been eliminated from the skeleton) or C_{45} or C_{50} carotenoids which have had one or more additional C_5 units added to their original skeleton (Britton et al., 2004).

The number of conjugated double bonds in the carotenoid backbone means that a number of *cis/trans* isomers are possible for each molecule; however, the all-*trans* form is predominant in nature. The molecules are lipophilic, accumulating in hydrophobic compartments such as membranes, lipid droplets and lipoproteins. This effects their absorption, excretion, and transport in the organism (Stahl & Sies, 2003).

Carotenoids in photosynthetic organisms have several important functions: acting as photosynthetic accessory pigments, and are important protectants against oxidative damage. The molecules interact with singlet molecular oxygen ($^1O_2$) and peroxyl radicals, and function as deactivators of electronically excited sensitizer molecules which are active in the production of radicals and singlet oxygen (Stahl & Sies, 2003). The interaction of carotenoids with $^1O_2$ is dependent on physical quenching: direct energy transfer between the molecules. The energy of the singlet molecular oxygen is transferred to the carotenoid, returning $^1O_2$ to ground state and yielding a triplet excited carotene which dissipates its energy to the surrounding solvent rather than to further chemical reactions (Fraser & Bramley, 2004). As the carotenoid remains intact throughout the process it can be used throughout several quenching cycles (Stahl & Sies, 2003).

The solar radiation profile at the surface of the earth is most intense at 450-550 nm (Stange, 2016). Due to the fact that carotenoids absorb light in the 400-500 nm region of the electro-magnetic spectrum they are capable of harnessing this excitation energy which can then be transferred to chlorophylls, significantly broadening the spectral range over which light can support photosynthesis. However, their action as photosynthetic accessory pigments is not what makes carotenoids essential for photosynthesis: it is their role as photo-protectants, guarding photosynthetic architecture from oxidative damage.

Photosynthetic architecture can be divided into Light Harvesting Complexes (LHCs) and Reaction Centres (RCs). There are two types of RC in photosynthetic organisms: type I RCs which have iron-sulphur complexes as their terminal electron acceptors and type II, which have quinones as their terminal electron acceptors. Oxygenic photosynthetic organisms possess both types: giving their names to photosystem I (PS I) and photosystem II (PS II) (Stange, 2016). Through reaction centres are broadly similar in structure, LHCs are incredibly varied. Figure 1.8 outlines the major LHC structures

found in photosynthetic eukaryotes. Most LHCs are composed of pigments and proteins arranged as complex antenna, however this is not the case for chlorosomes which contain only pigment. LHCs can also be classified into two major categories: core and peripheral. Usually all RCs possess associated core LHCs, with the number and arrangement of peripheral LHCs being determined by light intensity and quality. Carotenoids are found in all known RCs and LHCs (Figure 1.8).

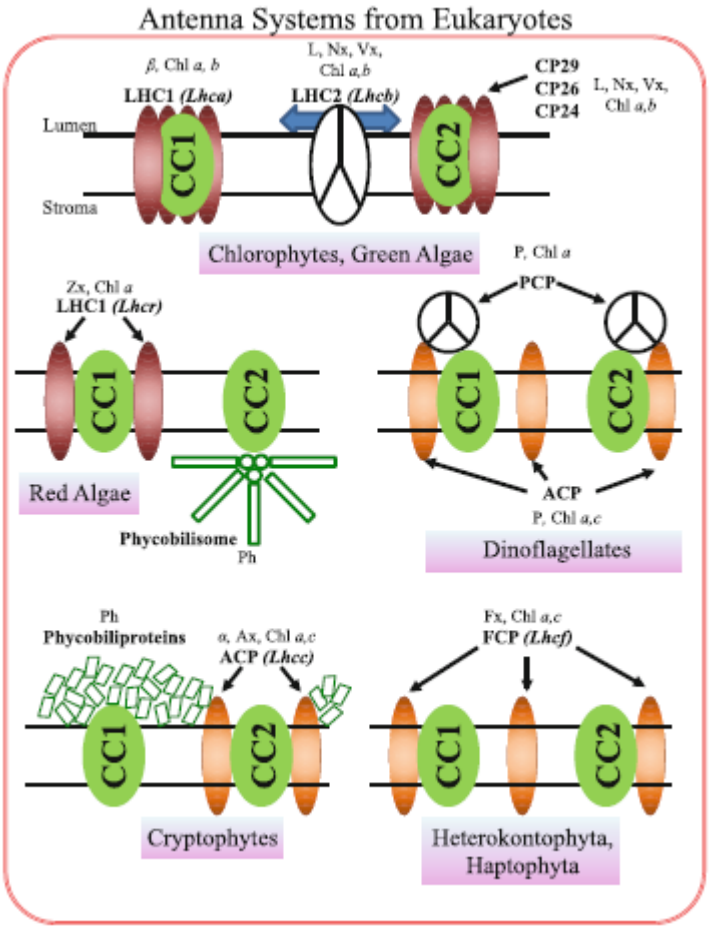


Figure 1.8. Schematic illustrations of antenna systems from prokaryotes and eukaryotes. The ellipses display the protein of LHC superfamily, and their genes are written in parentheses. The complexes that function with trimer unit are described with parting line in the circle. The complexes that proceed the state transition due to the light condition is described with two- way arrows. Abbreviations stand for as follows. β : β-Carotene; α: α-Carotene; Ax: Alloxanthin; Bchl: Bacteriochlorophyll; Chb; Chlorobactene; Chl: Chlorophyll; Diapon: Diaponeurosporene; Fx: Fucoxanthin; Iso: Isorenieratene; L: Lutein; Nx: Neoxanthin; P: Peridinin; Ph: Phycobiline; Sph: Spheroidene; Spx: Spirilloxanthin; Vx: Violaxanthin; Zx: Zeaxanthin; CC1,CC2: Core Complex 1,2; RC: Reaction Centre; PSI, II: Photosystem I, II; PCP: Peridinin-Chlorophyll a binding Protein; ACP: Chlorophyll a/c Protein; FCP: Fucoxanthin- Chlorophyll a/c binding Protein; LH1, LH2, LH3, LH4: Light Harvesting Antenna 1,2,3,4 (Stange, 2016)

Different compounds being associated with different photosystems. Violaxanthin and zeaxanthin are preferentially associated with PS II (Figure 1.9).

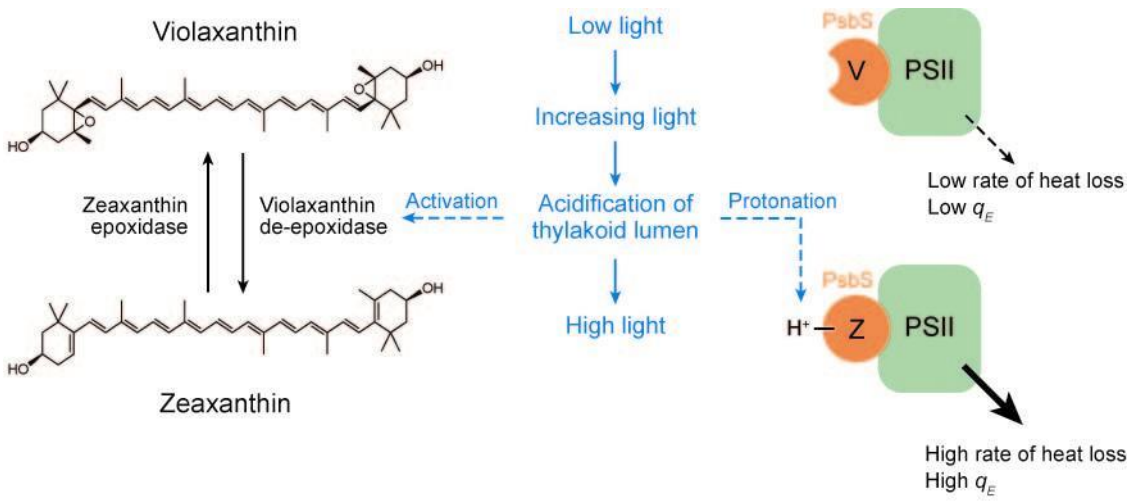


Figure 1.9. Mechanism of light-induced energy-dependent quenching of excitation energy in photosystem II (PSII). At low light that is limiting for photosynthesis a xanthophyll pigment, violaxanthin (V), is associated with the PSII antenna and PSII has a low rate of heat loss and consequently a low level of energy-dependent quenching, qE, which is an important component of nonphotochemical quenching (NPQ). At higher light intensities increased electron transport results in acidification of the thylakoid lumen. When the lumen pH drops below ca. 6 violaxanthin de-epoxidase is activated and converts violaxanthin to zeaxanthin (Z) and PsbS becomes protonated. The zeaxanthin associated with PSII is an efficient quencher of excitation energy in the PSII antenna and the rate of heat loss from PSII increases, which increases qE.When light intensity decreases deprotonation of PsbS occurs and zeaxanthin epoxidase converts zeaxanthin back to violaxanthin, which decreases qE. (Baker, 2008)

The xanthophyll cycle is regulated by the photosynthetic state of the plant: in intense light conditions, violaxanthin is converted to zeaxanthin, which is more efficient at dissipating excess light energy. Low light levels have the opposite effect, converting zeaxanthin back to violaxanthin (Hieber, Bugos, & Yamamoto, 2000; Yamamoto, 2006). β-carotene also has an important role to play in protecting PS II against damage: when P680 gets over-oxidised it can be dangerous, $P680^+$ can destructively oxidise amino acid side chains. However, β-carotene in the PS II RC can donate an electron to $P680^+$, thus preventing the reactions which would otherwise damage PS II (Horton & Ruban, 2005). Evidence for this is the fact that the β-carotene molecules in the RC only become oxidised if electron donation from water is perturbed; it is thought that a β-carotene molecule can mediate cyclic electron transfer via cytochrome *b*559 (Figure 1.10) (Ferreira, Iverson, Maghlaoui, Barber, & Iwata, 2004; Nelson & Ben-shem, 2004; Telfer, 2002; Zouni et al., 2001).

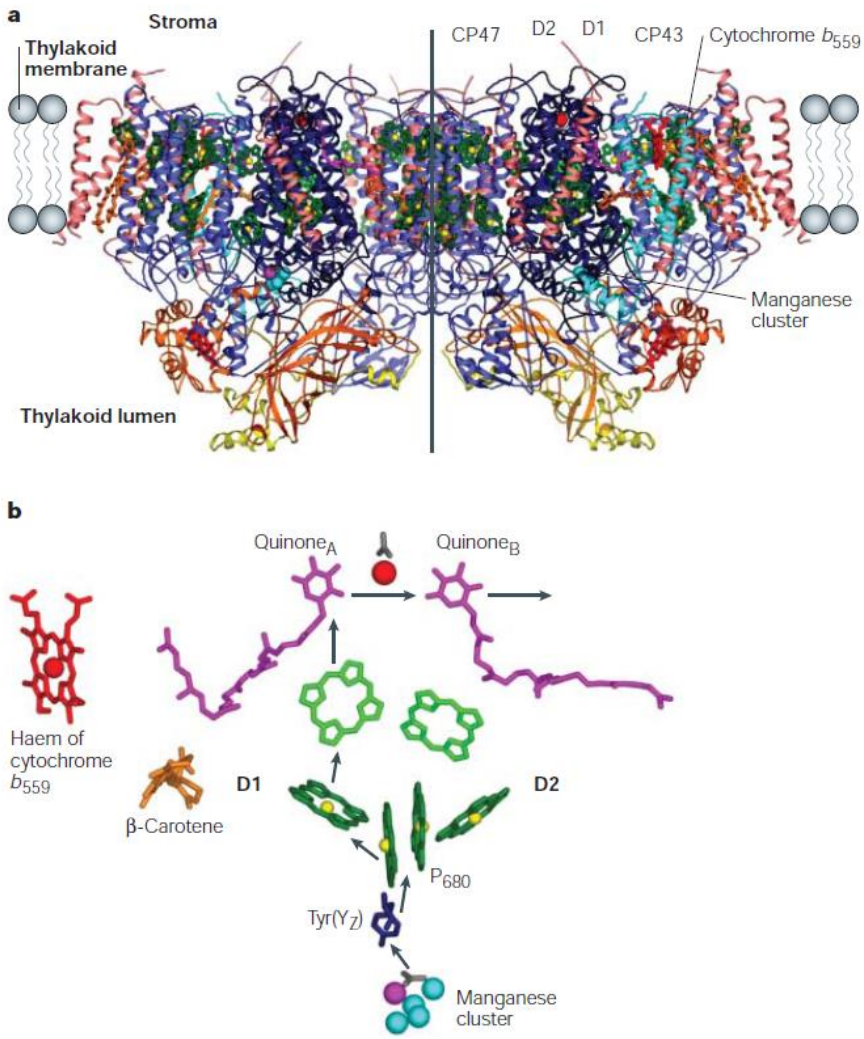


Figure 1.10. The structure of photosystem II and the cofactors that are involved in light- induced water oxidation and plastoquinone reduction. A. A view of the photosystem II (PSII) dimer perpendicular to the membrane normal. The D1 and D2 subunits are shown as dark-blue ribbon structures, CP43 and CP47 as purple ribbon structures, the cytochrome-b559 subunits as light-blue ribbon structures, and extrinsic subunits as both orange and yellow ribbon structures. The remaining subunits are shown as light-pink ribbon structures. Chlorophylls are shown in a dark-green stick representation (with the central magnesium ions shown as yellow spheres). The oxygen-evolving complexes/manganese clusters are shown as light-blue spheres (manganese ions) and dark-pink spheres (calcium ions). Haem groups are shown in a red stick representation, and red spheres represent iron ions. Orange stick-like structures represent β-carotene, and dark-pink stick-like structures represent the quinones. The axis of symmetry of the dimer is highlighted by a line. B. The cofactors that are involved in electron transport in PSII are shown in the same positions as in the left part of the PSII dimer in FIG. 2a. The arrows indicate the electron-transport pathway to the photooxidized P680 (two weakly coupled chlorophylls that function as the primary electron donor) and to the oxidized quinone. The colour scheme is the same as for part a, with the addition of pheophytins in light green, tyrosine 161 of D1 — Tyr (Yz) — in dark blue and bicarbonate ions in grey. The chlorophyll molecules are represented by symmetrical porphyrins. (Ferreira et al., 2004; Nelson & Ben-shem, 2004)

1.2.1.2 Health and nutrition

1.2.1.2.1 Provitamin A activity

Carotenes and xanthophylls play important roles in human nutrition: one of the best characterised functions of dietary carotenoids is their provitamin A activity. Retinal (provitamin A) is essential for proper function of the eye's retina; combing with the protein opsin to form the light-absorbing

molecule rhodopsin, which is necessary for both low-light and colour vision. Retinal can then be converted to retinol (the transport and storage form of vitamin A) and retinoic acid (which is also an important hormonal signalling molecule in vertebrates). Every year, millions of cases of irreversible blindness (mainly in Less Economically Developed Countries (LEDCs)) are caused by vitamin A deficiency, which could be alleviated by increasing dietary provitamin A content. Compounds which possess a β-ring end group (such as β-carotene, β-cryptoxanthin, and zeaxanthin) can be cleaved. One half of a β-ring containing carotenoid is converted into retinal through the action of the mono-oxygenase β-carotene oxygenase (BCO) isolated from the human intestine and liver (Lindqvist & Andersson, 2002).

In order to prevent blindness caused by a deficiency in provitamin A, enhancing crop carotenoid levels has become a major focus of research worldwide: the Golden Rice project introduced two ectopic enzymes (phytoene synthase (*Psy*)) from maize, and a phytoene desaturase (*CRTI*) from the bacterium *Erwinia uredovora*) to rice plants (*Oriza sativa*) which enabled them to successfully synthesise high levels of β-carotene, and store it in the grain (Beyer et al., 2002). The first attempts utilised phytoene synthase from daffodil; it took many years of experimenting with phytoene synthases from diverse plant species to conclude that the best results were from rice transformed with the maize enzyme, this is used in Golden Rice 2.0.

1.2.1.2.2 Roles in eye health

Age-related Macular Degeneration (AMD) and cateract are common problems across the world, contributing to large-scale decreases in vision-related quality of life. Specific carotenoids act to counter the harmful effects of these illnesses, probably by quenching active oxygen species. In cell cultures of human lens epithelial cells (HLECs) lycopene acts to significantly decrease vacuolisation. Lutein and zeaxanthin (two of the most high-value carotenoids) are important protectants against macular degeneration. They accumulate in the macula; protecting it against light-induced damage and scavenging free radicals. The NIH Eye Disease Case Control Study found a statistically significant positive association between the amount of antioxidant consumed and reduced risk of macular degeneration, with lutein and zeaxanthin having the strongest effect (Seddon et al., 1994). Other large-scale randomized control trials found positive associations between the amount of lutein and zeaxanthin consumed and improvements in vision-related quality of life: for example, the Collaborative Optical Macular Pigment Assessment Study (COMPASS), and the 2013 study which tested the visual performance of drivers in Shanghai (Nolan et al., 2011; Yao et al., 2013). The impact of lutein and zeaxanthin on eye health has been recently reviewed by Bernstein et al. (2017).

1.2.1.2.3 Roles in disease prevention

Carotenoids have been associated with alleviating cardiovascular disease (CVD): potential reasons for this are reductions in low-density lipoprotein (LDL) oxidation and oxidative stress occurring with plaque formation. It has been found that consumption of pressed tomato products reduces the sensitivity of lipoproteins to oxidative damage, and that low levels of serum lycopene increase the risk of atherosclerosis in middle-aged men (Hadley, Clinton, & Schwartz, 2003; Rissanen, Voutilainen, Nyyssonen, & Salonen, 2002). However, not all epidemiological studies have found positive effects: in 1994, an intervention study with β-carotene did not reduce CVD mortality in Finnish male smokers (Alpha-Tocopherol Beta Carotene Cancer Prevention Study Group, 1994). The impact of individual carotenoids on CVD is not clear; however, a common conclusion is that the combination of dietary carotenoids and other antioxidants is effective, rather than supplements of single components.

1.2.1.3 Market value and technological potential

As discussed above, one of the most important functional characteristics of carotenoids is their provitamin A activity. β-end group containing compounds can be cleaved to produce retinol (vitamin A). According to America's National Institutes of Health (NIH), most people over the age of fourteen require approximately 700-900 mg Retinol Activity Equivalents (RAE) per day, although this can vary depending on the individual's reproductive state and other health requirements (National Institutes of Health, 2014). Although most people in economically developed countries obtain enough retinol precursor through their diets, it may be necessary to supplement intake with medicinal products. Supplements usually contain retinol in the form of retinyl acetate or retinyl palmitate; some deliver it as provitamin A (β-carotene). Additionally, most multivitamin mineral supplements contain vitamin A (National Institutes of Health, 2014). A study investigating supplement use in the U.S. between 2003-2006 estimated that 28-37% of the population took vitamin A-containing supplements, most of whom were older than 71 or younger than 9 (Bailey et al., 2011).

Synthetic retinoids derived from vitamin A are also used in various pharmaceutical products; examples include actiretin (Soriatane ®) for the treatment of the skin condition psoriasis, and bexarotene (Targretin ®), which is used to treat the skin effects of T-cell lymphoma.

The advent of biotechnology has led to the development of biotechnological options for synthesising these high value compounds.

1.2.1.4 Biosynthesis in Capsicum

Carotenoids are synthesised in the plastids of photosynthetic organisms, deriving their precursors from the 2-C-methyl-d-erythritol 4-phosphate (MEP) pathway. This pathway also provides precursors for the pathway leading to the important phytohormone auxin. Although certain carotenoids are

present in the unripe green fruit as photosynthetic accessory pigments (the most prominent examples being neoxanthin, violaxanthin, lutein and β-carotene) ripening triggers a shift in composition: labelling studies in chilli peppers have revealed that, in the ripe-specific chromoplast α-carotene (found on the lutein branch of the biosynthetic pathway) becomes increasingly poorly labelled, demonstrating that the β-carotene branch is preferred by ripening fruit tissue (Camara, Bardat, & Moneger, 1982). These experiments also provided useful information about the intra-plastid locations of certain steps of the pathway: it was demonstrated that phytoene is synthesised in stroma, while the subsequent desaturation, isomerisation, cyclisation, and hydroxylation steps occur in the membrane fraction. This is supported by the previous finding that phytoene synthase (PSY) is located in the stroma in a variety of plants including *Capsicum annuum, Pisum sativum, Triticum aestivum* and *Zea mays* (Dogbo et al., 1987), and that enzymes catalysing downstream steps in the pathway (phytoene desaturase (PDS), and lycopene β-cyclase (LCY-B)) were located in the membrane fraction of isolated pepper chromoplasts (Camara & Dogbo, 1986; Hugueney, Romer, Kuntz, & Camara, 1992).

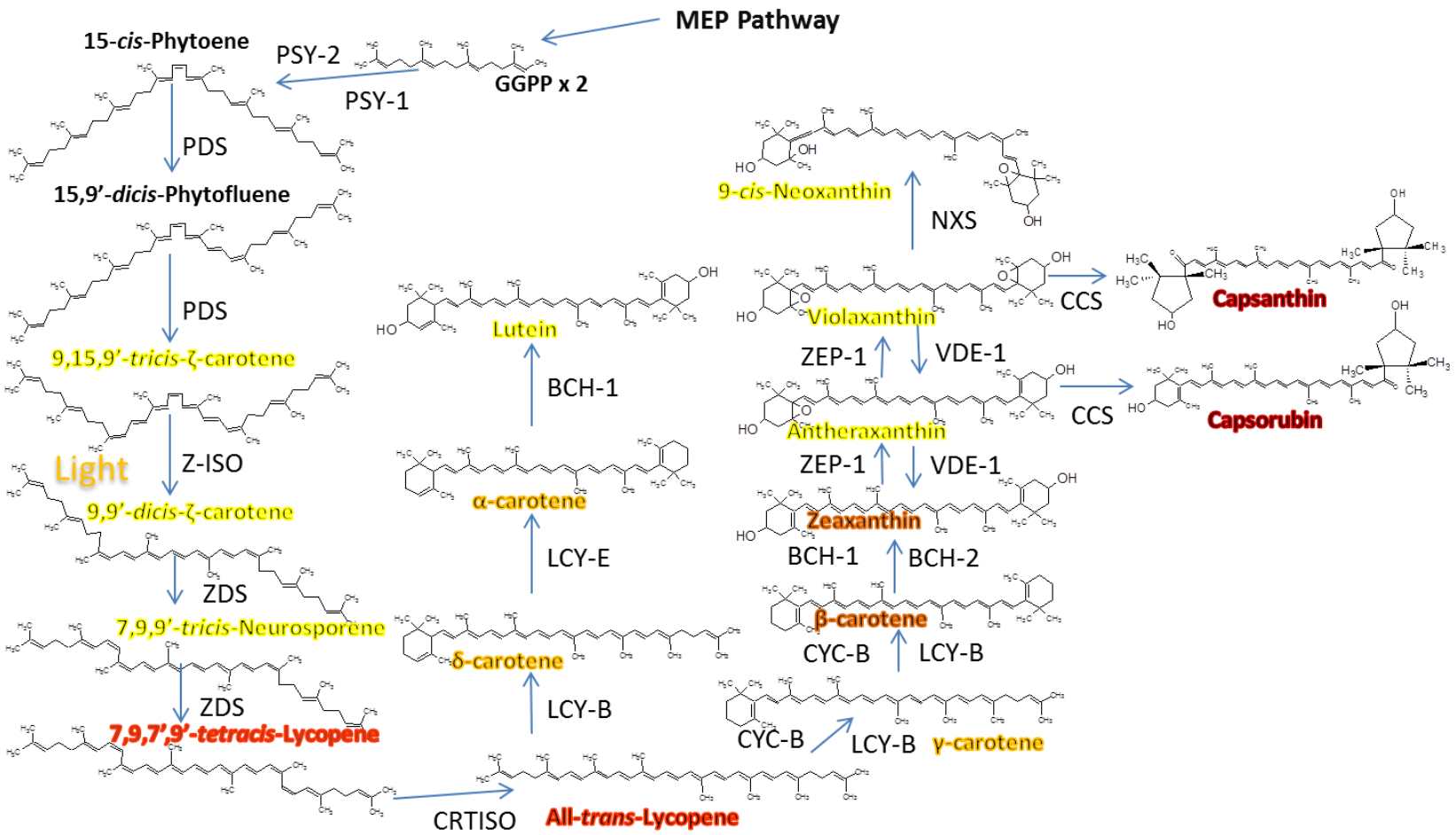


Figure 1.11. Carotenoid biosynthesis in wild-type pepper fruits. Abbreviations: MEP pathway: non-mevalonate pathway; GGPP: geranylgeranyl pyrophosphate; PSY-1/2: phytoene synthase 1/2; PDS: phytoene desaturase; Z-ISO: zeta-carotene isomerase; ZDS: zeta-carotene desaturase; CRTISO: carotene isomerase; LCY-B: Lycopene beta-cyclase (green tissue-specific); LCY-E: lycopene epsilon-cyclase; BCH-1/2: beta-carotene hydroxylase-1/2; CYC-B: lycopene beta-cyclase (fruit ripening specific); ZEP-1: zeaxanthin epoxidase-1; VDE-1: violaxanthin epoxidase-1; NXS: neoxanthin synthase; CCS: capsanthin capsorubin synthase. Compound names are highlighted in the colour they appear.

In relation to pepper lycopene cyclases, there has been much controversy about the level of homology between the enzyme capsanthin capsorubin synthase (CCS) and the tomato ripening-specific lycopene cyclase (CYC-B). CCS has a high level of identity with LCY-B and particularly CYC-B; furthermore, both CCS and CYC-B are located in similar regions of the genome, on chromosome 6 (Kim et al., 2014; Ronen, Carmel-goren, Zamir, & Hirschberg, 2000) In ripening tomato fruit, CYC-B is the chromoplast specific cyclase, however, (as mentioned above) LCY-B has been found in *Capsicum annuum* chromoplasts (Camara & Dogbo, 1986). Investigations into the expression patterns of *Capsicum LCY-B* revealed that the gene is expressed in leaves as well as throughout fruit development (Guzman, Hamby, Romero, Bosland, & O'Connell, 2010; Hugueney et al., 1995), which has led to the theory that the composition of cyclised carotenoids in red peppers is due to the action of the two proteins LCY-B and CCS.

The genetics of carotenoid biosynthesis in *Capsicum* species has been studied for decades. Although there is an established model of how three loci influence the inheritance of mature fruit colour in pepper (Hurtado-hemandez & Smith, 1985), the mechanisms underlying these phenotypes and the level of the gene, protein, and enzyme remain under investigation. According to the 1985 model, mature fruit colour is the result of three independently segregating loci: *y, c1* and *c2*. It appears that mutations in several different carotenoid biosynthetic genes can cause similar phenotypes. Conversely, mutations of different nature in the same gene causing changes at the level of the genome, transcript and protein, can lead to different fruit phenotypes. For example, several orange and yellow *C. annuum* lines possess phenotypes caused by mutations of different types in *Ccs*. It has been shown (through RFLP and specific PCR) that, in the F_2 progeny of a red x yellow cross, *Ccs* completely segregates with red fruit colour. In a study of 156 red and yellow accessions, each yellow line gave a negative result when tested with PCR primers specific to *Ccs*, leading the authors to the conclusion that the phenotype was caused by complete absence of *Ccs* transcript. However, one yellow line tested (CK7), gave the exact same banding pattern as the *Ccs*-positive red pepper lines. Sequencing of this PCR product revealed two types of mutation giving rise to early translation termination. A premature stop codon is caused by a 1 base pair deletion, causing a shift in the reading frame. Upstream of this mutation, a point mutation (C to G) results in a premature stop codon. Due to this mutation, a truncated protein (consisting of 334 amino acids) is synthesised, which lacks the 134 amino acid C-terminus present in the red cultivars. Comparison of the wild-type CCS protein with lycopene cyclases, revealed that the most conserved region was at the C-terminus; the CK7 ccs variants lacked a conserved domain at amino acids 430-460. Evidently, this caused the protein to be non-functional (Z. Li, Wang, Gui, Chang, & Gong, 2013). Another study also revealed that one of these mutations was present in the CCS variant of the orange coloured cultivars Fogo

(Guzman et al., 2010). In addition, yellow cultivars have been found to contain tandem repeats in their promoter regions, which prevent transcription of the gene (Li et al., 2013).

Yellow and orange cultivar phenotypes have also been related with the locus *c2*, which is associated with the gene coding for *Psy-1*; located on chromosome 4 (Thorup et al., 2000). The gene was first cloned from pepper in 1993 (Romer, Hugueney, Bouvier, Camara, & Kuntz, 1993). This study identified two phytoene synthases in pepper, which, unlike tomato, had transcripts present in both the leaf and the fruit. However, *Psy-1* appeared to be most strongly associated with fruit colour. Mutations in *Psy-1* result in orange or yellow peppers, depending on the state of the other two loci. Where *Ccs* is present and *Psy-1* is absent, an orange phenotype results; where both *Ccs* and *Psy-1* are absent, cultivars possess a yellow phenotype (Hurtado-hemandez & Smith, 1985). The yellow colour of these peppers is due to a dominant copy of the gene associated with the *c1* locus; the nature of which remains undetermined (Andrade-Souza, Costa, Chen, Gmitter, & Costa, 2011). Total carotenoid amount, and subsequently depth of fruit colour are both positively correlated with *Psy-1* expression levels.

There have been suggestions that the 1985 three loci model needs revising, and that genes which do not code for enzymes catalysing steps of the carotenoid biosynthesis pathway are involved in the determination of fruit colour. Rodriguez-Uribe et al. found that orange cultivars with wild type *Ccs* alleles failed to accumulate either *Ccs* transcript or protein. *Ccs* coding region and promoter analysis showed no differences between the cultivars Orange Grande and Oriole, which both accumulate no *Ccs* transcript, and those which do accumulate transcripts. The sequences from these lines were also identical to published sequences for red pepper *Ccs* coding and promoter regions (Ha, Kim, Park, Lee, & Cho, 2007). Another variety, Fogo, also with a wild type *Ccs* coding region, accumulated transcripts, but did not accumulate CCS protein (Rodriguez-Uribe, Guzman, Rajapakse, Richins, & O'Connell, 2012). These lines of evidence suggest that mutations in other genes, affecting multiple levels of regulation, should be incorporated into our thinking in relation to pepper colour.

1.2.1.5 Regulation of biosynthesis in Capsicum

Regulation of carotenoid biosynthesis during fruit development is complex due to the dramatic changes in amount and composition which occur with ripening. Many different methods of regulation act concurrently; examples are transcriptional control and post-transcriptional modifications of structural enzymes, and modifications to carotenoid sequestration mechanisms is also an important factor. Additionally, the existence of fruit-specific isoforms for key steps in the pathway allow for autonomous regulation of flux. Tomato has been the major model from which most aspects of carotenoid biosynthesis and regulation have been identified. However, novel

enzymatic activities or specific regulatory mechanisms have been characterised due to research into diverse fruit species.

The fruit-specific *Psy* isoform (*Psy-1*) is generally accepted as the most important regulatory step in carotenoid biosynthesis; it is tightly controlled by both endogenous factors and environmental cues (Cazzonelli & Pogson, 2010). These include high light, salt, drought, temperature, photoperiod and post-transcriptional feedback regulation. Two or more homologues have been identified in several plants including rice, poplar, bread wheat and maize. Their expression is tissue specific and shows unique responses to environmental stimuli in cereal roots. *Psy-3* is the homologue specifically expressed in roots; transcript levels have been correlated with drought and high salt conditions, this leads to increased carotenoid flux and ABA in maize roots (Li, Vallabhaneni, & Wurtzel, 2008). The same study found that *Psy-2* and *Psy-3* transcripts rapidly disappeared after watering, which implies tightly controlled mRNA stability or transcription (Li, Vallabhaneni, & Wurtzel, 2008). Another mechanism through which PSY enzymatic activity is controlled is allelic variation: alternative splicing of the *PSY-A1* allele is a major QTL determinant of fruit colour in bread wheat (Howitt et al., 2009). The alternative splicing results in the generation of four different transcripts only one of which is functional, thus dividing by four the abundance of functional *Psy* transcripts (Howitt et al., 2009).

There is considerable evidence to support metabolite feedback regulation, which modulates the supply of isoprenoids and the accumulation of carotenoids and ABA. Heightened expression of a *Psy* transgene led to increased carotenoid levels in Arabidopsis seedlings via post-transcriptional accumulation of *DXS* mRNA; this feedback mechanism initiated by *Psy* stimulates the supply of MEP intermediates (Rodriguez-Villalon, Gas, & Rodrıuez-Concepcion, 2009a; Rodrıguez-Villalon, Gas, & Rodrıuez-Concepcion, 2009b). The mechanism of increasing the action of DXS through PSY activity is no restricted to etiolated seedlings: it has been reported that elevated *Psy-1* expression levels increase the activity (albeit not the transcript level) of *Dxs* (Fraser et al., 2007). However, it was reported that overexpression of *Dxs* and not *Psy-1* did not increase carotenoid content, demonstrating that regulation of carotenogenesis through control of *Psy-1* expression is highly co-ordinated, and influenced by both source and sink metabolites (Rodrıguez-Villalon et al., 2009a).

Different transcription factors are related to the pathway, impacting different metabolic steps. The tomato gene *MADS1/RIN* was proposed as a negative regulator of fruit ripening, due to the fact that *Psy-1* expression and carotenoid content were enhanced in RNAi transgenic lines, this also occurred with genes related to ethylene biosynthesis and response (Dong et al., 2013; Liu, Shao, Zhang, & Wang, 2015). It has been established that massive increase in *Psy-1* expression at the onset of ripening determines carotenoid content, and therefore fruit colour. Downstream in the pathway,

ZISO catalyses the isomerisation of tri-cis ζ-carotene to di-cis ζ-carotene (Chen, Li, & Wurtzel, 2010; F. Li, Murillo, & Wurtzel, 2007); it has been observed that this gene is also highly induced during ripening in tomato. However, in fruit with lower carotenoid content, such as apple, a reduction in *ZISO* expression occurred during ripening; limiting flux through the pathway and preventing carotenoid accumulation (Ampomah-Dwamena et al., 2012).

The major carotenoids accumulated by yellow and orange peppers are lutein and other xanthophylls. Carotenoid biosynthesis bifurcates at the step following lycopene synthesis, producing epsilon- and beta-carotenoids via the action of two lycopene cyclase enzymes Lycopene ε-Cyclase (εLCY) and Lycopene β-Cyclase (β-LCY) (Figure 1.11). It has been well documented that lutein levels can be determined by *lycopene ε-cyclase* (*εLCY*) expression (Cuttriss, Chubb, Alawady, Grimm, & Pogson, 2007; B. J. Pogson & Rissler, 2000; B. Pogson, Mcdonald, Truong, Britton, & Dellapenna, 1996). However, in certain systems (*Brassica napus* for example) feedback from increased levels of β-carotene, zeaxanthin, and violaxanthin can also increase the level of lutein, even though they are β-carotenoids, resulting from the other lycopene cyclase branch (*β-LCY*) (Yu, Lydiate, Young, Schafer, & Hannoufa, 2008). Regarding composition of the β-xanthophylls, light stress leads to the synthesis of zeaxanthin from β-carotene (Depka, Jahns, & Trebst, 1998). *β-carotene hydroxylase* (*BCH*) and *violaxanthin de-epoxidase* (*VDE*) are respectively both highly light inducible and repressible (Rossel, Wilson, & Pogson, 2002). High light conditions also lead to post-translational modification of VDE activity: luminal pH and ascorbate content are crucial regulators of zeaxanthin levels under light stress (Rockholm & Yamamoto, 1996).

The accumulation of carotenoids in the chromoplast can be thought of as the net result of biosynthesis, degradation, and sequestration. Degradation inversely determines carotenoid content: this can be the result of random cleavage caused by photooxidation or peroxidase/lipoxygenase co-ordination, or the result of the action of the carotenoid cleavage dioxygenase family of enzymes (CCDs). These catabolise carotenoids into apocarotenoids, helping to maintain the carotenoid homeostasis of the plant (Auldridge et al., 2006; Floss & Walter, 2009; Vallabhaneni, Bradbury, & Wurtzel, 2010; Walter & Strack, 2011). Some CCDs are also responsible for controlling the synthesis of the phytohormones ABA and strigolactone (Gomez-Roldan et al., 2008; Schwartz, Tan, Gage, Zeevaart, & Mccarty, 1997; Umehara et al., 2008). It has been demonstrated that expression of *CCD1* and *CCD4* negatively correlates with carotenoid accumulation in diverse plant species and tissues, including chrysanthemum flowers, potato tubers, and strawberry fruits (Campbell et al., 2010; García-Limones et al., 2008; Ohmiya, Kishimoto, Aida, Yoshioka, & Sumitomo, 2006; Tanaka & Ohmiya, 2008; Zhou et al., 2011).

1.3 Plastids

Plastids are double membrane bound organelles in the cytoplasm of plant cells. Different classes of plastid take on different roles within the cell, but most are responsible for photosynthesis or specialised storage. Types of plastid include chloroplasts (dedicated to photosynthesis), chromoplasts (specialised for carotenoid storage), amyloplasts (for starch storage), etioplasts, and elaioplasts (for sterol storage). These are all derived from unspecialised proplastids (Figure 1.12). The most well-studied plastid type is the photosynthetic chloroplast, though the focus of the current work is the chromoplast as this is the site of carotenoid synthesis and storage.

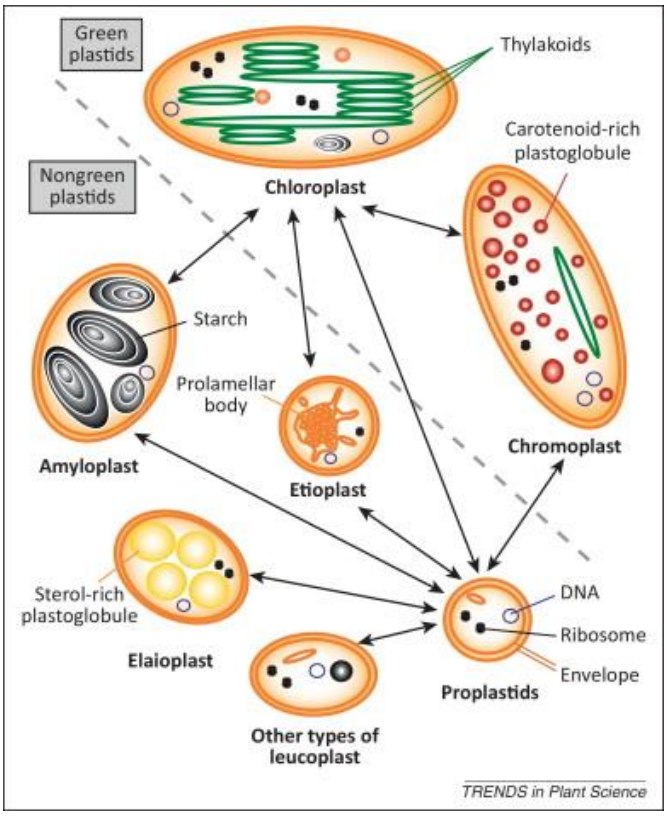


Figure 1.12. Plastid interconversion in angiosperms. In the zygote and meristematic cells, only embryonic plastids or proplastids are present. Proplastids can also occur in dedifferentiated cells and contain a circular DNA that is transmitted to all differentiated plastids, depending on the cell fate in the developmental program (such as the differentiation of chloroplasts in green tissues, chromoplasts in some organs such as tomato fruits, amyloplasts in some storage organs such as potato tubers, or elaioplasts in some floral parts, etc.) or on changes occurring in response to environmental changes (such as etioplasts in dark conditions). All green and nongreen plastids contain the relevant DNA and gene expression machinery (including ribosomes, tRNA, transcription system, etc.) and, thus, are considered to be semiautonomous organelles, with the ability of interconversion (reversible arrows). The differentiated function of the plastids can be dependent on the presence of plastoglobules with specific compositions. The surrounding envelope is the unique membrane structure conserved in all plastids. (Botte & Marechal, 2014)

1.3.1 The chloroplast to chromoplast transition

Although chromoplasts can develop from non-photosynthetic leucoplasts, the vast majority arise from chloroplasts through the chloroplast to chromoplast transition. Among non-photosynthetic plastids, chromoplasts have received the most research attention due to the fact that they are

specialised to synthesise and store large amounts of high-value pigment. It has been demonstrated that chromoplast biogenesis has a crucial role to play in determining carotenoid content as it determines biosynthesis and storage capacity. Many factors governing chloroplast development have been well characterised, however, genetic cues triggering the chloroplast to chromoplast transition are largely unidentified. The *Orange* (*Or*) gene isolated from the cauliflower orange curd mutant represents the only known molecular switch triggering the transition from non-coloured plastids to chromoplasts (Giuliano & Diretto, 2007; Lu et al., 2006). *Or* encodes a plastid associated protein containing a DnaJ cysteine-rich zinc finger domain and is highly conserved among taxonomically diverse plants. The OR protein does not contain the usual J-domain which defines DnaJ-like molecular chaparones; rather, it is a novel protein whose cellular function is triggering chromoplast differentiation (Lu et al., 2006; Paolillo, Garvin, & Parthasarathy, 2004). This function has been confirmed by studies which showed that ectopic expression of *Or* in both white cauliflower and potato tubers induced chromoplast biogenesis (Lu et al., 2006). Electron microscopy has confirmed the chromoplast morphology of plastids in tissues expressing *Or* (Paolillo et al., 2004). The *Or* homologue in melon is associated with a major QTL controlling β-carotene content (Cuevas, Staub, Simon, & Zalapa, 2009).

There are several metabolic pathways which are important for chromoplast biogenesis, particularly carotenoid biosynthesis, but also lipid and carbohydrate metabolism. Chromoplast development is linked to membrane proliferation; remodelling the internal membrane system to produce carotenoid-lipoprotein sequestration sub-structures as well as plastoglobuli is one of the most prominent changes occurring during chromoplast differentiation. Envelope membrane budding or fusion builds these sub-structures, therefore developing the plastid's biosynthetic capacity (Egea et al., 2010). Massive membrane proliferation during chromoplast biogenesis has been observed: in tomato fruits, membranes are newly synthesised from vesicles derived from plastid inner envelope (Simkin, Bramley, Fraser, & Kuntz, 2007). The same result has been reported in daffodil flowers (Kleinig & Liedvogel, 1980). Carotenoid lipoprotein sheet structure formation is associated with pigment accumulation in chromoplasts (L. Li et al., 2012). The synthesis of these membranes requires active fatty acid biosynthesis and lipid metabolism: correspondingly, a large number of proteins involved in lipid metabolism, including enzymes synthesising fatty acids, sulpholipids, and glycolipids, have been identified in chromoplast proteomes from diverse crop species (Barsan et al., 2010; Yong-qiang Wang et al., 2013). Transcriptomics studies have revealed that expression of the plastid-encoded gene *accD* is necessary for the massive increase in membrane lipid synthesis which the transition requires (Kahlau & Bock, 2008). A radiolabelling study revealed that most [^{14}C]-pyruvate incorporated into tomato chromoplasts was converted into lipophilic compounds,

supporting the hypothesis that lipid synthesis is very efficient in this cellular compartment (Angaman et al., 2012).

Carbohydrate metabolism also contributes to the biogenesis of chromoplasts. Several studies have provided evidence that sugars are important for the stimulation of chromoplast development and carotenoid accumulation. For example, a 2001 study found that sucrose content was positively associated with the rate of 'colour break' (caused by the chloroplast to chromoplast transition) in Satsuma mandarin (*Citrus unshiu*); when the plants were defoliated, natural sucrose build up was blocked, and colour break was prevented. It was also found that sucrose supplementation advanced the rate of colour change, taken together these results support the hypothesis that sucrose accumulation is necessary for the initiation of the chloroplast to chromoplast transition (Iglesias, Tadeo, Legaz, Primo-millo, & Talon, 2001). An older study also reported the re-greening of citrus pericarp upon incubation in low-sucrose media (Huff, 1983). It has been shown that sucrose levels have a direct impact on carotenoid accumulation: Tèlef and colleagues tested carotenoid content in excised mature green tomato pericarp discs incubated with varying concentrations of sucrose and found that sucrose limitation delayed and decreased the accumulation of phytoene and lycopene. This is explained by the fact that *Psy-1* expression was modulated by sucrose availability (Telef et al., 2006).

To conclude, it is clear that chromoplast biogenesis is a dynamic process which depends on several developmental and environmental cues. However, genetic factors influencing the transition remain uncharacterised. As the organelle is the site of synthesis and storage of high-value compounds, its unique characteristics deserve greater scrutiny.

1.3.1.1 Carotenoid sequestration in chromoplasts

Once synthesised in the *Capsicum* chromoplast, carotenoids are sequestered by two major structures: the plastoglobuli, and the fibrils. During the chloroplast to chromoplast transition thylakoid structure breaks down, producing spherical lipid globules in which the lipophilic carotenoids can be stored (Nogueira, Mora, Enfissi, Bramley, & Fraser, 2013). The colourless carotenoids phytoene and phytofluene, as well as esterified carotenoids and tocopherols have been located in this compartment through sub-chromoplast fractionation studies of tomato (Nogueira et al., 2013); however, such studies in pepper have yet to be published. In pepper, carotenoids are often found to be associated with long fibrillous proteins: 95% of carotenoids in the chromoplasts of ripe bell pepper are bound to these structures. A 32 kDa protein was identified and isolated from pepper chromoplasts in the 1990s, and named 'fibrillin' (Deruère et al., 1994a). The authors propose an arrangement of the protein-carotenoid complex, in which carotenoids collect in the centre of the

fibrils and are coated with a layer of polar lipid, which is then surrounded by a further layer of fibrillin. Association with fibrillin in this manner could be a long term storage strategy for ripe-fruit carotenoids: after being synthesised in the thylakoid-derived plastoglobuli, they could be transported to the fibrils, where they remain.

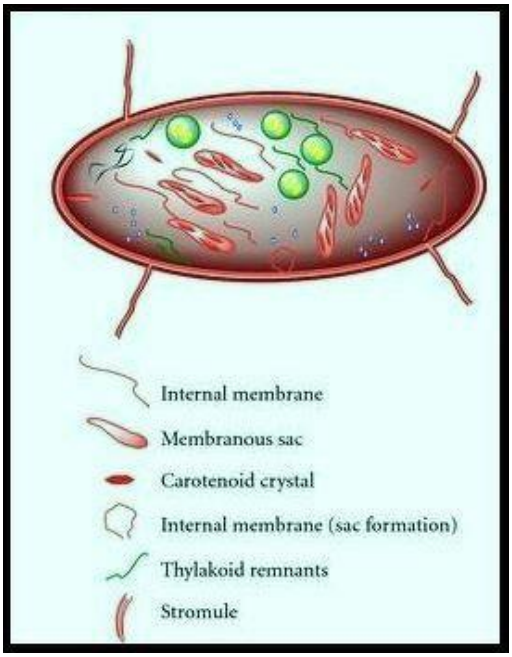


Figure 1.13. Chromoplast diagram, depicting internal membrane systems and carotenoid storage sites (Bian et al., 2011).

1.4 Lipids

1.4.1 Pepper fatty acid composition

Though fatty acids make up a small fraction of the edible portion of pepper, they are of significant structural importance: carotenoids are liposoluble and their esterification depends on the fruits' fatty acid content. Recent work has characterised the composition of fatty acids in *Capsicum* peppers: a 2015 study profiled pulps and seeds from six sweet and pungent *Capsicum* varieties and identified a total of 25 different fatty acids. Lipids in the samples were converted into fatty acid methyl esters (FAME), which were then methylated and separated with gas chromatography. The most abundant were found to be16:0 (palmitic acid), 18:1n-9 (oleic acid), and 18:2n-6 (linoleic acid). In all varieties tested, the ratios of polyunsaturated to saturated fatty acid were high due to the elevated amount of linoleic acid. It was revealed that the 25 fatty acids found in the samples included some with odd numbers of carbons in their chains. The omega-6 to omega-3 ratios were also determined; the varieties profiled ranged between 1.28 to 4.33. Combining the ratio of polyunsaturated to saturated fatty acids and the omega-6 to omega-3 ratio data suggested that the bell pepper had the most nutritionally valuable fatty acid composition (Sora et al., 2015).

In 2016 another study investigated fatty acid and tocopherol content in white, yellow, orange, red and purple sweet bell peppers with GC-MS and HPLC (Saini & Keum, 2016). Similarly to Sora et al., the authors found that the most common fatty acids in all varieties tested were C16:0 (palmitic acid) and C18:2 (linoleic acid). 17 distinct fatty acids were identified. Different colours differed in their ratios of accumulated fatty acids: the ratio of polyunsaturated fatty acid (PUFA) to monounsaturated fatty acid (MUFA) varied between 1.21 (white pepper) to 3.23 (red pepper).

1.4.2 Very Long Chain Fatty Acids (VLCFAs)

Very Long Chain Fatty Acids (VLCFAs) have chain lengths of C-18 or higher, and are essential for all eukaryotes. They are elongated by the elongase complex of the endoplasmic reticulum (ER), and can be incorporated into four lipid classes: triacylglycerols, waxes, phospholipids and complex sphingolipids. Their structural and functional diversity is a result of their chain length, polar head type and degree of unsaturation. Changes to the levels of VLCFAs in waxes, phospholipids and complex sphingolipids have profound effects on embryo, leaf, root and flower development. Alterations to VLFCA content in epiculicular waxes are critical for regulation of epidermal fusions during organogenesis (Bach & Faure, 2010). In phospholipids and sphingolipids, VLCFAs contribute to the structure and dynamics of membranes, and therefore to cell size and division. It has also been reported that they influence cell differentiation.

Their biosynthesis requires additions of successive pairs of carbons through four sequential reactions taking place in the elongase protein complex of the ER. This chain of reactions begins with the condensation of a long-chain acyl-CoA with a malonyl-CoA by the enzyme 3-keto-acyl-CoA synthase (KCS). The product, 3-keto-acyl-CoA, is then reduced by 3-keto-acyl-CoA reductase (KCR) to give 3-hydroxy-acyl-CoA. The following reaction dehydrates 3-hydroxy-acyl-CoA through the action of the enzyme 3-hydroxy-acyl-CoA dehydratase (HCD) to trans-2,3-enoyl-CoA, which in the final reaction of the sequence is reduced by trans-2,3-enoyl-CoA reductase (ECR) resulting in an acyl-CoA lengthened by two carbons. This product can then be incorporated into one of the different lipid classes. The substrates of VLFCA synthesis are C18 acyl-CoAs produced by the cytosolic fatty acid synthase complex (FAS). The elongation process is summarised in Figure 1.14.

In mammals and in yeast the KCS proteins are encoded by a family of elongase genes, known as the *ELO* family (Anthony A Millar & Kunst, 1997). In the Arabidopsis genome, KCS enzymes are encoded by an *ELO*-like gene family, consisting of four putative condensing enzymes and 21 fatty acid elongation genes (*FAE1*-like *KCS* genes) (Dunn, Lynch, Michaelson, & Napier, 2004; Joubès et al., 2008). Although FAE1-like enzymes are structurally unrelated to ELO enzymes, a number of FAE1-like proteins can rescue the lethal phenotype of the *elo2/elo3* double mutation in yeast. Two of the

condensing enzymes have been functionally characterised in detail: FATTY ACID ELONGATION 1 (FAE1) which encodes a condensing enzyme responsible for the C20 and C22 fatty acid elongation in seed triacylglycerol (Kunst, Taylor, & Underhill, 1992), and CER6, which is responsible for the catalysis of fatty acids longer than C22 in epidermal cells (Anthony A Millar et al., 1999; Paul et al., 2006).

The KCR protein is encoded by two genes in maize, named GL8A and GL8B. The *gl8a/gl8b* double mutant was found to be embryo lethal, providing evidence that VLCFAs are essential for plant development. Analysis of the single mutants revealed that the enzyme is necessary for the production of cuticle wax (Dietrich et al., 2005). In Arabidopsis, two *KCR* homologs (*KCR1* and *KCR2*) have also been identified, of which only one (*KCR1*) has been shown to be involved in the fatty acyl elongase complex. The *kcr1* mutation is also embryo lethal in Arabidopsis (Beaudoin et al., 2009).

The dehydratase gene was identified in yeast by Beaudoin and colleagues. Known as YJL097w/PHS1, the gene was found to be essential for cell viability: during cell division *phs1* mutants exhibited G2/M transition defects (Beaudoin et al., 2009). The Arabidopsis genome contains two *PHS1* homologs, *PASTICCINO2 (PAS2)* and At5g59770. Of these, *PAS2* alone was able to compensate for lack of the protein encoded by *PHS1* (Michaelson et al., 2008). Again, complete loss of PAS2 function was embryo lethal, highlighting the fact that VLCFAs are essential in plant development. However, when loss of function was only partial, a general decrease in VLCFA content in triacylglycerols, complex sphingolipids and cuticular waxes was observed, alongside a build-up of long-chain bases and sphingolipid precursors (Michaelson et al., 2008). Overexpression of dehydratase genes in Arabidopsis led to a VLCFA increase in epicuticular waxes, which lends evidence to the hypothesis that the elongase complex has several rate limiting enzymatic steps (Michaelson et al., 2008).

For the final reaction carried out in the elongase complex (catalysed by the trans-2,3-enoyl-reductase enzyme), only one gene has been identified in Arabidopsis: *CER10* (Zheng, Rowland, & Kunst, 2005). Its absence also leads to a reduction in VLCFA content. This gene has not been found to be essential for plant development, suggesting that functionally equivalent isoforms exist in Arabidopsis (Denic & Weissman, 2007).

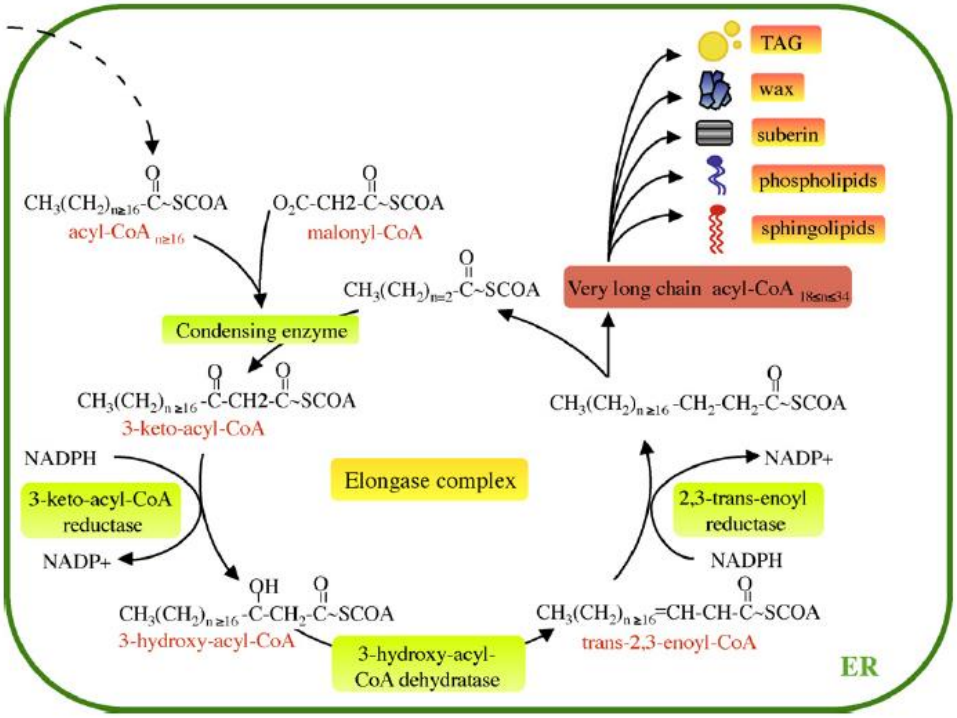


Figure 1.14. Very long fatty acid elongation cycle. VLCFAs are elongated from long chain acyl-CoA (n16) by the ER-associated elongase complex. Each two carbon cycle requires four successive enzymatic steps, the condensation of the malonyl-CoA with the acyl-CoA, reduction of the 3-keto-acyl-CoA, dehydratation of the 3-hydroxyacyl-CoA and the reduction of the trans-2,3-acyl-CoA. Finally, the enoyl-CoA gives rise to the formation of very long chain acyl-CoA. ER: enplasmic reticulum (Bach & Faure, 2010)

In membrane lipids, VLCFAs have important roles to play as signal molecules and structural components. Among plant phospholipids, phosphatidylethanolamine (PE) and phosphatidylcholine (PC) are the major components of the plasma membrane, and the membranes binding mitochondria and the endoplasmic reticulum. A third lipid class, phosphatidylserine (PS), accounts for less than 1%. The chloroplast envelope and thylakoid membrane are made up of glycolipids; mainly monogalactosyl-diacylglycerol (MGDG) and digalactosyl-diacylglycerol (DGDG). Additionally, phosphatidylglycerol (PG) is specific to the thylakoid membrane. PC, PE and PS contain significant amounts VLCFAs; these consist of 18-carbon species in combination with VLFCAs ranging from 20 to 26 carbons. Though it makes up only 1% of glycerolipids, PS is the class containing the highest proportion of VLCFAs, representing up to 69%. In addition to phospholipids, and sphingolipids are important structural membrane components, making up at least 40% of membrane lipids in some plants and found in the plasma membrane, Golgi body and in endosomes (Sperling, Franke, Lüthje, & Heinz, 2005). Heightened complex sphingolipid levels were associated with cell apoptosis, terminal differentiation, or cell cycle arrest (Pettus, Chalfant, & Hannun, 2002), and correspondingly, their reduction can lead to cell proliferation (Kohno et al., 2006).

VLCFAs influence membrane dynamics, including membrane bending, fusion and micro-domains. C26 has been shown to be important for the bending of membranes (Schneiter et al., 1996). It has also been shown that a decrease in 26:0 lipids are associated with defects in the nuclear pore, mainly linked to the separation of the envelope inner and outer membranes. The defects were the result of the absence of C26-phosphatidylinositol, which is associated with the stabilization of highly curved membrane domains (Schneiter et al., 2004). It has been reported that VLCFA content does have an impact on thylakoid membrane structure: when the fatty acid elongation gene *FAE1* was overexpressed (via linkage to a cauliflower mosaic virus 35S promoter) in *Arabidopsis* it was found that thylakoids had altered curvature morphology, appearing much more curved than those of wild-type plants (Millar, Wrischer, & Kunst, 1998). The authors found that leaf phospho- and glycerolipids contained no detectable VLCFAs, whereas transgenic 35S-*FAE1* plants contained VLCFAs in all phospholipid and galactolipid classes tested (these were: PC, PE, PG and phosphatidic acid (PA); as well as two galactolipids: MGDG and DGDG). Millar et al. observed changes in morphology in many different features of mutant plants over-expressing *FAE1*: lines which accumulated the most VLCFAs were bushy in appearance due to incomplete elongation of shoot and floral meristems, root elongation was also impaired. Additionally, 35S-*FAE1* plants had irregular rosette formation; leaf distribution appeared disorganised and asymmetrical. Morphology of the leaves and flowers themselves was also affected: leaves were darker, shorter and slightly more curled than those of wild type plants, while flowers were rounder and possessed shorter petals. In lines exhibiting the most extreme phenotypes, fertility was also reduced (Millar et al., 1998). Although fruit phenotypes were not examined in this study, it is plausible that fruit development and plastid morphology could be impacted given the wide ranging effects of alterations to VLCFA accumulation described.

1.5 Plant breeding strategies

1.5.1 Marker assisted selection (MAS)

Marker assisted selection (MAS) can be defined as: "*The use of DNA markers to improve selection in a population. The markers will be closely linked to one or more target loci, which may often be quantitative trait loci*" (Bujnicki et al., 2017). The development of DNA analysis techniques meant that MAS was first used in 1978; its use has continued into the 21st century (Collard & Mackill, 2008).

The approach has been successfully implemented in *Capsicum*, yielding several quantitative trait loci (QTL) associated with important consumer traits which add value to fruit produced. Important loci for fruit colour include the previously discussed *y* and *c2* loci (on chromosomes 4 and 6 respectively), and the *A* locus on chromosome 10 (corresponding to the gene encoding Anthocyanin2 (AN2), and associated with purple mature fruit (Borovsky, Oren-Shamir, Ovadia, De Jong, & Paran, 2004)).

Another is the *chlorophyll retainer* (*cl*) locus on chromosome 1; this encodes the protein STAY-GREEN (SGR), an amino acid substitution in which results in green mature fruit (Borovsky & Paran, 2008; Efrati, Eyal, & Paran, 2005). Major loci for other traits are the fruit ripening *S* locus on chromosome 10 (associated with the *Polygalacturonase* gene (Rao & Paran, 2003)), and the pungency locus *Pun1,* the underlying gene of which encodes the acyltransferase AT3 (Stewart et al., 2005).

1.5.2 Genetic engineering

With the advance of recombinant DNA techniques in the second half of the 20th century came the possibility of modifying organisms to give desirable phenotypes. It became possible to cut and splice individual DNA molecules together, yielding entirely novel ones. This makes it possible to alter genomes by integrating DNA from other organisms. It was first developed in bacteria and the viruses which infect them, but since then it has been translated to multi-cellular organisms, including plants and vertebrates. It has proved a powerful tool in studies of gene function as well as being used to modify organisms to give desired phenotypes, for example increased shelf-life or enhanced colour.

However, there are limitations. Two major drawbacks of transgenesis were highlighted by the Nuffield Council on Bioethics in 2016: *"(i) most phenotypic traits are complex, and require more than a single gene and (ii) it offers no control over where the added genes are inserted into the genome".*

1.5.3 New breeding techniques: CRISPR/Cas 9

Genome editing techniques have attracted a great deal of attention in recent years. Editing using these techniques facilitates the alteration of genomes in a specific and site-targeted way. Genome editing enables the selective mutation of a single gene or a small set of genes. It can also precisely replace entire genes with copies from other organisms.

The development of Site-Directed Nucleases (SDN) has been hugely important. These are capable of cutting DNA at selected target sites, and producing double-stranded breaks (DSBs). SDNs have been divided into three categories: randomly inserted mutations (SDN1) non-randomly inserted mutations, which are targeted to a specific location (SDN2), and the insertion of large sections of genetic material (for example, whole genes) into targeted locations (SDN3). DSBs are repaired by pre-existing mechanisms in the cell: homology-directed repair or non-homologous end joining.

Genome editing has advanced rapidly following the development of RNA-directed SDNs based on the bacterial CRISPR (clustered regularly interspaced short palindromic repeats) system and its associated nucleases (Cas9). This has quickly become the technology of choice for most genome editing projects. Major advantages are its simplicity, efficiency, and wide range of application. The

system has been utilised to introduce genome modifications for improvement of specific crop quality traits.

The first reports of successful gene editing in plants with CRISPR/Cas9 were published in 2013, proving that the technique could be used for both transient expression and recovery of stable transgenic lines. The first species to be transformed with CRISPR were *Arabidopsis*, and *Nicotiana benthamiana*, followed by rice (*Oryza sativa*) and wheat (*Triticum aestivum*) (Li et al., 2014)(Nekrasov et al., 2013) The delivery into plant cells of DNA constructs which encode CRISPR/Cas9 is often by *Agrobacterium tumifaciens* mediated T-DNA transfer. Occasionally, biolistic bombardment with plasmid DNA is used in its place. After introduction into the organism, these constructs are expressed, and the resulting CRISPR/Cas9 products direct cleavage of target sites, yielding the desired mutations. This way, the CRISPR/Cas9 constructs are integrated into the plant genome, but are eliminated from the following generation. Off-target effects are therefore minimised; their stable integration increases the possibility of cleavage and mutation at unintended sites due to persistent nuclease activity.

To date, simultaneous or staggered introduction of several mutations at different sites within the genome has been reported. This is made possible with the use of several different RNA guides. In Arabidopsis, mutations in 14 separate genes simultaneously have been reported, with no detection of off-target effects (Peterson et al., 2016). A plant-specific advantage of the technique is its ability to simultaneously inactivate all the alleles in polypoid plants; this has been demonstrated in hexaploid wheat and tetraploid potato (Andersson et al., 2016; Liang et al., 2017; Peterson et al., 2016; Yanpeng Wang et al., 2014).

CRISPR/Cas9 has been successfully implemented in tomato; if transformation techniques in pepper improve, the similarities between the two species will facilitate its use in *Capsicum* varieties. The first report of CRISPR/Cas9 being used in tomato was published in 2014 (Brooks, Nekrasov, Lippman, & Eck, 2014). Since then the technique has been employed to enhance fruit quality traits; editing of the Ripening Inhibitor transcription factor (RIN) has been reported (Ito, Nishizawa-yokoi, Endo, Mikami, & Toki, 2015).

1.6 Statement of Purpose

1.6.1 Project aims

The aims of the current work are divided into two distinct sections: first, to confirm the contribution of *Psy-1* (associated with the *c2* locus) to mature fruit colour through investigation at multiple levels; and second, to identify potential gene candidates for association with the *c1* locus. The first section is detailed in Chapters Three and Four, and the second in Chapters Five and Six.

Intentions regarding characterisation of the role of PSY-1 are twofold: firstly, to validate the enzyme's role in determining mature pepper colour, and thus obtain tools which breeders could employ to fine-tune the phenotype of their products. Secondly, to expand understanding of the nature of the biosynthesis and storage of these high value compounds, and further establish pepper as a model for their study. Multiple lines of investigation have been used to address the question: determining the pigment profiles associated with colour phenotypes ranging from red to pale yellow; revealing the gene expression patterns of key carotenoid biosynthetic genes during ripening in lines of diverse colour phenotype; investigating plastid structure to reveal carotenoid sequestration mechanisms and sub-chromoplast locations of proteins of interest; and testing PSY-1 enzyme activity in different sub-plastid locations.

1.6.2 Aims and experimental programme

Each of the two major objectives of the project was divided into several aims, which are detailed below. The experimental strategy is summarised in Figure 1.15.

Objective one: confirm the contribution of PSY-1 to fruit colour, characterising a new molecular marker for absence of the gene.

- Determine the carotenoid profiles of lines in a diversity panel for the marker associated with the *c2* locus.
- Sequence *Psy-1* in lines of diverse phenotype to determine the nature of the mutation in genomic DNA.
- Confirm the absence of *Psy-1* transcripts from lines possessing the marker with quantitative real time PCR.
- Assess the impact of the absence of *Psy-1* gene products on transcript levels of other genes encoding enzymes in the carotenoid biosynthesis pathway.
- Profile fruit metabolites at different stages of ripening in lines with and without functional *Psy-1* gene products, using GC-MS.
- Assess the impact of the absence of *Psy-1* on the structural components of the organelles synthesising and storing carotenoids at different stages of fruit ripening, using microscopy, sub-chromoplast fractionation, western blots, and chromatography to characterise both pigment and protein content.

Objective two: generate gene candidates for association with the *c1* locus.

- Determine the pigment profiles from lines making up a second diversity panel for the possession of the *c1* locus, lack of which is associated with white ripe fruit.
- Employ GC-MS to profile fruit metabolites at different stages of fruit ripening in accessions with and without functional copies of the *c1* locus.
- Assess the impact of absence of the *c1* locus on the structure of plastids at different stages of fruit ripening, using electron microscopy.
- Use RNA-seq to generate the first white and yellow bell pepper transcriptomes, and exploit this data to select potential gene candidates for association with the *c1* locus.

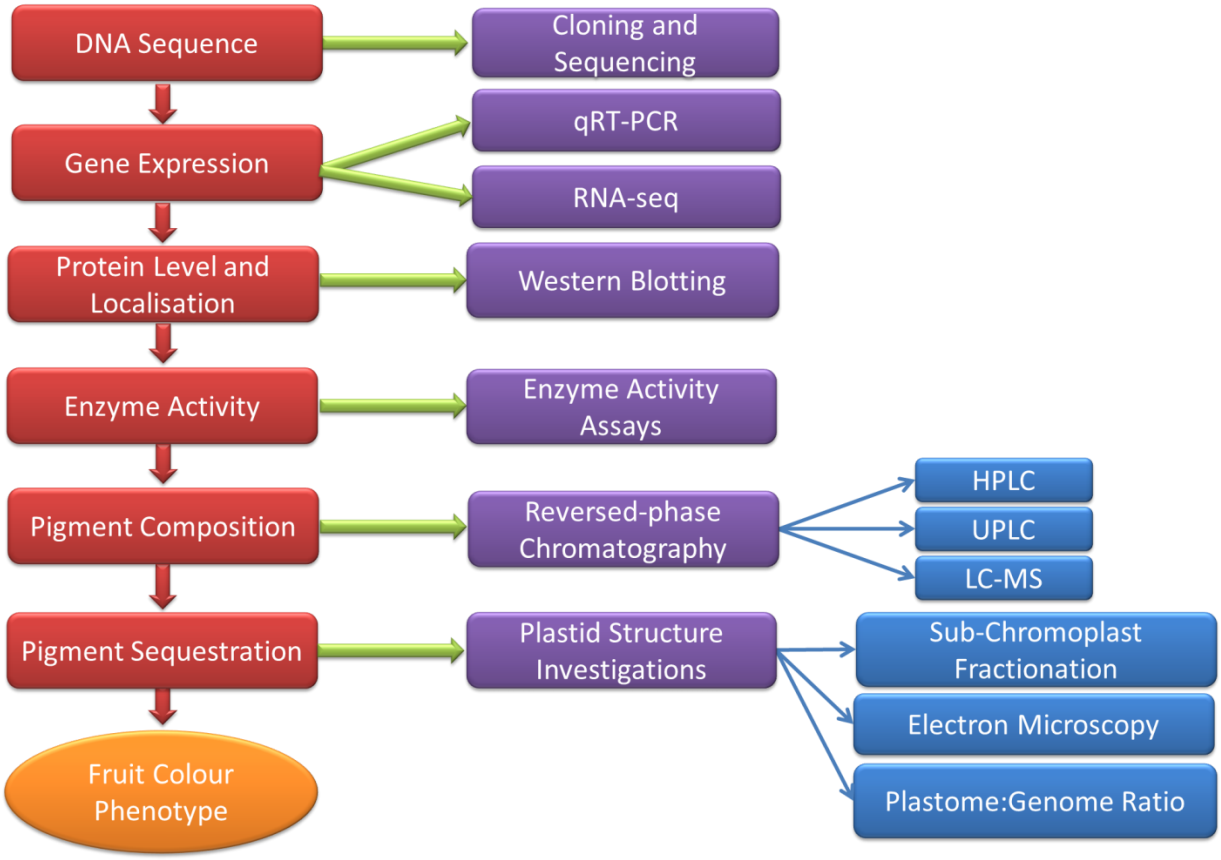


Figure 1.15. Outline of the experimental strategy used to address the aims of the project, detailing investigations into DNA and RNA sequence, protein content, enzyme activity, pigment and metabolite content, and organelle structure. Abbreviations: qRT-PCR: quantitative real-time polymerase chain reaction; RNA-seq: ribonucleic acid sequencing; HPLC: high-pressure liquid chromatography; UPLC: ultra-high-pressure liquid chromatography; LC-MS: liquid chromatography-mass spectrometry.

2 Materials and Methods

2.1 Plant Material and Growth Conditions

Seeds of two groups of sweet bell pepper (*Capsicum annuum*) were sown. The first group consisted of fifteen lines representing a diversity panel for a marker associated with the *c2* locus (*Psy-1*), these were obtained from Syngenta (Material Transfer Agreement (MTA) dated September 2013). The second were a diversity panel for the *c1* locus, consisting of twelve lines, also obtained from Syngenta (MTA dated October 2014). These were planted in a growth facility in Levington compost, and grown under a long day light regime (16h light, 8h dark). Fruit from all lines were harvested at the following stages: mature green (50 days post anthesis), breaker (the first point of colour visible on the fruit), turning (when fruit are half green and half their ripe colour), and ripe (fruit are firm, and a uniform ripe colour). After harvesting, fruit were immediately chopped and placed in liquid nitrogen to ensure metabolic quenching. Post-quenching, fruit were stored at -80°C.

2.2 Pigment Identification and Quantification

2.2.1 Pigment Extraction

Pigments were extracted with methanol and chloroform; work was carried out in a fume hood under low light. All solvents were Fisher HPLC grade. Methanol (250 µl) was added to 10mg freeze-dried, ground tissue, pooled from three biological replicates (three technical replicates were used for each pooled sample), this was then vortexed heavily. Chloroform (500 µl) was added, and samples were incubated on ice for 20 minutes. Water (250 µl) was added and samples were spun at 12,000 rpm for 3 minutes. The chloroform phase was then removed to a fresh 1.5 ml microcentrifuge tube. Chloroform (500 µl) was added to the initial sample, and the subsequent steps were repeated, with the chloroform phase from the first and second extractions being pooled. The polar phase was discarded. The non-polar, chloroform phase was dried with a GeneVac rotary evaporator. Samples were stored in the dark at -20°C until analysis.

2.2.2 Chromatography

2.2.2.1 High-Performance Liquid Chromatography (HPLC)

Extracted pigments were resuspended in ethyl acetate (100 µl), and centrifuged at maximum speed for 10 minutes before analysis with a Waters Alliance High Performance Chromatography (HPLC) separations module containing a C30 column maintained at a temperature of 25°C. An injection volume of 10 µl was run with a solvent gradient optimised for carotenoid analysis. Solvents were: A: 100% methanol; B: 0.2% ammonium acetate, 80% methanol, 20% water; C: tert-butyl-methylamine. The gradient was as follows: 65% A: 30% C: 5% B for 10 minutes; then stepped to 30% A: 65% C: 5% B. The total run time was 60 minutes. Absorption was detected with an online photodiode array detector scanning between the wavelengths of 250-600 nm. Standard carotenoids, xanthophylls,

and chlorophylls were purchased from Sigma; quantification of sample pigments was obtained from a standard curve of amount plotted against peak area.

2.3 Metabolite Profiling with Gas Chromatography-Mass Spectrometry (GC-MS)

2.3.1 Metabolite Extraction

Metabolites for analysis with gas chromatography-mass spectrometry (GC-MS) were extracted with a methanol-chloroform extraction method. 10 mg of freeze dried fruit pericarp tissue was weighed from three biological replicates per condition. Three technical replicates were also used. Methanol (400 µl) and water (400 µl) were added to freeze-dried tissue, and samples were mixed vigorously with a vortex before being incubated with continuous circular inversion for one hour at room temperature. Chloroform (800 µl) was added and mixed with vortexing. This was followed by centrifugation at 10,000 rpm for 5 minutes. The sample separated into two phases; the upper phase consisted of methanol and water, and the lower phase consisted of chloroform. The lower phase was collected in a fresh microcentrifuge tube. An internal standard of myristic acid-d27 (10 µg) was added, and the sample was dried in a GeneVac rotary evaporator. Derivatisation was carried out a maximum of 24 hours before samples were added to the GC-MS autosampler for analysis. This consisted of addition of methoxyamine-HCl (30 µl) prepared at a concentration of 20 mg/ml in pyridine. This mixture was then incubated at 40°C for one hour in glass tubes, before the addition of methyltrimethylsilytrifluoroacetamide (70 µl) was added and incubated for at 40°C for a further two hours.

2.3.2 Analysis with Gas Chromatography-Mass Spectrometry (GC-MS)

GC-MS was carried out on non-polar metabolites of selected lines as described previously (Perez-fons et al., 2014). The machine was an Agilent HP6890 gas chromatograph with a 5973MSD for mass detection. An injection volume of 1 µl was injected with a split/splitless injector with a 20:1 split at 290°C. The oven for containing the gas chromatography column was held for 4 minutes at 70°C before ramping at 5°C per minute to 310°C, which was held for a further ten minutes. The total run time was 60 minutes. The connection to the MS was set to 290°C and MS performed in full scan mode with 70 eV EI+, scanning from 10-800 D. An MS library was adapted to identify compounds present in pepper samples, adding to existing libraries with information from the NIST08 MS library. Using retention indices and MS, identification was performed by comparison with the MS library. Compounds in each sample were quantified relative to the internal standard. Statistical analyses (principal components analysis (PCA) and partial least squares discriminant analysis (PLS-DA)) were carried out using the computer platform Smica P (2014).

2.4 Nucleic Acid Extraction

2.4.1 DNA Extraction

Total DNA was extracted from plant cells using a DNeasy Plant Kit (Qiagen), according to the instruction manual supplied by the manufacturer. After harvesting, chopped fruit pericarp tissue was immediately frozen in liquid nitrogen. This frozen tissue was ground in liquid nitrogen with a pestle and mortar. A maximum of 100 mg was weighed into a liquid-nitrogen cooled, sterile, 2 ml Eppendorf tube. Buffer AP1 (400 µl) and RNase A (4 µl) were added to the tube, and vigorously mixed with a vortex. Subsequently, sample tubes were incubated in a 65°C water bath; being mixed through inversion three times during this period. Buffer AP2 (130 µl) was added, sample tubes were incubated for 5 minutes on ice. The resulting lysate was transferred to a fresh QIAshredder Mini spin column in a 2 ml collection tube, this was then centrifuged for 2 minutes at 14,000 rpm (5424 bench-top centrifuge (Eppendorf)) The resulting flow-through was transferred to a new sterile 2 ml Eppendorf tube, and Buffer AP3 (1.5 volumes) was added. This mixture was then transferred to a fresh DNeasy Mini spin column and centrifuged for 1 minute (14,000 rpm). After discarding the flow-through, the column was transferred to a fresh, sterile 2 ml Eppendorf tube and washed with Buffer AW (500 µl) twice, being centrifuged at 14,000 rpm for 1 minute for each wash. To elute the DNA, the column was transferred to a sterile 1.5 ml Eppendorf tube and the elution buffer (Buffer AE (50 µl)) was added. After centrifugation at 14,000 rpm for 1 minute, a further 50 µl Buffer AE was added and the centrifugation was repeated. The quality and quantity of the resulting DNA was measured with a NanoDrop Spectrophotometer, before samples were stored in the dark at -20°C.

2.4.2 RNA Extraction

RNA was extracted from chopped, liquid nitrogen-quenched fruit pericarp tissue, with an RNeasy Plant Kit (Qiagen) using the standard protocol provided by the manufacturer (RNeasy handbook, 2001). Tissue was ground in liquid nitrogen in a pre-chilled pestle and mortar. A maximum of 100 mg was weighed into a liquid nitrogen cooled 2 ml Eppendorf tube, and Buffer RLT (450 µl) was added. The sample tube was mixed thoroughly with a vortex. The resulting lysate was transferred to a QIAshredder Mini spin column in a 2 ml collection tube; this assembly was centrifuged for 2 minutes at the centrifuge's maximum speed (Eppendorf 5424 bench-top centrifuge). The flow-through was then transferred to a fresh, sterile 2 ml Eppendorf tube to which 100% ethanol (0.5 volumes) was added. This was placed in a RNeasy Mini spin column inside a 2 ml collection tube and centrifuged for 15 seconds (14,000 rpm), with the flow through being subsequently discarded. Buffer RW1 (350 µl) was added and the column centrifuged (15 seconds at 14,000 rpm). After discarding the flow-through DNaseI incubation mixture (80 µl) was added directly to the membrane of the column assembly, and incubated on the bench at room temperature for 15 minutes. Buffer RW1 (350 µl)

was used to wash the membrane through centrifugation at 14,000 rpm for 15 seconds. This was followed by a second wash: Buffer RPE (500 µl), also centrifuged at 14,000 rpm for 15 seconds. Buffer RPE (500 µl) was used for a final wash, with the column being centrifuged for 2 minutes at 14,000 rpm. The flow through of each of the washes was discarded. After the final wash, the column was transferred to a new sterile 1.5 ml Eppendorf tube; the RNA was eluted by two rounds of centrifugation for one minute at 14,000 rpm with RNase-free water (30 µl each time). The quality and quantity of the resulting RNA was measured with a NanoDrop, and single-use aliquots (25 ng/µl; 15 µl) were stored in the dark at -80°C.

(URL: https://www.qiagen.com/gb/resources/resourcedetail?id=14e7cf6e-521a-4cf7-8cbc-bf9f6fa33e24&lang=en)

2.5 Cloning

2.5.1 Amplicon Design

Amplicons for RT-qPCR with lengths of 80-150 base pairs were designed using cDNA sequence data obtained from Syngenta's internal *Capsicum* genome (version 7). Where possible, the amplicon spanned an intron, thus making gDNA contamination easily recognisable. Primers were designed using Primer3Plus.

(URL: http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi)

2.5.2 Cloning

Amplicons were cloned into a TOPO TA vector (Promega, UK). Purified PCR products were ligated into the vector, and transformed into TOP10 chemically competent *E. coli* cells (Promega, UK) according to the protocol provided by the manufacturer. Lysogeny Broth (LB) medium was prepared and autoclaved, this consisted of: tryptone (10 g/l), yeast extract (5 g/l), and NaCl (5 g/l). The PCR product (2 µl) was added to the TOPO vector (1 µl), a salt solution (1 µl) and molecular biology grade sterile water to a final volume of 6 µl. To allow ligation, the cloning reaction mixture was incubated at room temperature for 5 minutes before being placed on ice. The cloning reaction (2 µl) was added to a vial of One Shot TOP10 chemically competent cells, which had been thawed from storage at -80°C on ice. The competent cells were transformed by a heat shock of 42°C for 30 seconds, before immediate incubation on ice. SOC medium (250 µl) at room temperature was added to the transformation, and samples were incubated at 37°C for 1 hour, with horizontal shaking at 200 rpm. 20 and 40µl of each transformation were spread onto LB-agar plates (pre-warmed at 37°C) with kanamycin (50 µg/ml) as a selection agent. Plates were incubated at 37°C for 16 hours. Four single colonies from each plate were then selected and grown in liquid LB medium with kanamycin (50

µg/ml) for 16 hours at 37°C with shaking. Cells were then pelleted, and plasmid DNA was purified with a Wizard *Plus* SV Miniprep DNA Purification System (Promega), according to the protocol provided. Purified plasmids were tested by PCR with the appropriate primers, followed by agarose gel electrophoresis to ensure they contained the correct product.

2.6 Gene Expression Studies

2.6.1 Reverse-Transcription PCR

Reverse-transcription PCR (RT-PCR) was used to create a complementary DNA (cDNA) copy of the amplicons of interest. Illustra Ready-to-go RT-PCR beads (GE Healthcare, UK) were used to perform the reaction. Reactions contained oligo (dT) primer (2.5 pmol), each of the respective forward and reverse primers (5 pmol) and template RNA (200 ng). Molecular biology grade sterile water was added to make up the final volume (50 µl). Reactions were initially denatured at 55°C for 5 minutes, followed by a reverse transcription step at 65°C for 30 minutes, with a subsequent 5 minute incubation to denature the reverse transcriptase. This was followed by the addition of gene-specific primers 30 cycles of PCR amplification (94°C for 30 seconds; annealing at 52°C; and extension at 72°C for 30 seconds). The final step was a fixed extension at 72°C for 5 minutes. Reactions were carried out in a thermos cycler. Products were analysed with gel electrophoresis and cloned into TOPO vectors to be used in the production of standard curves in real-time qPCR.

2.6.2 Quantitative Real-Time PCR

Transcript level in the sampled tissue was analysed with quantitative real-time PCR (qRT-PCR), using a Rotor-Gene and a QuantiFast SYBR Green florescence marker (Qiagen, UK) and a Rotor-Gene RG-3000 thermal cycler (Corbett Life Sciences, UK). The level of transcript of a gene of interest in 30ng cDNA per sample was compared to a standard curve of that gene, obtained from dilutions of the purified plasmid. The transcript level of the gene of interest was also compared to the transcript level of a reference (housekeeping) gene in each tissue sample. AtpA was chosen as the reference gene for *C. annuum* due to its stability across tissue samples. Per sample, in addition to the 30 ng cDNA, each reaction contained a forward and reverse primer (1 µM), 2X QuantiFast SYBR Green RT-PCR Master Mix (10 µl), QuantiFast RT Mix (0.2 µl) and RNA-free water to final volume of 20 µl. The thermal cycle was 5 minutes at 95oC for PCR activation, followed by 35 cycles of 95oC for 5 seconds and 60oC for 10 seconds for PCR amplification. Calibration curves for each gene of interest were run simultaneously with samples of unknown concentration. The curves contained 5 points of dilution, covering the complete range of expression, and determined the assay's linearity, reaction sensitivity, efficiency, and reproducibility. Reactions were considered acceptable when PRC efficiency, produced by the Rotor-Gene software, was between 90 and 110% and R^2 values of at least 0.985. From raw

cycle threshold (Ct) values, data was processed through comparison to the housekeeping gene to give a ΔΔCt value for each sample.

2.6.3 Transcriptome Analysis with RNA-seq

Total RNA was extracted from pepper fruits at three different ripening stages (mature green, breaker plus five days, and ripe). This was then prepared for sequencing in an Illumina flow cell with the addition of adapter sequences which bound RNA strands to tiles on the cell. 100 base pair, paired end sequences were produced for each sample in an Illumina HighSeq flow cell, by Syngenta at their facility at Research Triangle Park. Raw sequences were delivered as FastQ files for downstream informatic processing.

Firstly, the quality of sequences was checked with the software package FastQC. The programme reports on per-tile sequence quality, %GC content, sequence length, and over-represented sequences. Samples passed the quality checks and were transferred to the next stage of the analysis. Adapter sequences were masked with the programme Cutadapt, before being removed from samples with the programme Ernefilter. Following the removal of adapter sequences, the Cufflinks software package was used . Sequences were aligned to a reference genome (*Capsicum annuum* Zunla version 2.0) with the programme Tophat2. Aligned sequences were then annotated against the reference annotation (for *Capsicum annuum* Zunla version 2.0) with Cufflinks. Statisitcal analysis to identify differentially expressed transcripts was carried out with the programme Cuffdiff. The R package CummeRbund was used to generate plots from Cufdiff outputs. Differentially expressed transcripts were annotated against a vascular plant database from NCBI using the sotware platform Blast2GO; the programme also generated associated cellular components and metabolic functions were for each transcript.

2.7 Plastid Structure Investigations

2.7.1 Sub-Chromoplast Fractionation

To separate different membrane structures present within chromoplasts of ripe fruit, the organelles were extracted, broken, and suspended on a stepwise sucrose gradient. This was carried out at 4°C, working on ice. Chopped fruit kept in the dark for 16 hours were blended in an extraction buffer consisting of the following components: 0.4 M sucrose, 50 mM Tris, 1 mM DTT, and 1 mM ethylenediaminetetraacetic acid (EDTA) at pH 7.8. Blended tissue was passed through layers of muslin to extract intact chromoplasts. Plastids were pelleted and subsequently broken with a hand held Potter homogeniser. The broken chromoplasts were suspended on the discontinuous sucrose gradient, with the following steps: 45% (8 mL), 38% (6mL), 20% (6mL), 15% (4 mL), 5% (6 mL) (w/v). To prepare this gradient, sucrose was added to a buffer consisting of 50 mM Tricine, 2 mM EDTA, 2

mM DTT, 5 mM Sodium bisulphite at pH 7.9. The broken chromoplasts on the gradient were centrifuged at 4°C for 18 hours at 100,000 x g. 1mL fractions were collected from the sucrose gradient. Carotenoids and proteins were extracted from each fraction for further analysis (Nogueira et al., 2016, 2013).

2.7.2 Transmission Electron Microscopy

2 mm cubes of pepper pericarp tissue from mature green or ripe fruit were cut and placed in cold fixative solution (CAB; 2.5% glutaraldehyde in 100 mM sodium cacodylate buffer, pH 7.2). Tissue cubes in 2 ml cold fixative were incubated overnight at 4°C. Samples were washed in CAB twice (10 minutes each) before being post-fixed in osmium tetroxide in CAB for 1 hour at room temperature. Following this, samples were washed twice in milliQ water for 10 minutes each time, and dehydrated in increasing concentrations of ethanol (50, 70, 90 and 100%, for 10 minutes each). A transition solvent of propylene oxide was then used to wash samples. Tissue cubes were then moved for 30 minutes to 50% propylene oxide, 50% agar low viscosity resin (ALVR), before being placed in 100% ALVR for 1.5 hours twice. A vacuum was applied four times during the incubation. Samples were transferred to silicone moulds and polymerised at 60°C for 24 hours. Polymerised blocks were sectioned into 70 nm sections, and collected on 400 mesh copper grids. Sections were counterstained with uranyl acetate (4.5%) in acetic acid (1%) for 45 minutes, and lead citrate for 7 minutes. Finally, sectioned samples were viewed in a Jeol 1230 TEM with an accelerating voltage of 80 kV.

2.8 Protein Level Investigations

2.8.1 Protein Extraction

Proteins were extracted and separated as described previously (Nogueira et al., 2013). After carotenoids had been extracted and the organic phase removed, methanol (500 µl) was added. The sample was vortexed and centrifuged for 5 minutes at 14,000 rpm. After discarding the supernatants, protein pellets were air-dried and stored in the dark at -20°C.

2.8.2 Separation with SDS-PAGE

Protein pellets were suspended in SDS loading buffer (50 µl; 0.5 M Tris-HCl (pH 6.8), 10% (v/v) glycerol, 10% (w/v) SDS, 0.05% (w/v) bromophenol blue, 1.4% (v/v) β-mercaptoethanol). Samples were incubated in boiling water for 5 minutes. Dissolved protein in SDS loading buffer were loaded in 0.5 cm SDS-PAGE gel (12.5%) wells. The SDS-PAGE gel had two components: a stacking gel (acrylamide/bis-acrylamide solution (16.7% v/v), Tris-HCl (0.5 M; 12.6% (v/v), pH 6.8), SDS (0.1% w/v) ammonium persulphate (0.1% w/v) and tetramethylethylenediamine (temed) (0.1% v/v)) and a running gel (acrylamide/bis-acrylamide solution (40% v/v), Tris-HCl (1.5 M; 12.6% (v/v), pH 8.8), SDS

(0.1% w/v) ammonium persulphate (0.1% w/v), and tetramethylethylenediamine (temed) (0.1% v/v)). A Hoefer Scientific Instrument gel-casting apparatus was used to shape the gels. They were run in a running buffer (Tris (0.25 M), glycine (1.92 M), and SDS (1% w/v)), using a ladder (High-Range Rainbow Molecular Weight Markers (GE Healthcare)). The gels were run in a Hoefer protein electrophoresis apparatus at 100V for 2.5 hours.

2.8.3 Silver Staining of SDS-PAGE Gels

A ProteoSilverTM Plus Silver Stain Kit (Sigma-Aldrich, UK) was used to stain proteins separated with SDS-PAGE, according to manufacturer's instructions. Firstly, gels were incubated in a fixing solution (50% (v/v) ethanol, 10% (v/v) acetic acid) for a minimum 20 minutes. Following this, gels were washed in ethanol (30% v/v), then distilled water. They were then incubated for 10 minutes in a sensitising solution (ProteoSilver Sensitizer (1% v/v)), followed by two ten minute washes in distilled water. Gels were equilibrated for 10 minutes in Silver solution (ProteoSilver Solution, 1% v/v) before being washed for 1.5 minutes in distilled water. This was followed by gel development: Developer solution (ProteoSilver Developer 1 (5% v/v) and ProteoSilver Developer 2 (0.1% v/v)). Development was halted when protein bands on the gels had reached the required colour intensity with a Stop solution (ProteoSilver Stop Solution (5 ml).

2.8.4 Western Blotting

After proteins had been separated through SDS-PAGE, they were blotted onto a polyvinyldiene fluoride (PVDF) membrane via electrophoresis at 100 V. Western blots with antibodies raised against tomato PSY-1, *Capsicum annuum* Fibrillin and tomato Plastoglobulin (PG35) were developed.

Initially, the transfer membrane was incubated in methanol (20% v/v) for 10 seconds, followed by incubation in Transfer Buffer (25 mM Tris, 192 mM glycine, and 20% (v/v) methanol) for 10 minutes inpreparation for blotting. The SDS-PAGE gels intended for blotting were rinsed with Transfer Buffer and overlaid with transfer membranes. The gel-membrane pairs were placed between 2 pieces of equilibrated filter paper, and two equilibrated foam pads, in a cassette designed for membrane transfers. This was placed into a Mini Trans-Blot Electrophoretic Transfer Cell (Bio-Rad), together with an ice pack to maintain a low temperature. The Transfer Cell was flooded with Transfer Buffer and run at 100V for approximately 1 hour.

After proteins had been transferred to the PVDF membrane, this was incubated overnight in Bovine Serum Albumin (BSA) (1% w/v) at 4^0C with gentle agitation (200 rpm). Solutions of primary antibodies were prepared, by diluting BSA in a TBST solution: Tris-HCl (5 mM; pH 7.5), NaCl (15 mM), Tween20 (0.01% v/v), and BSA (0.1% v/v). Primary antibody solutions were then incubated with the appropriate membrane for 1 hour with gentle agitation (200 rpm). Following this, membranes were

washed in TBST three times for 10 minutes with gentle agitation (200 rpm). The secondary antibody was bound to the membranes: this was Anti-Rabbit IgG Alkaline Phosphatase Conjugate for Fibrillin and PG35, and Anti-Mouse IgG Alkaline Phosphatase Conjugate for PSY-1 in a 1:5000 (v/v) dilution in TBST. After three washes for 5 minutes in TBST and two washes for one minute in a TBT buffer (Tris-HCl (5 mM; pH 7.5), NaCl (15 mM)) with gentle agitation (200 rpm). Blots were developed using 5-Bromo-4-chloro-3-indolyl phosphate/Nitro blue tetrazolium (*FAST*TM BCIP®/NBT) tablets (Sigma-Aldrich, UK). One tablet was dissolved in 20 ml of distilled water twice for 2 minutes.

2.8.5 Enzyme Activity Assays

Isolated chromoplasts were broken and divided into stroma and membrane fractions. Stroma and membrane samples were incubated with radiolabelled compounds: ^{3}H-GGPP and ^{14}C-IPP. Carotenoids were extracted with methanol/diethyl ether/chloroform, and separated with Thin Layer Chromatography (TLC) using alumina plates and a solvent system of 6% toluene in petroleum ether, visualised with iodine. Bands with the same *Rf* as the phytoene standard were scraped into scintillation vials for the measurement of radioactivity. Radioactivity in the band corresponding to phytoene was measured with a scintillation counter. A Bradford Assay was used to determine the amount of protein per sample (Bradford, 1976).

2.9 Statistical Analysis

Three biological replicates and three technical replicates were analysed for each experiment unless stated otherwise stated. Principal components analysis (PCA) and partial least squares discriminant analysis (PLS-DA) were carried out using the software platform SIMCA-P v 14.0 (Umetrics, UK). Student's t-tests were calculated using Excel (Microsoft) embedded algorithms.

3 *Psy-1* diversity panel: population characterisation

3.1 Introduction

As colour is such an economically important fruit quality trait for peppers, it is essential that breeders have reliable tools with which to control it. Mature fruit colour is largely dependent on carotenoid content; flux through this pathway is controlled by the enzyme phytoene synthase (PSY) (Pathway, Bartley, Viitanen, Bacot, & Scolniks, 1992). In ripening fruit, this is carried out by a specific paralogue of the enzyme, PSY-1.

To quantify the contribution of PSY-1 to ripe fruit colour in yellow and orange peppers, a diversity panel of fifteen lines was generated by Syngenta. In relation to the three loci model, all homozygous recessive for *y (Ccs)*, four possess a gene encoding functional copies of the enzyme PSY-1 (associated with the *c2* locus) and the gene associated with the *c1* locus *(unknown)* (RH 1-4; Figure 3.3), ten have functional copies of the gene associated with *c1* only, (RH 5-14; Figure 3.3) and one is homozygous recessive for all three loci (RH 15; Figure 3.3). *Psy-1* sequencing primers were designed in order to confirm the nature of the mutation in genomic DNA; carotenoids were profiled and expression of other carotenoid biosynthetic genes was studied in order to investigate the impact of the lack of the enzyme on the pathway it controls.

3.2 Results

3.2.1 Determination of carotenoid content through HPLC-PDA

Determining the relationship between Psy-1 expression status and carotenoid content was a principal aim of the project. High Pressure Liquid Chromatography (HPLC) on a C-30 column is a well-established and reliable method of quantifying this class of compounds in plant material (Fraser et al., 2000)

Extracted carotenoids were re-suspended in ethyl acetate and subjected to reverse phase High Pressure Liquid Chromatography (HPLC) on a C-30 column. The resulting chromatograms were complex, separating free carotenoids and their mono- and di-esters, and contained up to forty peaks for ripe fruit when interrogated at the wavelength corresponding to carotenoid absorption (450 nm) (Figure 3.1; Table 3.1). Interrogating the chromatogram at different wavelengths revealed amounts of compounds absorbing energy in the corresponding regions of the electromagnetic spectrum. For example, analysing the chromatogram at 286 nm shows peaks for phytoene and α-tocopherol (vitamin E) (Figure 3.2; Table 3.2). Mature green samples contained fewer mono- and di-esters and generally consisted of twenty peaks. It has been shown with Liquid Chromatography-Mass Spectrometry (LC-MS) on non-polar *Capsicum* extracts that mono-esters elute before the 28th minute of the HPLC run, and di-esters elute after this (Schweiggert, Kammerer, Carle, & Schieber, 2005). Fruit colour relates to carotenoid amount: accessions expressing Psy-1 are a darker yellow or orange than those which do not (Table 3.3; Figure 3.3; Figure 3.4).

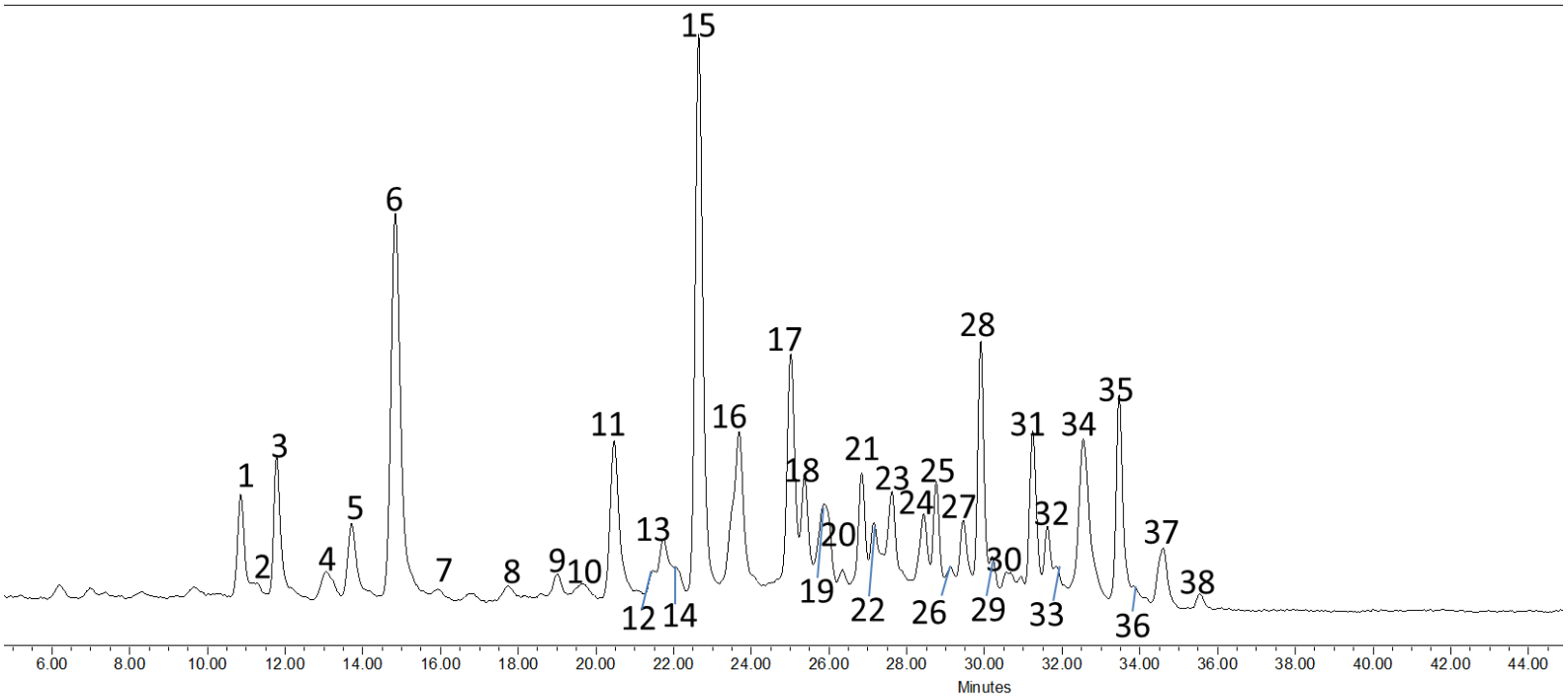


Figure 3.1 HPLC chromatogram showing separation of carotenoids in ripe fruit of line RH 14 at 450 nm. A C30 column was used, in reverse phase with the solvents methanol and tert-butyl-methylamine. Numbers relate to numbers in Table 3.1.

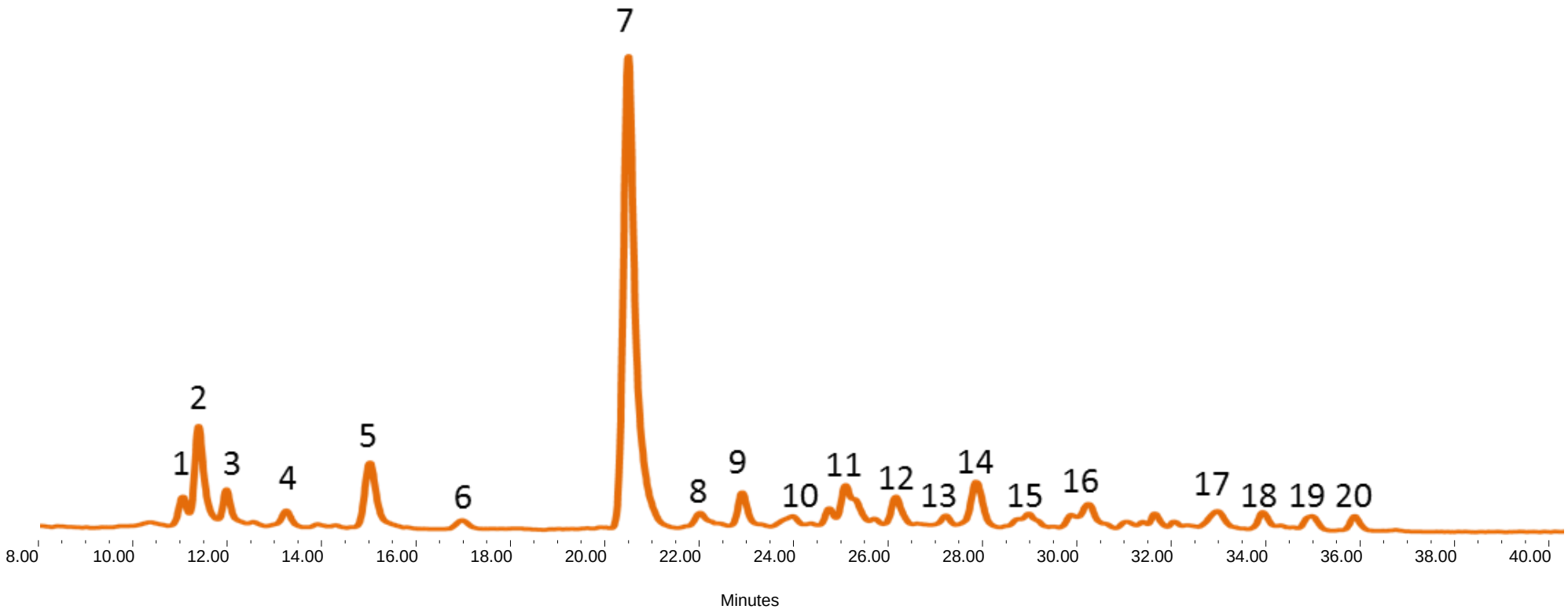


Figure 3.2. HPLC chromatogram showing separation of a non-polar ripe fruit extract from RH 1 at 286 nm. A C30 column was used in reverse phase, with the solvents methanol and tert-butyl-methylamine. Numbers relate to those in table 3.2..

Number	Retention Time (minutes)	Peak Identification	Spectral Characteristics (nm)
1	10.873	Violaxanthin	439.1, 468.1
2	11.313	Neoxanthin	436.7, 463.2
3	11.783	Luteoxanthin	422.2, 448.8
4	12.245	Auroxanthin	399.4, 424.6
5	13.709	Antheraxanthin	443.9, 471.7
6	14.789	Lutein	445.1, 472.9
7	15.873	Violaxanthin ester 1	442.2, 469.3
8	17.662	Violaxanthin ester 2	440.3, 469.3
9	18.898	Violaxanthin ester 3	440.3, 471.7
10	19.601	Neoxanthin ester 1	435.5, 463.2
11	20.351	Violaxanthin ester 4	441.5, 470.5
12	21.361	Luteoxanthin ester 1	421.0, 448.8
13	21.613	Luteoxanthin ester 2	423.4, 448.8
14	21.912	Neoxanthin ester 2	435.5, 464.5
15	22.502	Violaxanthin ester 5	441.5, 470.5
16	24.428	Luteoxanthin ester 3	419.8, 443.9
17	24.543	Luteoxanthin ester 4	423.4, 447.5
18	25.206	Lutein ester 1	443.9, 472.9
19	25.755	Luteoxanthin ester 5	423.4, 448.8
20	26.165	Violaxanthin ester 6	439.1, 469.3
21	26.971	Lutein ester 2	443.9, 474.1
22	27.416	Lutein ester 3	443.9, 472.9
23	27.628	Violaxanthin ester 7	441.5, 468.9
24	28.26	Neoxanthin ester 3	437.9, 465.7
25	28.587	Violaxanthin ester 8	441.4, 470.5
26	28.878	Lutein ester 4	445.1, 469.3
27	29.276	Neoxanthin ester 4	437.9, 465.7
28	29.738	Violaxanthin ester 9	441.5, 470.5
29	30.037	Violaxanthin ester 10	441.5, 470.5
30	30.412	Neoxanthin ester 5	439.1, 465.7
31	30.758	Luteoxanthin ester 6	423.4, 450.0
32	31.63	Lutein ester 5	445.1
33	31.832	Luteoxanthin ester 7	423.4, 448.8
34	32.318	Lutein ester 6	445.1, 474.1
35	33.495	Lutein ester 7	443.9, 474.1
36	33.809	Violaxanthin ester 11	442.9, 470.5
37	34.235	Antheraxanthin ester 1	448.8, 474.1
38	35.136	Antheraxanthin ester 2	450.0, 475.3

Table 3.1. Peak identification of carotenoids in ripe fruit of line RH 14. Numbers correspond to those in figure 3.1.

Number	Retention Time (minutes)	Peak Identification	Spectral Characteristics (nm)
1	10.846	Cis-Violaxanthin	266.3, 439.1, 469.3
2	11.167	Alpha tocopherol	292.3
3	11.746	Cis-Luteoxanthin	252.1, 422.2, 448.8
4	12.319	Auroxanthin	382.0, 400.6, 425.8
5	14.739	Lutein	297.5, 420, 445.1, 472.9
6	16.653	Zeaxanthin	275.7, 451.2, 477.8
7	20.086	Phytoene	286.4, 300.0
8	21.576	Unidentified	321.0, 348.2, 365.8
9	22.44	Violaxanthin monoester	416.2, 440.3, 470.5
10	23.463	Luteoxanthin monoester	399.4, 422.2, 448.8
11	24.554	Auroxanthin monoester	382.5,399.4, 423.4
12	25.563	Luteoxanthin monoester	423.4, 448.8
13	26.581	Lutein monoester	423.0, 445.1, 474.1
14	27.198	Lutein monoester	428.0, 448.8, 474.1
15	28.231	Neoxanthin diester	423.0, 439.1, 465.7
16	29.547	Violaxanthin diester	417.1, 441.5, 470.5
17	32.184	Violaxanthin diester	420.0, 441.5, 471.7
18	33.179	Antheraxanthin diester	424.1, 446.3, 474.1
19	34.239	Antheraxanthin diester	425.4, 446.3, 475.3
20	35.176	Antheraxanthin diester	425.4, 446.3, 475.3

Table 3.2 . Peak identification of compounds in ripe fruit of line RH 1. Numbers correspond to those in figure 3.2.

Accession	*Psy-1* Genotype	Compound (Amount µg/g DW)													
		Capsanthin	SE	Violaxanthin	SE	Antheraxanthin	SE	Lutein	SE	β-carotene	SE	Neoxanthin	SE	Phytoene	SE
RH 0	**+**	23674.68	1779.113	56.36	1.94	81.15	5.31	359.44	74.28	576.29	45.64	27.72	4.18	1030.76	146.62
RH 1	**+**	n.d	n.a	638.966	82.417	62.575	10.170	577.997	30.788	n.d	n.a	67.909	9.216	1973.490	60.065
RH 2	**+**	n.d	n.a	151.802	5.673	21.920	0.257	200.359	17.236	342.238	49.685	35.821	0.375	2203.317	138.423
RH 3	**+**	n.d	n.a	267.624	31.400	20.875	1.124	296.095	10.781	340.999	35.274	n.d	n.d	888.950	13.769
RH 4	**+**	n.d	n.a	386.798	87.558	23.174	0.438	352.965	2.566	584.235	8.693	n.d	n.d	1632.183	54.341
RH 5	**-**	n.d	n.a	106.928	2.585	15.831	0.049	183.996	1.207	235.489	0.668	38.150	0.276	321.920	0.565
RH 6	**-**	n.d	n.a	50.336	0.815	15.943	1.086	99.555	30.097	n.d	n.a	n.d	n.a	n.d	n.a
RH 7	**-**	n.d	n.a	130.686	4.413	17.563	0.234	230.395	2.501	n.d	n.a	n.d	n.a	n.d	n.a
RH 8	**-**	n.d	n.a	67.322	17.801	15.180	0.323	110.111	40.653	n.d	n.a	35.817	0.677	257.383	2.404
RH 9	**-**	n.d	n.a	60.491	1.739	15.997	0.144	122.082	2.416	n.d	n.a	36.003	0.395	n.d	n.a
RH 10	**-**	n.d	n.a	56.959	0.703	17.950	1.400	132.370	4.417	n.d	n.a	35.935	0.191	n.d	n.a
RH 11	**-**	n.d	n.a	58.827	3.414	15.937	0.223	109.256	5.691	n.d	n.a	36.007	0.480	n.d	n.a
RH 12	**-**	n.d	n.a	55.283	0.887	15.002	0.013	104.191	0.798	n.d	n.a	35.862	0.252	n.d	n.a
RH 13	**-**	n.d	n.a	83.332	16.919	15.103	0.135	96.583	13.562	n.d	n.a	n.d	n.a	244.191	6.183
RH 14	**-**	n.d	n.a	94.997	9.841	17.629	0.379	103.150	2.263	n.d	n.a	37.278	0.598	n.d	n.a
RH 15	**-**	n.d	n.a	n.d	n.a	n.d	n.a	n.d	n.a	n.d	n.a	n.d	n.a	n.d	n.a

Table 3.3. Carotenoid amounts in ripe fruit determined by HPLC-PDA (C30 column) as described in Methods (Section 2.2.2). Quantification was carried out using calibration curves from authentic standards (SE = Standard Error).

Figure 3.3. Ripe fruit of the fifteen lines making up the *Psy-1* diversity panel. Scale bar = 5 cm. RH 1-4: homozygous: *C1, C2, y*; RH 5-14: homozygous *C1, c2, y*; RH 15: homozygous: *c1, c2, y*..

Accession	RH 0		RH 1		RH 2		RH 4		RH 9		RH 10		RH 14		RH 15	
Ripening Stage	**MG**	**R**	**MG**	**R**	**MG**	**R**	**MG**	**R**	**MG**	**R**	**MG**	**R**	**MG**	**R**	**MG**	**R**
***Psy-1* Expression level (ΔΔCt ± SE)**	n.d.	258.04 ±91.71	n.d.	1519.9 ±583.10	n.d.	1016.66 ±120.30	n.d.	362.6 ±72.38	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Carotenoid Amount (µg/g DW ± SE)																
Phytoene	**293.04 ± 13.99**	**1030.76 ± 146.62**	**269.37 ± 4.81**	**1973.49 ± 60.06**	**240.91 ± 0.60**	**2203.32 ± 138.42**	**271.29 ± 7.57**	**1632.18 ± 54.34**	n.d.	n.d.	176.09 ± 2.23	n.d.	445.13 ± 55.84	n.d.	n.d.	n.d.
Lutein	505.14 ± 36.61	359.44 ± 74.28	406.49 ± 61.31	578.00 ± 30.79	197.34 ± 0.70	200.36 ± 17.24	200.99 ± 6.47	352.97 ± 2.57	239.23 ± 8.77	122.08 ± 2.42	432.11 ± 7.74	146.82 ± 0.46	228.13 ± 37.46	104.19 ± 0.80	126.02 ± 5.05	n.d.
β-carotene	595.58 ± 64.92	576.29 ± 45.64	248.41 ± 35.76	n.d.	333.34 ± 0.94	379.95 ± 28.54	**274.62 ± 16.93**	**584.24 ± 8.69**	284.87 ± 24.98	n.d.	204.76 ± 8.97	n.d.	90.10 ± 12.99	n.d.	98.60 ± 7.90	n.d.
Zeaxanthin	n.d.	251.53 ± 22.87	n.d.	100.12 ± 15.52	n.d.	86.89 ± 0.56	n.d.	n.d.	37.94 ± 7.58	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Violaxanthin	**115.07 ± 7.63**	**56.36 ± 1.94**	**96.57 ± 11.00**	**638.97 ± 82.42**	**97.80 ± 0.43**	**151.80 ± 5.67**	**106.39 ± 1.06**	**386.80 ± 87.56**	**52.39 ± 1.10**	**60.49 ± 1.74**	98.14 ± 1.46	47.39 ± 10.34	84.47 ± 7.36	55.28 ± 0.89	63.70 ± 2.11	
Antheraxanthin	n.d.	81.15 ± 5.31	**19.90 ± 1.60**	**62.57 ± 10.17**	17.40 ± 0.05	21.92 ± 0.26	n.d.	23.17 ± 0.44	n.d.	16.00 ± 0.14	n.d.	19.12 ± 1.05	n.d.	15.00 ± 0.01	n.d.	n.d.
Neoxanthin	**59.75 ± 2.03**	**27.72 ± 4.18**	50.18 ± 3.25	67.91 ± 9.22	**124.84 ± 0.17**	**119.40 ± 1.25**	58.73 ± 0.43	n.d.	**67.29 ± 0.62**	**36.00 ± 0.40**	41.91 ± 1.87	24.98 ± 0.64	**36.96 ± 0.16**	**35.86 ± 0.25**	38.42 ± 0.27	n.d.

Table 3.4. *Psy-1* expression levels and carotenoid amounts determined through HPLC-PDA. Bold text denotes values which are significantly different: p-value < 0.05 determined by Student's t-test (n=3 biological replicates).

Lines expressing *Psy-1* were darker coloured, and accumulated higher levels of pigment in ripe fruit. The lines RH 0 (red fruit control), RH 1, and RH 2 were the only ones to contain detectable levels of zeaxanthin in ripe fruit; the same is true of these lines with respect to β-carotene, with the exception of RH 1, where the compound could only be detected in mature green fruit. The line which produced white mature fruit (RH 15) did not contain any detectable pigment when ripe. As expected, the most abundant pigment in the red control line was esterified capsanthin, reaching amounts of over 2000 µg/g DW (Table 3.3, Table 3.4). This pigment was missing from yellow, orange and white accessions which did not possess a functional copy of *Ccs* (Table 3.3, Table 3.4).

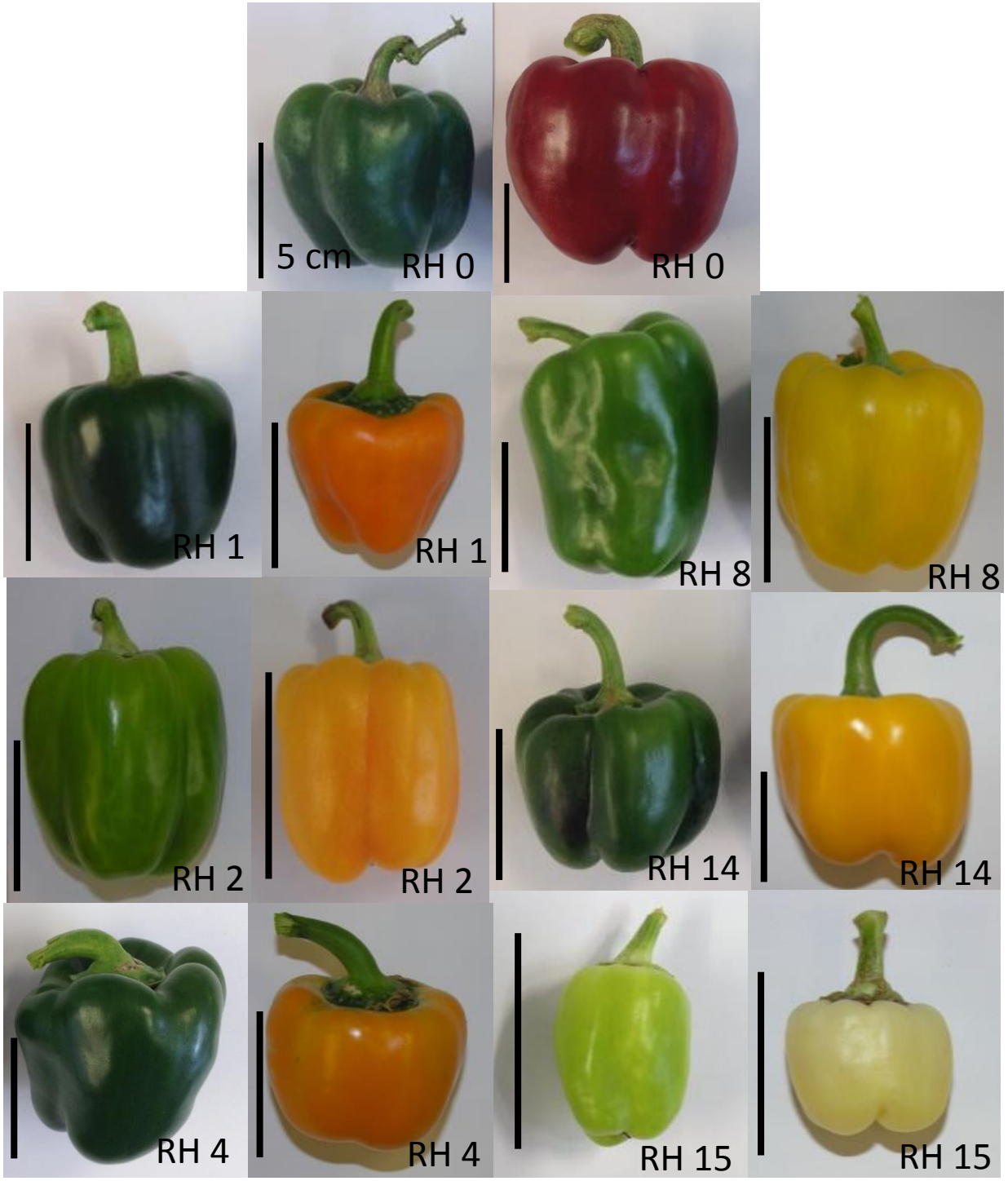


Figure 3.4. Mature green and ripe fruit from the accessions selected for further analysis. Genotypes of the accessions: RH 0: *C1, C2, Y*; RH 1-4: *C1, C2, y*; RH 8-14: *C1, c2, y;* RH 15: *c1, c2, y*. Scale bar = 5 cm.

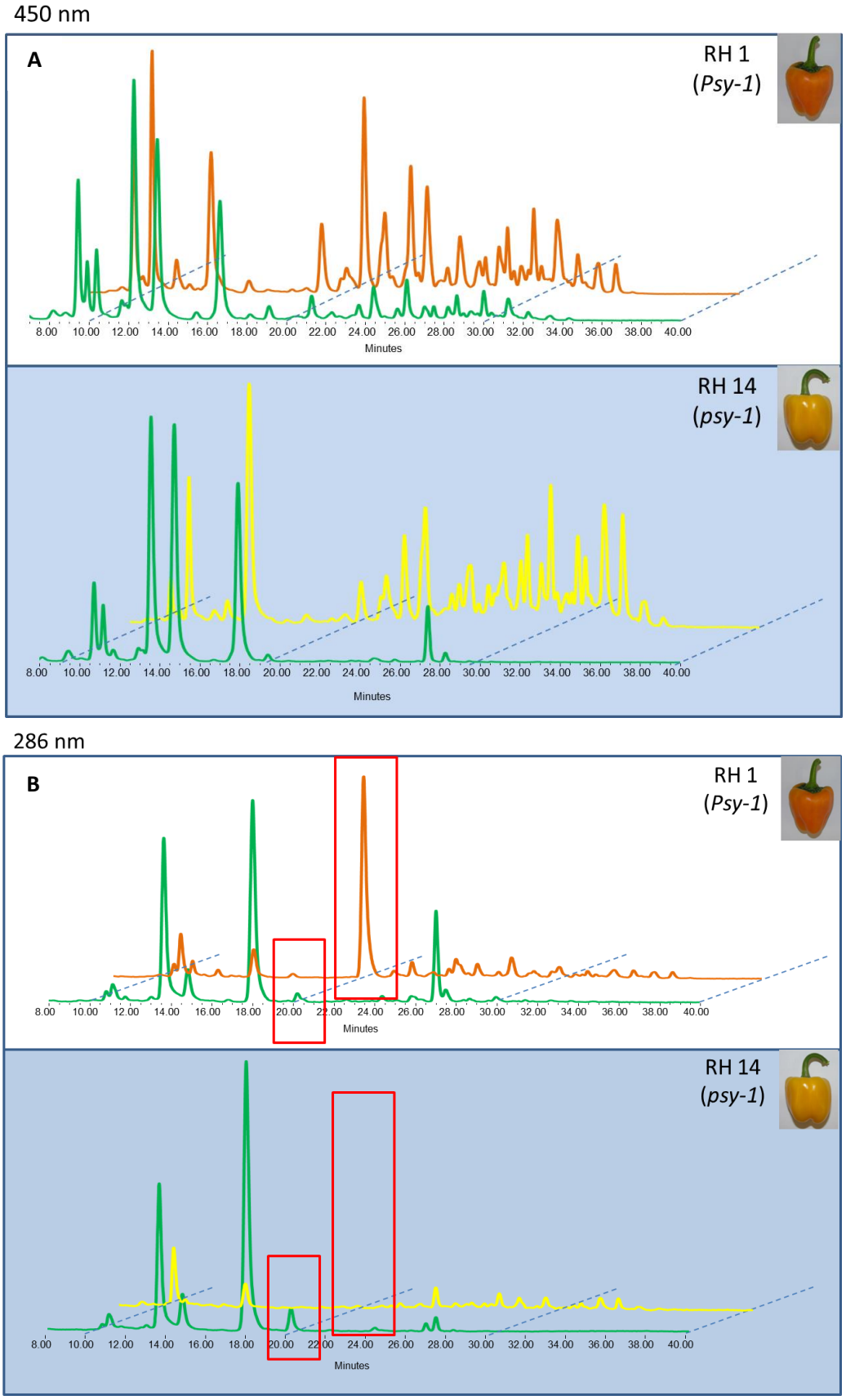


Figure 3.5. HPLC chromatograms from mature green and ripe fruit (lines RH 1 and RH 14). Green chromatogram: mature green fruit. Orange chromatogram: ripe fruit, RH 1. Yellow chromatogram: ripe fruit, RH 14. Red box: phytoene. A. Separation at 450 nm. B. Separation at 286 nm.

The chromatograms from ripe fruit are complex, featuring up to fifty peaks at 450 nm (Figure 3.1, Table 3.1). Violaxanthin, neoxanthin, antheraxanthin, lutein, zeaxanthin, β-carotene were identified in chromatograms from orange and yellow fruit and verified with authentic standards, as were capsanthin and capsorubin from red fruit. In addition the violaxanthin epoxidation products luteoxanthin and auroxanthin were putatively identified. These compounds were synthesised in house by acid treatment of violaxanthin; the resulting compounds had the same retention times and spectra as sample peaks. These sample peaks were then labelled as free luteoxanthin and auroxanthin (Table 3.1).

Chromatograms from mature green fruit contained free carotenoids: neoxanthin, violaxanthin, lutein, β-carotene, and phytoene as well as chlorophyll a and b. The major change between the two stages at 450 nm was the formation of esters over the course of ripening (Figure 3.5). At 286 nm, the most significant change was the massive increase in phytoene in fruit possessing functional *Psy*-1. This was not observed in fruit with the mutated version of the gene (Figure 3.5).

The complexity of the ripe fruit chromatogram is a result of the accumulation of mono- and di-esters, formed through the addition of fatty acids to hydroxyl groups on xanthophyll end rings. In order to simplify the profile by removing fatty acids from carotenoids, saponification and digestion with a cholesterol esterase were trialled (Figure 3.6). However, both approaches led to overall loss of pigment. The loss of detail about the samples was also deemed important: their removal would make the study of this method of pigment sequestration impossible. Therefore, peaks with the same characteristic spectra as free xanthophylls, but eluting later in the run were labelled mono- or di-esters (Figure 3.6).

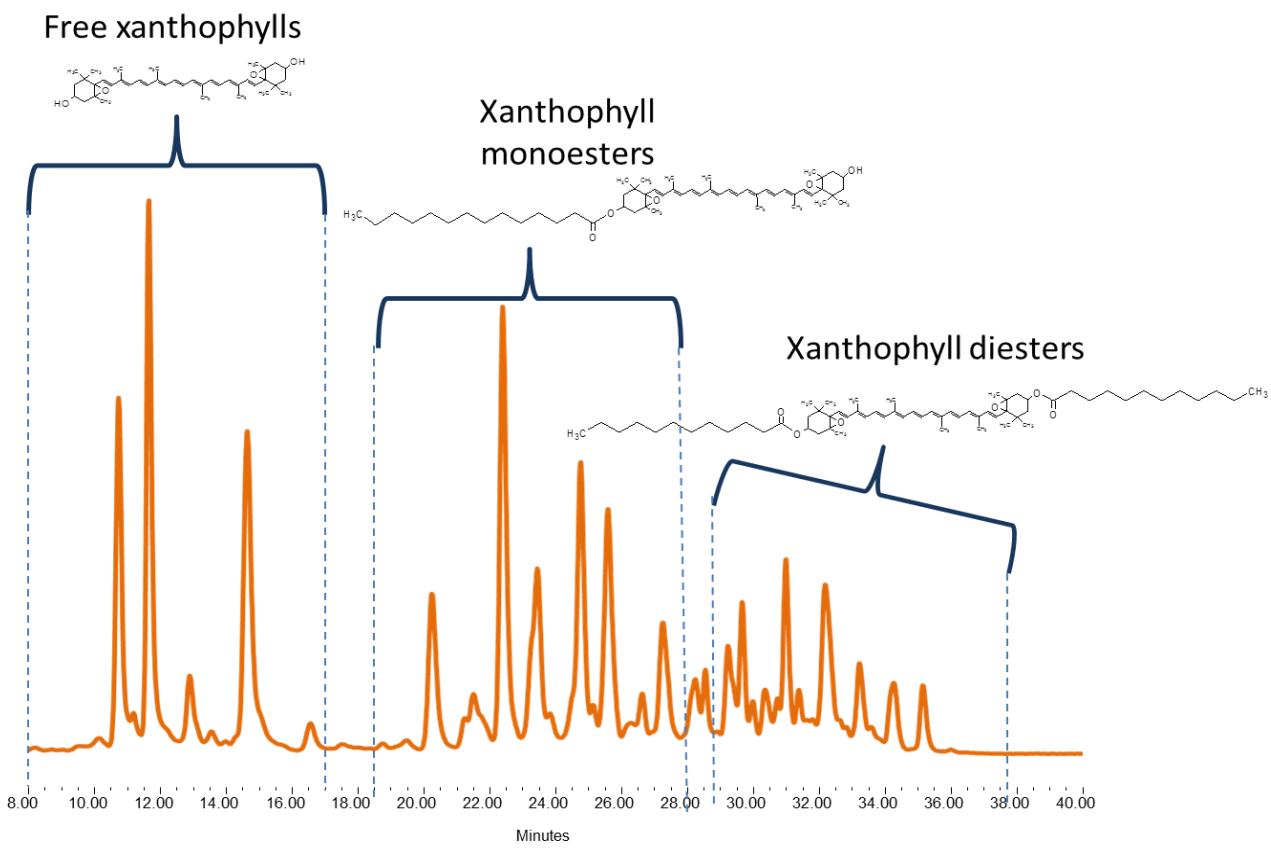


Figure 3.6. Ripe fruit chromatogram showing approximate elution times of carotenoids and their mono- and di-esters (450 nm).

Eight accessions from the original sixteen were chosen for further analysis, due to the fact that their carotenoid contents were representative of each genotype. Carotenoid and *Psy-1* gene expression levels were profiled from mature green fruit from these lines (Table 3.4).

Yellow fruited lines which did not express *Psy-1* contained similar carotenoid levels at mature green and ripe stages. A major change in carotenoid composition between mature green and ripe fruits is the esterified pigment content: for example, in RH 14 moving from one violaxanthin ester in mature green to eleven in ripe tissue. Esters were identified with LC-MS, with the most abundant across yellow and orange lines being violaxanthin ester 5.

3.2.2 Phytoene Synthase-1 Sequencing

It was necessary to determine the nature of the mutation in *Psy-1* present in accessions of this panel in order to identify the cause of lowered carotenoid levels in certain lines. This way, changes in gene expression level and protein level could be linked back to a genomic difference. A sequencing strategy was designed in order to test the presence or absence of the promoter and gene in accessions with varying carotenoid contents.

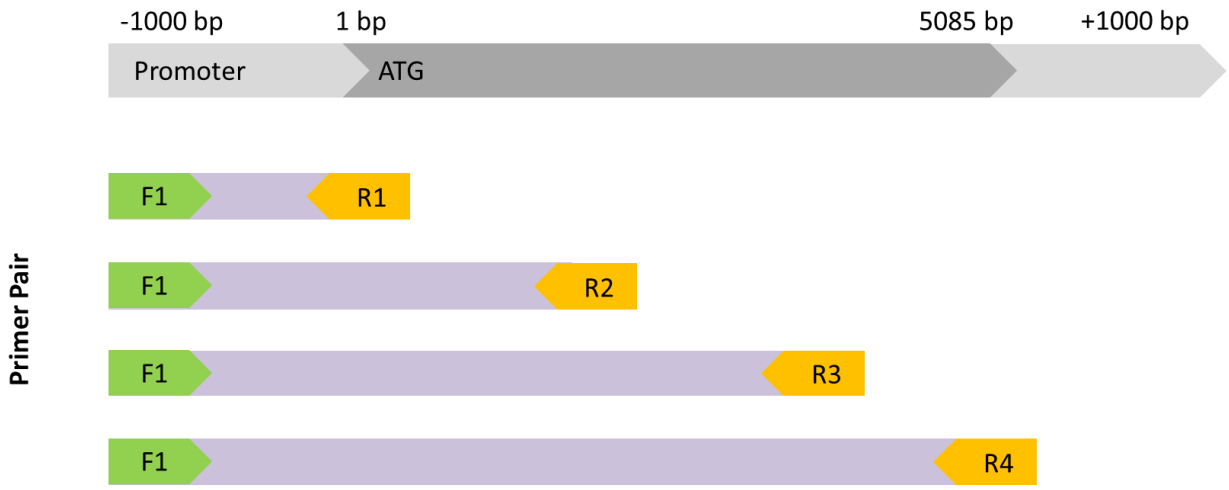


Figure 3.7. *Psy*-1 sequencing primer pair positions, and sizes of the resulting fragments. Primers were designed with Primer 3 Plus software, and tested on genomic DNA of accessions RH 0, RH 1, RH 2, RH 4, RH 9, RH 10, RH 14 and RH 15. Comparison to a 1 kb ladder confirmed fragment sizes.

Line	Primer Pair			
	F1 R1	F1 R2	F1 R3	F1 R4
RH 0	Y	Y	Y	Y
RH 1	Y	Y	Y	Y
RH 2	Y	Y	Y	Y
RH 4	Y	Y	Y	Y
RH 9	N	N	N	N
RH 10	N	N	N	N
RH 14	N	N	N	N
RH 15	N	N	N	N

Table 3.5. Amplification of products from primer pairs, amplified with PCR. Y = positive; N = negative. Genotypes of the tested accessions: RH 0: homozygous: *C1, C2, Y;* RH 1-4: *C1, C2, y*; RH 9 -14: *C1, c2, y*; RH 15 *c1, c2, y*.

Sequencing primers to genomic *Psy-1* sequence were designed, giving overlapping fragments covering over 3000 base pairs of the gene (Figure 3.7). Amplification of these fragments in lines

which did not possess the marker for a non-functional PSY-1 enzyme confirmed that the marker was associated with a large-scale deletion of the gene from genomic DNA (Table 3.5).

3.2.3 Expression levels of related genes

To investigate patterns of *Psy-1* expression throughout ripening, qPCR was carried out on representatively coloured orange, yellow and white accessions at four stages throughout the course of fruit development. The impact of the absence of *Psy-1* transcript on the expression of other genes in the carotenoid biosynthesis and preceding MEP pathways was also analysed.

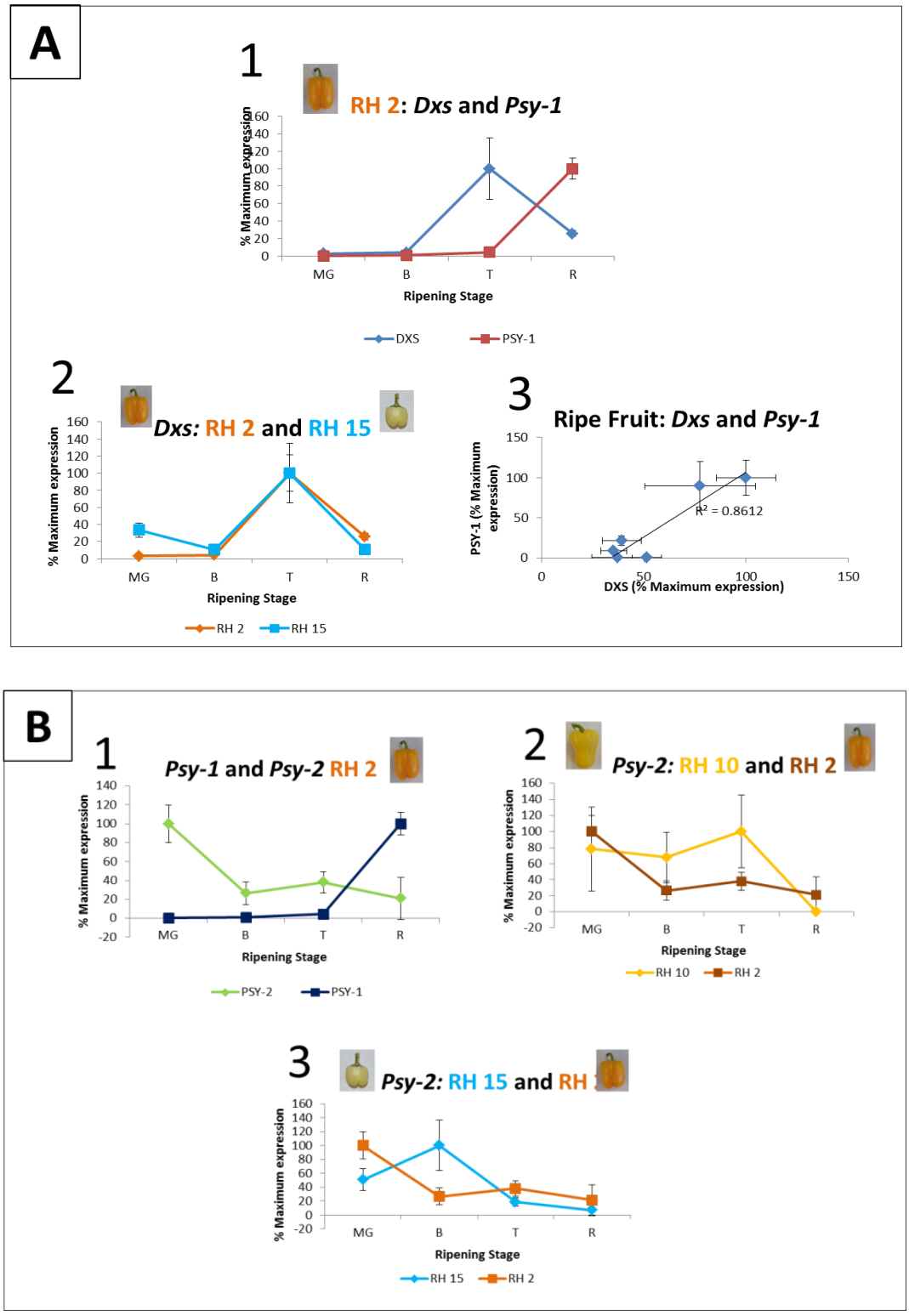


Figure 3.8. Transcript levels of selected genes over the course of ripening (mature green (MG), breaker (B), turning (T), and ripe (R)) in orange, yellow and white lines, determined with qRT-PCR. A. (1) *Psy-1* and *Dxs* expression in RH 2. (2) *Psy-1* and *Dxs* expression in RH 15 and RH 2. (3) *Psy-1* and *Dxs* expression in ripe fruit. B. (1) *Psy-1* and *Psy-2* expression levels in RH 2 (an orange, high phytoene line). (2) *Psy-2* expression levels in RH 10 and RH 2. (3) *Psy-2* expression levels in RH 15 and RH 2.

When the expression levels of the second phytoene synthase (*Psy-2*) (thought to be green tissue specific) was tested with qRT-PCR it became apparent that there was no significant difference between orange, yellow and white fruited lines; and that transcript levels were substantially higher in green tissue. However, transcript levels in non-*Psy-1* accessions were not significantly different to those in lines expressing *Psy-1*, at any of the ripening stages tested. Furthermore, the expression pattern of *Psy-2* was a mirror image of its fruit-specific paralogue: highest at mature green and decreasing as ripening progressed. *Psy-2* in pepper appears to have the same expression pattern observed for the gene in tomato: it is approximately 28-fold higher in leaf tissue than it is at any stage of fruit ripening (Figure 3.8).

The expression of genes in the MEP pathway, immediately preceding carotenoid biosynthesis, was also tested with qRT-PCR. Deoxyxylulose-5-phosphate (DXP) synthase (DXS) catalyses the first step of the pathway; producing DOXP from pyruvate and glyceraldehyde-3-phosphate. When its expression is tested over the course of ripening, it is apparent that it peaks in expression at the stage preceding the peak of *Psy-1* transcript level. It displayed the same pattern in lines expressing *Psy-1* as in lines which did not. Its expression was positively correlated with that of *Psy-1*, with an R^2 value of 0.8612 (Figure 3.8).

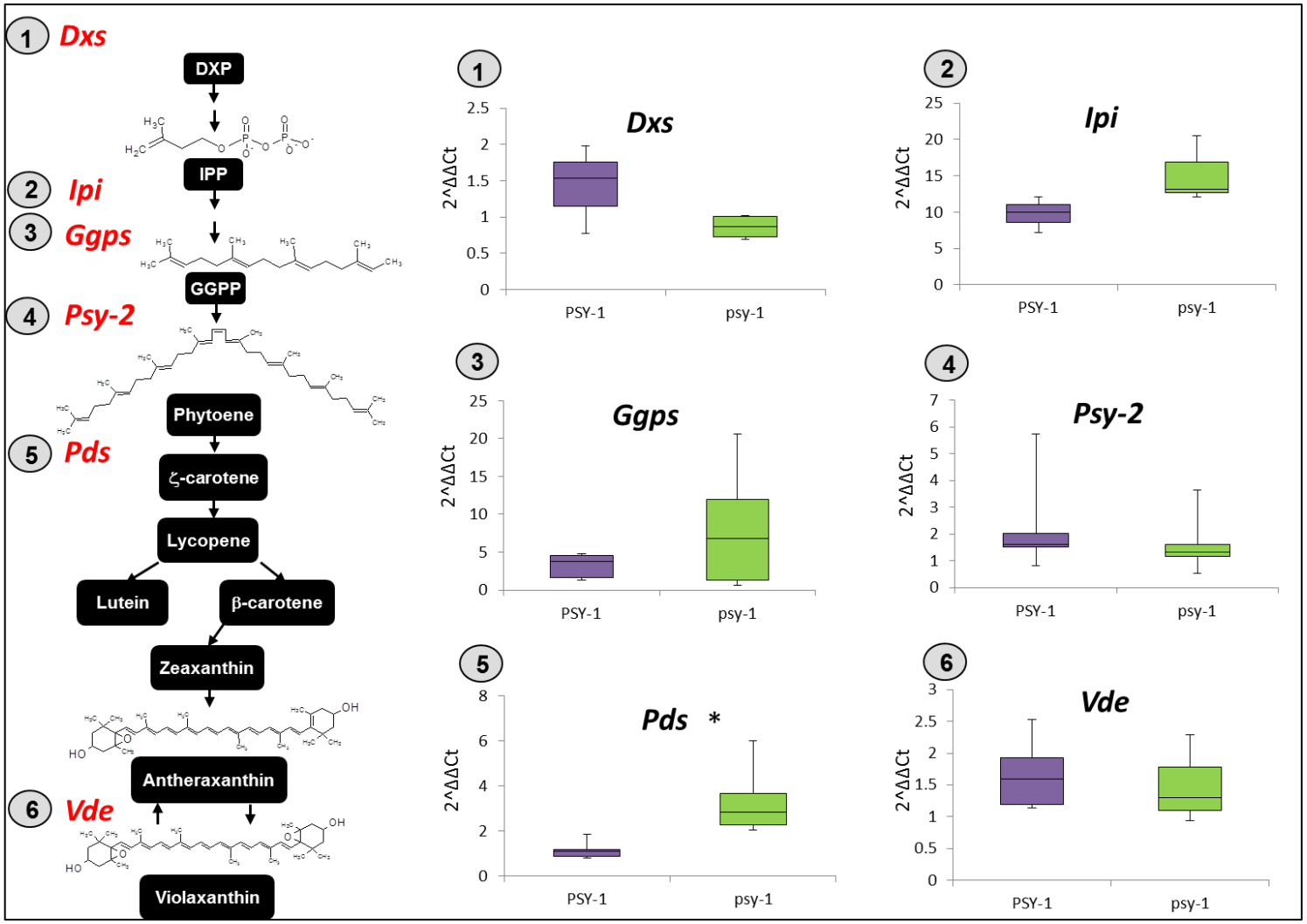


Figure 3.9. Transcript levels of MEP pathway and carotenoid biosynthetic genes in *Psy-1* expressing and non-expressing accessions, determined with qRT-PCR. Three biological replicates from each *Psy-1* expressing line were pooled and divided into three technical replicates (RH 0, RH 1, RH 2, and RH 4). The same was done for lines which did not express *Psy-1*. (RH 8, RH 14, and RH 15). Abbreviations: DXP: deoxyxylulose-5-phosphate; IPP: isopentenyl pyrophosphate; GGPP: geranylgeranyl pyrophosphate; *Dxs*: DXP synthase; *Ipi*: IPP isomerase; *Ggps*: GGPP synthase; *Psy-2*: Phytoene synthase 2; *Pds*: Phytoene desaturase; *Vde*: Violaxanthin de-epoxidase.

When differences in the expression levels of selected genes between lines which express *Psy-1* and those which do not are examined, it was revealed that levels of *Phytoene desaturase* (*Pds*) are significantly higher in accessions lacking *Psy-1* transcript (p=0.0114). *Dxs* expression was elevated in *Psy-1* expressing lines, albeit non-significantly; transcript levels of *Dxs* were approximately 10-fold lower than those of other MEP pathway genes tested. As violaxanthin was a major pigment determining the colour of these orange and yellow accessions, the transcript levels of *Violaxanthin de-epoxidase* (*Vde*), which catalyses its conversion to antheraxanthin in the xanthophyll cycle were tested. It was revealed that expression levels of *Vde* and *Psy-2* were not noticeably different in the two classes, and were low across all accessions tested (Figure 3.9).

3.3 Discussion

3.3.1 Pigment profiles in relation to *Psy-1* expression

Characteristic non-polar pepper HPLC chromatograms are complex, featuring free carotenes and xanthophylls, as well as xanthophyll mono- and di-esters. An interesting difference between yellow or orange and red accessions is the diversity of xanthophylls present in accessions without capsanthin and capsorubin. In addition to the well characterised xanthophylls zeaxanthin, violaxanthin, antheraxanthin, and neoxanthin, yellow and orange accessions profiled here contained the violaxanthin epoxidation products luteoxanthin and auroxanthin.

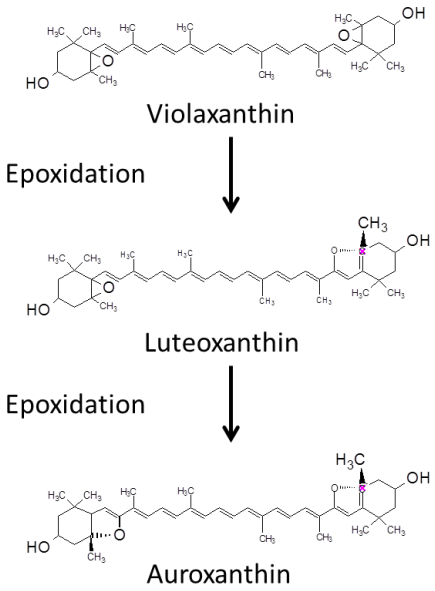


Figure 3.10. Conversion of violaxanthin to luteoxanthin and auroxanthin.

The formation of both has been associated with acidity and high light (Matus, Baranyai, Toth, & Szabolcs, 1981). According to Britton's Carotenoid Handbook, luteoxanthin was first isolated as a natural carotenoid from calamondin fruits in 1979, and has been identified in *Capsicum* spp. (Britton, Liaaen-Jensen, & Pfander, 2004; Gross, Eckhardt, Sonnen, & Lenz, 1979; Zoltan Matus, Deli, & Szabolcs, 1991) and several other crop species including sweet orange, mango, and apple (Dugo et al., 2008; Gross et al., 1979; Mercadante & Rodriguez-Amaya, 1998; Pott, Breithauptb, & Carle, 2003), and commercially grown flowers like *Narcissus* (Berset & Pfander, 1985). Britton (2004) states that auroxanthin was isolated as early as 1945, and, as would be expected, has been found in similar crops to luteoxanthin (mango, orange, loquat) and some additional species, for example banana (Karrer & Jucker, 1945; Subagio, Morita, & Sawada, 1996). Both have been found in their free and esterified forms.

Separation of carotenoids with HPLC showed that lines expressing *Psy-1* were darker coloured, and accumulated higher levels of pigment in ripe fruit. There were clear differences between carotenoid

amounts in mature green and ripe tissue in these accessions, whereas yellow *c2* mutant lines had similar pigment amounts at both stages. The white fruited line which does not retain pigments from mature green is lacking a functional copy of the gene associated with the *c1* locus, the identity of which remains undetermined.

Yellow fruited lines which did not express *Psy-1* contained similar carotenoid levels at mature green and ripe stages, suggesting that pigment remaining from green fruit is responsible for their colour phenotype. Violaxanthin and lutein are the most abundant pigment in ripe peppers of this panel, whether or not they possess a functional copy of *Psy-1*. This is to be expected of peppers which do not possess a functional copy of *Ccs*; violaxanthin is one of the substrates for the enzyme. In *Psy-1*-positive lines, the amount of total carotenoid increases dramatically over the course of ripening. In particular, the content of esterified pigments rises substantially. The number of different violaxanthin esters in line RH 4 (identified by retention time) increases from none at mature green, to thirteen in ripe tissue. The diversity and total amount of carotenoid esters in lines without a functioning PSY-1 enzyme also increases; however, total carotenoid amount does not. Yellow *c2* mutant lines could be analogous to the *r* mutant in tomato: their pale colour is a result of photosynthetic accessory carotenoids being retained throughout fruit development (Fray & Grierson, 1993; Ray et al., 1992).

Previous work has profiled carotenoids in *c2* mutant sweet pepper accessions: as early as 1970 it was recognised that in yellow-fruiting *Capsicum* varieties, ripening was not accompanied by a massive increase in carotenoid content (Davies, Matthews, & Kirk, 1970). In this study red peppers were reported as having a 100-fold increase in carotenoid level, whereas a 5-fold increase was observed for a yellow-fruited accession. However, a similar ripening increase was reported for an orange-fruiting variety, and genotypes for the lines are not characterised. In accordance with the current work, lutein and violaxanthin were seen to be the most abundant pigments in yellow peppers. Matus et al (1991) carried out a detailed analysis of the carotenoids in yellow pepper at six different ripening stages. In their ripe fruit chromatogram of 55 peaks, they identified 32 different compounds (accounting for more than 90% of the total carotenoid content). It was found that violaxanthin and antheraxanthin were the most abundant carotenoids in ripened fruit, accounting for 34% and 11% respectively. These were followed by lutein and zeaxanthin, which accounted for approximately 9% each. The Syngenta accessions presented here differed in their profiles; although violaxanthin was the most abundant pigment, lutein was much more abundant than antheraxanthin in all lines, and zeaxanthin was consistently at lower levels than lutein.

The presence of esterified pigments was one of the major differences between mature green and ripe fruit chromatograms, suggesting that this is an important mechanism employed by *Capsicum* fruits to store high levels of pigment as they ripen. However, it is interesting to observe that there seems to be no minimum pigment level required to initiate the formation of carotenoid esters: yellow fruits from lines lacking *Psy-1* transcript had just as many ester peaks as lines containing high levels of pigment (the red-fruited line RH 0 for example). Therefore, we may be observing an evolutionary overhang in yellow and orange lines: the pepper employs mechanisms for dealing with the extremely high carotenoid levels of ancestral fruit, even when levels are substantially lower.

Accessions show consistent pigment profiles across seasons and locations, showing that the Psy-1 marker is effective in a wide variety of contexts. Peppers grown in the field in Morocco showed similar pigment levels relative to each other to those grown under glass in the UK (Data in appendix; Section 9.2).

3.3.2 Deletion of *Psy-1* from the genome

PSY-1 is absent from peppers possessing the maker associated with this mutation.

Sequencing shows a deletion of several kilobases in lines possessing a marker associated with no functional PSY-1 enzyme. It has long been established that mature fruit colour in pepper is controlled by three independently segregating loci: *c1*, *c2*, and *y* (Hurtado-Hernandez & Smith, 1985).Two of the genes associated with these loci have been identified: *y* corresponds to *Capsanthin-capsorubin synthase* (*Ccs*) the gene catalysing the conversion of violaxanthin and antheraxanthin to the red ketocarotenoids capsanthin and capsorubin respectively. It has also been established that *Phytoene synthase-1* (*Psy-1*) was the gene associated with the *c2* locus (Huh et al., 2001; Thorup et al., 2000). Since then, several *Psy-1* mutations have been identified in *Capsicum*, however, no deletion of several kb of genomic DNA has been reported previously. Huh et al. looked at the F_2 progeny of a cross between an orange (*Capsicum chinense* cv Habanero) and a red cultivar (*Capsicum annuum* cv TF68). On finding that the plants in this generation almost perfectly adhered to the Mendelian ratio for a trait controlled by a single gene (25 out of 103 were orange, giving a ratio of 3:1), candidate genes from carotenoid biosynthesis and its precursor pathway were sequenced. Amplification of *Psy-1* genomic DNA and mRNA gave fragments of the same size from both orange Habanero and red TF68, however, Southern blotting with a *Psy-1* probe and a template of *Eco*RI-digested DNA revealed a polymorphism which segregated perfectly with orange and red fruit phenotypes (Huh et al., 2001). The *Psy-1* mutation identified by Huh et al. is of a different nature to that analysed in the present work: though the copy of the gene possessed by orange-

fruited Habanero plants did not produce a functional enzyme, no large-scale deletion of DNA from the genome was found to have occurred.

In addition to being unique in pepper, the mutation is novel compared to non-functioning *Psy-1* mutations in other crops – namely, *r* mutant, which still possesses part of the gene.

Mutations in *Psy-1* are also well studied in tomato, though no large-scale deletion of the gene has been reported to date in this crop. Perhaps the best characterised of these is the *r* mutant, which produces yellow mature fruit (Fray & Grierson, 1993). Unlike the pepper *Psy*-1 mutation examined here, the gene is transcribed by mutant tomatoes; however resulting transcripts are approximately 750 nucleotides shorter than wild type (850 base pairs in total). The mutants possess no functional PSY-1 protein, and therefore do not synthesise ripening-related carotenoids as they mature. Similarly to the pepper lines examined here, their yellow colour is conferred by carotenoids remaining from the mature green stage, and flavonoids in the pericarp skin (Fray & Grierson, 1993).

Research into control of fruit colour in pepper has focussed on genes encoding enzymes involved in the biosynthesis of carotenoids; however, the impact of polymorphisms in genes encoding degradative enzymes can be equally great. In addition to polymorphisms in *Psy-1* leading to diversity in carotenoid content and therefore colour, several commercially important crops possess mutations in other genes related to the biosynthesis and degradation of coloured compounds. For example, peaches are commercially available in various colours ranging from orange to white. Both yellow and white fleshed peaches are desired by consumers, however, yellow fruited varieties contain much higher carotenoid levels (particularly β-carotene and β-cryptoxanthin), and are therefore are considered more nutritionally valuable(Vizzotto, Cisneros-zevallos, & Byrne, 2007). Flesh colour in peach is controlled by a single locus (also known as *Y*); in contrast to pepper, white flesh is dominant over yellow (Bliss et al., 2002). A polymorphism in the peach *carotenoid-cleavage dioxygenase-4* (*CCD4*) has been associated with retention of carotenoid from yellow-fleshed varieties: a point mutation leads to a premature stop codon, and subsequent loss of the N-terminus of the protein. The resulting polypeptide is non-functional, resulting in peaches retaining higher carotenoid levels at the point of harvest (Falchi et al., 2013).

3.3.3 The second phytoene synthase in pepper does not contribute to ripe fruit carotenoid content

A second phytoene synthase enzyme in tomato was cloned, sequenced and mapped in 1993. The sequence was 83% identical to that of *Psy-1*, and included a highly similar partial transit peptide; revealing that the second phytoene synthase is also targeted to the plastid. To demonstrate that the gene encoded an active phytoene synthase, a complementation assay was carried out: transforming bacteria (*Rhodobacter capsulatus*) with a plasmid containing *Psy-2* cDNA was sufficient to

complement the crtB phenotype (lack of carotenoids due to a mutation in the bacterial *Psy*-1 homologue). Using restriction fragment length polymorphism (RFLP) analysis, the gene was mapped to chromosome 2 (Bartley & Scolnik, 1993).

The *Capsicum* homologue of *Psy-2* has been poorly characterised in the literature thus far, at both the transcript and protein level. Until recently, the gene had not been cloned, and its expression never tested quantitatively in *Capsicum*. Evidence for the existence of two phytoene synthase genes in pepper has been provided in the form of the appearance of two separate bands on a northern blot (Romer et al., 1993) and correspondingly two bands on a genomic southern blot (Thorup et al., 2000). However, as these investigations were not quantitative, the contribution of the second phytoene synthase to ripe fruit carotenoid content could not be determined.

Sequences of the two phytoene synthases in both tomato and pepper reveal an insert of 24 base pairs, approximately 170 base pairs from the start codon of *Psy-2*. This insert encodes a peptide region of eight amino acids. Comparative analysis of the promoters of the two phytoene synthases would reveal binding sites responsible for their different expression patterns over the course of ripening.

The role of the second phytoene synthase has been investigated in tomato. Through digestion with restriction enzymes and reverse transcriptase PCR, it was revealed that *Psy-1* transcripts were present in ripening fruit, whereas *Psy-2* transcripts were not observed at detectable levels. *Psy-1* and *Psy-2* cDNAs were digested with *AvaI* and *HindIII,* resulting in differentially sized fragments of 602 and 508 base pairs respectively. The 508 base pair fragment was only observed in leaf, and at very low levels in immature green fruit (Bartley & Scolnik, 1993). However, results of this experiment were only semi-quantitative, and did not give numerical values for transcript level making statistical analysis impossible.

The role of PSY-2 in the absence of a functional PSY-1 enzyme has been investigated in tomato; using the *r,r* and *Psy-1* antisense mutants in the context of *Psy-1* mutant yellow peppers (Bramley, Teulieres, Blain, & Schuch, 1992; Fray & Grierson, 1993). The *r,r* mutant is so named because it is homozygous recessive for the locus *R* (meaning 'Red', from its association with red tomatoes) which encodes the gene *Psy-1* (Fray & Grierson, 1993). It was reported that, though *Psy-2* transcripts and protein were detected throughout ripening in both *r,r* and *Psy-1* antisense fruit, the enzyme did not synthesise carotenoids over the course of ripening(Fraser, Kiano, Truesdale, Schuch, & Bramley, 1999). This is in concurrence with the results presented here: when lines do not possess a functional copy of *Psy-1*, carotenoid levels are similar in mature green and ripe fruit. It can therefore be concluded that the pepper PSY-2 enzyme does not contribute to mature fruit colour phenotype.

The role of the enzyme in pepper is incredibly similar: it transpired that *Psy-2* transcripts did not accumulate at higher levels in the absence of *Psy-1,* suggesting that the green-tissue specific copy of the gene does not take on a compensatory role. Here, a quantitative assay is performed which shows that the amount of *Psy-2* transcript in the fruit compared to the leaf is miniscule (it is at least 28-fold higher in the leaf than it is at any stage of fruit ripening), and that expression of *Psy-1* in the same tissue does not make a significant difference to this. It has been postulated that low levels of carotenoid in the *c2* mutant could be due to the action of the second phytoene synthase enzyme (Paran & Knaap, 2007), however, the gene was expressed in white fruit tissue (containing no detectable carotenoid) at similar levels to the yellow *c2* mutant, suggesting that the enzyme does not contribute to ripe fruit pigment level

3.3.4 Correlation between *Psy-1* and *Dxs* expression

When expression levels of other isoprenoid biosynthetic genes were tested over the course of ripening in accessions with and without functional copies of *Psy-1,* it was found that the gene encoding DOXP synthase (DXS), the first committed step of the MEP pathway (the plastid-specific pathway immediately preceding carotenoid biosynthesis), correlated strongly with expression of *Psy-1* (in accessions were *Psy-1* transcripts were detectable) (Figure 3.8). *Dxs* transcripts were much less abundant in lines which did not express *Psy-1*, suggesting that the products of carotenoid biosynthesis could be upregulating early stages of the preceding pathway, acting as a positive feedback loop. However, the exact identity and mechanism of the proposed signal remains undetermined. It has previously been suggested that a similar regulatory system exists in tomato.

The peak of expression for *Dxs* transcripts occurred at an earlier stage of ripening than that of *Psy-1*, due to the fact that the final product of the MEP pathway (GGPP) is the substrate for the first committed step in the carotenoid biosynthesis pathway. Therefore, if expression of *Dxs* is responding to that of *Psy-1*, it is sensitive to the low levels of transcript present at the early stages of ripening (Figure 3.8).

3.4 Conclusion

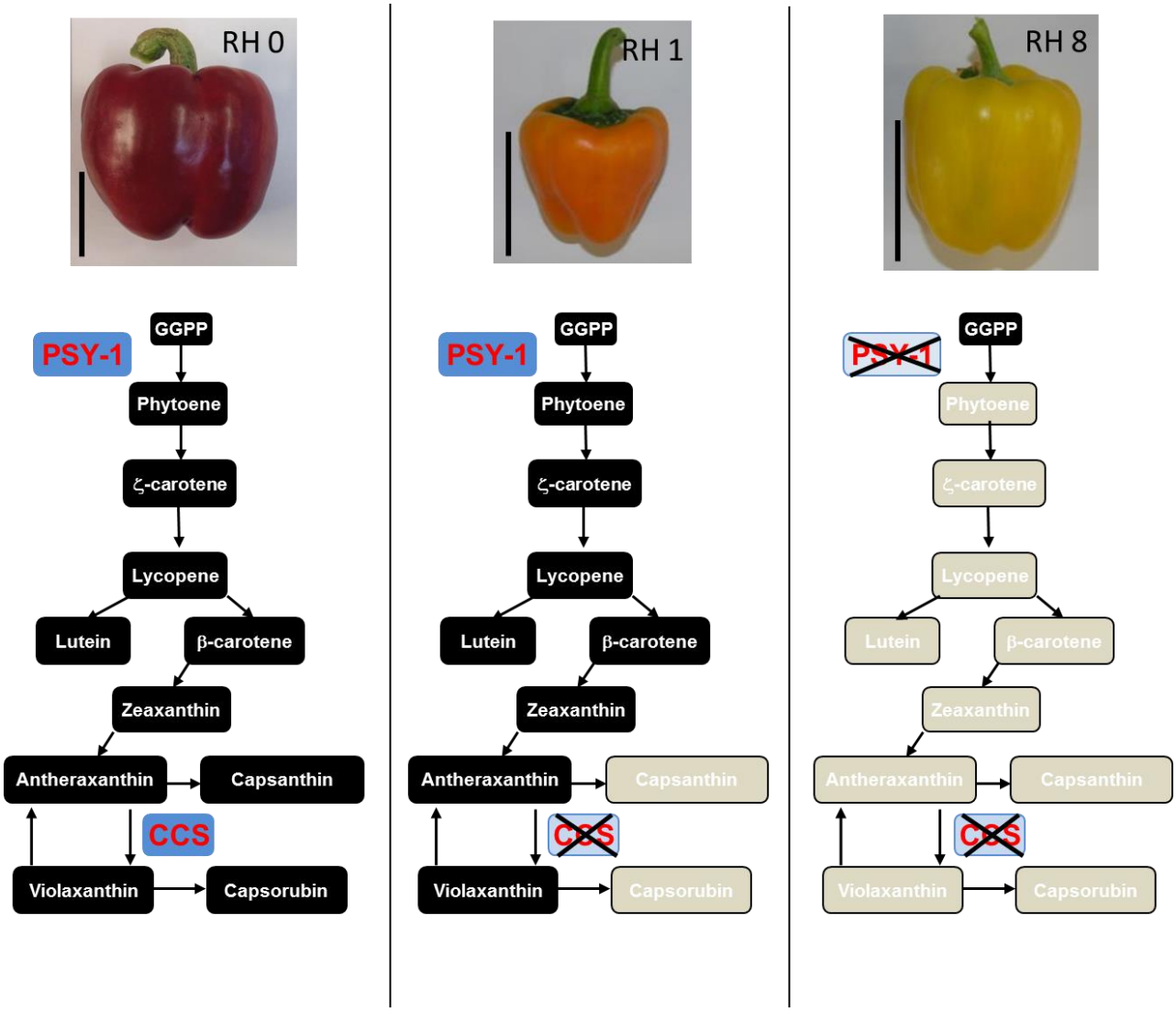


Figure 3.11. Scheme to show the impact of *Ccs* and *Psy*-1 deletions on carotenoids synthesised by ripening fruit. RH 0: functional CCS and PSY-1 enzymes; RH 1: functional PSY-1 enzyme only; RH 8: neither PSY-1 or CCS are functional.

To conclude; it is clear that a functioning fruit-specific phytoene synthase enzyme is necessary for intensely pigmented ripe fruit. It is clear that the presence of PSY-1 enzyme leads to dramatically increased pigment levels and subsequently colour in mature pepper fruit, as its expression level and enzyme activity both correlate with carotenoid content. Despite the large impact of PSY-1 on colour phenotype, there are other aspects which also contribute: these include plastid structure and broader fruit metabolism, as well as the *c1* locus which remains uncharacterised.

4 Impact of the *Psy-1* mutation on plastid structure; carotenoid sequestration and associated storage architecture; and the broader fruit metabolic network

–

4.1 Introduction

It is important to retain a sense of the pathway under investigation in the context of the entire metabolic network and its associated organellar architecture. Carotenoids are accumulated in the plastid; present at the mature green stage (before the fruit begin to ripen) as photosynthetic accessory pigments, and synthesised in high levels throughout the course of ripening as the chloroplast develops into a chromoplast. The chromoplast is a specialised plastid, well adapted for the accumulation of high levels of carotenoid (L. Li & Yuan, 2013). Therefore, the phytoene synthase (*Psy-1*) diversity panel provided a unique opportunity of studying the structure of the chromoplast in the absence of the enzyme catalysing the first committed step of carotenoid biosynthesis.

Several experimental techniques were employed in order to do this: firstly, transmission electron microscopy (TEM) was used to generate images of whole plastids and sub-plastid structures in mature green and ripe fruit in accessions from the collection with and without *Psy-1*. Secondly, chromoplasts from ripe fruit were isolated, broken, and suspended on a stepwise sucrose gradient. One millilitre fractions were collected from the gradient, from which carotenoids and proteins were collected for further analysis with high pressure liquid chromatography (HPLC), and Western blotting with antibodies to both carotenoid biosynthetic enzymes and structural proteins. This analysis enabled putative functional characterisation of chromoplast sections settling at different steps of the sucrose gradient, and comparison between plastids from the two conditions. In order to gain a better understanding of the sub-chromoplast location of carotenoid biosynthesis, a radiolabelled enzyme activity assay to PSY-1 was carried out.

In order to measure the impact of a lack of PSY-1 protein on the broader fruit metabolic network, gas chromatography-mass spectroscopy (GC-MS) was used to profile metabolites in a collection of accessions from the panel. These results were interpreted in the light of information provided by the plastid structure investigations.

4.2 Results

4.2.1 Whole Plastids Visualised with Transmission Electron Microscopy (TEM)

It was established in Chapter 3 that carotenoid content and resulting ripe fruit colour were impacted by the level of Psy-1 expression. Therefore, it is possible that the organelles and sub-organellar structures which carotenoids are stored would be altered accordingly. To gain an understanding of how a lack of *Psy-1* gene products would impact carotenoid storage architecture, accessions of representative colours were selected to have plastids from mature green and ripe fruit visualised with transmission electron microscopy (TEM). Chloroplasts and chromoplasts from RH 2, RH 10, and RH 15 were imaged.

4.2.1.1 Chloroplasts: mature green fruit (RH 2)

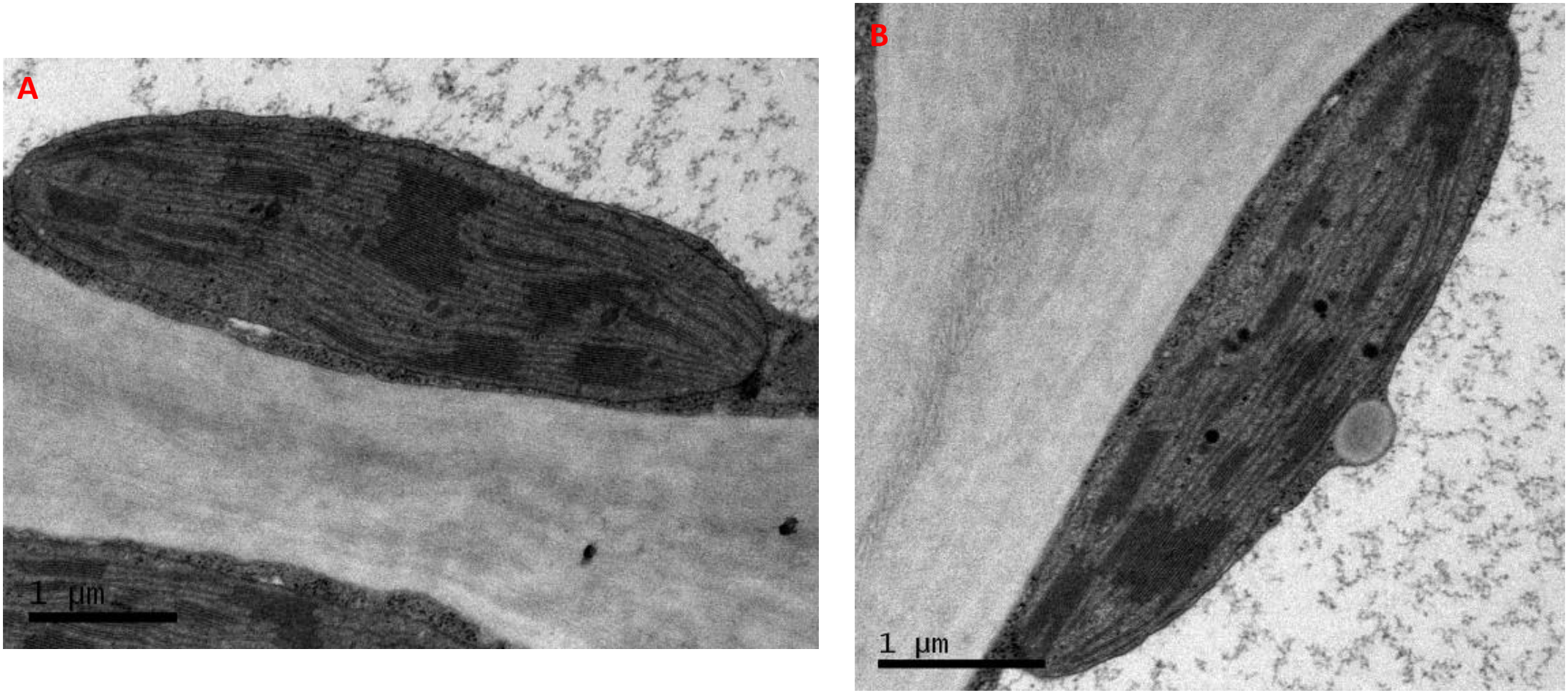


Figure 4.1. Whole chloroplasts visualised with TEM. Fruit pericarp was sectioned, fixed and counterstained before visualisation (mature green fruit, RH 2, n=3 biological replicates).

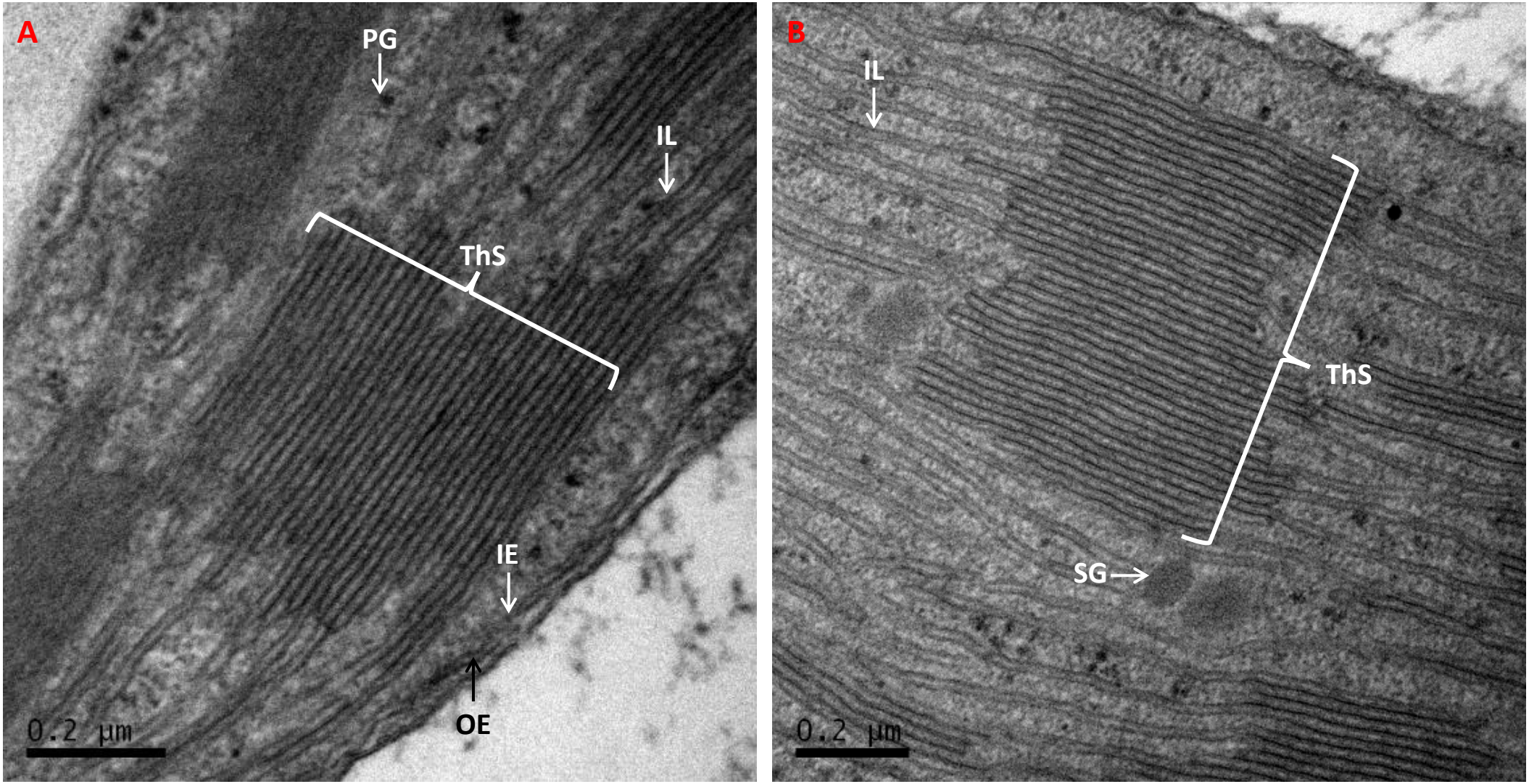


Figure 4.2. Chloroplast sections visualised with TEM. Fruit pericarp was sectioned, fixed and counterstained before visualisation (mature green fruit, RH 2). Abbreviations: ThS: thylakoids making up a granal stack. IL: intergranal lamella. PG: plastoglobule. IE: chloroplast inner envelope. OE: chloroplast outer envelope. SG: starch granule.

4.2.1.2 Chloroplasts: mature green fruit (RH 10)

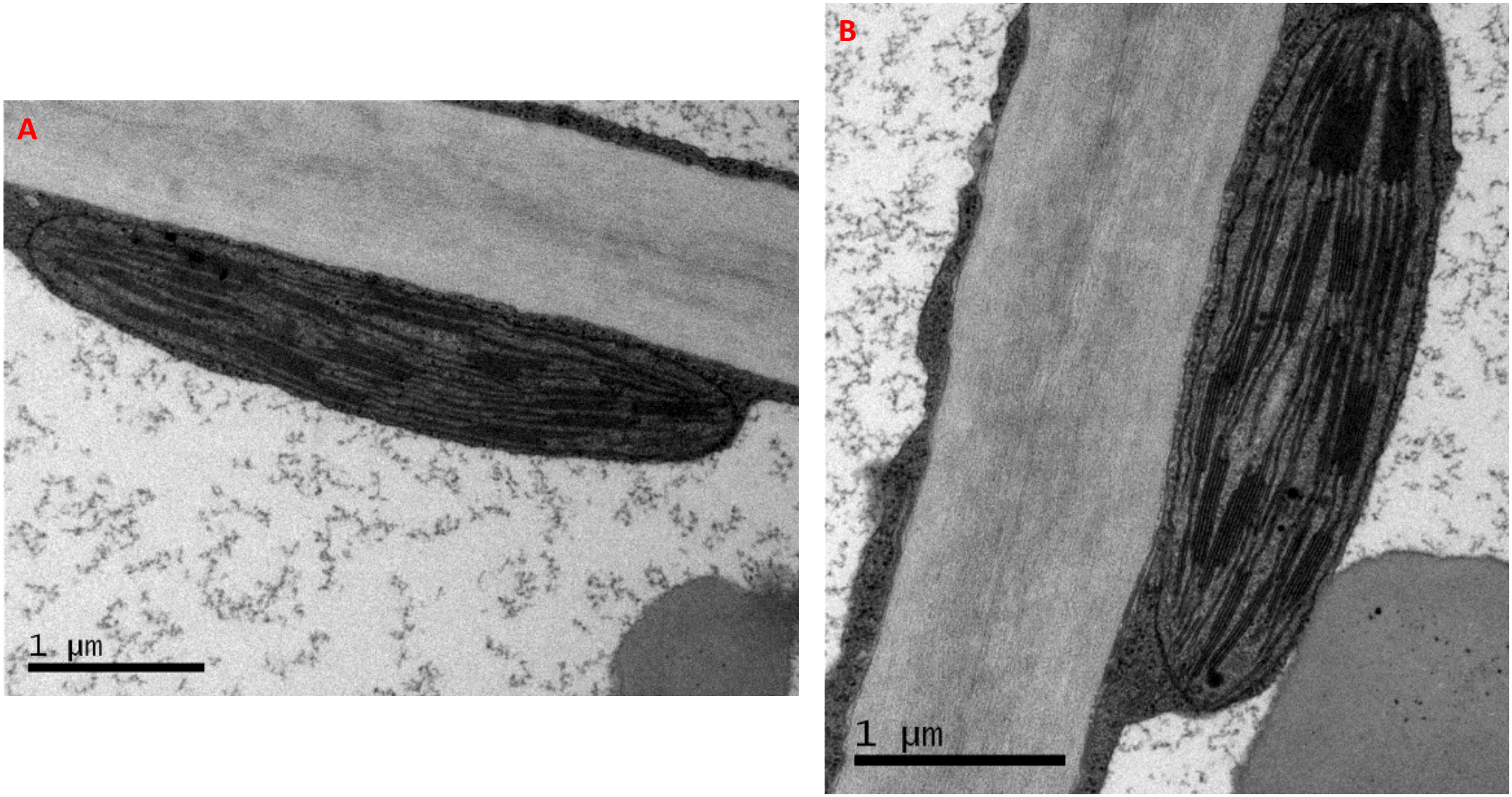


Figure 4.3. Whole chloroplasts visualised with TEM. Fruit pericarp was sectioned, fixed and counterstained before visualisation (mature green fruit, RH 10; n = 3 biological replicates)

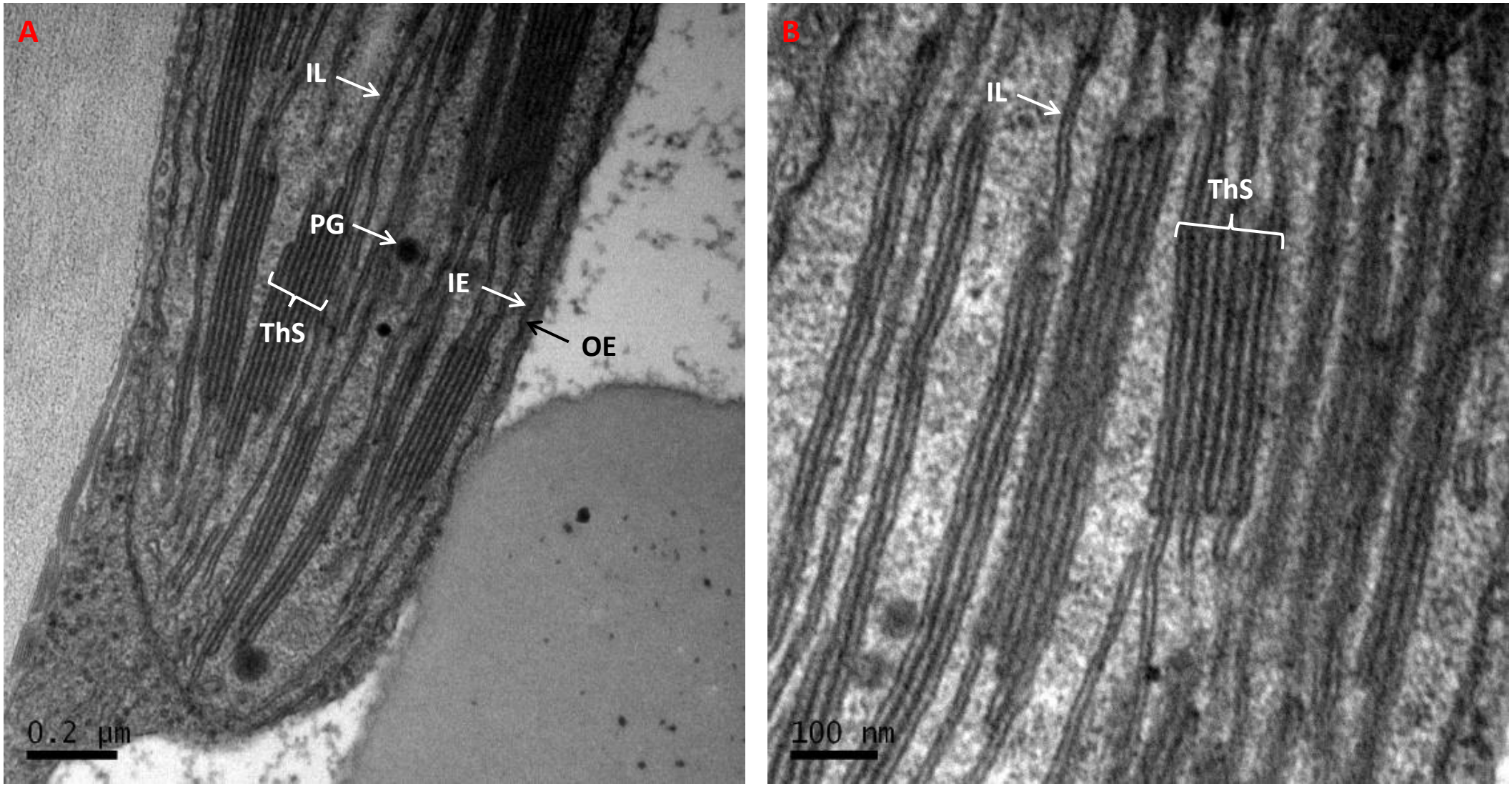


Figure 4.4. Chloroplast sections visualised with TEM. Fruit pericarp was sectioned, fixed and counterstained before visualisation (mature green fruit, RH 10). Abbreviations: ThS: thylakoids making up a granal stack; IL: intergranal lamella; PG: plastoglobule; IE: chloroplast inner envelope; OE: chloroplast outer envelope.

Transmission electron microscopy (TEM) images of chloroplasts from mature green fruit from RH 2 (possessing a functional copy of *Psy-1*) and RH 10 (no *Psy*-1 transcripts detected) revealed typical structures for both lines. Narrow, linear thylakoid sacs can be observed; these are stacked in grana containing between seven and ten sacs (RH 10), and between twenty and thirty sacs (RH 2). The number of thylakoids per granum corresponds to the darkness of the mature green fruit phenotype, and correlates with chlorophyll content. Intergranal lamellae can be observed: these are the long, linear, thylakoid sacs connecting separate grana together. Small spherical plastoglobuli can also be seen in chloroplasts of both lines (Figure 4.1-Figure 4.4).

4.2.1.2.1 Carotenoid Content in RH 2 and RH 10

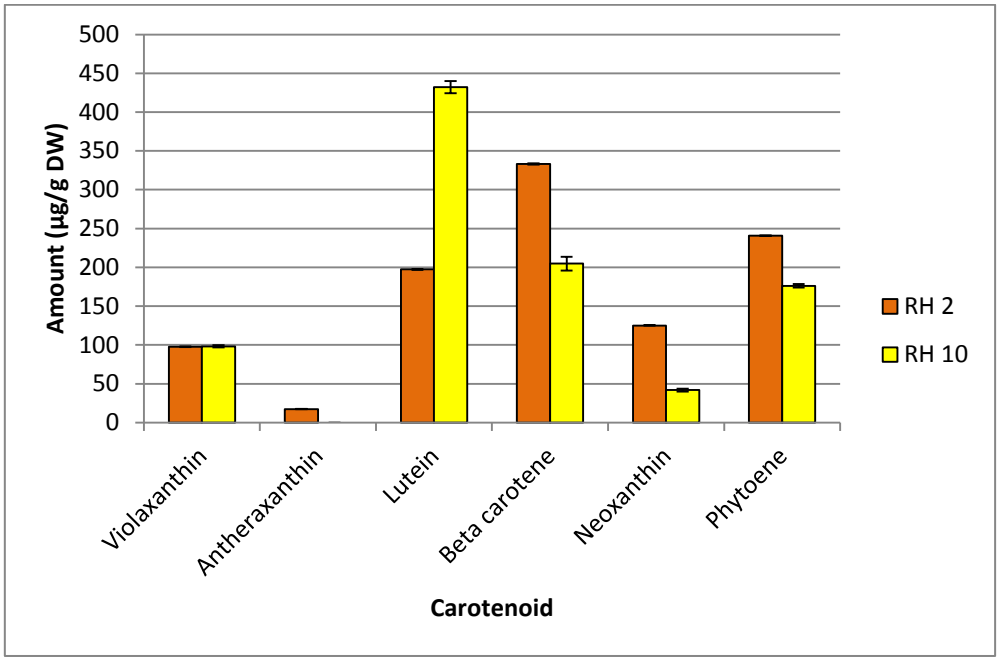


Figure 4.5. Carotenoids in mature green fruit of lines RH 2 and RH 10 determined with HPLC (n=3 biological replicates).

Carotenoids are differentially associated with the two light harvesting complexes (LHCs) as photosynthetic accessory pigments. β-carotene is preferentially associated with LHC I, whereas the xanthophylls are found in association with LHC II; the carotenoid composition of the two lines could affect the efficiency of light transfer to the photosystems in each one. The higher level of β-carotene in RH 2 would give this line a more effective LHC I, while increased amounts of lutein in RH 10 would favour LHC I (Figure 4.5).

4.2.1.3 Chromoplasts: ripe fruit (RH 2)

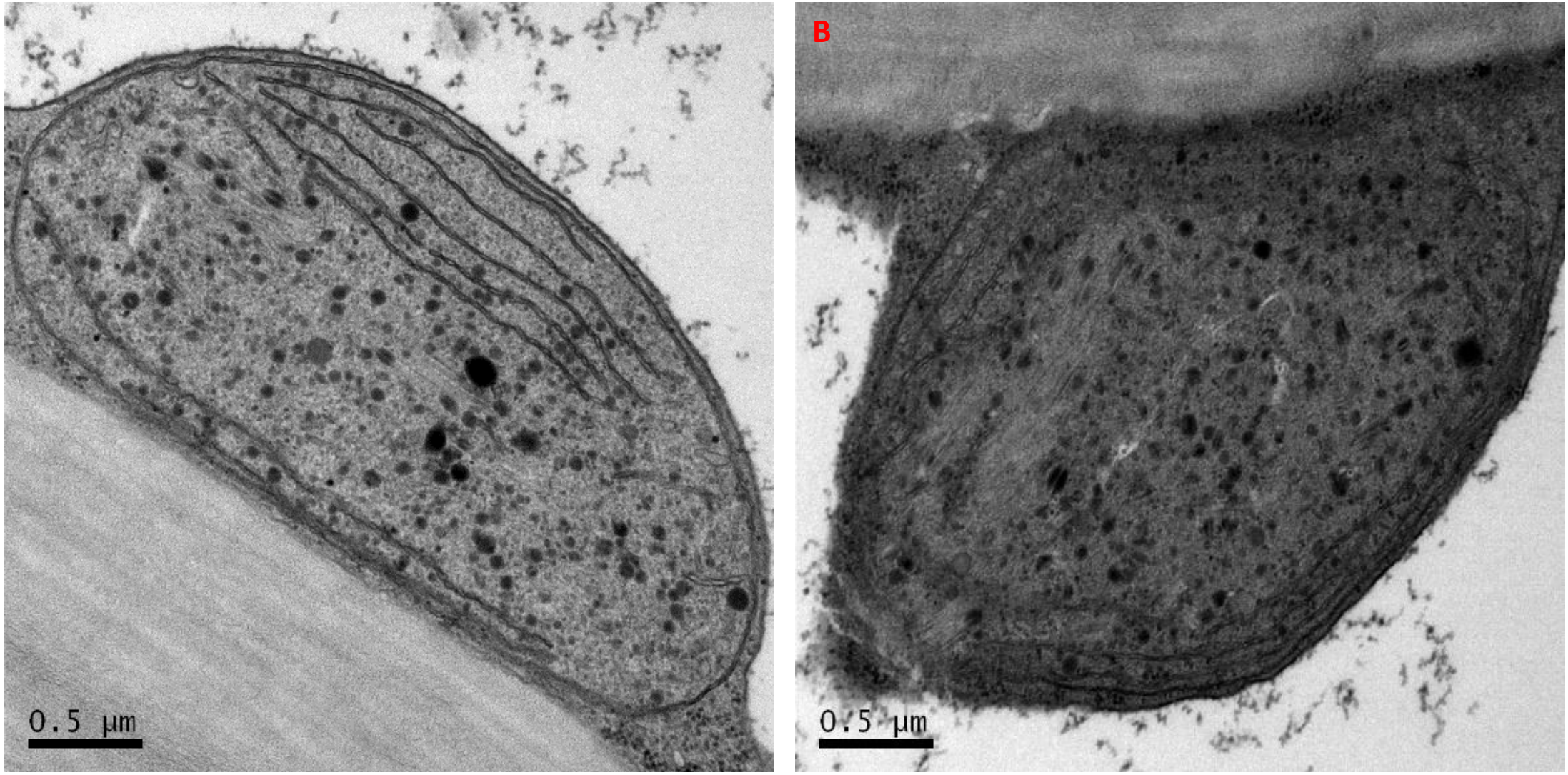


Figure 4.6. Whole chromoplasts from ripe fruit (RH 2), visualised with TEM. Fruit pericarp was sectioned, fixed and counterstained before visualisation (n = 3 biological replicates).

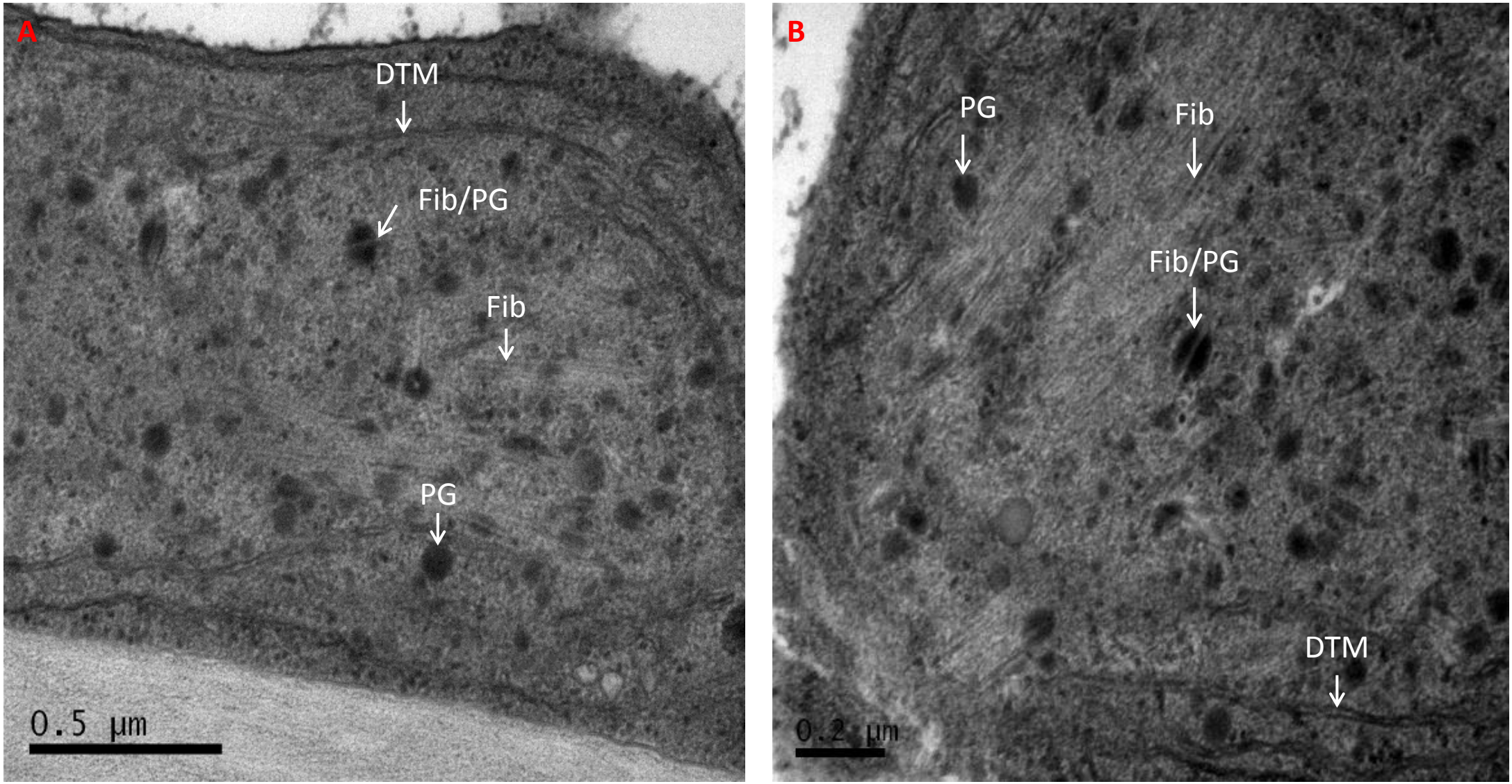


Figure 4.7. Chromoplast sections showing fibrils, plastoglobuli, and decomposing thylakoid membranes. Abbreviations: Fib: fibrils; PG: plastoglobuli; Fib/PG: fibrils intersecting plastoglobuli; DTM: decomposing thylakoid membranes. Genotypes: RH 2: homozygous: *C1, C2, y*; RH 10: *C1, c2 y*.

4.2.1.4 Chromoplasts: ripe fruit (RH 10)

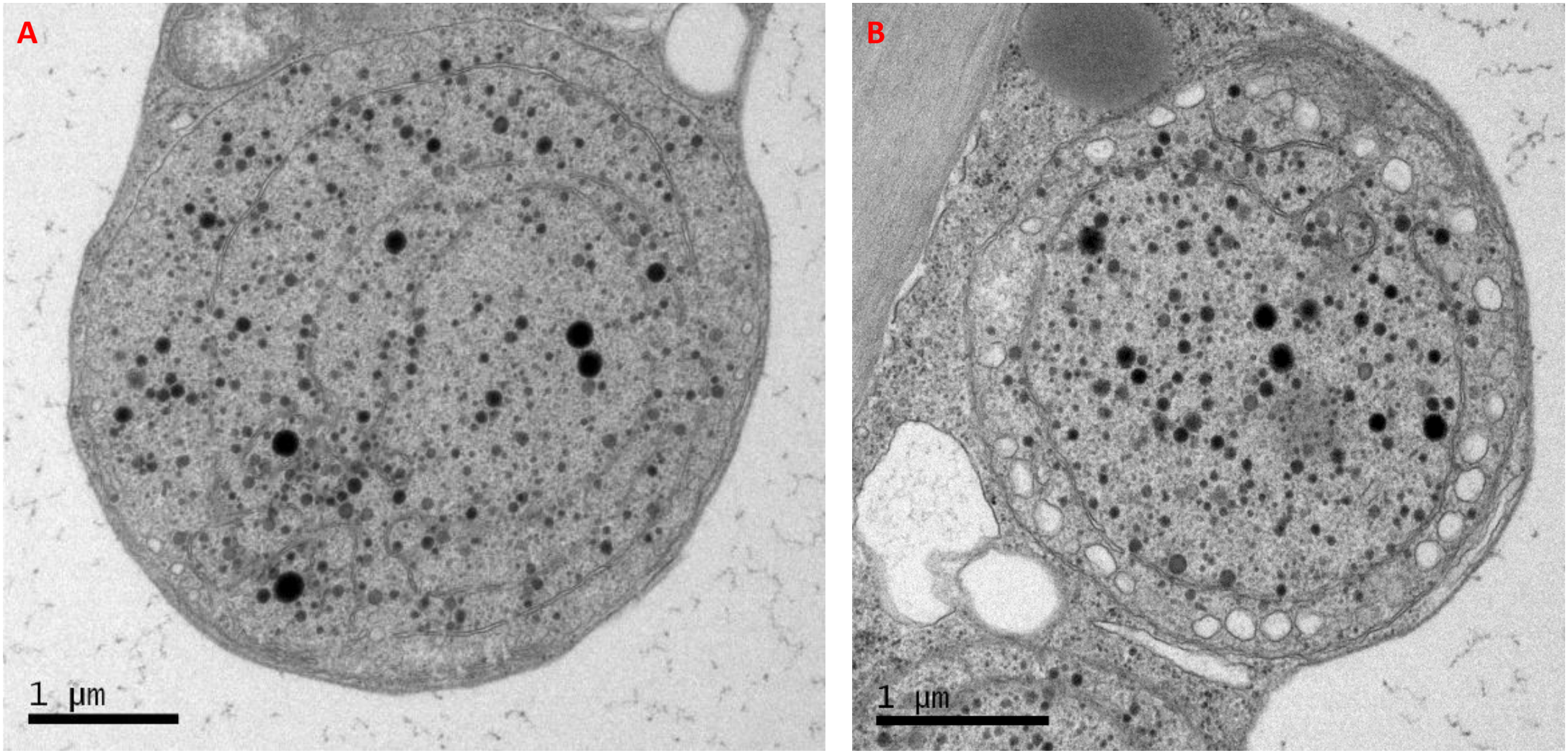


Figure 4.8 Whole chromoplasts from ripe fruit (RH 10), visualised with TEM. Fruit pericarp was sectioned, fixed and counterstained before visualisation. Sections were taken from three fruits and the plastids shown are representative.

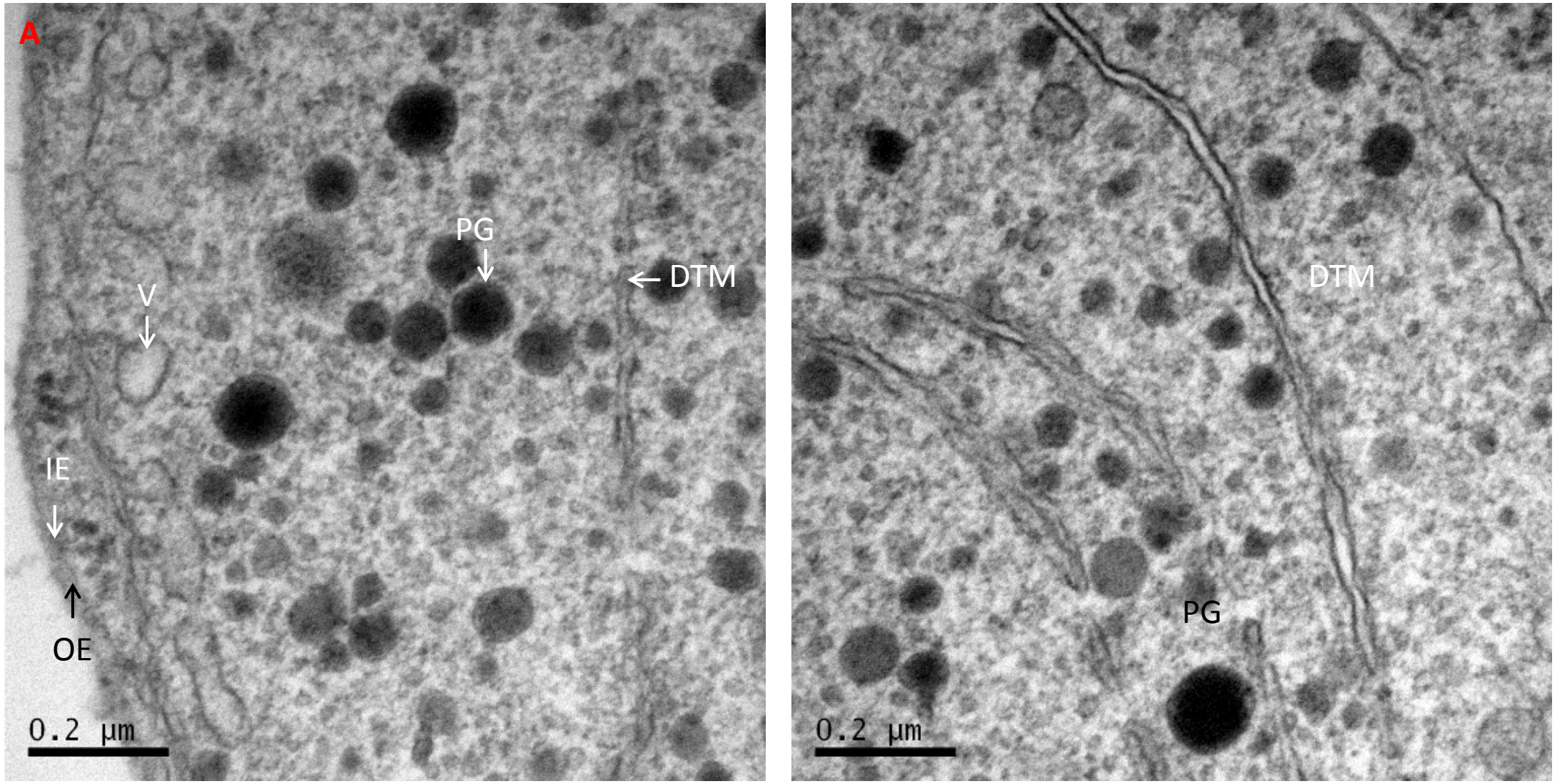


Figure 4.9. Chromoplast sections showing plastoglobuli and decomposing thylakoid membranes. PG: plastoglobuli; DTM: decomposing thylakoid membranes; V: vesicle budding from the thylakoid membrane.

The ripe fruit chromoplasts visualised with electron microscopy exhibit several organelle-specific features: thylakoids from mature green chloroplasts can be seen decomposing; lipid-membrane derived plastoglobuli are visible; as are the long, linear fibrillous proteins. Vesicles budding from the chromoplast inner envelope and the decomposing thylakoids can also be seen. There are clear differences between the two lines; most notably, the high density of fibrils in RH 2 is not seen in any of the chromoplasts from RH 10 (Figure 4.6-Figure 4.9). The differing carotenoid contents of the lines could account for some of the differences in plastid structure; for example, RH 2 contains a large amount of phytoene, while RH 10 contains none (Figure 4.10).

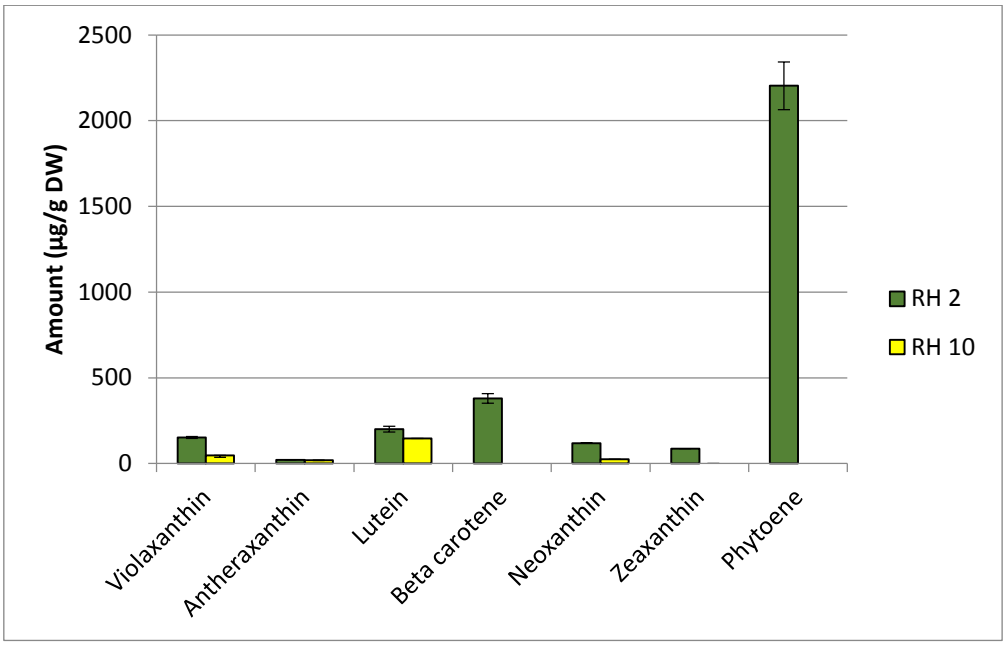


Figure 4.10 Carotenoid amounts in ripe fruit of lines RH 2 and RH 10 determined with HPLC (n=3 biological replicates).

.

4.2.2 Sub-chromoplast Fractionation

To further examine the differences in sub-chromoplast architecture between accessions of different genotype, membrane types within the plastid were separated via sub-chromoplast fractionation. Ripe fruit chromoplasts were isolated, broken, and suspended on a stepwise sucrose gradient. Carotenoids and proteins were then extracted for analysis with HPLC and western blotting.

Lines RH 0 and RH 8 were selected: RH 0 is a red line possessing functional copies of all three loci in the colour inheritance model (*c1* (unknown), *c2* (*Psy-1*), and *y* (*Ccs*)), RH 8 is a yellow line which only possesses a functional copy of the *c1* locus.

The two accessions showed different patterns of banding on the gradient (Figure 4.11). Both showed a coloured band at the top of the gradient, followed by colourless space until the 15% sucrose step. RH 8 showed diffuse aggregates in the next two steps of the gradient, and a thick band of concentrated colour in the 38% sucrose gradient.

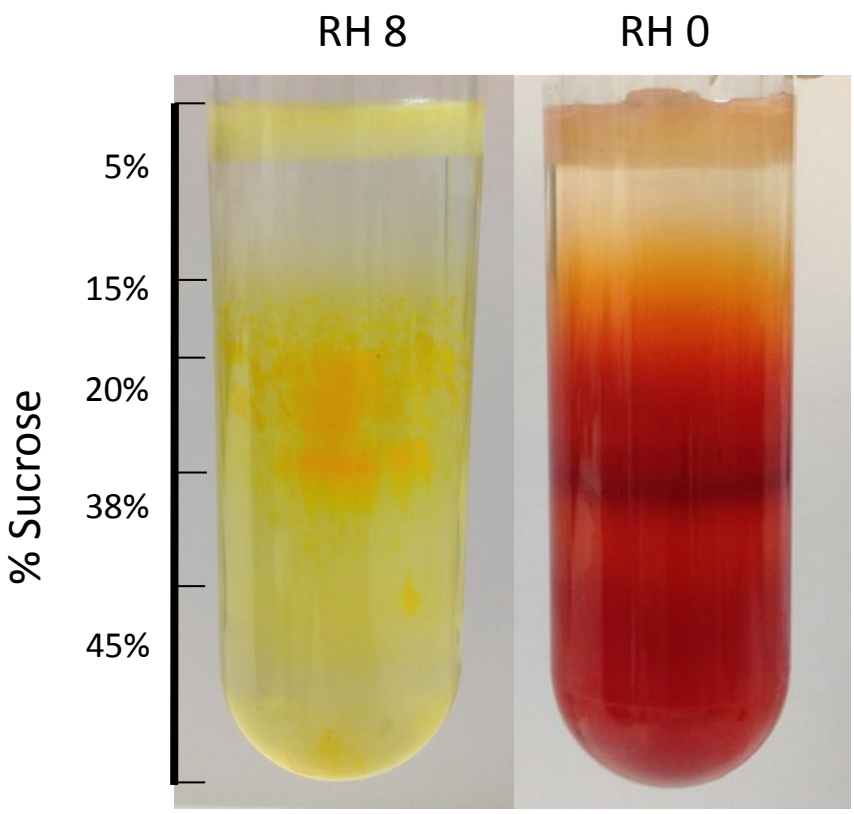
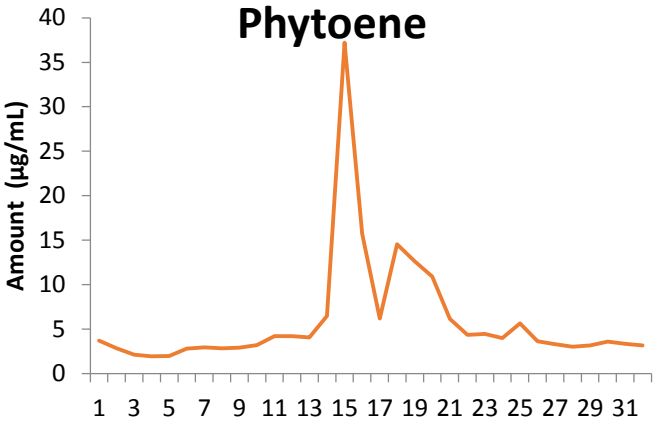
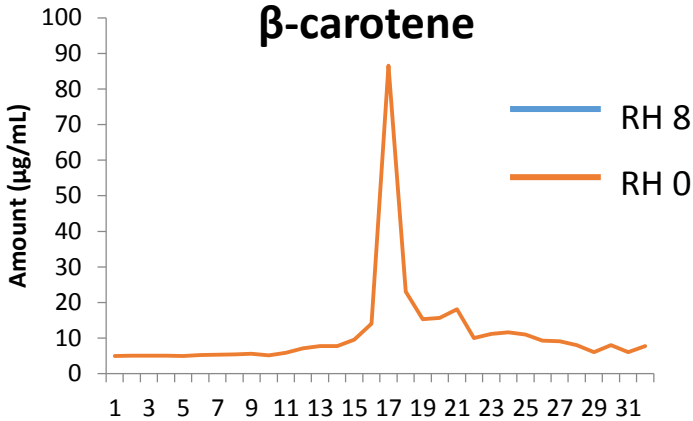


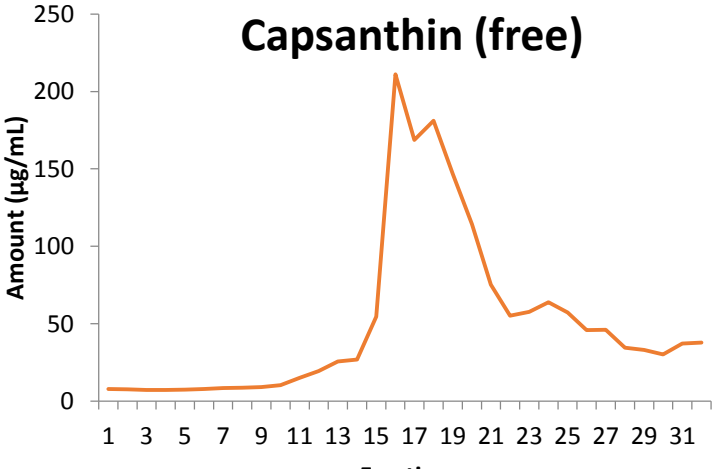
Figure 4.11. Isolated, broken chromoplasts from RH 8 and RH 0, suspended on a stepwise sucrose gradient.

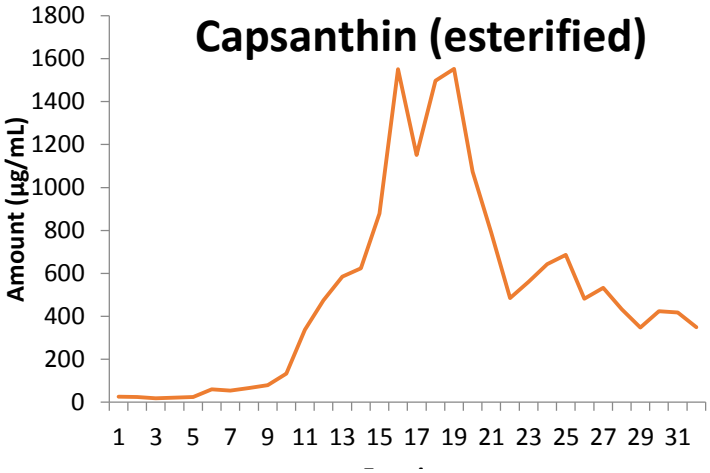
4.2.2.1 Pigment Profiles

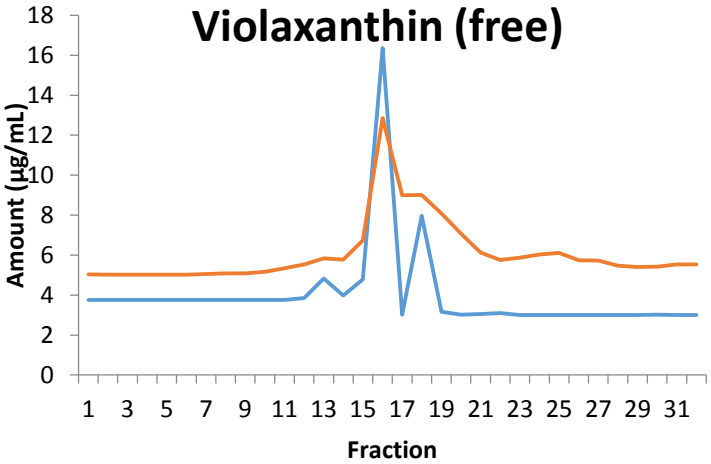
Carotenoids were extracted from each one millilitre fraction collected from the stepwise sucrose gradient and subjected to HPLC-PDA.

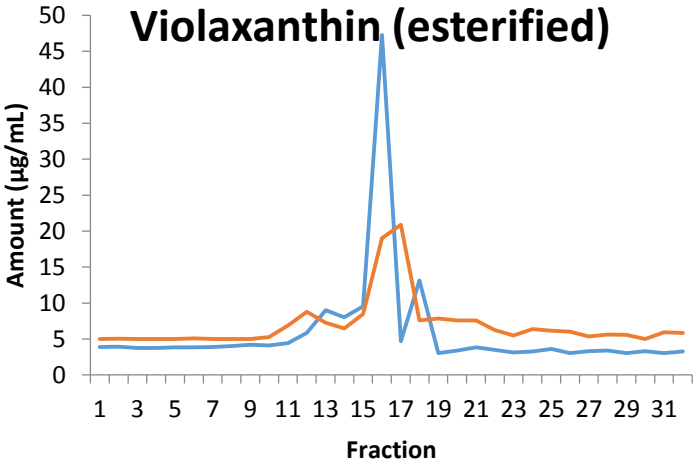












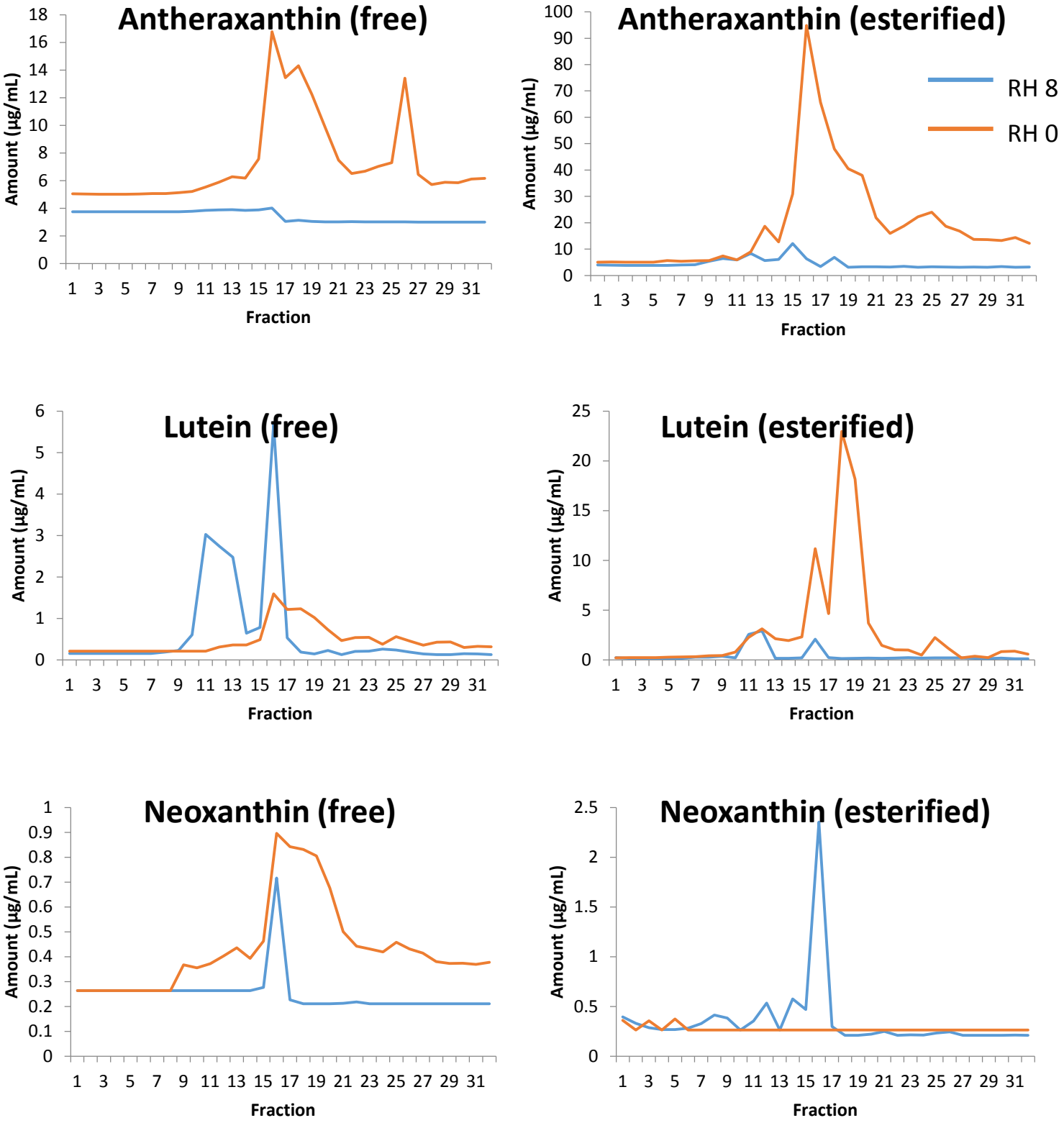


Figure 4.12. Carotenoid amounts from sub-chromoplast fractions of a *Ccs* and *Psy-1* positive line (RH 0), and a *Ccs* and *Psy-1* negative line (RH 8) determined by HPLC with a C30 column.

The pigment amounts across the fractionation gradient reveal striking differences between the two lines. Both exhibit a sharp peak in carotenoid amount in the middle of the gradient, however, individual pigments have separate peaks. There are also noticeable changes based on esterification. Phytoene was not detected in line RH 8, as would be expected from the lack of PSY-1 enzyme in this line. It was present in RH 0 and peaked in amount in fraction 15, reaching 37.2 µg/mL. In addition to the major spike there is a lower but wider peak in phytoene amount in fractions 18-21, with amounts reaching 10.9-14.6 µg/mL. Although amounts are low compared to the highest level fractions, phytoene is also present in the first fractions (1-4); at levels varying between 3.7-1.9 µg per mL fraction (Figure 4.12).

Another compound detected only in RH 0 samples was β-carotene, which reached a sharp peak in amount slightly lower in the gradient than phytoene, in fraction 17 (86.5 µg/mL). There is also a second peak in amount for β-carotene in fractions 19-21, though is this not high; reaching only 18.07 µg/mL.

The red pigment capsanthin was exclusively found in RH 0; the free carotenoid reached a peak of 211.2 µg/mL in fraction 16, with a close secondary peak of 181.1 µg/mL in fraction 18. As the final product of carotenoid biosynthesis in fruit possessing *Psy-1* and *Ccs*, esterified capsanthin was present at much higher levels than any other carotenoid: its peak is in a lower fraction than that of free capsanthin, 1553.0 µg/mL in fraction 19. There is a pronounced lower peak in fraction 25, with amounts reaching 686.6 µg/mL.

In RH 8, the most highly accumulated pigment is esterified violaxanthin, which peaks at 47.3 µg/mL in fraction 17. As with the compounds discussed above, there is a second peak lower in the gradient (13.1 µg/mL in fraction 18). Levels are substantially lower in RH 0; a single peak in fraction 16 reaches 19.0 µg/mL. These patterns are mirrored at lower levels by free violaxanthin, which has pronounced double peaks in RH 8: 16.3 µg/mL and 8.0 µg/mL in fractions 16 and 18 respectively. RH 0 has a single peak of 12.0 µg/mL in fraction 16.

Antheraxanthin shows drastically different distributions across the gradient for the two lines in both its free and esterified forms. In RH 8, levels of free antheraxanthin are low across all fractions; varying between 3-4 µg/mL and decreasing to below 3.5 µg/mL for the second half of the gradient. RH 0 shows an interesting pattern for the free pigment: amounts reach a peak of 16.8 µg/mL in fraction 16, and a second sharp peak of 13.4 in fraction 26. This is the only case of peaks in amount being separated by more than a couple of fractions. Esterified antheraxanthin is seen in higher amounts in RH 0, where it reaches its highest level in fraction 15 (94.9 µg/mL) as part of a wide peak with a shoulder (fractions 18-21). There is a second low peak in a similar region of the gradient to

that of RH 8: amounts reach 24.0 µg/mL in fraction 25. For RH 8, levels of esterified antheraxanthin reach a peak of 12.1 µg/mL in fraction 16.

Lutein shows differing patterns for its free and esterified states. The free carotenoid has pronounced double peaks in RH 8, though levels are low, reaching 3.0 and 5.7 µg/mL in fractions 11 and 16 respectively. Amounts are lower still in RH 0, reaching only 1.6 µg/mL in fraction 16. Esterified lutein is present only in very small amounts in both lines: in RH 8 it varies between 0.2-2 µg/mL, and in RH 0 between 0.2 and 11.1 µg/mL.

Levels of free neoxanthin were low for both lines across the entire gradient, varying between 0.2-0.7 µg/mL in RH 8, and 0.3-0.9 µg/mL in RH 0. The esterified form of the pigment varied in amount between 0.2-2.6 µg/mL in RH 8, reaching a peak in fraction 16. In RH 0 levels ranged between 0.3-0.4 µg/mL, and were at their highest in the first fractions. This is the only example of a compound peaking in amount in this region of the gradient.

More information can be gained about carotenoid distribution throughout the steps of the sucrose gradient when amounts are expressed as percentages per fraction. When these proportions are broken down by sucrose percentage steps (5, 15, 20, 38, and 45% sucrose) it is clear that esterified capsanthin is the highest accumulated pigment across all steps for RH 0, and that esterified violaxanthin and antheraxanthin are most abundant in all steps for RH 8 (Figure 4.13). When this is broken down by individual carotenoid, the differences between regions of the gradient become more pronounced (Figure 4.14). The percentages per fraction of β-carotene and phytoene are substantially higher in the first step (5% sucrose), than in any of the subsequent steps. All free and esterified xanthophylls are present in much higher percentages per fraction at all steps of the gradient in RH 8 than in RH 0. This is due to the fact that free and esterified capsanthin account for such high proportions of carotenoid in all fractions in RH 0. Interestingly, free capsanthin is at its highest percentage in the 5% sucrose step of the gradient, whereas esterified capsanthin is at its lowest in this region.

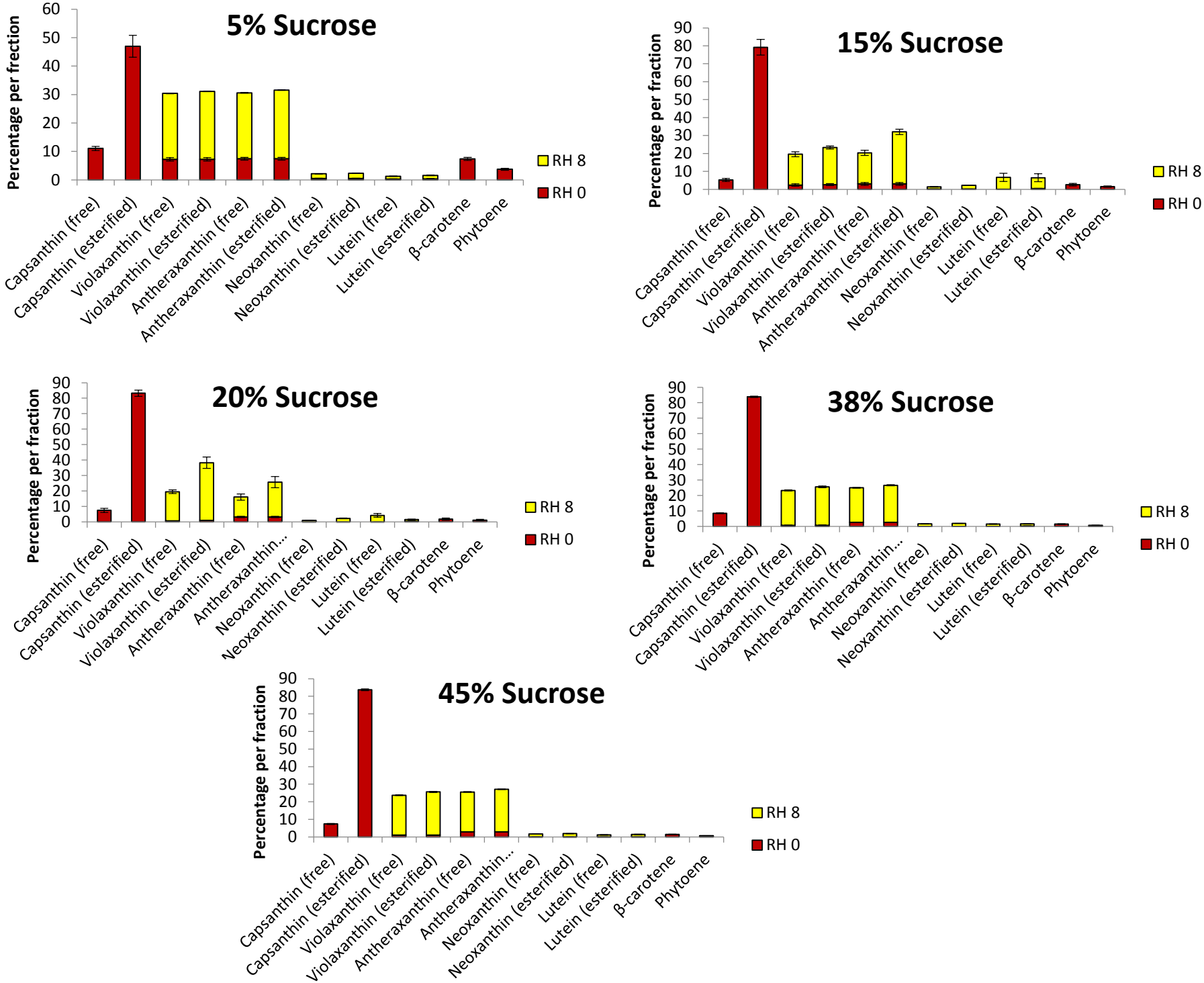
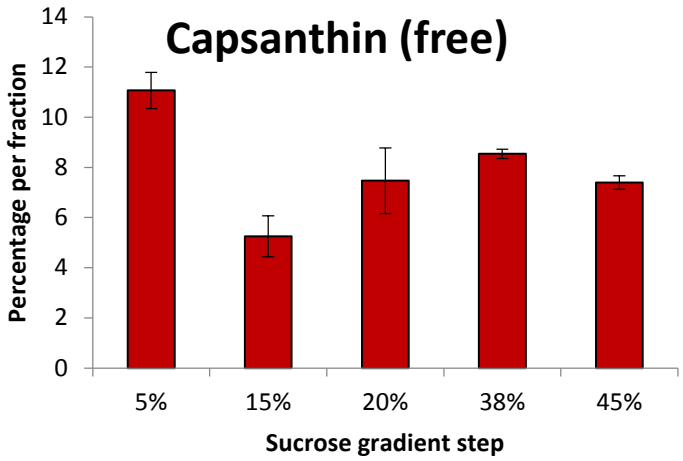
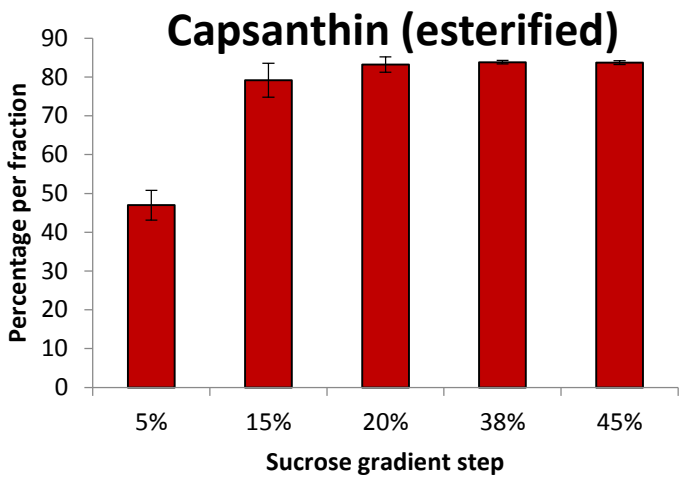
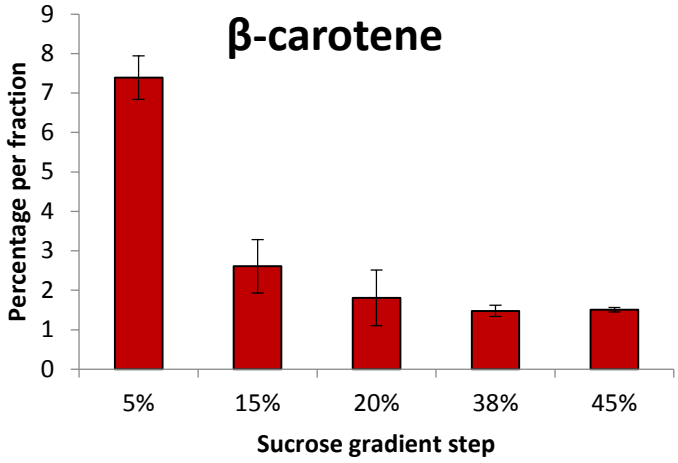
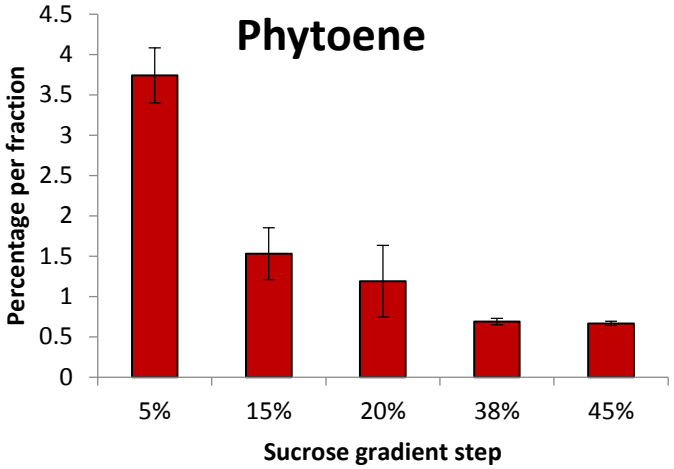


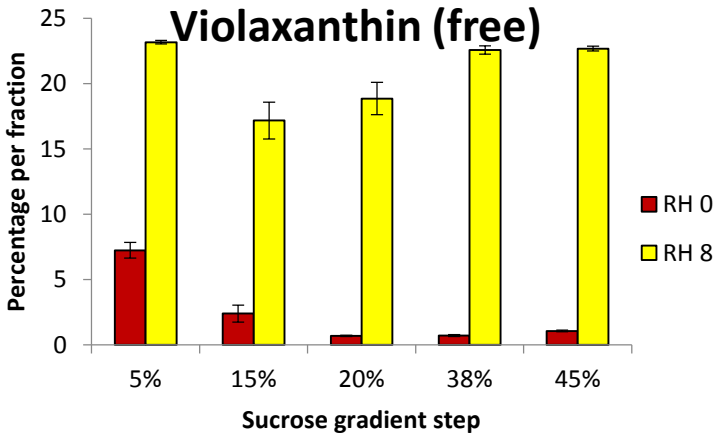
Figure 4.13. Carotenoid amounts (determined through HPLC) displayed as percentages of total carotenoid per sucrose gradient step (n=2 fractionation tube replicates)

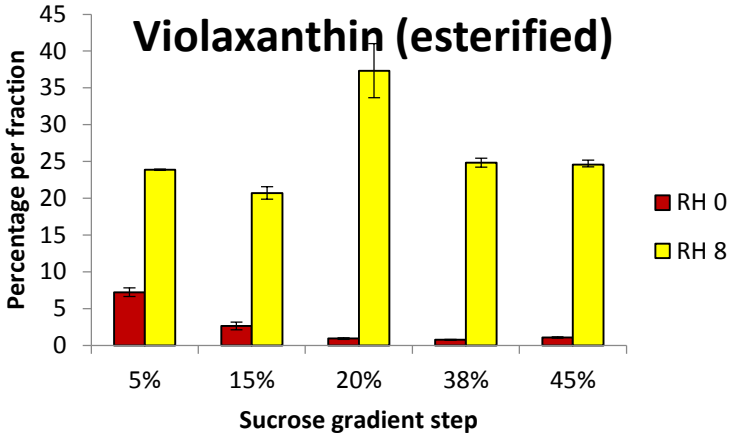

Capsanthin (free)
Percentage per fraction
0
2
4
6
8
10
12
14
5%
15%
20%
38%
45%
Sucrose gradient step


Capsanthin (esterified)
Percentage per fraction
0
10
20
30
40
50
60
70
80
90
5%
15%
20%
38%
45%
Sucrose gradient step


β-carotene
Percentage per fraction
0
1
2
3
4
5
6
7
8
9
5%
15%
20%
38%
45%
Sucrose gradient step


Phytoene
Percentage per fraction
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5%
15%
20%
38%
45%
Sucrose gradient step


Violaxanthin (free)
Percentage per fraction
0
5
10
15
20
25
5%
15%
20%
38%
45%
Sucrose gradient step
RH 0
RH 8


Violaxanthin (esterified)
Percentage per fraction
0
5
10
15
20
25
30
35
40
45
5%
15%
20%
38%
45%
Sucrose gradient step
RH 0
RH 8

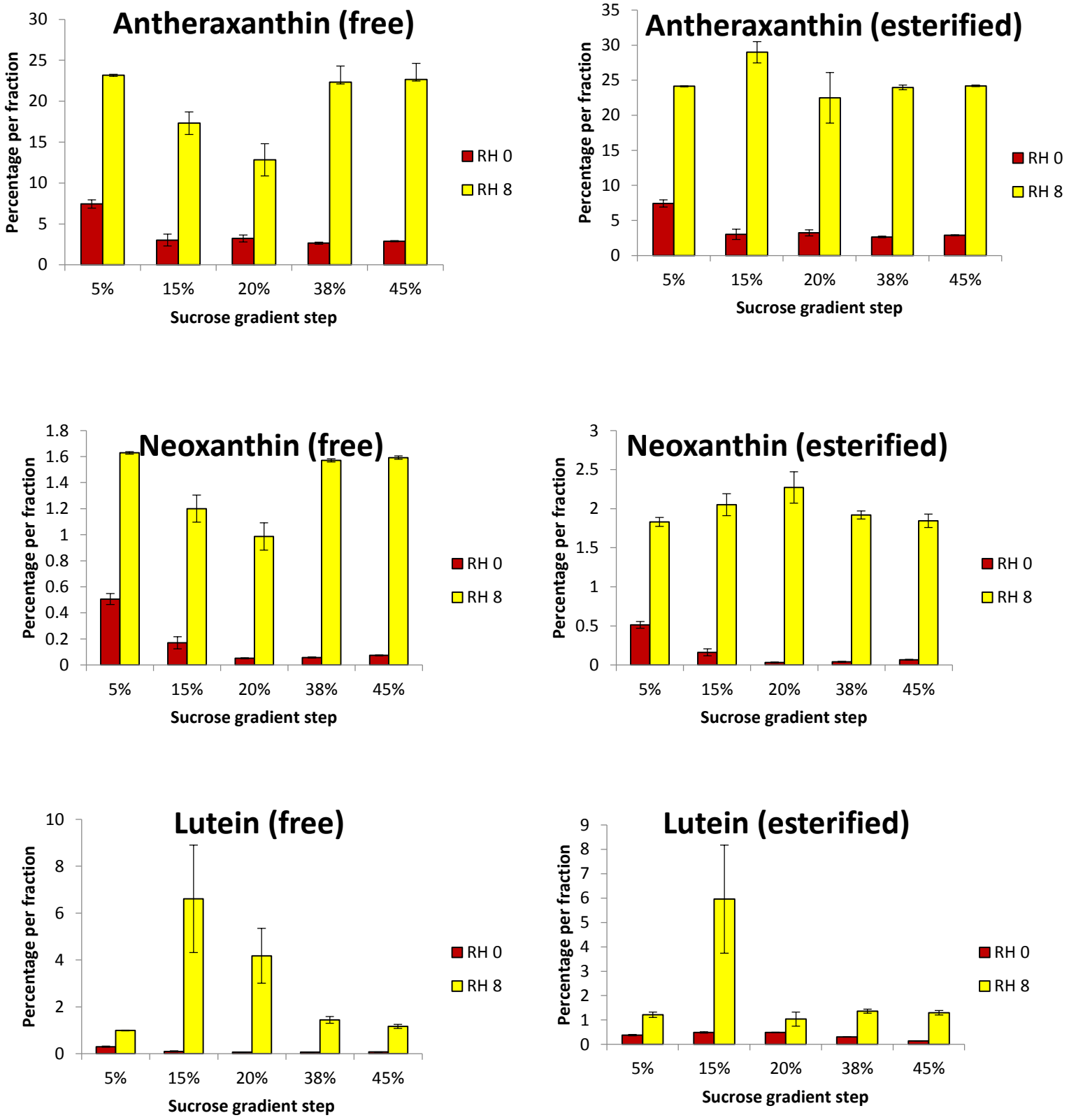


Figure 4.14. Carotenoid amounts (determined with HPLC) expressed as percentages per fraction, shown for each carotenoid across the five steps of the sucrose gradient (n=2 fractionation tube replicates).

4.2.2.2 Protein-level Fraction Characterisation

In addition to carotenoids, proteins were extracted from each one mL fraction collected from the gradient. To gain an overall impression of how the protein profile changed across the gradient and between different samples, the extracts were silver stained. To characterise the functional identity of the fractions, Western blots were carried out with monoclonal antibodies to a collection of important structural and biosynthetic proteins.

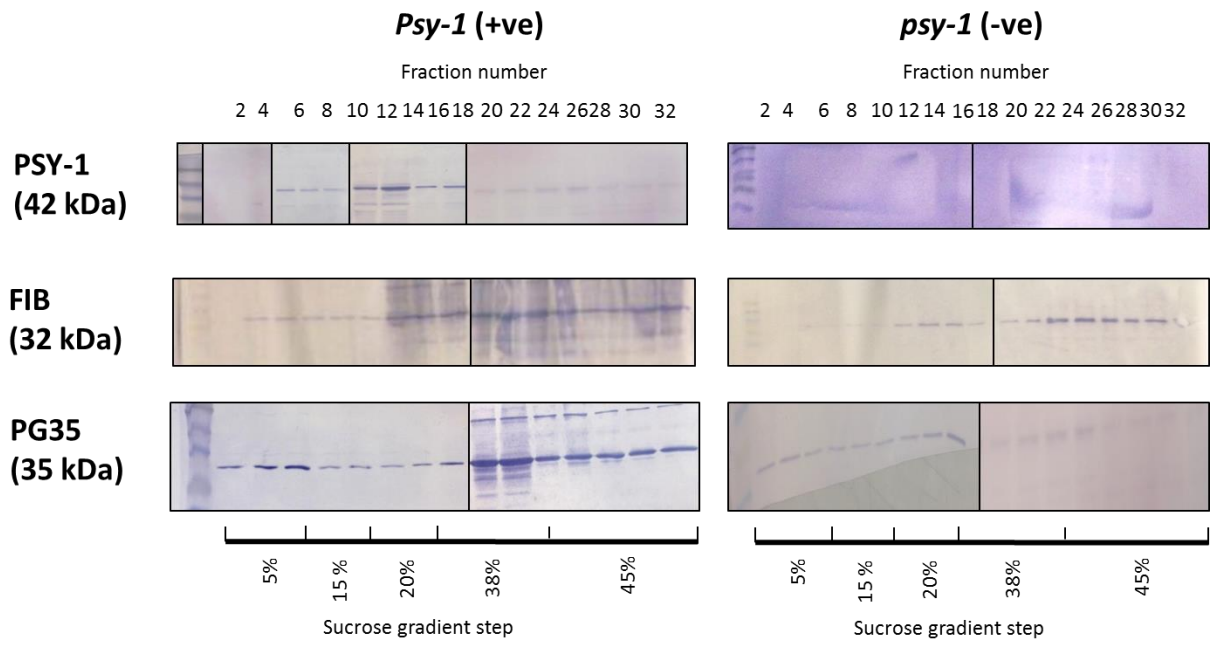


Figure 4.15. Western blots to carotenoid biothetic enzymes and structural proteins in fractions from the sucrose gradients of RH 0 and RH 8. PSY-1: *Solanum lysopersicum* Phytoene Synthase-1. FIB: *Capsicum annuum* Fibrillin. PG35: *Solanum lycopersicum* Plastoglobulin-35. Black outlines denote separate blots; protein bands are for visual rather than quantitative comparison.

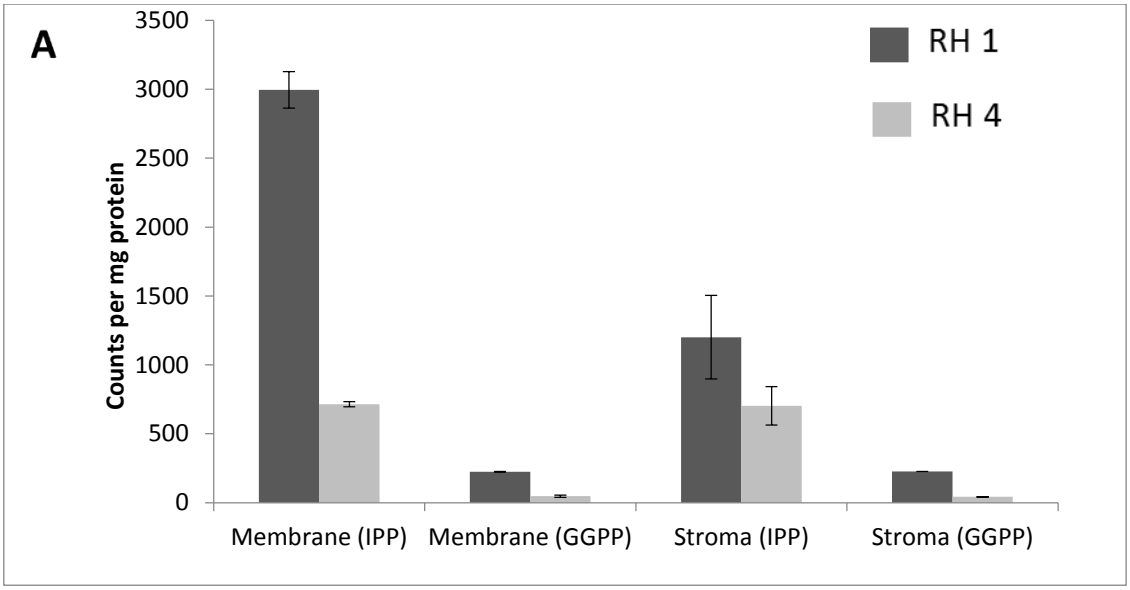
PSY-1 was not detected in the yellow *Psy-1* negative line (RH 8) in any fraction (Figure 4.15). It was detected as a strong band in fractions from the 20% sucrose step in RH 0, suggesting it is associated with membranes which are visible as a clear band in the gradient in this step. It is also present in the stromal fractions of this line: visible as a weak band in fractions 22-30. Fibrillin is clearly visible as a strong band from fraction 12-32 in RH 0. However, the antibody appears to be reacting with proteins of different sizes, smearing on either side of the main band is visible. In RH 8 fibrillin gives only a single band, again in the second half of the gradient. The major band is in a lower fraction than that of RH 8. Plastoglobulin (PG35) gives a stronger band in the first six fractions than in fractions 8-16, suggesting that the 5% sucrose step is the location of the plastoglobuli in this gradient (Figure 4.15).

4.2.3 Sub-plastid Phytoene Synthase activity

To test whether the enzyme was active in the sub-chromoplast regions where the protein was located, an activity assay using radiolabelled geranylgeranyl pyrophosphate (^{3}H-GGPP) and isopentenyl pyrophosphate (^{14}C-IPP) was carried out. These are the two immediate precursors of

phytoene synthesis; the final products of the MEP pathway. Two molecules of GGPP combine in a head-to-tail condensation reaction to form phytoene Insight on whether transcript level could be related to levels of active enzyme and subsequent fruit carotenoid content was gained from this study. It provides a link between different levels of regulation: enabling connections to be made between transcript level, protein level, carotenoid content and resulting fruit colour. Testing both IPP and GGPP also provides information about the orientation of the enzymes PSY-1 and GGPS in relation to each other: if radiation from IPP rather than the GGPP is incorporated into the resulting phytoene, this means GGPS and PSY-1 are tightly configured, with metabolites being efficiently channelled between them. If more radiation from GGPP was incorporated into the resulting phytoene, the tight configuration between the enzymes would be unlikely.

Chromoplasts from selected *Psy-1* expressing lines were isolated and broken into stroma and membrane fractions. These fractions were incubated with one of the labelled compounds, and the amount of radioactive label accumulated in phytoene produced during the incubation was tested.



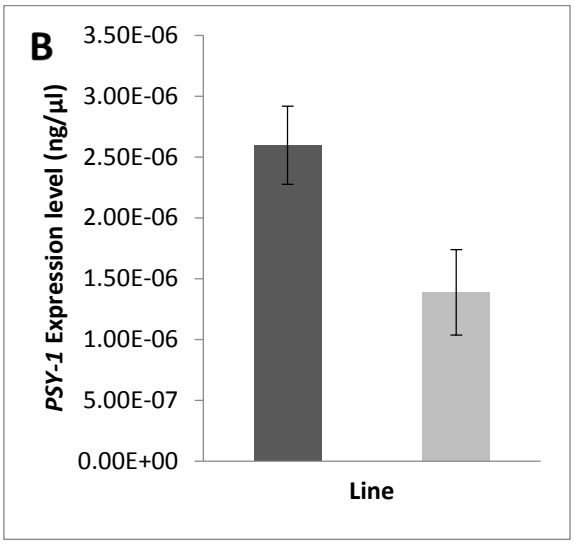


Figure 4.16. (A)) Level of radioactivity found in phytoene extracted from stroma or membrane of isolated chromoplasts incubated with either ^{14}C-IPP or ^{3}H-GGPP. (B) *Psy-1* transcript level, determined through qRT-PCR.

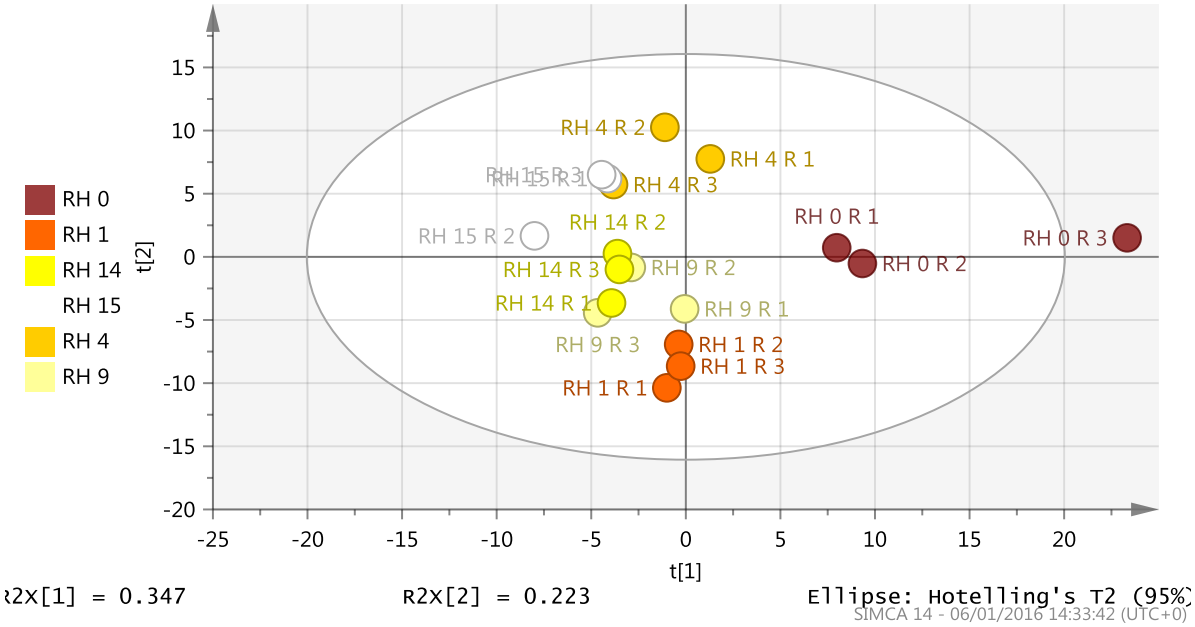
It was found that, corresponding to the expression levels, activity of PSY-1 was highest in RH 1. Greater levels of labelling were observed in fractions incubated with ^{14}C-IPP than with ^{3}H-GGPP, suggesting that a metabolon streamlining the reactions forming GGPP and phytoene exists. This would explain the lack of labelling from GGPP incorporated into phytoene (Figure 4.16).

4.2.4 Fruit Metabolome Profiling

In light of the differences in plastid structure being caused by the lack of the PSY-1 enzyme from certain lines, the broader metabolism of pepper fruits was investigated with gas chromatography-mass spectroscopy (GC-MS). As the mutation has implications for the overall structure of the plastid, areas of metabolism other than carotenoid biosynthesis may be affected. Knowing exactly how the

global fruit metabolism is altered by this mutation could be important for future decisions about how these lines could be used by Syngenta in their breeding programme.

A.



B.

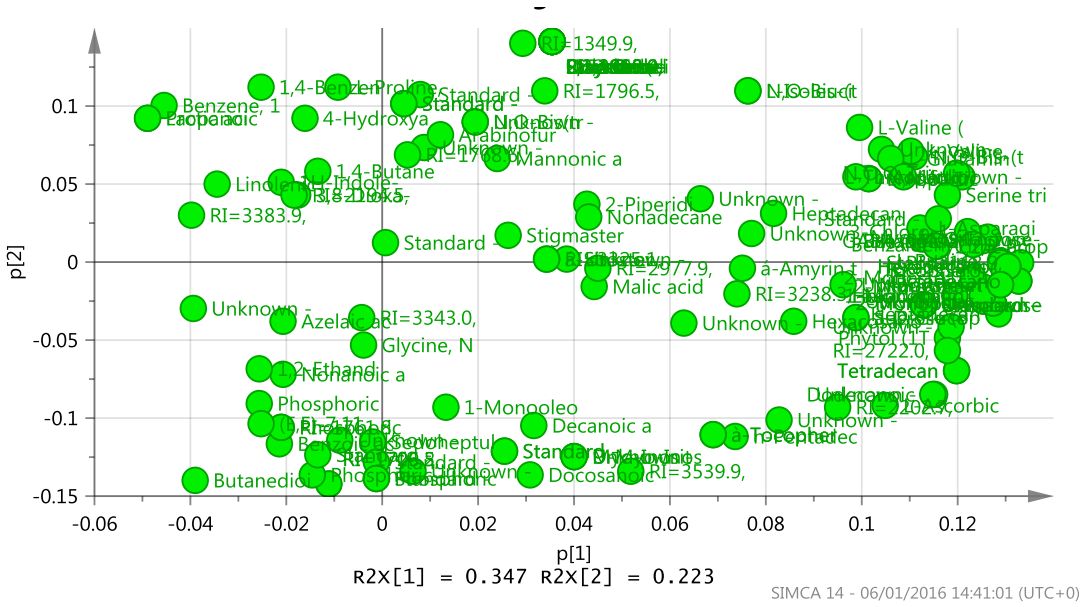
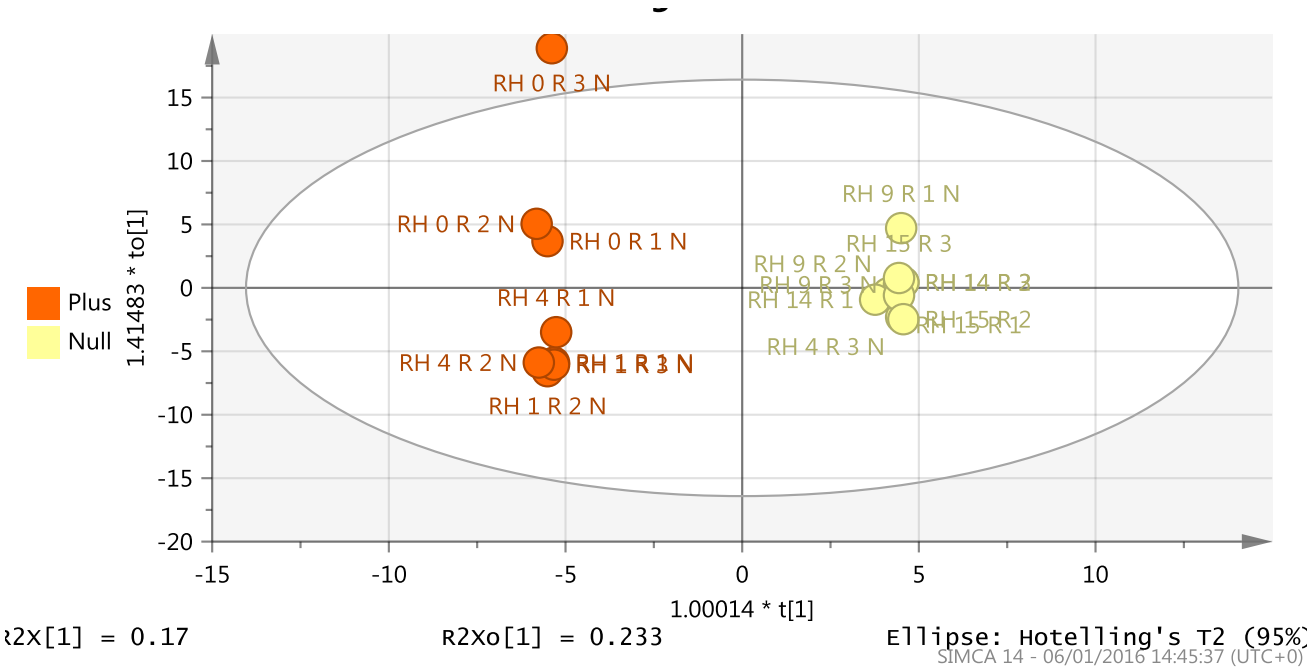


Figure 4.17. Principal Components Analysis (PCA) showing the separation of accessions based on analysis with GC-MS. A. Score plot, showing separation of accessions. B. Loadings plot, showing compounds driving the separation between accessions. Statistics were carried out with embedded equations in SIMCA-P; n=3 biological replicates.

A.



B.

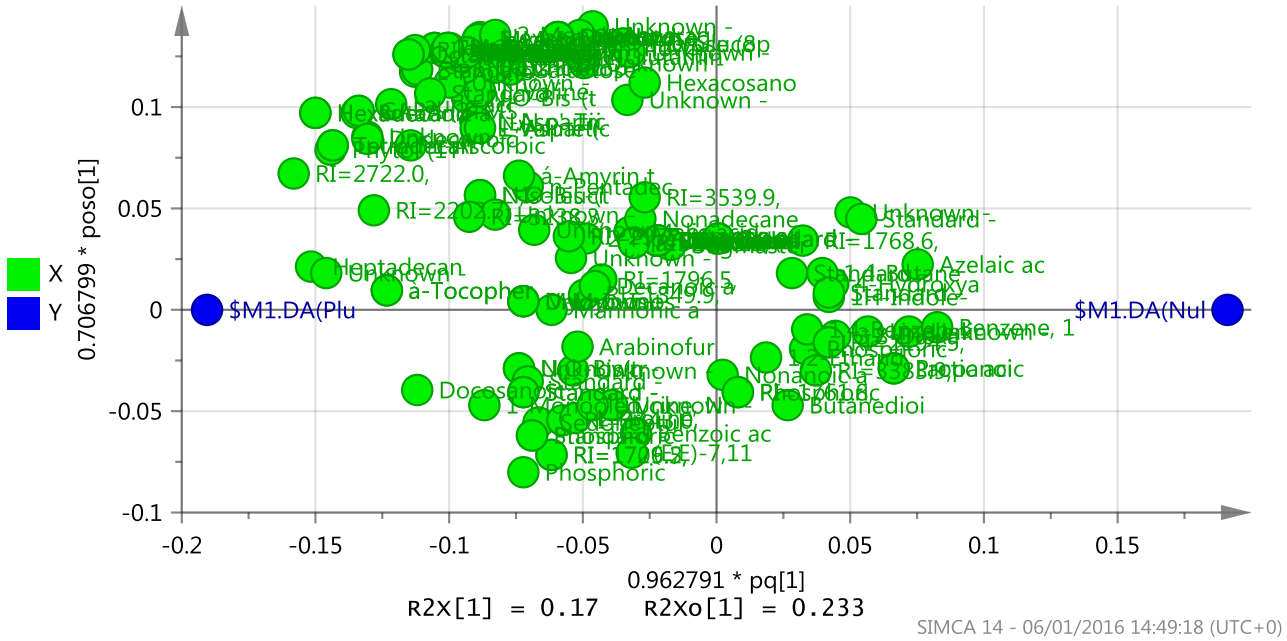


Figure 4.18. Partial Least Squares Discriminant Analysis (PLS-DA), showing separation of accessions based on possession of a functional copy of *Psy-1*. A. Score plot, showing separation between accessions. B. Loadings plot, showing compounds driving the separation. Statistics were carried out with SIMCA-P; n=3 biological replicates.

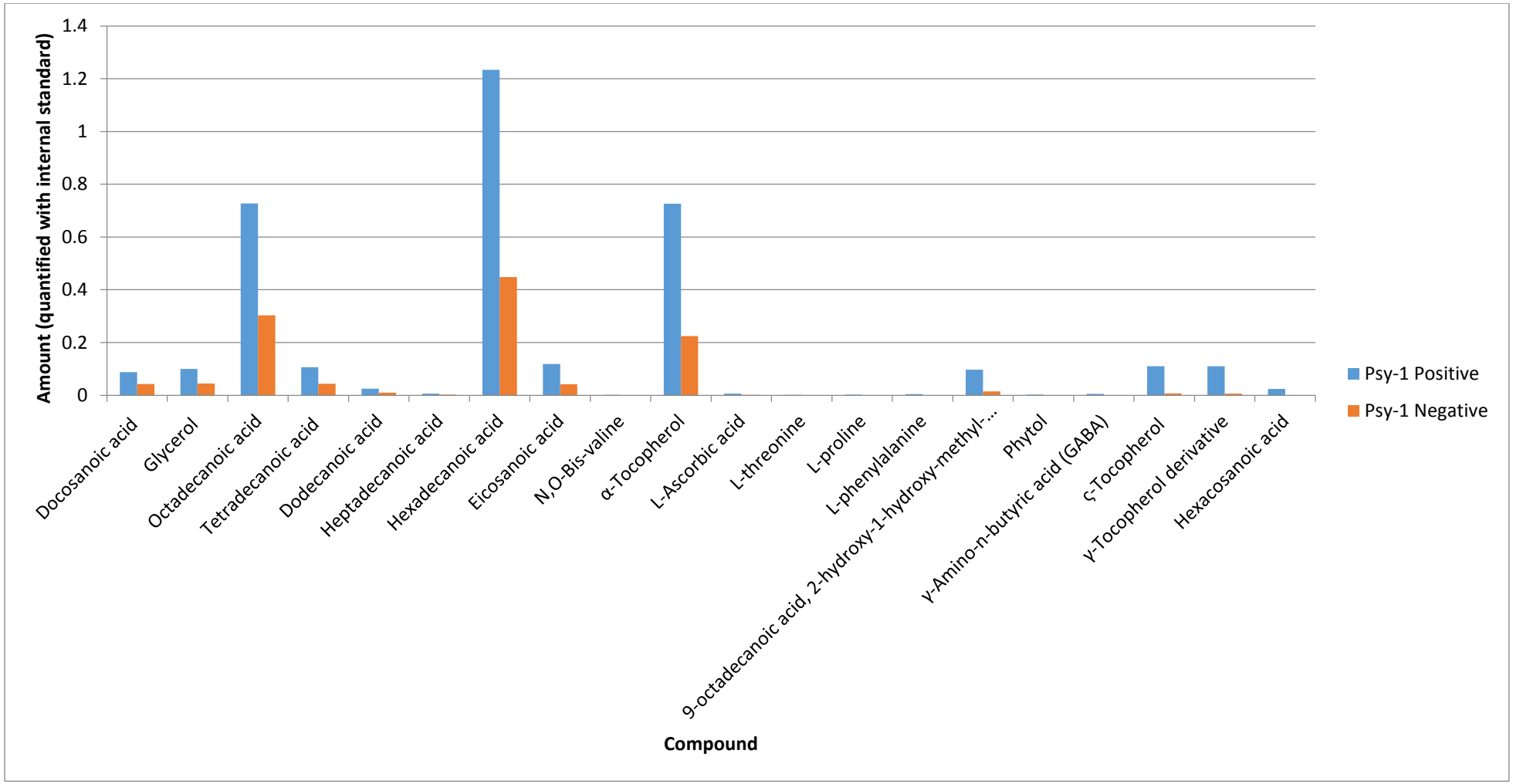


Figure 4.19. Amounts of significantly different compounds showing the largest changes between the two classes when accessions are separated by *Psy-1* possession, relative to an internal standard.

Principal components analysis (PCA) revealed that the most intensely pigmented lines (particularly RH 0 and RH 1, but to a lesser extent RH 4) do separate from the paler yellow and white accessions, without the separation being driven by assigned classes (as is the case for Partial Least Squares-Discriminant Analysis (PLS-DA)) (Figure 4.17, Figure 4.18) According to the corresponding loadings plot (Figure 4.18 B), this separation is driven largely by fatty acids, particularly hexadecanoic acid (C16:0) and octadecanoic acid (C18:0). This loadings plot also shows that the RH 0, which accumulates a huge amount of pigment in ripe fruit, separates from the other lines based on many more compounds than those driving separations between the other lines. The same can be seen on the loadings plot for the PLS-DA: a dense cluster of compounds are visible in the region of the plot where RH 0 lies. Interestingly, two of the orange lines (RH 1 and RH 4, expressing *Psy-1*) separate more from each other than they do from the yellow and white lines which lack a functional copy of the gene (Figure 4.18). According to the PCA loadings plot, the fatty acids decanoic and docosanoic acid drive the separation of RH 1 from the other accessions (Figure 4.17). These plots show that possession of a functional copy of *Psy-1* does not explain a substantial proportion of the variation in the data. Forcing separation between the classes in this way does not explain most of the variation in the data: according to PLS-DA only 32.9% is explained by defining the first component as *Psy-1* expression status (Figure 4.17, Figure 4.18).

When classes are assigned to accessions based on their *Psy-1* expression status, compounds which are significantly different between the two groups can be examined further: the largest fold changes were seen in octadecanoic acid, docosanoic acid, and tetradecanoic acid. Octadecanoic and tetradecanoic acids (C18:0, and C14:0 respectively) can be attached to hydroxylated carotenoids as part of the esterification process, but those with longer chains, such as docosanoic acid (C22:0) cannot. Significant changes were also seen in the antioxidants α- and γ-tocopherol (Figure 4.19).

4.3 Discussion

4.3.1 Electron Microscopy: Visible differences between *Psy-1* positive and negative lines

As plastids are the site of carotenoid synthesis and storage in mature green fruit and during ripening, it is important to investigate the variation between structures visible in a line possessing the PSY-1 enzyme, and a line which does not. Before the onset of ripening, carotenoids are synthesised in the photosynthetic membranes of the chloroplast; during ripening the organelle undergoes the chloroplast-chromoplast transition, with carotenoids continuing to be synthesised in increasingly high amounts, and stored in the specialised structures of the fibrils and plastoglobuli. *Capsicum* chromoplasts employ novel sequestration mechanisms, marking their chromoplasts out from those of their close relatives.

They have a long history of visualisation: the earliest studies date from the late 1950s (Frey-Wyssling and Kreutzer, 1958). A decade later, a comprehensive study of chloroplast and chromoplast ultrastructure in several different sweet and pungent *Capsicum* accessions was carried out by Spurr and Harris (Spurr & Harris, 1968). This was one of the first pieces of work examining chromoplast organisation in detail; previously, plastid structure studies and focussed on chloroplasts and etioplasts. The insight into chromoplast structure shed new light on the potential plasticity of the plastid: the dynamic changes in membrane organisation particularly changed the way the plastid was perceived. Plastids in fruits from seven cultivars ranging in colour from yellow to scarlet were imaged with TEM at anthesis, mature green, mid-ripening, and ripe. Important features of the chloroplast to chromoplast transition were observed for the first time: selective lysis of granal stacks as the fruit turn from green to yellow, orange or red, and development of carotenoid-storing 'crystalloids' (identified as fibrils in later work). No grana were observed in chromoplasts of fully ripe fruit, however, other novel thylakoid structures are present. Spurr and Harris identify the 'thylakoid plexus': a system of inter-connected intergranal lamellae which have not undergone the lysis of the thylakoids in the granal stacks. Similar structures can be observed in the chromoplasts in the current work (Figure 4.6, Figure 4.7, Figure 4.8).

4.3.1.1 Chloroplast structure

The chloroplasts of lines RH 2 and RH 10 contained similar structures: thylakoids with narrow lumens, organised into grana which are connected by intergranal lamellae. The thylakoid envelope is visible, as are a small number of osmiophillic lipid droplets (Figure 4.1, Figure 4.2, Figure 4.3, Figure 4.4). The most noticeable difference between chloroplasts from the two lines is the number of thylakoids making up each granum: on average, grana of RH 2 contain approximately twice the number of thylakoids as those in RH 10 (Figure 4.2, Figure 4.4). This corresponds with the darker green fruit phenotype of RH2; a result of its increased chlorophyll content.

4.3.1.2 Chromoplast structure

Thylakoid membranes are the site carotenoid biosynthesis in the chloroplast; during the transition they break down, resulting in the lipid droplets of the plastoglobuli. These are then incorporated into the tubular fibrils, 32 kDa linear proteins with a lipophilic core which interacts with carotenoids (Deruère et al., 1994b). The model proposed by Deruère et al suggests that hydrophobic lipid tails hold carotenoids in place at the centre of the complex, with their polar heads interacting with the fibrillin protein which is arranged in a cylinder around the phospholipid. It was reported that bicyclic carotenoids formed the core of the molecule and that, where hydroxyl groups were present, 80-100% of the pigments were esterified. Fibrils have previously been observed in electron microscopy images (Deruère et al., 1994b). However, recent publications have focussed on red pepper fibril assembly; the panel under investigation here provides an important opportunity to study the nature of fibrils in orange and yellow pepper fruits and to compare these to the ancestral red pepper.

The images of RH 2 and RH 10 ripe fruit show many more visible fibrils in the *Psy-1* possessing line, RH 2. This corresponds with the significantly greater pigment levels in ripe fruit from this line. There are clear differences in shape between chromoplast shapes between RH 2 and RH 10; those in RH 2 appear an elongated oval shape, closer to that of the chloroplast, wheres RH 10 chromoplasts are much more circular. This could be the result of the increased fibril levels in RH 2: their efficiently packed long, linear, structures could determine the shape of the whole organelle.

Plastoglobuli are visible in both lines; however, they appear more abundant in RH 2, again corresponding to the elevated carotenoid level of this line.

In order to assess whether plastid number had an impact on colour phenotype, the ratio of a plastid encoded gene to a nuclear encoded gene was estimated with a PCR-based assay. The number of copies of genome-encoded *Phytoene desaturase* (*Pds*) was compared to that of plastome-encoded *Rubisco large-subunit* (*RbcL*). It was found that the paler line (RH 10) had a higher plastome-genome ratio, telling us that, in this case, the darkness of mature fruit colour is not the result of increased plastid number (data in appendix; Section 9.3).

4.3.2 Sub-chromoplast Fractionation: Insights into carotenoid sequestration in *Psy-1* positive and negative lines

4.3.2.1 Location of carotenoid synthesis in the chromoplast and fibril-based carotenoid sequestration

When the carotenoid amounts in the *Psy-1* positive and negative sub-chromoplast fractions were examined, it was clear that almost all the free and esterified pigments reached a peak in amount around fraction 16 (Figure 4.12). This must therefore be the final destination of carotenoids present in ripe fruit. As stated above, the long-term storage sites for pigments in the *Capsicum* chromoplast

are the fibrils, therefore this data suggest that the fibrils are found at fraction 16 of the sucrose gradient, forming a band in the 20% sucrose step. Further evidence was provided by analysis with Western blotting: when proteins from the fractions were tested with antibodies to fibrillin, the strongest bands were in fractions with the highest pigment levels, and lower fractions (18-32) (Figure 4.15).

PSY-1 was found to be membrane associated, and present in the stroma. This result is similar to previous work which placed *Capsicum* PSY-1 in the chromoplast stroma (Romer et al., 1993). However, the two experiments are not directly comparable due to the fact that Romer et al. used a different technique to divide regions of the chromoplast. PSY-1 from tomato was found to be loosely membrane bound and present in the stroma (Fraser, Truesdale, Bird, Schuch, & Bramley, 1994). The PSY-1 activity assay supported the hypothesis that the enzyme was active in the sub-chromoplast location suggested by the Western blots: more activity was measured in the membrane fraction than in the stroma. Samples incubated with ^{14}C-IPP accumulated much more radiolabelled phytoene than those with ^{3}H-GGPP, suggesting that GGPS and PSY-1 are tightly configured, making it difficult for extrinsic GGPP to enter the metabolon. The activity levels also corresponded to gene expression data for RH 1 and RH 4.

4.3.2.2 Role of the plastoglobuli in carotenoid sequestration

Due to the fact that Western blot band for tomato plastoglobulin (PG35) in the first four millilitre fractions of the gradient was stronger than in the proceeding fractions, this region was designated as the location of the plastoglobuli from the broken chromoplast. The same result had previously been observed in tomato (Nogueira et al., 2013). Phytoene and β-carotene were found in in fractions 1-4, as well as low levels of esterified xanthophylls (Figure 4.12, Figure 4.13, Figure 4.14). Information can be inferred about the role of this compartment from its lack of free carotenoid. As a relic of chloroplast thylakoid membrane, does it store photosynthetic accessory pigment and transport it to the developing fibrils as the fruit ripens? This seems unlikely, as red fruit plastoglobuli contain esterified capsanthin, which is synthesised in ripening tissue and not present in the chloroplasts of immature fruit. As the plastoglobuli did not give a positive result when tested with anti-PSY-1, we can deduce that they are not the original sites of carotenoid biosynthesis in the chromoplast (there are reports of PSYs in the plastoglobuli of other plant species, such as *Zea mays* (Shumskaya, Bradbury, Monaco, & Wurtzel, 2012)). They could be overflow compartments, intended for the storage of carotenoids when total levels are high. Esterification must be a necessary modification for storage in this very hydrophobic environment. In addition to the first fractions at the top of the gradient, Western blots to plastoglobulin gave positive results in the fractions corresponding to the membranous fibrils; given reports that this protein is restricted to the plastoglobule membrane

(Austin, Frost, Vidi, Kessler, & Staehelin, 2006), this result implies that the two compartments are not discrete entities.

The proteome of the plastoglobule in *Capsicum* chromoplasts has been characterised previously (Ytterberg, Peltier, & van Wijk, 2006); analysis with mass spectrometry identified 28 proteins in this compartment. The plastoglobule undergoes substantial changes as the plastids transitions from chloroplast to chromoplast: Ytterberg et al compared it to that of plastoglobuli from *Arabidopsis* leaf chloroplasts, and found an overlap of less than half of the proteome (12 proteins).One of this study's main conclusions was that fibrillin proteins were highly structurally important for the plastoglobule: seven fibrillin homologues were found in the Arabidopsis chloroplast, and four in the *Capsicum* chromoplast. Four enzymes involved in carotenoid biosynthesis were also identified in the chromoplast (Ytterberg et al., 2006).

The presence of biosynthetic enzymes in the plastoglobule has inspired debate about whether the compartment is metabolically active or not. Though their sole function was previously thought to be lipid storage, some recent studies have associated them with lipid biosynthesis and repair (Eugeni Piller, Abraham, Dormann, Besagni, & Kessler, 2012; Rottet, Besagni, & Kessler, 2015; Spicher & Kessler, 2015). The observation by Shumskaya et al. (2012) that phytoene synthases were present supported the idea that (in some species at least) plastoglobuli were involved in biosynthetic processes, however, the activity of plastoglobule PSYs was not confirmed. Due to the fact that plastoglobuli are formed from the blistering of thylakoid membranes from the chloroplast, the presence of some membrane-associated proteins would be expected. This alone does not confirm that the plastoglobule is a metabolically active compartment. Recently it has been proposed that plastoglobuli could be a site of carotenoid cleavage: as well as being found in the plastoglobule proteome (Ytterberg et al., 2006), it was demonstrated that CCDs were active in the lipid droplets (Rottet et al., 2016). The study found CCD4 located in the plastoglobuli of *Arabidopsis* leaves, and observed that *ccd4* mutant plants accumulated carotenoids in plastoglobuli as leaves senesced.

4.3.3 Fruit Metabolome Profiling: Insights from GC-MS data

4.3.3.1 Higher carotenoid levels are associated with accumulation of esterifying fatty acids

Membranes in the chromoplast harbour both free and esterified carotenoids, as well as being the site of some of the early steps in the biosynthetic pathway. Therefore, membrane constituents must have an influence on carotenoid sequestration. Indeed, the data from analysis with GC-MS brings out differences in fatty acid composition between *Psy-1* expressing and non-expressing lines. When the two groups are compared, some of the significantly different compounds were Very Long Chain Fatty Acids (VLCFAs), with acyl chain of lengths greater than twenty carbons: docosanoic acid and

hexcosanoic acid (C20:0, and C26:0). These have been previously detected in pepper (Conforti, Statti, & Menichini, 2007).They are normally associated with cuticular or epicuticular waxes; in Arabidopsis, knocking out a key gene in VLCFA biosynthesis led to embryo fatality, suggesting that the compounds play a crucial developmental role involving patterning the plant through polar auxin transport (Roudier et al., 2010). In fruit, VLCFAs are involved in water retention: it has been shown in tomato that β-ketoacyl-coenzyme A synthase (a key gene in VLCFA biosynthesis) knockout mutants experienced a 3-8 fold increase in water loss when fruit began to ripen (immature fruits remained unaffected) (Leide, Hildebrandt, Reussing, Riederer, & Vogg, 2007). This could imply that, with significantly less VLCFA *Psy-1* non-expressing pepper fruits are at risk of losing water at a greater rate. The largest fold changes between the classes in this study are observed in hexadecanoic acid (C16:0) and octadecanoic (C18:0) acid, both of which can esterify carotenoids with hydroxyl groups. As *Psy-1* expressing accessions contain more esterifiable pigments, the higher C14:0-C18:0 fatty acid content in these lines is explainable.

4.3.3.2 High carotenoid levels correlate with increased tocopherol levels

Tocopherols were also found to be significantly different between *Psy-1* expressing and non-expressing lines (α-tocopherol: p=0.019; γ-tocopherol: p=0.004). Their biosynthesis depends on the cytosolic metabolism of aromatic amino acids for the cyclic head group, whilst production of the tail group is from either the plastidic MEP pathway (resulting in GGPP, the tail group of the tocotrienols) or the mevalonate pathway (giving phytyl-diphosphate (phytyl-PP), tail of the tocopherols) (Dellapenna & Pogson, 2006). As GGPP is an important branch intermediate in tocochromanol biosynthesis, its accumulation as a result of the absence of PSY-1 could be expected to lead to an increase in levels of these compounds (albeit in tocotrienols, rather than tocopherols). However, this is not observed: tocopherols are significantly higher in lines expressing *Psy-1*. One of the important properties of tocopherols is their ability to scavenge free radicals, and therefore protect membrane lipids; they terminate polyunsaturated fatty acid (PUFA) chain reactions by donating a hydrogen atom from their ring hydroxyl group. This results in a tocopherol radical which is rapidly recycled back to the corresponding tocopherol, and is thus able to repeat the cycle multiple times (Dellapenna & Pogson, 2006) As accessions lacking in *Psy-1* transcript possess significantly lower levels of α- and γ-tocopherol, they may be at slightly greater risk of membrane damage from reactive oxygen species.

However, despite these significant differences in certain metabolites, it is apparent that the lack of *Psy-1* transcript is not causing a huge change in fruit metabolism: less than half the variation in metabolites was accounted for by this change.

4.4 Conclusion

To conclude, it is clear that while there are some differences in plastid structure between lines with and without a functioning PSY-1 enzyme, chloroplast and chromoplasts are broadly similar. As would be expected from a mutation in a gene expressed exclusively during fruit ripening, the major differences observed are in the chromoplast, with the major difference being the abundance of fibrils in lines with PSY-1. This is clear from both EM pictures, and Western blots to fibrillin, (Figure 4.6, Figure 4.7, Figure 4.15). It is interesting to observe that, although fruit without *Psy-1* do not synthesise carotenoids throughout ripening, the chromoplast still develops the same storage mechanisms to accommodate pigment remaining from the mature green stage; xanthophylls are esterified, and fibrillin is present. A model of the differences in carotenoid biosynthesis and the sequestration mechanisms employed is proposed in Figure 4.20. Western blotting also revealed that PSY-1 was not present in the line possessing the marker for absence of the gene from the genome (RH 8). It was detected in RH 0, and appears to be membrane associated and present in the stroma (Figure 4.15).

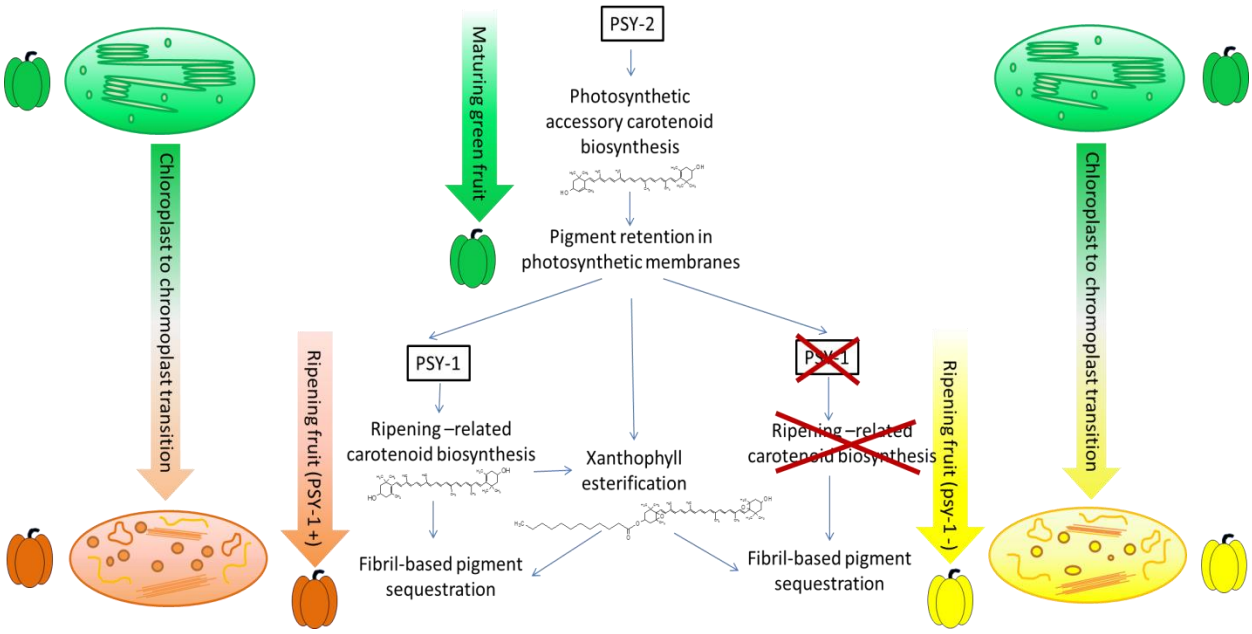


Figure 4.20. Proposed model of carotenoid biosynthesis and storage throughout ripening in *Psy-1* positive and negative peppers.

Profiling the lines with GC-MS revealed differences in fatty acid composition which are likely to be associated with carotenoid accumulation; fatty acids of lengths less than C18 are capable of esterifying carotenoids. It is interesting to observe that, in a panel bred for colour diversity, the major differences in metabolism can be related to pigment sequestration.

5 Impact of the *c1* locus on pigment composition, broader fruit metabolism and plastid structure

5.1 Introduction

Of the three loci making up the model of mature fruit colour in pepper proposed in 1985, one remains unidentified (Gómez-García & Ochoa-Alejo, 2013; Hurtado-Hernandez & Smith, 1985). This is the *c1* locus. *Capsicum* accessions lacking functional copies of all three of the loci produce white mature fruit; those which possess a functional copy of the *c1* alone produce lemon yellow mature fruit (Hurtado-Hernandez & Smith, 1985). Mapping of the locus carried out in-house at Syngenta has provided enough information to generate markers close to the locus; however, the functional identity of the gene associated with the locus is still uncharacterised. Syngenta provided seed for a diversity panel of fourteen lines, seven of which were *c1*-positive and seven which were *c1*-negative.

To characterise accessions from the panel, carotenoids were profiled with high pressure liquid chromatography (HPLC), and metabolites were profiled with gas chromatography-mass spectroscopy (GC-MS). In order to assess differences in plastid structure between lines possessing a functional copy of the *c1* locus and those which do not, transmission electron microscopy (TEM) was used to generate images of chloroplasts and chromoplasts. Sub-chromoplast fractionation was also carried out on plastids from ripe fruit, enabling comparisons between the architecture of yellow and white accessions.

5.2 Results

Twelve accessions segregating for the c1 locus were grown for analysis with HPLC, GC-MS, and electron microscopy. In the absence of functional copies of Psy-1 and Ccs, the c1 locus causes the difference between yellow and white mature fruit.

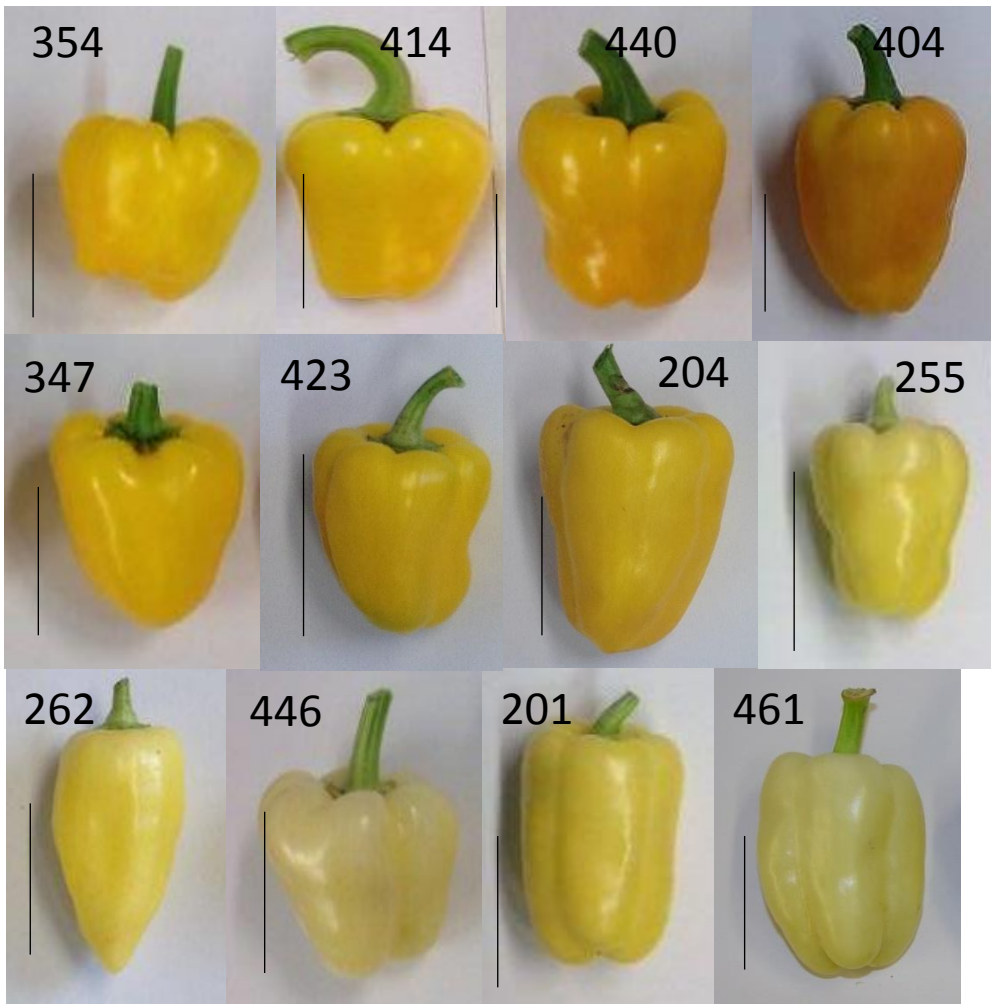


Figure 5.1. Ripe fruit from the *c1* diversity panel provided by Syngenta. Plants were grown under glass with a long day light regime. 354, 414,440,404,347,423: possessing functional copies of the *c1* locus. 204, 255, 262,446,201,461: not possessing a functional copy of the *c1* locus. Scale bar: 5cm.

5.2.1 Carotenoid composition

Line	Violaxanthin (free)		Violaxanthin (monoesters)		Violaxanthin (diesters)		Antheraxanthin (free)		Antheraxanthin (monoesters)		Antheraxanthin (diesters)	
	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M
204	37.527	0.027	37.534	0.034	37.641	0.141	38.752	0.620	n.d		37.840	0.340
355	38.511	0.516	56.997	1.657	61.871	0.992	40.251	0.121	38.642	1.142	48.773	4.506
354	60.510	8.906	72.248	8.417	69.779	10.674	44.781	0.989	n.d		49.352	0.527
414	40.103	0.706	68.488	5.769	67.202	1.318	40.205	0.037	38.246	0.746	45.386	1.066
440	90.225	2.082	86.021	4.477	66.399	2.434	42.606	0.248	n.d		46.848	2.420
262	40.470	0.761	42.684	1.092	41.048	0.975	39.279	0.112	37.884	0.384	38.824	0.543
423	52.708	0.798	87.034	2.590	77.621	2.321	40.185	0.252	n.d		44.176	0.631
201	38.178	0.079	39.575	0.788	40.046	0.782	42.089	0.897	n.d		38.281	0.501
446	37.560	0.060	n.d		n.d		37.810	0.066	n.d		n.d	
461	n.d		n.d		n.d				n.d		n.d	
255	40.003	1.256	39.314	1.021	39.246	0.887	38.825	0.669	n.d		38.078	0.295
404	45.056	2.436	71.820	17.294	89.848	26.175	42.790	0.144	n.d		41.556	2.876

Line	Neoxanthin (free)		Neoxanthin (monoesters)		Neoxanthin (diesters)		Lutein (free)		Lutein (monoesters)		Lutein (diesters)	
	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M
204	n.d		n.d		n.d		2.189	0.424	n.d		1.576	0.013
355	2.707	0.037	2.951	0.159	5.876	1.313	16.924	5.978	8.234	2.245	9.219	2.438
354	3.145	0.344	3.837	0.157	6.934	0.866	23.714	5.069	4.207	0.011	7.134	0.875
414	2.902	0.024	3.828	0.124	5.415	1.086	9.780	0.166	3.096	0.202	3.968	1.323
440	4.962	0.096	4.936	1.155	7.061	0.925	39.877	0.489	2.310	0.747	6.973	1.587
262	2.689	0.007	2.953	0.104	3.455	0.223	4.783	0.800	1.847	0.104	2.077	0.185
423	3.093	0.039	4.328	0.209	6.012	0.315	9.857	0.814	2.813	0.209	5.448	0.431
201	2.652	0.009	2.913	0.073	3.044	0.163	4.740	1.184	3.105	0.441	2.480	0.222
446	n.d		n.d		n.d		1.907	0.084	n.d		n.d	
461	n.d		n.d		n.d		n.d		n.d		n.d	
255	2.695	0.060	2.780	0.100	2.947	0.177	2.110	0.547	1.880	0.160	n.d	
404	2.983	0.020	4.211	0.902	6.986	2.727	19.432	0.662	11.814	9.061	n.d	

Line	Luteoxanthin (free)		Luteoxanthin (monoesters)		Luteoxanthin (diesters)		Auroxanthin (free)		Auroxanthin (monoesters)		Auroxanthin (diesters)	
	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M
204	37.544	0.044	37.890	0.029	38.095	0.104	n.d		37.549	0.019	n.d	
355	39.735	0.956	49.965	3.323	38.727	0.154	37.532	0.032	40.547	3.047	n.d	
354	48.677	3.984	45.212	3.939	n.d		40.124	2.624	n.d		n.d	
414	47.293	0.428	83.369	4.179	45.211	2.115	39.405	0.721	n.d		n.d	
440	n.d		n.d		n.d		n.d		n.d		n.d	
262	39.383	0.383	39.597	0.467	37.783	0.053	n.d		n.d		n.d	
423	44.447	0.742	49.731	0.503	38.654	0.480	n.d		n.d		n.d	
201	38.046	0.276	40.625	0.414	38.274	0.214	n.d		37.681	0.181	n.d	
446	37.591	0.091	n.d		n.d		n.d		n.d		n.d	
461	n.d		n.d		n.d		n.d		n.d		n.d	
255	38.309	0.405	38.098	0.327	37.541	0.041	n.d		n.d		n.d	
404	42.680	2.611	44.151	4.445	42.077	2.306	37.595	0.095	n.d		n.d	

Line	Phytoene		Zeaxanthin (free)		Zeaxanthin (monoesters)		Zeaxanthin (diesters)	
	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M	Mean (µg/g DW)	S.E.M
204	n.d		n.d		n.d		1.602	0.039
355	n.d		1.758	0.109	n.d		n.d	
354	n.d		1.951	0.165	n.d		2.287	0.455
414	n.d		n.d		n.d		n.d	
440	n.d		1.915	0.028	2.695	1.132	3.811	1.887
262	n.d		n.d		n.d		n.d	
423	n.d		1.728	0.021	n.d		n.d	
201	n.d		n.d		n.d		n.d	
446	n.d		n.d		n.d		n.d	
461	n.d		n.d		n.d		n.d	
255	n.d		n.d		n.d		n.d	
404	n.d		1.768	0.013	n.d		n.d	

Table 5.1. Carotenoid amounts in ripe fruit determined through HPLC-PDA on a reverse phase C-30 column. Three biological replicates per line were used. S.E.M: Standard error of the mean.

Table 5.1 shows carotenoid amounts in ripe fruit of this panel, determined through reverse phase HPLC-PDA with a C30 column. Of all the lines tested, only one had no detectable carotenoids at all: line 446. All other accessions contained at least one pigment detectable through this system. The most abundant carotenoids were violaxanthin and its mono- and diesters (reaching levels of 90 µg/g DW in line 440), and the violaxanthin epoxidation product luteoxanthin and its mono-esters (reaching 83 µg/g DW in line 414). Phytoene was not detected in any of the lines tested, as would be expected from the absence of *Psy-1* transcripts. Certain carotenoids were preferentially esterified with one or two fatty acids: overall, luteoxanthin mono-esters were more abundant than luteoxanthin di-esters, auroxanthin mono-esters were more abundant than di-esters. Conversely, antheraxanthin and zeaxanthin di-esters were present in more lines than their mono-esters were.

Figure 5.2. Carotenoid amounts in mature green and ripe fruit and amplification of *Psy-1* (*c2*) and *Ccs* (*y*) amplicons in ripe fruit from *Psy-1* positive lines of panel one (RH 0, RH 1, and RH 4), *c1* positive lines of panel two (404, 423, and 414), and *c1* negative lines from panel two (461 and 446). Amplicons are the results of rt-PCR reactions. Green bars: mature green fruit. Red/orange/yellow bars: ripe fruit. Red box: lines lacking functional copies of *y* and *c2*, but with a functional copy of *c1*. Blue box: fruit with non-functional copies of *y*, *c2*, and *c1*. Error bars: standard error of the mean. n= 3 biological replicates.

Syngenta provided accessions differing in their possession of a marker associated with the *c1* locus. Lines in this diversity panel were either yellow or cream coloured when ripe, and lacked functional copies of the other two loci influencing mature fruit colour in pepper (*y, Ccs*; and *c2, Psy-1* (Hurtado-Hernandez & Smith, 1985)). Amplicons designed to test the presence of *Psy-1* and *Ccs* transcripts gave negative results for all of the accessions in this panel. In

Figure 5.2, the amplicon amplifications of *Ccs* and *Psy-1* from selected lines are shown alongside pigment amounts from mature green and ripe fruit. These are compared with the results of *Ccs* and *Psy-1* positive lines from the panel described in Chapters Five and Six. This demonstrates that, in the absence of both *Ccs* and *Psy-1*, colour differences are clear between lines with and without copies of the *c1* locus. Therefore, as hypothesised by Hurtado-Hernandez and Smith (1985), a third factor influencing pepper colour remains to be identified.

5.2.2 Metabolite profiling

To investigate the way a lack of gene products associated with the *c1* locus influenced fruit metabolism more broadly, extracts from lines of the diversity panel were profiled with GC-MS.

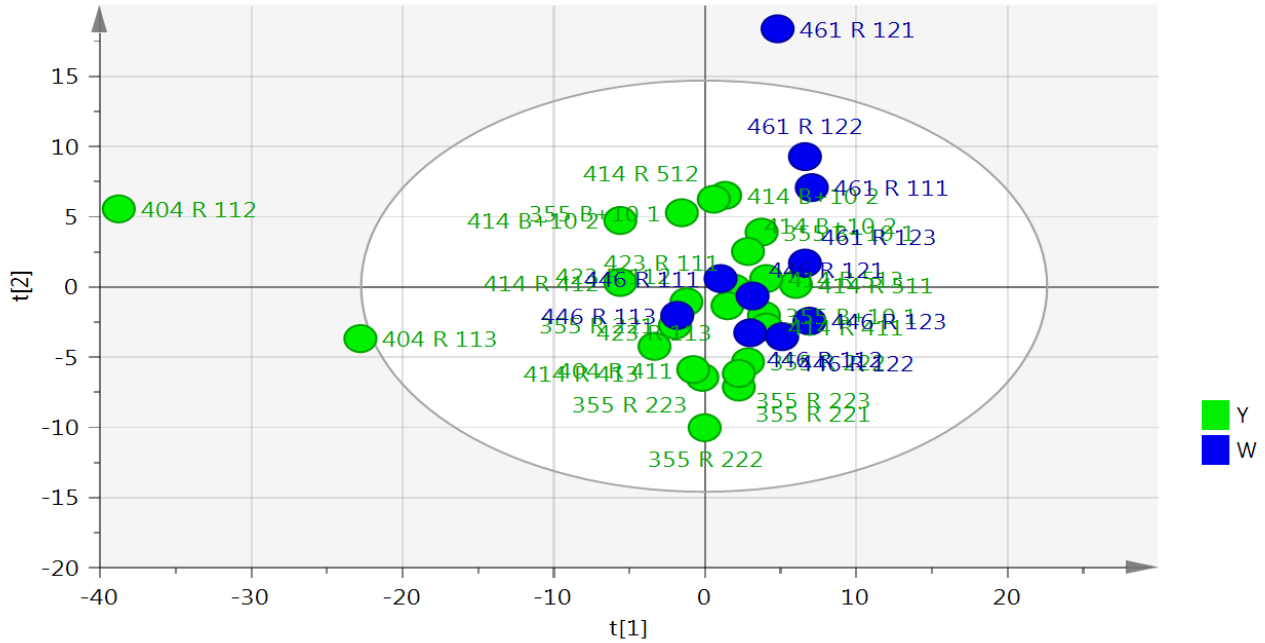


Figure 5.3.Principal Components Analysis (PCA) showing separation of accessions based on metabolite composition. Plots were drawn with SIMCA-P. (n=3 biological replicates)

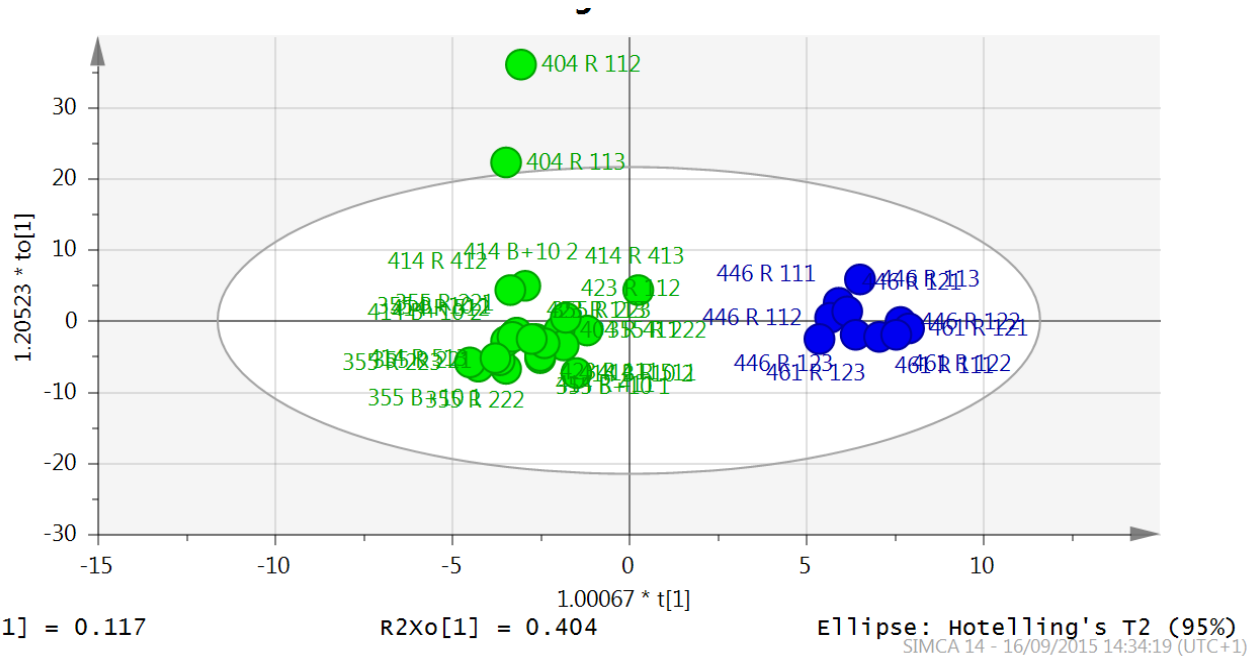


Figure 5.4. Partial Least Squares Discriminant Analysis (PLS-DA) showing separation of accessions based on metabolite composition. Plots were drawn with SIMCA-P. (n=3 biological replicates)

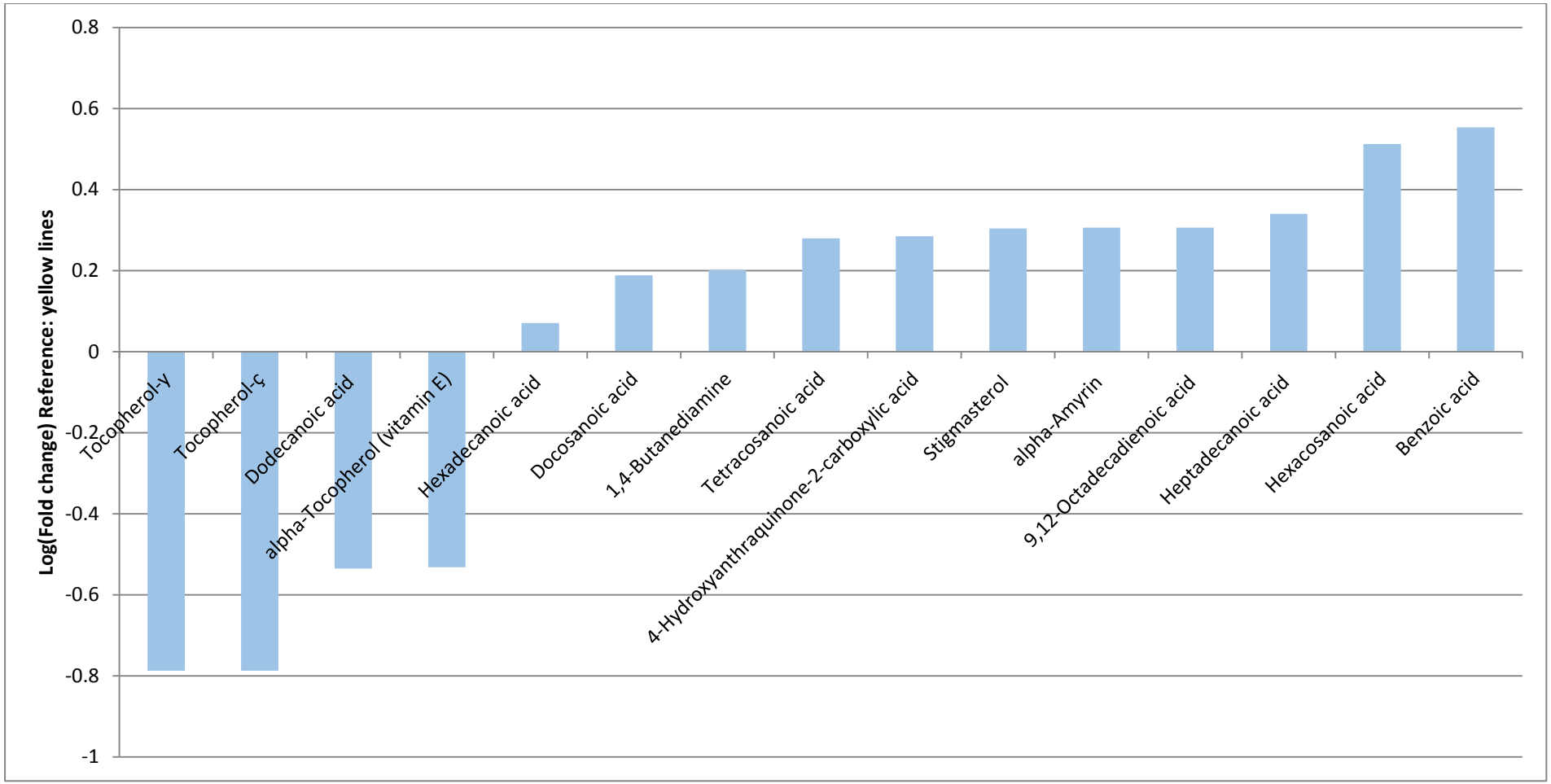


Figure 5.5. Fold changes of significantly different compounds on a log scale, where classes are assigned as yellow (possessing a functional copy of *c1*) and white (without a functional copy of *c1*). Yellow lines are the reference (above or below zero signifies deviation from the amount in yellow lines).

Through GC-MS profiling of accessions from the *c1* diversity panel 110 compounds were identified. When separation between the conditions based on metabolite content was investigated with Principal Components Analysis (PCA) (Figure 5.3), it was apparent that the accessions did not separate clearly with possession of a functional copy of the *c1* locus. The first component only explained 27.2% of the variation in the data. However, to discover which compounds were the main differences between *c1* possessing and non-possessing lines, the data was examined with Partial Least Squares Discriminant Analysis (PLS-DA). With this technique, separation between the samples was forced based on *c1* possession (Figure 5.4). Descriptive statistics show compounds which are significantly different between the two classes, and their fold changes. This information tells us which metabolites are associated with the difference in colour phenotype, and therefore may indicate the function of the gene associated with the *c1* locus.

Figure 5.5 shows the fold changes of compounds which are significantly different between white and yellow accessions. Of the fifteen compounds six are very long chain fatty acids (VLCFAs), three are tocopherols, two are sterols, are the remainder are a long chain fatty acid (LCFA), an amine, an anthraquinone and benzoic acid. There is a decrease in white lines in the content of one VLCFA (dodecanoic acid), however, all other VLCFAs are present in higher levels in white lines (Figure 5.5). This could be important for two reasons: firstly, fatty acid content could be influencing the structure of membranes in the plastids; and secondly, as carotenoids are only esterified by fatty acids with chain lengths of C18 or less, an increase in fatty acids with chain lengths of C20 or more could mean that these lines are less capable of esterification, which would have negative implications for their carotenoid storage ability.

5.2.3 Transmission Electron Microscopy (TEM)

As a lack of a functional copy of the *c1* locus results in an almost complete absence of pigment, a clear direction of investigation is to examine the plastid-located structures and substructures which ordinarily accumulate carotenoids. Mature green and ripe fruit sections were analysed with transmission electron microscopy (TEM) to gain insight into the impact of the *c1* mutation on pigment storage architecture. As any carotenoid in ripe fruit of lines with a functional copy of *c1* in this panel would be pigment retained by ripening plastids over the chloroplast to chromoplast transition, differences in mature green plastid structures could be equally important in determining final fruit colour.

5.2.3.1 Chloroplast structure (RH 15)

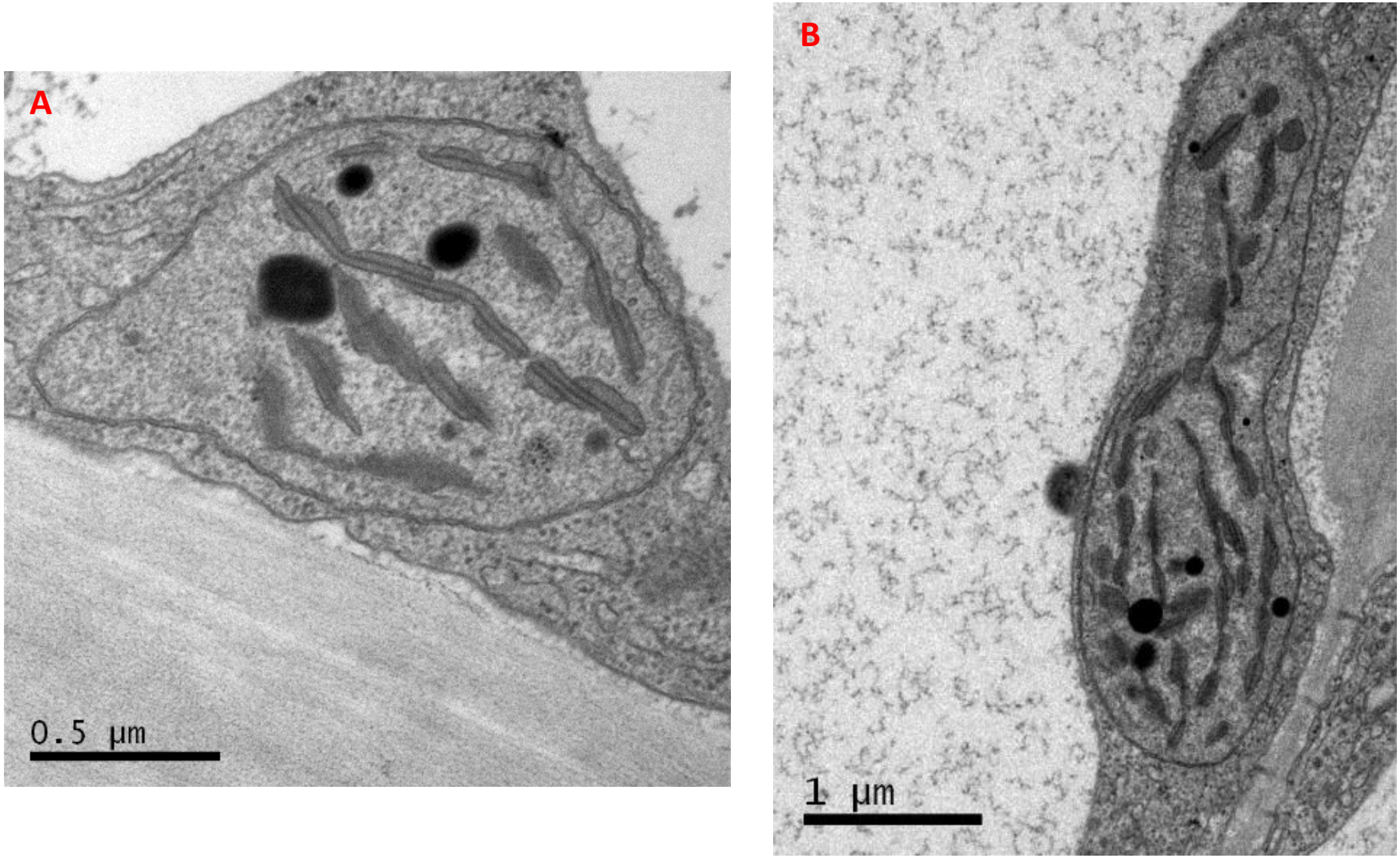


Figure 5.6. Whole chloroplasts from a *c1* negative line (RH 15). Fruit were sectioned before being fixed and counterstained. Sections were taken from three fruit, and chloroplasts A and B are representative.

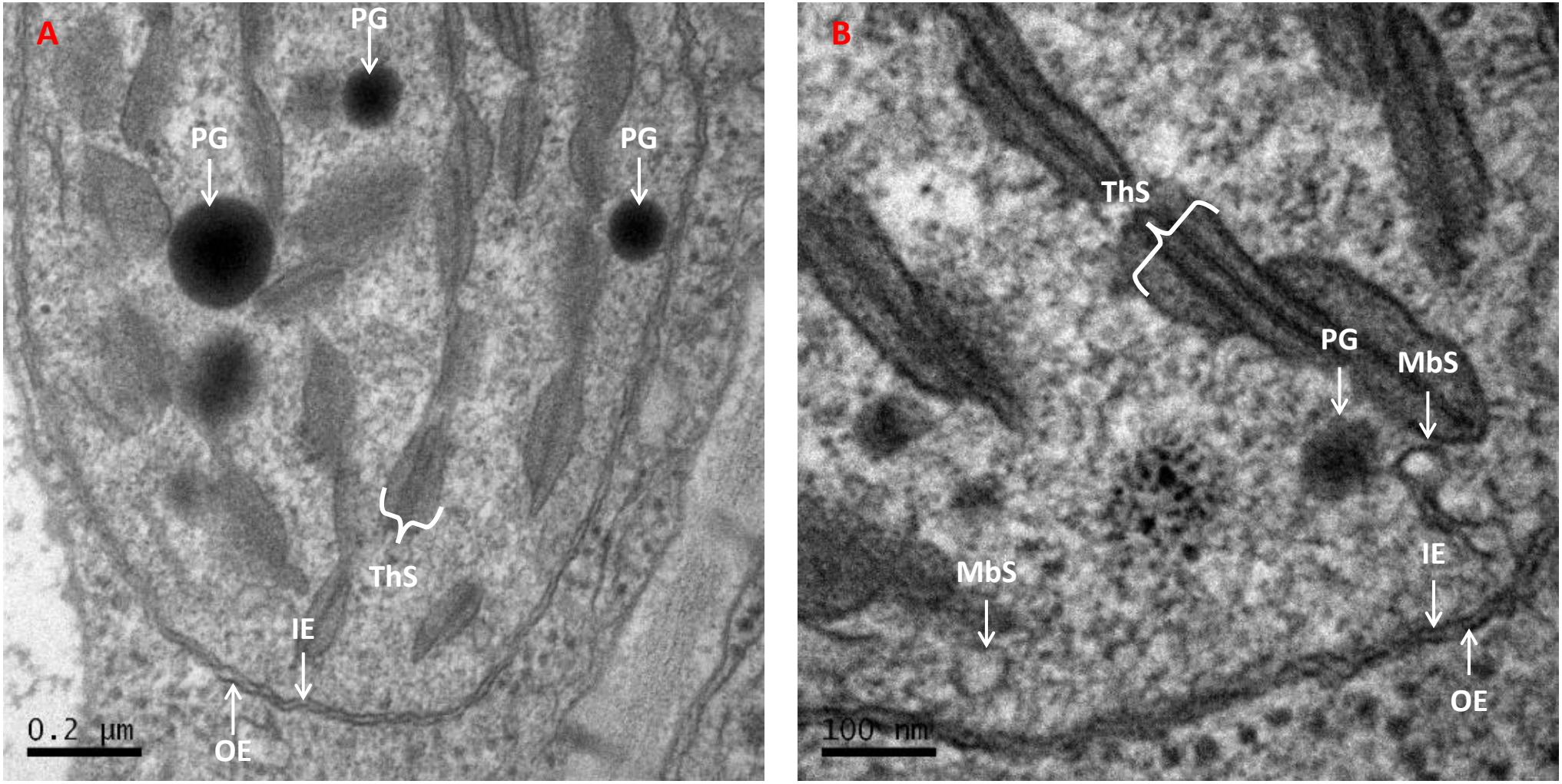


Figure 5.7. Chloroplast sections from a *c1* negative line (RH 15). Three fruit were sectioned and visualised with TEM. A and B are representative sections. PG: plastoglobule. ThS: thylakoid stack. MbS: membranous sac. IE: inner chloroplast envelope. OE: outer chloroplast envelope.

5.2.3.2 Chromoplast structure (RH 15)

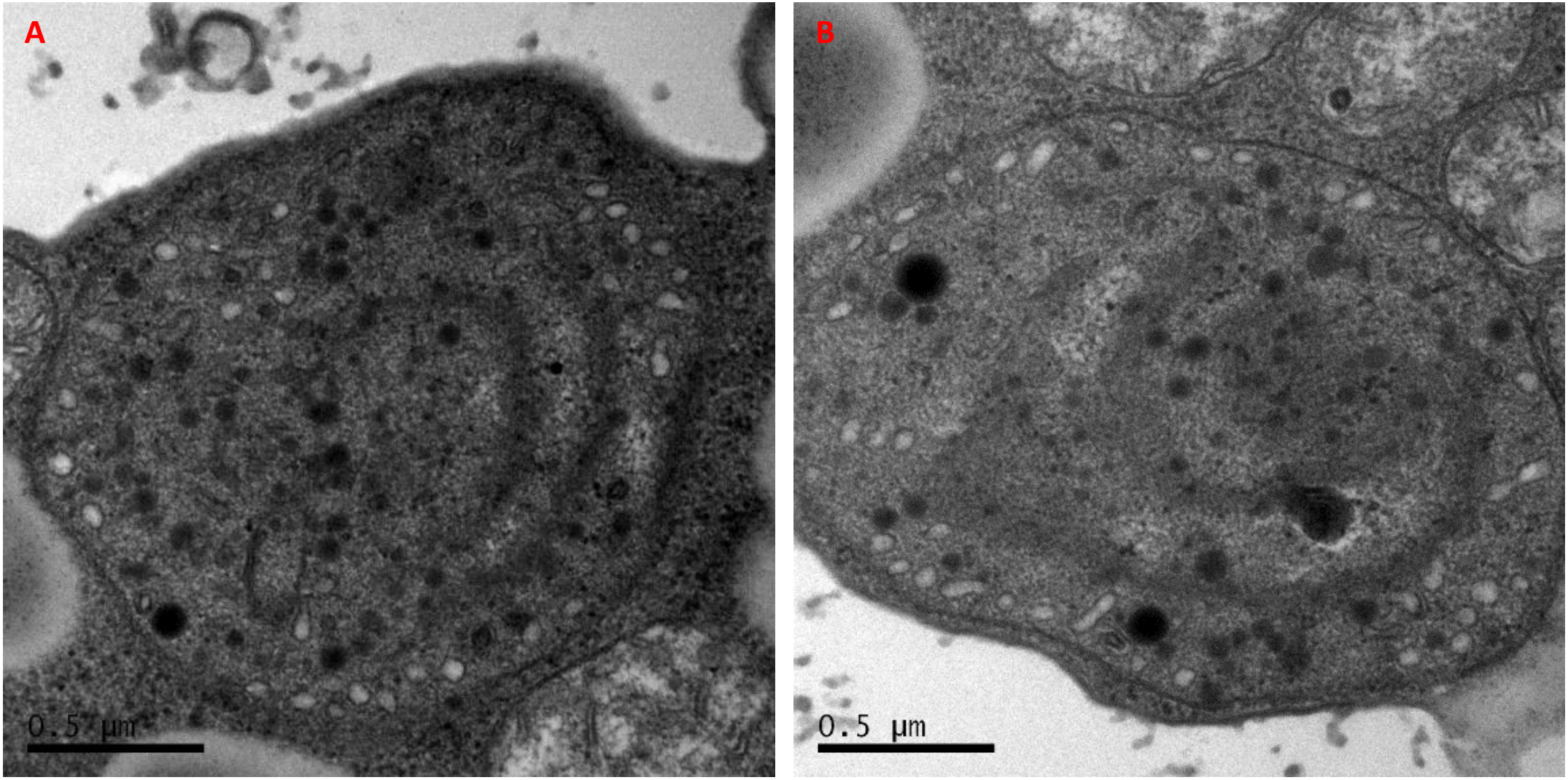


Figure 5.8. Whole chromoplasts from a *c1* negative line (RH 15). Fruit were sectioned before being fixed and counterstained. Three fruit were sectioned and visualised with TEM. A and B are representative plastids.

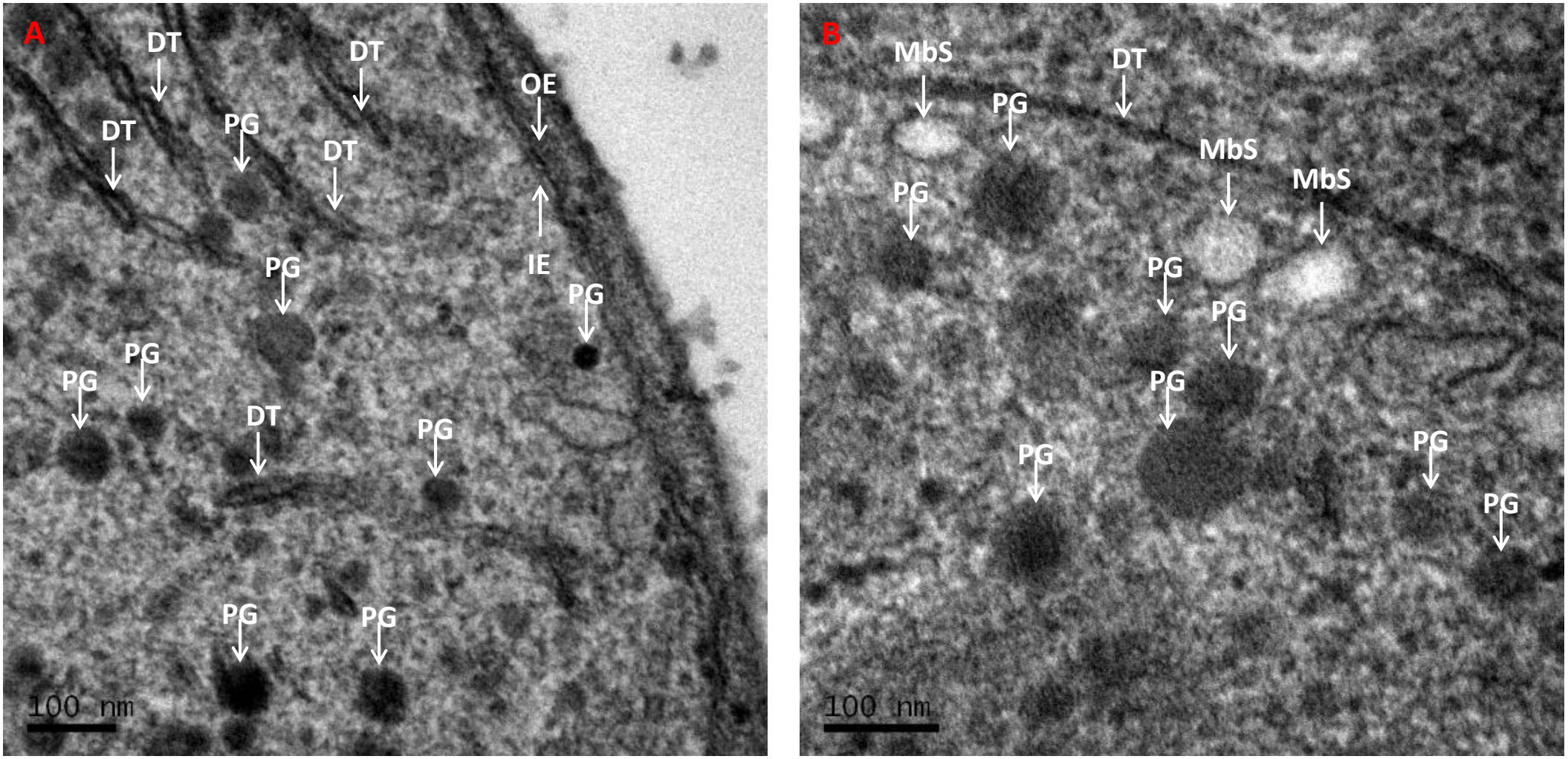


Figure 5.9. Chromoplast sections from ripe fruit of a *c1* negative line (RH 15). Three fruit were sectioned. A and B are representative. PG: plastoglobule. DT: decomposing thylakoid. MbS: membranous sac. IE: inner chloroplas envelope. OE: outer chloroplast envelope.

Chloroplast Structure (414)

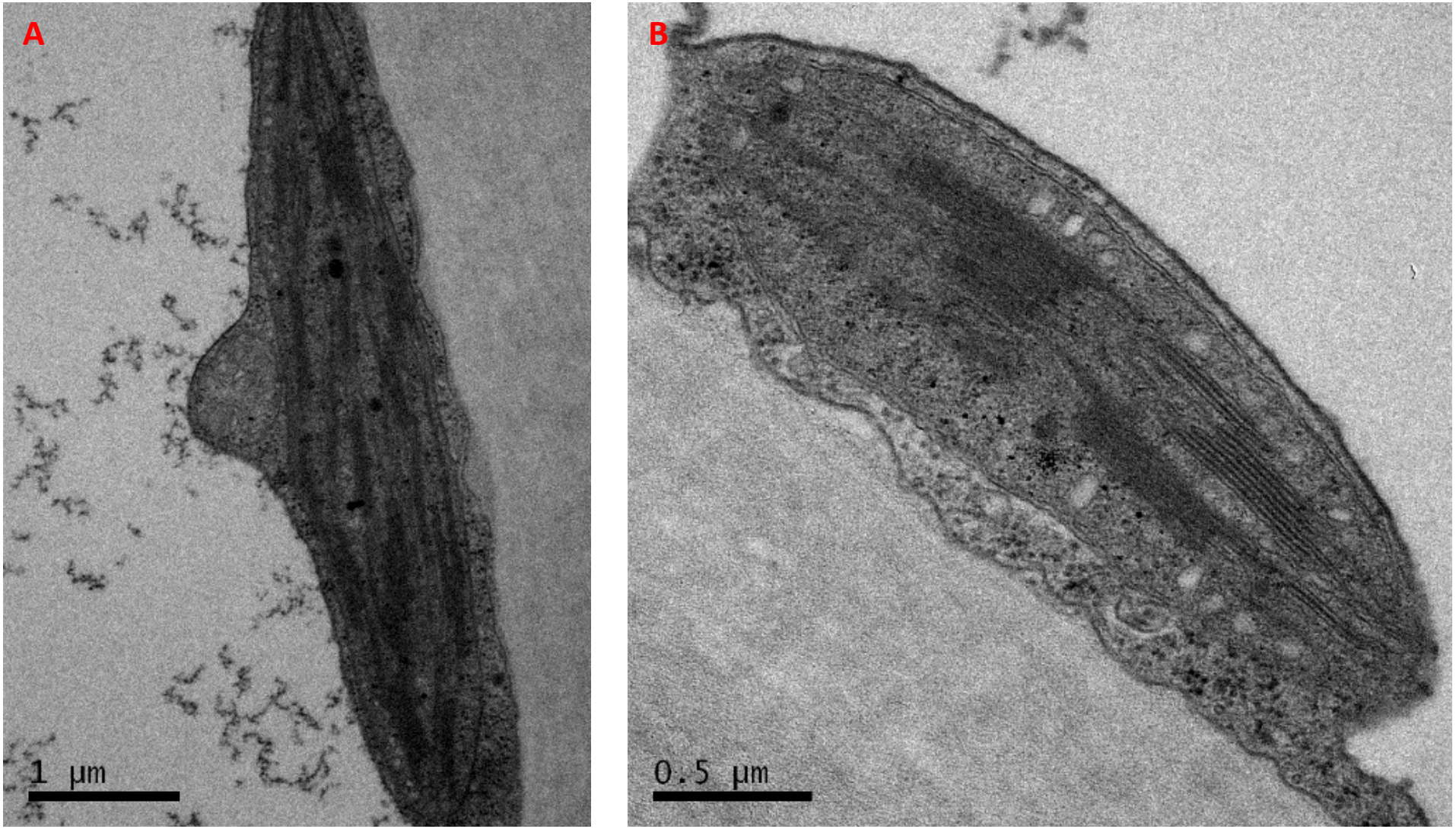


Figure 5.10. Whole chloroplasts from a *c1* positive line (414). Fruit were sectioned before being fixed and counterstained. Sections were taken from three fruit, and chloroplasts A and B are representative.

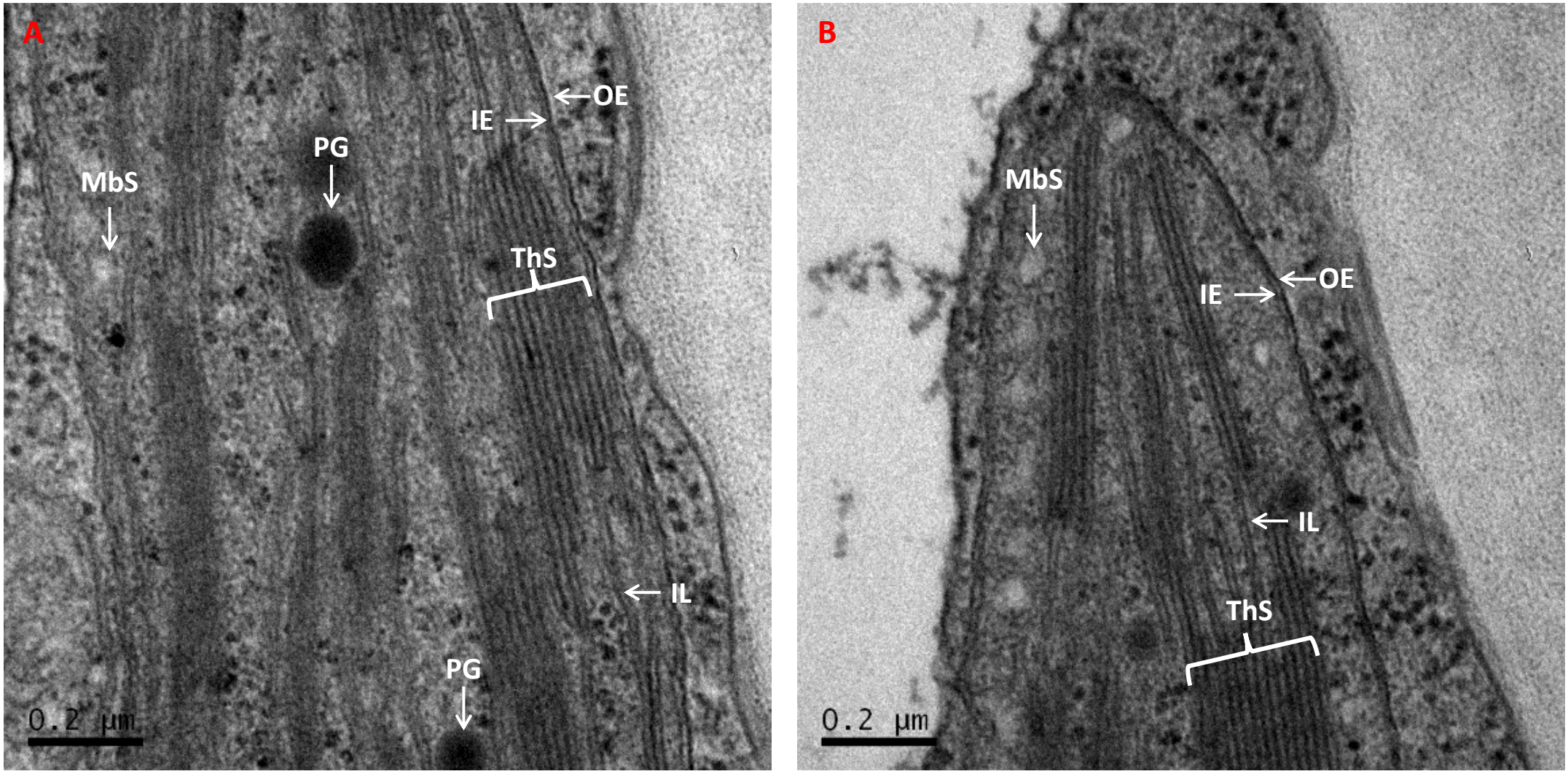


Figure 5.11. Chloroplast sections from a *c1* positive line (414). Three fruit were sectioned and visualised with TEM. A and B are representative sections. PG: plastoglobule. ThS: thylakoid stack. MbS: membranous sac. IL: intergranal lamella. IE: inner chloroplast envelope. OE: outer chloroplast envelope.

5.2.3.3 Chromoplast Structure (414)

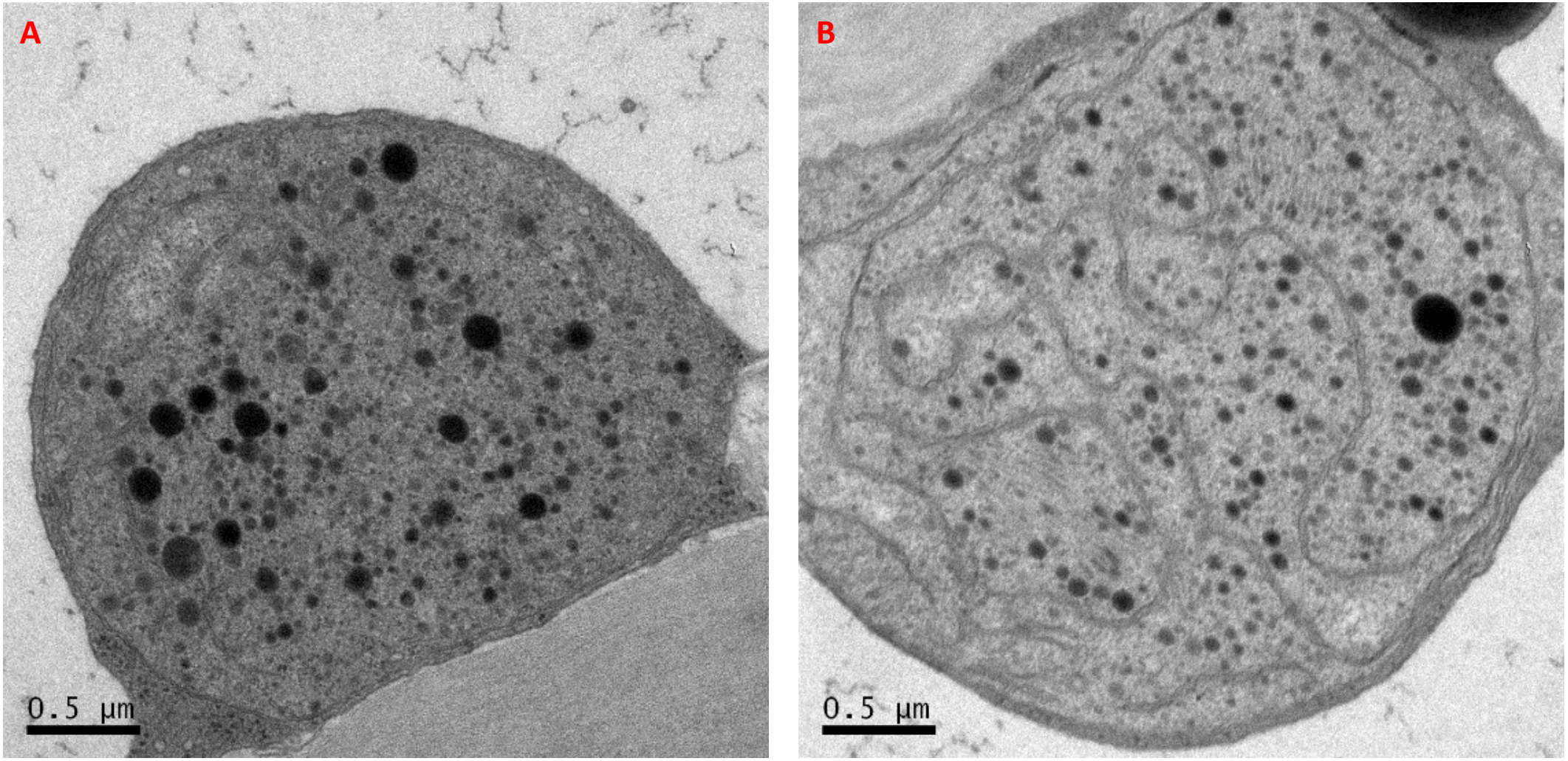


Figure 5.12. Whole chromoplasts from a *c1* positive line (414). Fruit were sectioned before being fixed and counterstained. Three fruit were sectioned and visualised with TEM. A and B are representative plastids.

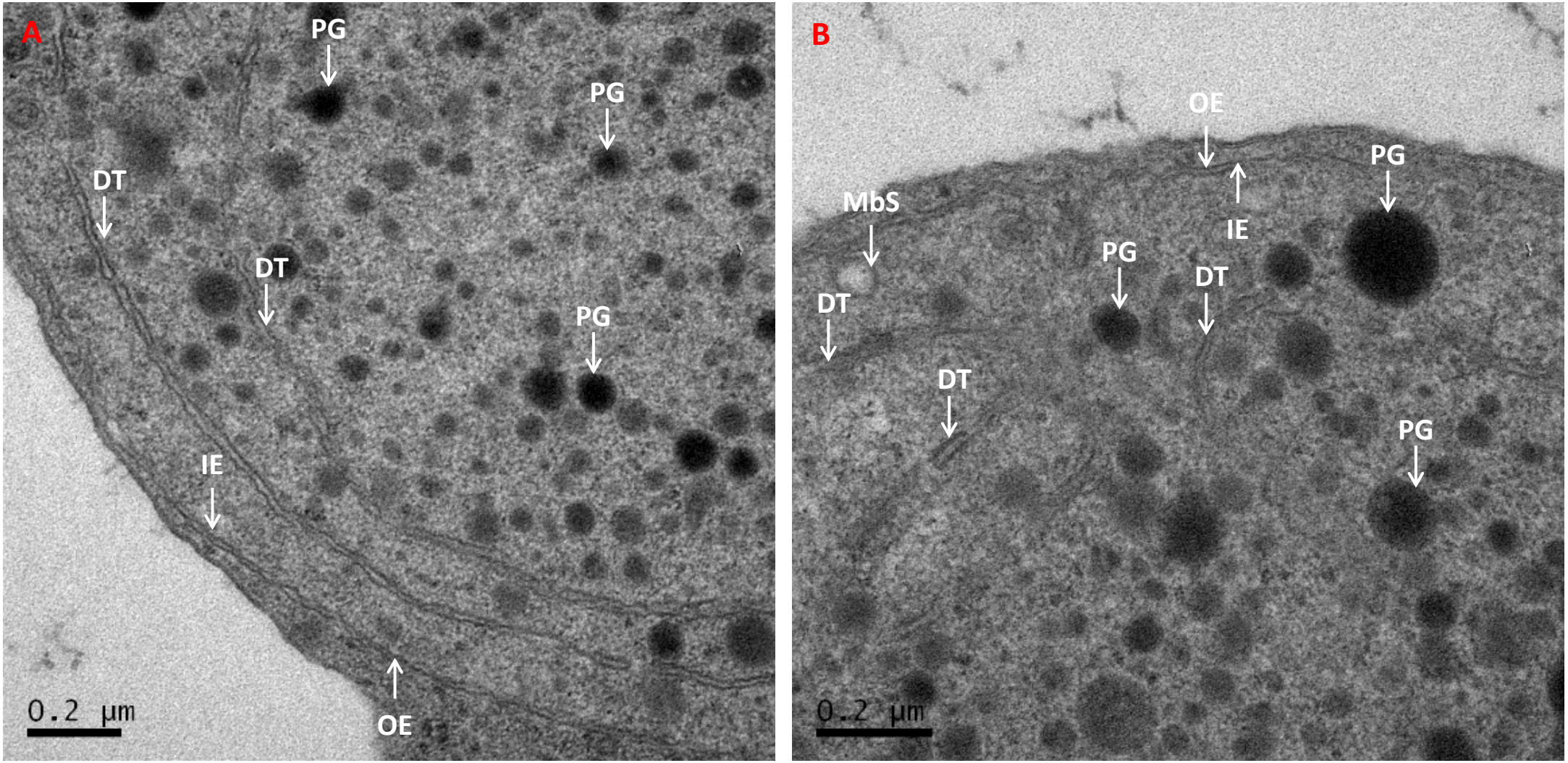


Figure 5.13. Chromoplast sections from ripe fruit of a *c1* positive line (414). Three fruit were sectioned. A and B are representative. PG: plastoglobule. DT: decomposing thylakoid. MbS: membranous sac. IE: inner chloroplas envelope. OE: outer chloroplast envelope

Transmission electron microscopy (TEM) images of chloroplasts from lines with and without a functional copy of the *c1* locus reveal interesting structural differences: chloroplasts in the *c1* negative line (RH 15) have much wider thylakoid lumens, than lines which possess a functional copy of the gene associated with the locus and grana containing only two thylakoid sacs (Figure 5.6, Figure 5.7). Chloroplasts from mature green fruit of the *c1* positive line 414 appeared typical: thylakoids had narrow lumens, and narrow intergranal lamellae connected stacks of grana containing at least five thylakoid sacs.

The images of ripe fruit chromoplasts from the two types of accession also revealed differences: *c1* negative chromoplasts contained many more membranous sacs, fewer and larger plastoglobuli, and fewer decomposing thylakoids than those from *c1* positive fruit.

5.2.4 Sub-chromoplast Fractionation: *c1* negative chromoplasts

In order to gain a better understanding of sub-organelle organisation in the absence of pigment, sub-chromoplast fractionation was carried out on ripe fruit lacking a functional copy of the *c1* locus. Chromoplasts from a *c1* negative line (RH 15) were extracted, broken and suspended on a stepwise sucrose gradient. Millilitre fractions from the gradient were collected, from which carotenoids and proteins were extracted. Carotenoids were analysed with HPLC.

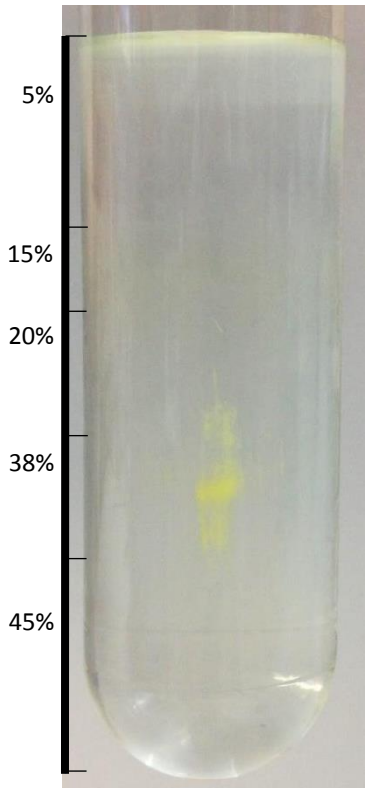


Figure 5.14. Isolated and broken chromoplasts from white ripe fruit (RH 15) suspended on a stepwise sucrose gradient. Sucrose percentage steps are shown to the left of the gradient.

HPLC-PDA analysis showed that no pigments were detectable at any step of the gradient. Integrating chromatograms at 450nm gave no recognisable peaks.

5.3 Discussion

5.3.1 Effect of the *c1* locus on carotenoid composition

Lines which lack both *Ccs* and *Psy-1* transcripts, but which possess a functional copy of the *c1* locus contain xanthophylls in both mature green and ripe fruits (

Figure 5.2). Some lines lacking a functional copy off the *c1* locus contain trace levels of lutein, and free violaxanthin and antheraxanthin. As with the previous panel (detailed in Chapters Three and Four), lines possessing *c1* only (the *Psy-1* negative lines from panel one) contain similar levels of pigment at mature green and ripe stages, due to the fact that no pigment is synthesised *de novo* by PSY-1 as the fruit ripen. Like yellow-fruited lines in panel one, the *c1* positive lines in panel two contain an array of xanthophyll mono- and di-esters, including those of violaxanthin, antheraxanthin, neoxanthin, luteoxanthin, auroxanthin, and zeaxanthin. This supports the hypothesis that xanthophyll esterification is crucial for the storage of xanthophylls in mature pepper fruits. It could be the case that the lack of esterification capability in lines which produce white fruit contributes to their inability to retain pigments throughout the course of ripening (Table 5.1).

5.3.2 Effect of the *c1* locus on broader fruit metabolism

This is the first study to profile metabolites in carotenoid-deficient sweet *Capsicum* fruits. Previous studies have focussed on coloured and pungent cultivars. Examples of metabolomics studies on other colourless fruit species are also lacking in the literature, though carotenoid, tocopherol, sugar and organic acid content has been profiled in white bell pepper previously (Matsufuji, Ishikawa, Nunomura, Chino, & Takeda, 2007). However, white peppers used by Matsufuji et al. (2007) were not mature fruit; they belonged to a variety which were white before the onset of ripening and therefore are not directly comparable to the white ripe fruit used in the current work.

Previous metabolomics studies in pepper have focussed on comparing climacteric and non-climacteric fruit species. For example, Osorio et al integrated transcriptomic and metabolomics data, specifically to compare tomato to a non-climacteric member of the Solanaceae family. Some *Capsicum* species are climacteric, but *C. chinense*, which is not, was used in this study (Osorio et al., 2012). A similar comparison is made in (Klie et al., 2014). They use the metabolomics platform to establish the fact that a few important metabolites separate climacteric from non-climacteric ripening fruit. These include the three amino acids of the Asp family (Asp, Ser, and Thr) and malate.

(Wahyuni, Ballester, Sudarmonowati, Bino, & Bovy, 2011)– Metabolite profiling 32 *Capsicum* accessions, looking at health related compounds and linking to potential implications for breeding programmes. This is a good example of metabolite analysis being useful for breeders and therefore enhancing economic value of the crop.

In the current investigation, it was found that accessions lacking a functional copy of the *c1* locus accumulated more very long chain fatty acids (VLCFAs) and tocopherols. The result of accumulation of VLCFAs is particularly interesting in the light of the altered membrane structures seen in TEM images of plastids from *c1* negative fruit. The fact that the thylakoid envelopes in *c1* negative lines could contain different ratios of fatty acids may contribute to the explanation of their aberrant shape and stacking behaviour. The accumulation of VLCFAs could also be related to a lack of xanthophyll esterification capability: fatty acids with chains longer than C18 are not observed as the chains added to the hyroxy groups of xanthophylls. If fatty acids with chain lengths longer than this are being accumulated by these lines, esterification during the course of ripening could be affected.

5.3.3 Insights into chloroplast structure from TEM

Accessions which did have a functional copy of the *c1* locus had ordinary structures in mature green fruit. Lines without the *c1* locus possess paler mature green fruit and accumulate lower carotenoid levels at this stage. The images generated by TEM show that fruit which are carotenoid-deficient when they ripen possess atypical chloroplast structures at the mature green stage. Most notably, the plastids contain thylakoids with enlarged lumens, and grana with no more than three thylakoids per stack (Figure 5.6, Figure 5.7).

Alterations in some genes controlling chloroplast structure in *Capsicum* have been characterised previously. For example; mutations in *Golden-2 like* transcription factors (GLKs) have profound implications for membrane-bound structures in the plastid. The GLKs were first described in maize over 90 years ago (Jenkins, 1926), and homologues have subsequently been identified in taxonomically diverse plant species: *Arabidopsis*, rice (*Oryza sativa*), moss (*Physcomatrella patens*), tomato (*Solanum lycopersicum*), and *Capsicum* (Brand et al., 2014; Bravo-Garcia, Yasumura, & Langdale, 2009; Fitter, Martin, Copley, Scotland, & Langdale, 2002; Powell et al., 2012). *GLK* transcription factors function in differentiation of photosynthetic cell types in both C_3 and C_4 types, but also have a sub-cellular differentiation role in controlling chloroplast development (Langdale & Kidner, 1994; Rossini, Cribb, Martin, & Langdale, 2001). It has been demonstrated that GLKs control levels of galactolipids synthesised and incorporated into thylakoid membranes; galactolipid-synthesis genes are under direct transcriptional control from *GLK*s (Kobayashi et al., 2014). In *Capsicum* and tomato, two *GLK* orthologues are present: *GLK1* and *GLK2*. It was demonstrated in

tomato that *GLK1* is expressed mainly in vegetative tissues such as the leaf, whereas GLK2 expression fruit localised; the same was found to be true in pepper (Brand et al., 2014; Nguyen et al., 2014). Mutations in a fruit-specific copy of a *GLK2* result in a very pale, almost white mature green fruit phenotype. It was found that *GLK2* mutants possessed fruit chloroplasts with dramatically altered compartment size, similar to the phenotype observed in the present work (Brand et al., 2014). However, when *GLK2* expression was tested in *c1* positive and negative lines, no quantitative difference between the lines was observed, therefore, it is unlikely that this gene is a potential candidate for association with the *c1* locus. However, sequencing studies could be carried out to verify this conclusion as tests for expression do not reveal point mutations. If the mutation for c1 was a point mutation, it would not have been found with this method.

Ruling out *GLK2* as a candidate for the *c1* locus does not exclude genes which are under its transcriptional control. Thylakoid biogenesis is dependent on a supply of galactolipids, mono- or digalactosyldiacylglycerol (MGDG and DGDG), which account for 50 and 25% of plastid lipids respectively. They are synthesised by two pathways in the plant cell: the prokaryotic pathway (localised entirely in the plastid) and the eukaryotic pathway (involving some steps in the endoplasmic reticulum) (Holzl & Dormann, 2007; Whitaker, 1986). Plants possessing mutations in the gene encoding the enzyme catalysing the final step in the galactolipid biosynthesis pathway (taking place in the thylakoid membrane), *MGD synthase 1* (*MGD1* and *MGDG1*), are either albino (with seedlings germinating only in the presence of sucrose) or have seriously compromised chloroplast development and unusual thylakoid structures (Jarvis et al., 2000; Kobayashi, Kondo, Fukuda, Nishimura, & Ohta, 2007). Digalactosyldiacylglycerol synthesis is also crucial: mutations in the gene encoding the final step in its biosynthesis have altered lipid trafficking and development of photosynthetic complexes (Kobayashi et al., 2013; Mizusawa & Wada, 2012). However, the altered sub-plastid structures do not resemble those observed in the *c1* diversity panel: thylakoid lumen diameter remains unaffected. Most investigations into galactolipid biosynthesis have focussed on *Arabidopsis*; examples of galactolipid biosynthesis mutations in non-leaf tissues are sparse, therefore implications for the chloroplast-chromoplast transition and subsequent chromoplast morphology are unknown. Amplicons to *MGDG1* and *2* and DGDG1 and 2 were designed and tested in c1 positive and negative lines, and no differences in amplification level were observed.

Additionally, the *lutescent1* and *2* (*lut1* and *2*) mutations in tomato alter chloroplast development, but have yet to be characterised in pepper (Barry et al., 2012). Phenotypes of the thylakoids in tomato *lut* mutants are different to those observed here, and the mutations (despite delaying ripening) do not prevent development of a functional chromoplast and the associated carotenoid

accumulation (Barry et al., 2012). However, it is conceivable (though unlikely) that the protein could have a slightly different role in pepper.

5.3.4 Insights into chromoplast structure from TEM and sub-chromoplast fractionation

Examples of chromoplasts from carotenoid-deficient fruit and flowers are sparse in the literature. This panel of yellow and white peppers therefore provided a unique opportunity to investigate structures in the organelle adapted specifically for carotenoid accumulation in the absence of detectable carotenoids.

5.4 Conclusion

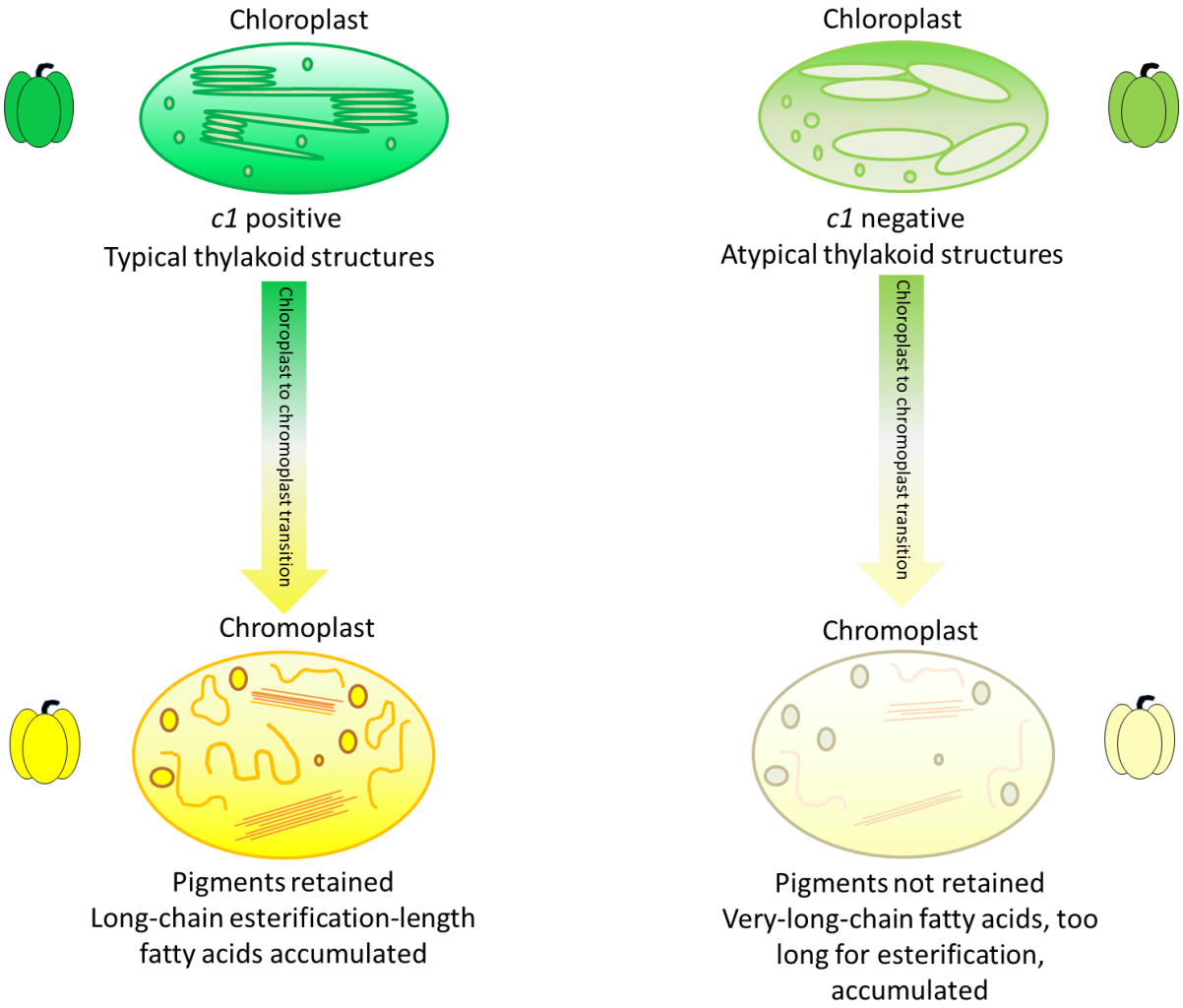


Figure 5.15. Model summarising the differences between chloroplast and chromoplast structure between *c1* positive and negative lines, developed with data from TEM, HPLC, and GC-MS.

To conclude, it is clear that when they are also *y* and *c2* deficient, accessions without a functional copy of the *c1* locus do not retain pigments during the chloroplast to chromoplast transition to the same degree as those which possess one (summarised in Figure 5.15). Amplifying cDNA fragments of

Ccs (*y*) and *Psy-1* (*c2*) and combining the information with carotenoid amounts determined through HPLC confirmed this. Investigating the structures of the plastid before the onset of ripening and when the fruit were mature gave some insight into the reasons for loss of the pigments: thylakoid sacs were different in shape from those in *c1* positive lines, possessing much wider lumens. They also did not form thick grana stacks like those observed in lines of the opposite condition. Subjecting samples from *c1* positive and negative accessions to GC-MS revealed striking differences in fatty acid composition, with higher amounts of VLCFAs being found in *c1* negative accessions. This could relate to the aberrant structures observed in their chloroplasts, or could be related to carotenoid esterification. In the absence of carotenoids to be esterified, diversion of LCFAs from fatty acid elongation to xanthophyll esterification would not occur, therefore VLCFAs would accumulate. However, the converse could also be true: xanthophylls are not esterified because fatty acids of the correct chain length are not accumulated. It could be speculated that lack of esterification capability is the causing the loss of carotenoids between mature green and ripe stages. Further investigation into acyltransferase activity in lines accumulating VLCFAs would be necessary to verify this hypothesis.

6 Transcriptome analysis to identify candidates for association with the *c1* locus

–

6.1 Introduction

Three lines were selected for analysis with RNA-seq, based on fruit phenotypes at the mature green and ripe stages. These lines were 414, 255, and 446.

Figure 6.1. Phenotpyes of mature green and ripe fruit of the lines selected for transcriptome profiling with RNA-seq. Three biological replicates were used for each line at each stage.

Using the Cufflinks software package data were processed and differentially expressed transcripts were generated (Trapnell et al., 2012). From this information, differentially expressed genes were narrowed down to those in the region of the genome were markers for the *c1* locus have been identified by Syngenta. Based on differential expression and function of their translated proteins, genes were identified for potential further analysis. These were: the lipid transfer protein *GDL62*, the NAC transcription factor *NAC37*, the *Capsicum*-specific acyl transferase *Catf-2*, and the plastid specific kinase *ABCK*. Furthermore, investigations into gene expression levels in pathways in different areas of metabolism related to colour phenotype: carotenoid metabolism, photosynthetic metabolism, and fatty acid metabolism. These revealed differences between the line possessing a functional copy of the *c1* locus and the two lines that did not which support the hypothesis that the gene associated with the c1 locus influences the development of plastid compartments.

6.2 Results

6.2.1 Validation of the RNA-seq method with qPCR

In order to confirm that transcript levels produced with RNA-seq were accurate, selected genes were tested with qPCR for quantification against a standard curve. Relative amounts were compared to relative amounts of the same transcripts quantified with RNA-seq. Trends in relative amount were similar from both methods of quantification, meaning that accurate conclusions can be drawn from the large-scale transcriptome study (Figure 6.2).

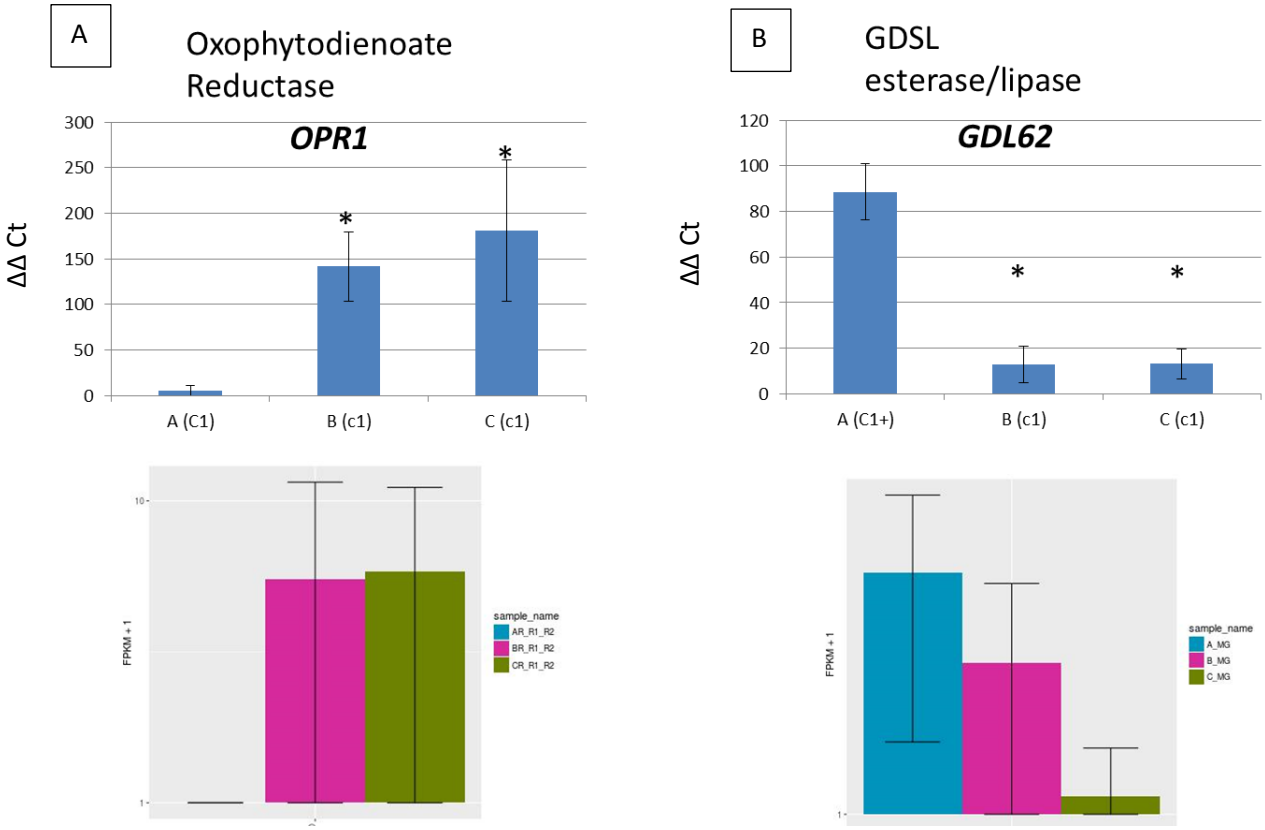


Figure 6.2. Transcript levels of (A) *OPR1* and (B) *GDL62* quantified with RT-qPCR (n=3 biological replicates), and RNA-seq (n=3 biological replicates). * = p-value < 0.05 from Student's t-test. For qPCR reactions, transcript levels were tested relative to a housekeeping gene with stable expression (*ATP-alpha* (*ATPA*)). ΔΔCt values are relative to *ATPA* expression.

6.2.2 Global transcriptomic statistics

6.2.2.1 Mature green fruit

Transcripts were identified in mature green fruit samples, of which 2946 were infinitely different between lines 414 and 255 (1794 being higher in line 414, and 1152 higher in 255), 2909 between 414 and 446 (1423 higher in 414 and 1487 in 446), and 1221 between 255 and 446 (394 higher in 255 and 828 in 446). Transcript levels were defined as infinitely different when a transcript was completely absent from one condition. Volcano plots show the difference between fold change and significance (Figure 6.3). From these it can be seen that, above the significance level assigned during analysis with the Cuffidiff module of the Cufflinks package (defined on the y-axis as $-\log_{10}$(p value) of

above 2.5 using default Cufflinks parameters), significance is unaffected by fold change. Fold change is on the x-axis, and it can be seen from the part of the plot above p-values of 2.5 (denoted on the y-axis) that points ranging from zero to over 20 are all significant, although most points lie between logged fold changes of -10 and 10. In the lower part of the plot, below the significance level of 2.5, there is a greater range of points between -10 to -20 and below, and 10-20 and above. Transcripts with a fold change of close to zero are equally likely to be significantly different between samples as transcripts with infinite fold changes. This highlights the sensitivity of the technique.

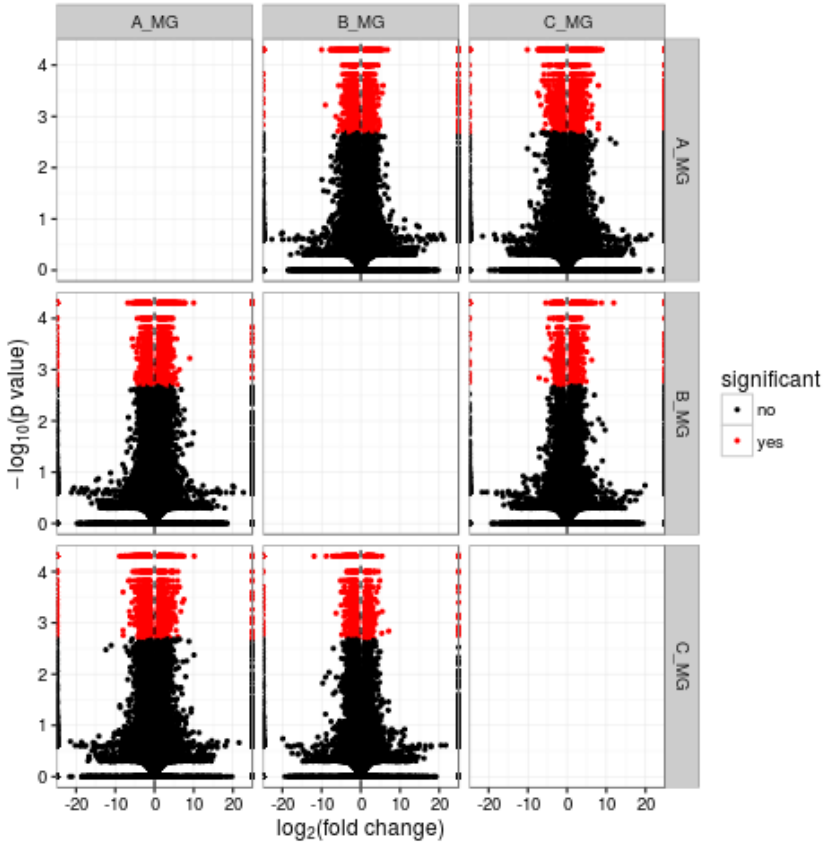


Figure 6.3. Volcano plots showing the relationship between fold change and significance for mature green samples. The plot was generated with the R package CummeRbund (n=3 biological replicates)

6.2.2.2 Ripe fruit

Transcripts were detected in ripe fruit, 319 of which were infinitely differentially expressed between lines 414 and 255 (227 being higher in 414, and 92 in 255), 306 between 414 and 446 (246 being higher in 414, and 60 in 446), and 122 between lines 255 and 446 (74 being higher in 255, and 49 being higher in 446).

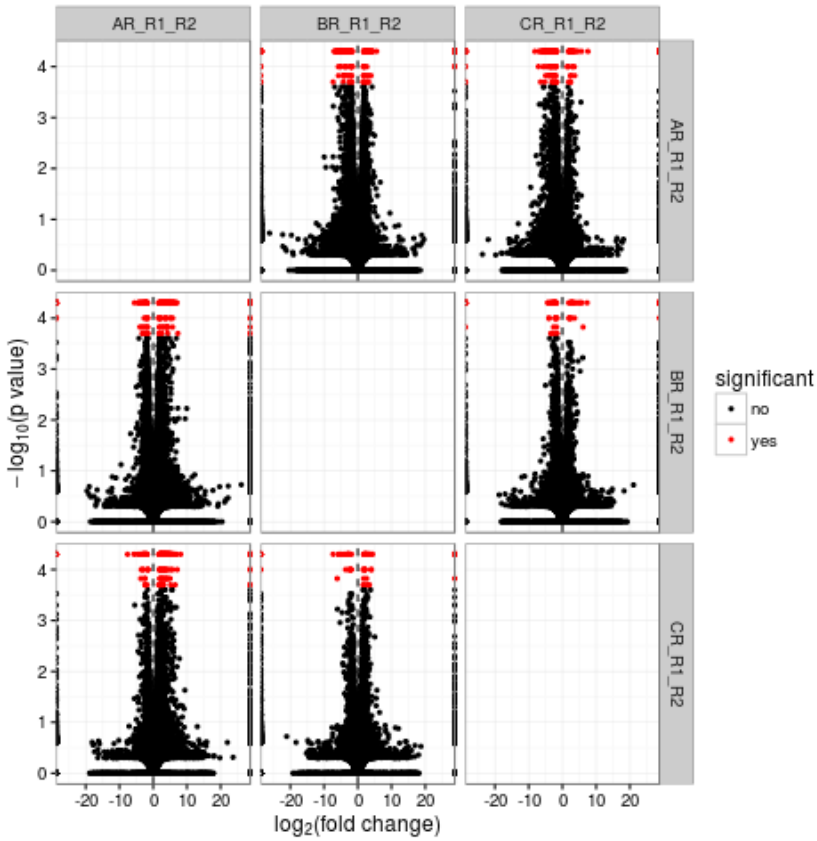


Figure 6.4. Volcano plots showing the relationship between fold change and significance for ripe fruit samples. The plot was generated with the R package CummeRbund (n=3 biological replicates).

The depiction of fold change and significance in the volcano plots for ripe fruit samples shows that, there are many fewer significantly different genes than at the mature green stage (Figure 6.4).

.

6.2.3 Differentially expressed transcripts

6.2.3.1 Mature green fruit

6.2.3.1.1 Biological functions of differentially expressed genes

Differentially expressed transcripts which were infinitely more highly expressed in either the *c1* positive line (414) or the two *c1* negative lines (225 and 446) were identified, and their corresponding sequences were extracted from the genome (*C. annuum* cv. *Zunla* (Qin et al., 2014)). Only infinitely differentially expressed genes were selected to limit the number of genes in the analysis. It is probable that to have such a profound effect on fruit phenotype as a large colour change, the change in gene expression would be large. However, it is not definite, and further analysis on genes with smaller expression level changes could be carried out in future. To investigate the biological significance of differences between transcriptomes of *c1* positive and negative fruits, the sequences were functionally annotated with Blast2GO, which blasted them against a vascular plant database from NCBI (Conesa et al., 2005; Conesa & Götz, 2008; Götz et al., 2008).

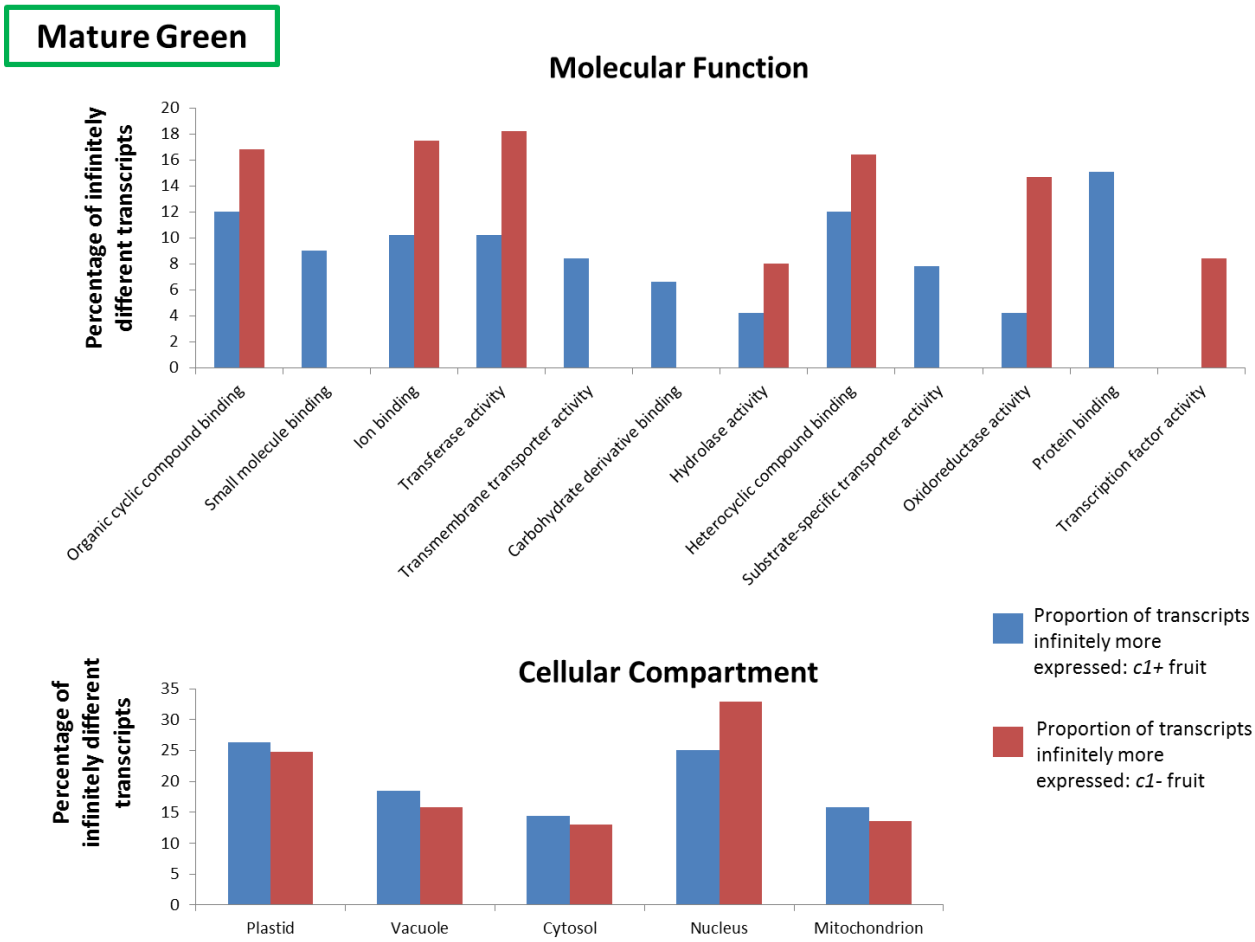


Figure 6.5. Annotations of infinitely different transcripts from *c1* positive and negative lines, generated and assigned to specific molecular functions and cellular components with Blast2GO. (n=3 biological replicates.)

Investigating the molecular functions of differentially expressed genes in *c1* positive and negative fruit revealed a wide range of biological roles for transcripts which are infinitely more expressed in one of the two conditions. In mature green fruit, several molecular functions were only present in the set of infinitely more expressed genes in *c1* positive fruit (Figure 6.5). Possessing a functional copy of the *c1* locus causes an infinite increase in expression levels of genes involved in small molecule binding, transmembrane transporter activity, carbohydrate derivative binding, substrate-specific transporter activity, and protein binding. This wide range of molecular functions influence a huge number of biological processes in the fruit. Only one molecular function was present in the set of genes infinitely more expressed in *c1* negative fruit; this was transcription factor activity. There were several molecular functions with transcripts in the infinitely different sets for both conditions, these include: organic cyclic compound biding, ion binding, transferase activity, hydrolase activity, heterocyclic compound binding, and oxidoreductase activity. Interestingly, for all of these, a greater proportion was found in the *c1* negative fruit. Differences in cellular compartments for differentially expressed transcripts from the two conditions were less marked than those for molecular function (Figure 6.5).

6.2.3.1.2 Differentially expressed genes within the *c1* marker region

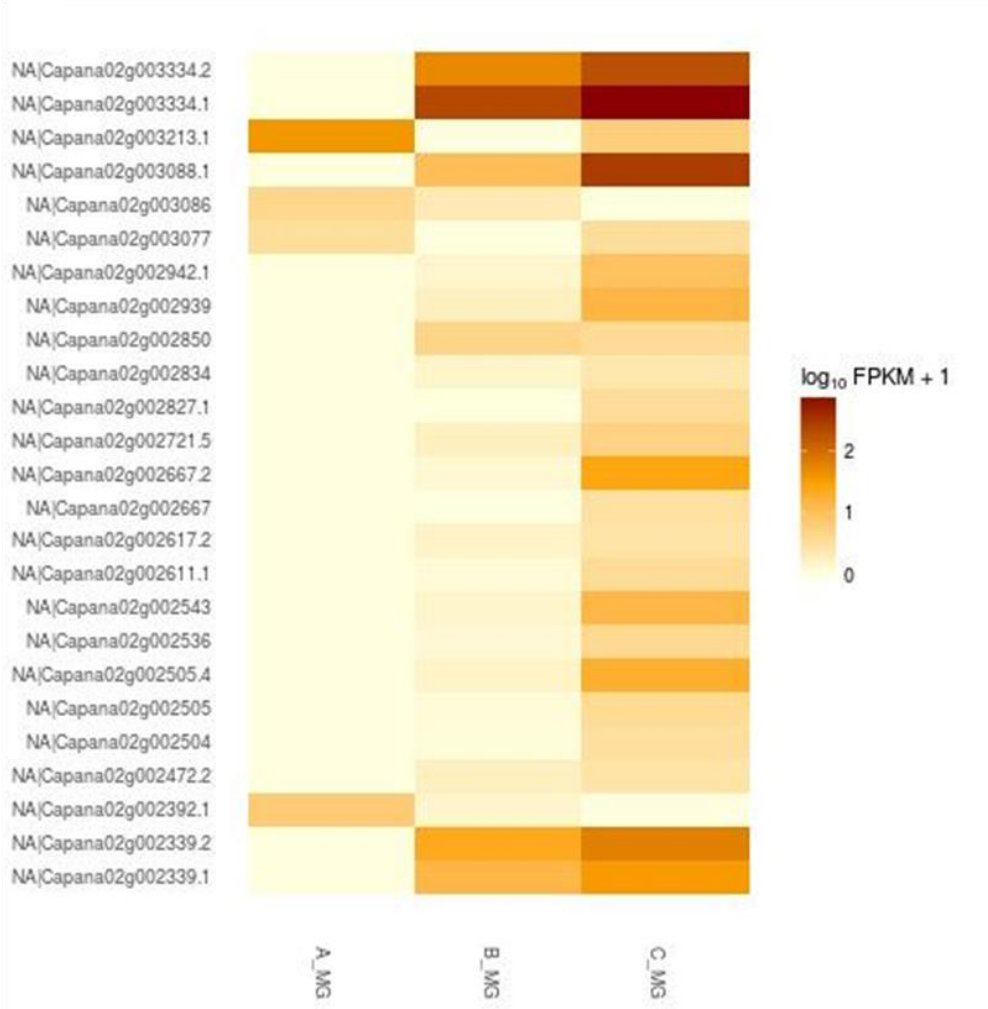


Figure 6.6. A heatmap showing infinitely different transcripts located in the region of the molecular marker associated with the *c1* locus, in lines not possessing a functional copy of the locus (A: 414; B: 255; C:446). Log_{10}FPKM+1 values for expression levels of the genes were calculated by the Cuffdiff module of the Cufflinks software suite. The heatmap was generated with the R package CummeRbund. (n=3 biological replicates.)

Gene ID	Gene name (NCBI)	Biological function	Reference
Capana02g002339.1	*Capsicum annuum Catf2* gene for acyl-transferase, partial cds	Uncharacterised acyl-transferase; a homologue (*catf1*) is involved in capsaicinoid accumulation in pungent pepper.	(Lang, Yanagawa, Sasanuma, & Sasakuma, 2006)
Capana02g003088.1	PREDICTED: *Capsicum annuum* protein RALF-like 34 (LOC107861394), mRNA	Rapid alkalinisation factor.Cell signaling peptide; mediates an increase in cytoplasmic Ca^{2+} concentration, causing activation of intracellular mitogen activated kinases.	(Mundy, Olsen, Mundy, & Skriver, 2002)
Capana02g002504	PREDICTED: *Capsicum annuum* peroxidase 64-like (LOC107859813), transcript variant X1, mRNA	Hydrogen peroxide breakdown; oxidation of toxic reductants; biosynthesis and degradation of lignin and suberin; auxin catabolism; xylem secondary cell wall formation.	(Cosio & Dunand, 2009; Yokoyama & Nishitani, 2006)
Capana02g002536	PREDICTED: *Capsicum annuum* GDSL esterase/lipase At5g33370-like (LOC107861307), mRNA	Glycine-aspartic acid-serine-leucine motif lipase/hydrolase; possible role in cutin biosynthesis.	(Chepyshko, Lai, Huang, Liu, & Shaw, 2012; Hong, Brown, Segerson, Rose, & Adrienne, 2017; Tabata, 2000)
Capana02g002543	PREDICTED: *Solanum pennellii* E3 ubiquitin-protein ligase RNF13-like (LOC107010709), mRNA	E3 ubiquitin ligase; may play a role in cell proliferation.	(Zhang, Meng, Zhang, Chen, & Zhu, 2009)
Capana02g002611.1	PREDICTED: *Capsicum annuum* NAC domain-containing protein 37-like (LOC107861319), mRNA	NAC family transcription factor; plays a role in responses to biotic and abiotic stress and cell wall biosynthesis.	(Bhattacharjee, Das, Mandal, & Kundu, 2017; Nuruzzaman, Sharoni, & Kikuchi, 2013; Olsen, Ernst, Leggio, & Skriver, 2005)
Capana02g002617.2	PREDICTED: *Capsicum annuum* probable serine/threonine-protein kinase abkC (LOC107859902), mRNA	May play a role in the regulation of plant growth; has a possible role in plant defence signalling.	(Moran & Walker, 1993)

Capana02g002827.1	PREDICTED: *Capsicum annuum* zinc finger CCCH domain-containing protein 14 (LOC107861343),mRNA	RNA binding protein; has a regulatory role in mRNA processing; may be stress responsive.	(Peng et al., 2012)
Capana02g002834	PREDICTED: *Capsicum annuum* acylsugar acyltransferase 3-like (LOC107861344), mRNA	Esterifying glandular trichome acylsugars; Herbivore resistance.	(Schilmiller, Gilgallon, Ghosh, Jones, & Last, 2016; Weinhold & Baldwin, 2011)
Capana02g003334.1	PREDICTED: *Capsicum annuum* lamin-like protein (LOC107860496), mRNA	Nucleoskeleton structure.	(Ciska, Masuda, Moreno, & De, 2013)
Capana02g002392.1	PREDICTED: *Capsicum annuum* probable purine permease 4 (LOC107859739), mRNA	Purine nucleobase transmembrane transport.	(Gillissen et al., 2000)
Capana02g003086	PREDICTED: *Capsicum annuum* mitogen-activated protein kinase kinase kinase NPK1-like (LOC107861392), mRNA	Regulatory role in the oxidative stress signalling pathway.	(Kovtun, Chiu, Tena, & Sheen, 1999)

Table 6.6. Differentially expressed genes found to be infinitely more expressed in mature green fruit of lines without a functional copy of the *c1* locus. (n=3 biological replicates)

A particular focus was given to genes which were infinitely more highly expressed in mature green fruit of lines which did not possess a functional copy of the *c1* locus, due to the fact the strikingly different chloroplast structures observed in these lines. Additionally, functional analysis proved that there were several genes with functions influencing fatty acid metabolism and plastid compartment structure in the region of the genome which the *c1* locus had been mapped to. These are detailed in Table 6.6 and Figure 6.6.

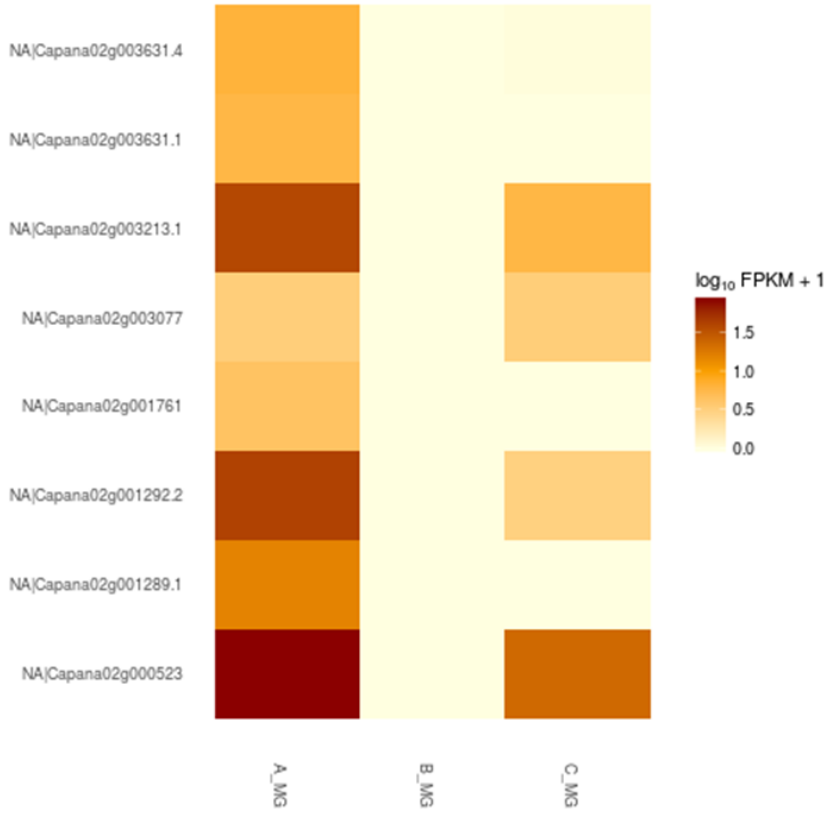


Figure 6.7. A heatmap showing transcripts more highly expressed located in the region of the molecular marker associated with the *c1* locus, in lines possessing a functional copy of the locus (A: 414; B: 255; C:446). Log_{10}FPKM+1 values for expression levels of the genes were calculated by the Cuffdiff module of the Cufflinks software suite. The heatmap was generated with the R package CummeRbund. (n=3 biological replicates.)

Due to the large number of infinitely differentially expressed genes identified in mature green fruit samples, analysis was limited to genes in the location in the genome where the *c1* locus has been mapped to (Figure 6.7). It can be seen that most of the genes whose expression is not detected in line 255 (B), are also lower in expression in 466 (C) (which is also negative for a copy of the *c1* locus) than they are in A (414) (which does possess a functional copy of the locus) (Figure 6.7).

There are three different genes expressed in the accession possessing a functional copy of the *c1* locus, which have no detectable transcripts in either of the accessions lacking functional copies of the locus. These are Capana02g001761 (*abscisic acid receptor (PLY12-like)*), Capana02g001289.1 (*Capsicum annuum fe(2+) transport protein 1-like*), and Capana02g003631.1 and 4 (*protein LYK2*) (Figure 6.7).

6.2.3.2 *Ripe fruit*

6.2.3.2.1 Biological functions of differentially expressed genes

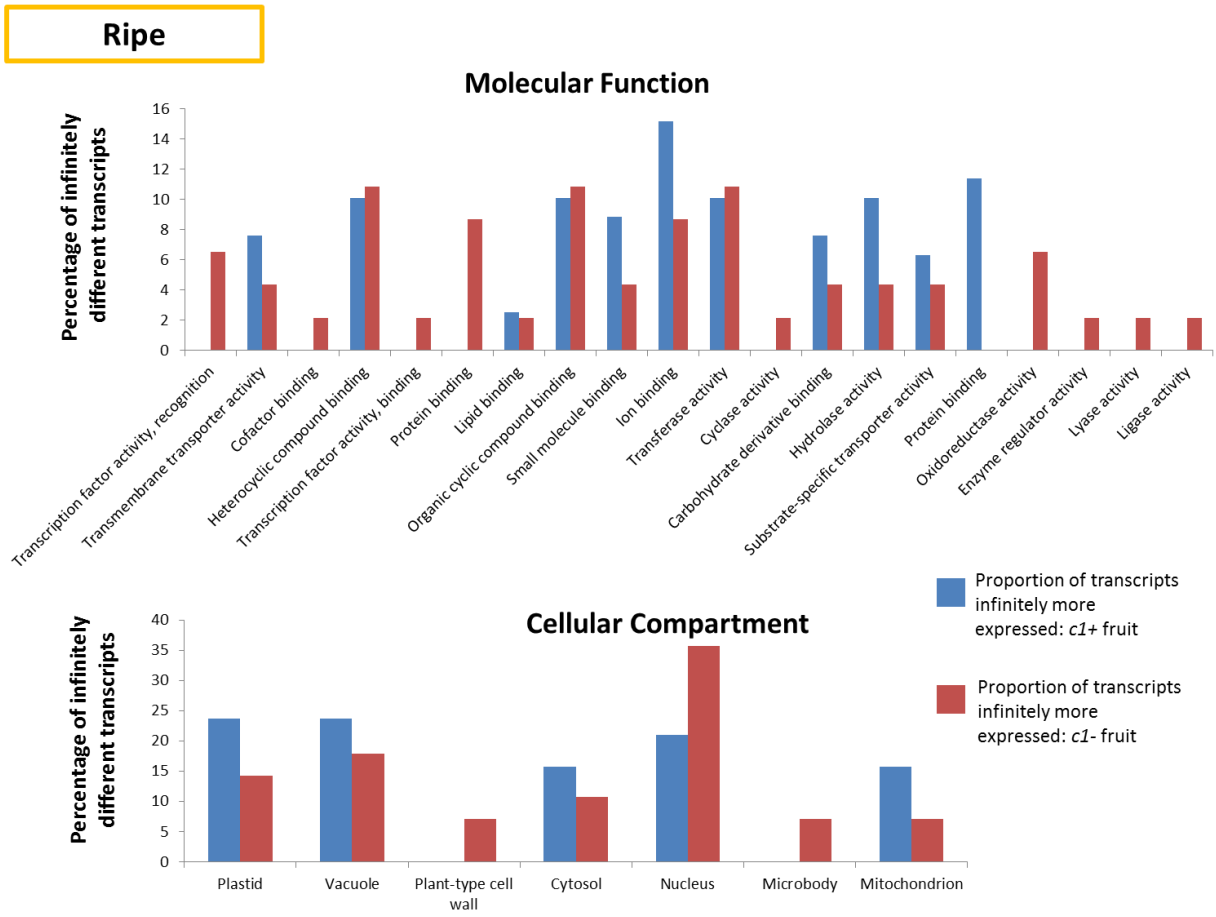


Figure 6.8. Annotations of infinitely different transcripts from *c1* positive and negative lines, in ripe fruit generated and assigned to specific molecular functions and cellular components with Blast2GO. (n=3 biological replicates.)

The software package Blast2GO was used to carry out large-scale blasts of differentially expressed transcript sequences against a vascular plant NCBI database. Gene products were identified, and associated with biological functions and cellular compartments. Infinitely differentially expressed transcripts which were higher in *c1* positive fruit were found to be involved in ion binding, protein binding, carbohydrate binding, transmembrane transporter activity, substrate-specific transporter activity, and hydrolase activity. Furthermore, they were found in diverse cellular compartments: the plastid, the vacuole, the cytosol and the mitochondrion, however this could be simply due to the fact that more annotated genes are found in these locations. Further statistics would be necessary to draw meaningful conclusions from this finding. There were several more metabolic functions which were infinitely more highly expressed in *c1* negative fruit, suggesting the mutation in the gene associated with the locus could confer a gain of function. Analysis of the cellular compartments which the differentially expressed transcripts were found in revealed that there were a greater

number in the nucleus for this condition, which could suggest an increase in transcriptional activity also relating to a gain of function mutation.

6.2.3.2.2 Differentially expressed genes within the *c1* marker region

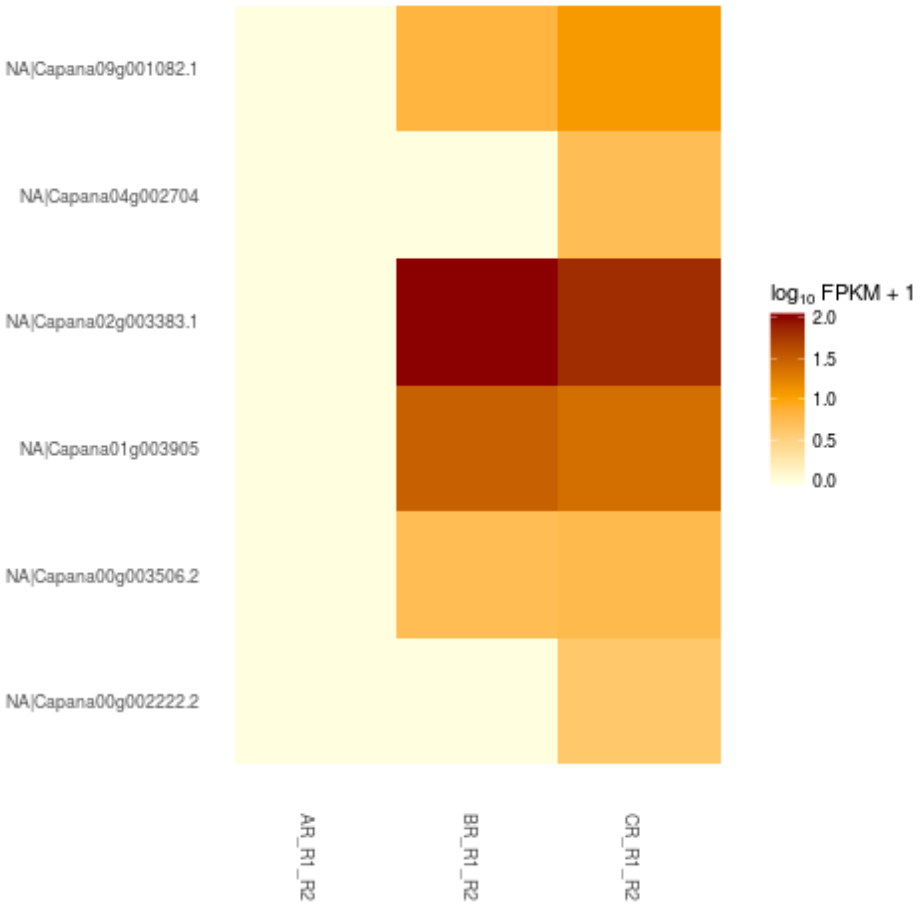


Figure 6.9. A heatmap showing transcripts infinitely more highly expressed in either B (255) or C (446) lines not possessing a functional copy of the locus. Genes are located in across the genome rather than only in the region of the molecular marker associated with the *c1* locus due to the small number of infinitely different genes identified in ripe fruit. (A: 414; B: 255; C:446). Log_{10}FPKM+1 values for expression levels of the genes were calculated by the Cuffdiff module of the Cufflinks software suite. The heatmap was generated with the R package CummeRbund (n=3 biological replicates).

Due to the fact that a small number of genes were identified as being differentially expressed in ripe fruit, it was possible to include genes located throughout the genome in the analysis, rather than only genes in the area which the *c1* locus has been mapped to (Figure 6.9). When transcript levels of genes which are undetectable in line 414 (A) (possessing a functional copy of the *c1* locus) are analysed in lines lacking a functional copy of the locus, it is clear that, with the exception of two genes which are not detected in B (255), all have similar expression patterns in both B (255) and C (446) (Figure 6.9).

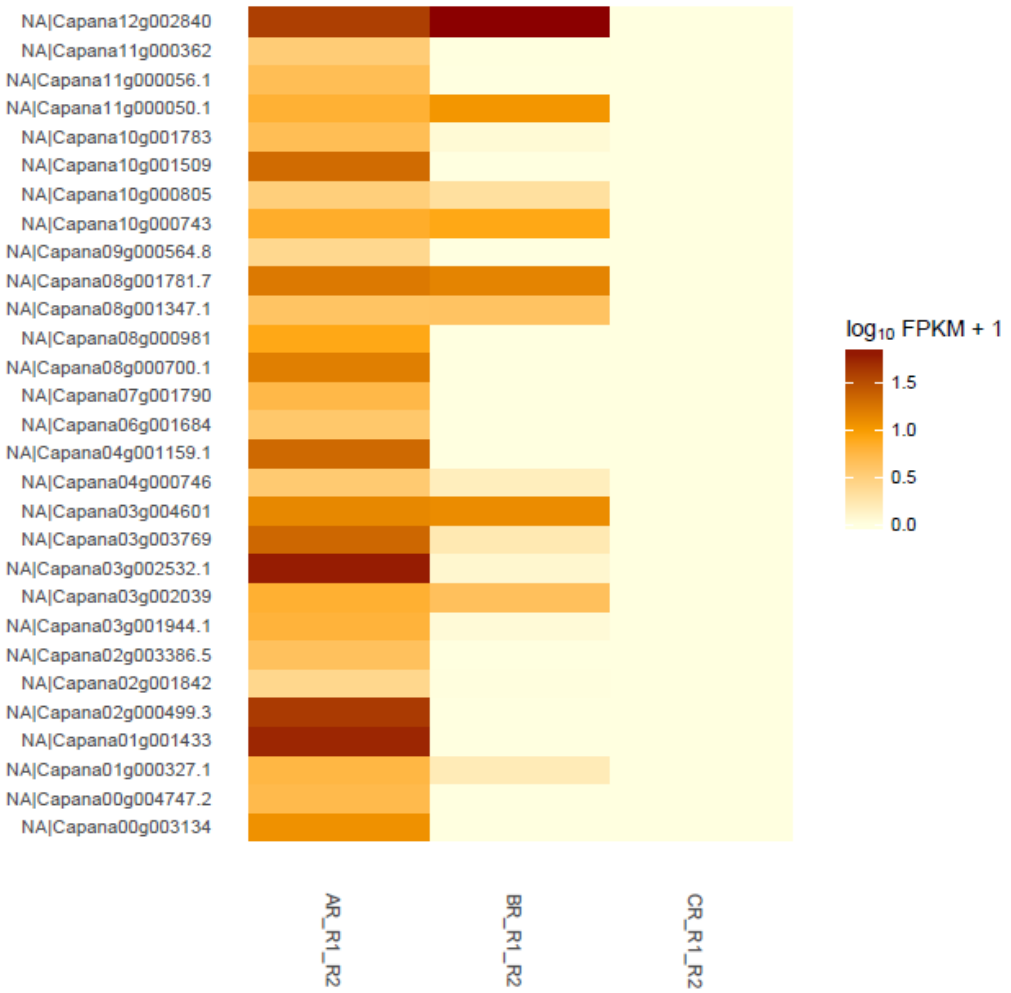


Figure 6.10. Heatmap showing transcripts more highly expressed in lines possessing a functional copy of the *c1* locus (A: 414; B: 255; C:446). Log_{10}FPKM+1 values for expression levels of the genes were calculated by the Cuffdiff module of the Cufflinks software suite. The heatmap was generated with the R package CummeRbund. (n=3 biological replicates.)

When transcript levels of differentially expressed genes which are more highly expressed in the line possessing a functional copy of the *c1* locus (414; A), it is apparent that, in most cases, genes are absent in both lines which do not possess functional copies of the *c1* locus (Figure 6.10). Some genes are expressed in line 255 (B), and not in line 446 (C). This aligns with the gradient of colour across the three lines: the colour phenotype of 466 (C) is paler than that of 255 (B) (Figure 6.10). However, the genes could be involved in processes unrelated to colour; they are located throughout the genome, not restricted to the mapping location of the *c1* locus. This is due to the fact that the number of differentially expressed genes in ripe fruit was small enough that genes throughout the genome could be included.

Gene ID	NCBI Match
Capana12g002840	RER1A_ARATH Protein RER1A OS=Arabidopsis thaliana GN=RER1A PE=2 SV=1
Capana11g000362	PUB23_ARATH E3 ubiquitin-protein ligase PUB23 OS=Arabidopsis thaliana GN=PUB23 PE=1 SV=1
Capana11g000056	KIWEL_ACTDE Kiwellin OS=Actinidia deliciosa PE=1 SV=1
Capana11g000050	RN139_MOUSE E3 ubiquitin-protein ligase RNF139 OS=Mus musculus GN=Rnf139 PE=1 SV=1
Capana10g001783	SUBL_ARATH Subtilisin-like protease OS=Arabidopsis thaliana GN=ARA12 PE=1 SV=1
Capana10g001509	EF1A_SOLLC Elongation factor 1-alpha OS=Solanum lycopersicum PE=2 SV=1
Capana09g000564	POT5_ARATH Potassium transporter 5 OS=Arabidopsis thaliana GN=POT5 PE=1 SV=1
Capana08g000700	CRR38_ARATH Cysteine-rich repeat secretory protein 38 OS=Arabidopsis thaliana GN=CRRSP38 PE=2 SV=1
Capana07g001790	ATL41_ARATH E3 ubiquitin-protein ligase ATL41 OS=Arabidopsis thaliana GN=ATL41 PE=1 SV=1
Capana06g001684	PUP3_ARATH Purine permease 3 OS=Arabidopsis thaliana GN=PUP3 PE=2 SV=1
Capana04g001159	ATL51_ARATH RING-H2 finger protein ATL51 OS=Arabidopsis thaliana GN=ATL51 PE=2 SV=2
Capana04g000746	Y4729_ARATH G-type lectin S-receptor-like serine/threonine-protein kinase At4g27290 OS=Arabidopsis thaliana GN=At4g27290 PE=2 SV=4
Capana03g004601	MB21_ARATH Myrosinase-binding protein-like At2g25980 OS=Arabidopsis thaliana GN=At2g25980 PE=2 SV=1
Capana03g003769	YOAA_BACSU Uncharacterized N-acetyltransferase YoaA OS=Bacillus subtilis GN=yoaA PE=3 SV=2
Capana03g002532	21KD_DAUCA 21 kDa protein OS=Daucus carota PE=2 SV=1
Capana03g002039	GDL62_ARATH GDSL esterase/lipase At4g10955 OS=Arabidopsis thaliana GN=At4g10955 PE=2 SV=1
Capana03g001944	BCS1_SCHPO Probable mitochondrial chaperone bcs1 OS=Schizosaccharomyces pombe (strain 972 / ATCC 24843) GN=SPAC644.07 PE=2 SV=1
Capana02g003386	GALM_PIG Aldose 1-epimerase OS=Sus scrofa GN=GALM PE=2 SV=1
Capana02g001842	GLR29_ARATH Glutamate receptor 2.9 OS=Arabidopsis thaliana GN=GLR2.9 PE=2 SV=1
Capana02g000499	BURP3_ORYSJ BURP domain-containing protein 3 OS=Oryza sativa subsp. japonica GN=BURP3 PE=2 SV=1
Capana01g001433	BAP2_ARATH BON1-associated protein 2 OS=Arabidopsis thaliana GN=BAP2 PE=1 SV=1
Capana01g000327	YD338_YEAST Uncharacterized transporter YDR338C OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=YDR338C PE=1 SV=1
Capana00g004747	Y1743_ARATH Probable LRR receptor-like serine/threonine-protein kinase At1g74360 OS=Arabidopsis thaliana GN=At1g74360 PE=1 SV=1
Capana00g003134	CML44_ARATH Probable calcium-binding protein CML44 OS=Arabidopsis thaliana GN=CML44 PE=2 SV=2

Table 6.7. NCBI matches for infinitely different transcripts in ripe fruit of lines possessing a functional copy of the *c1* locus. Differential expression statistics were calculated with the Cuffdiff module of the Cufflinks software package. Gene IDs relate to the *Capsicum annuum* Zunla genome. (n=3 biological replicates.)

6.2.4 Selected pathway transcript expression levels

6.2.4.1 Carotenoid metabolism

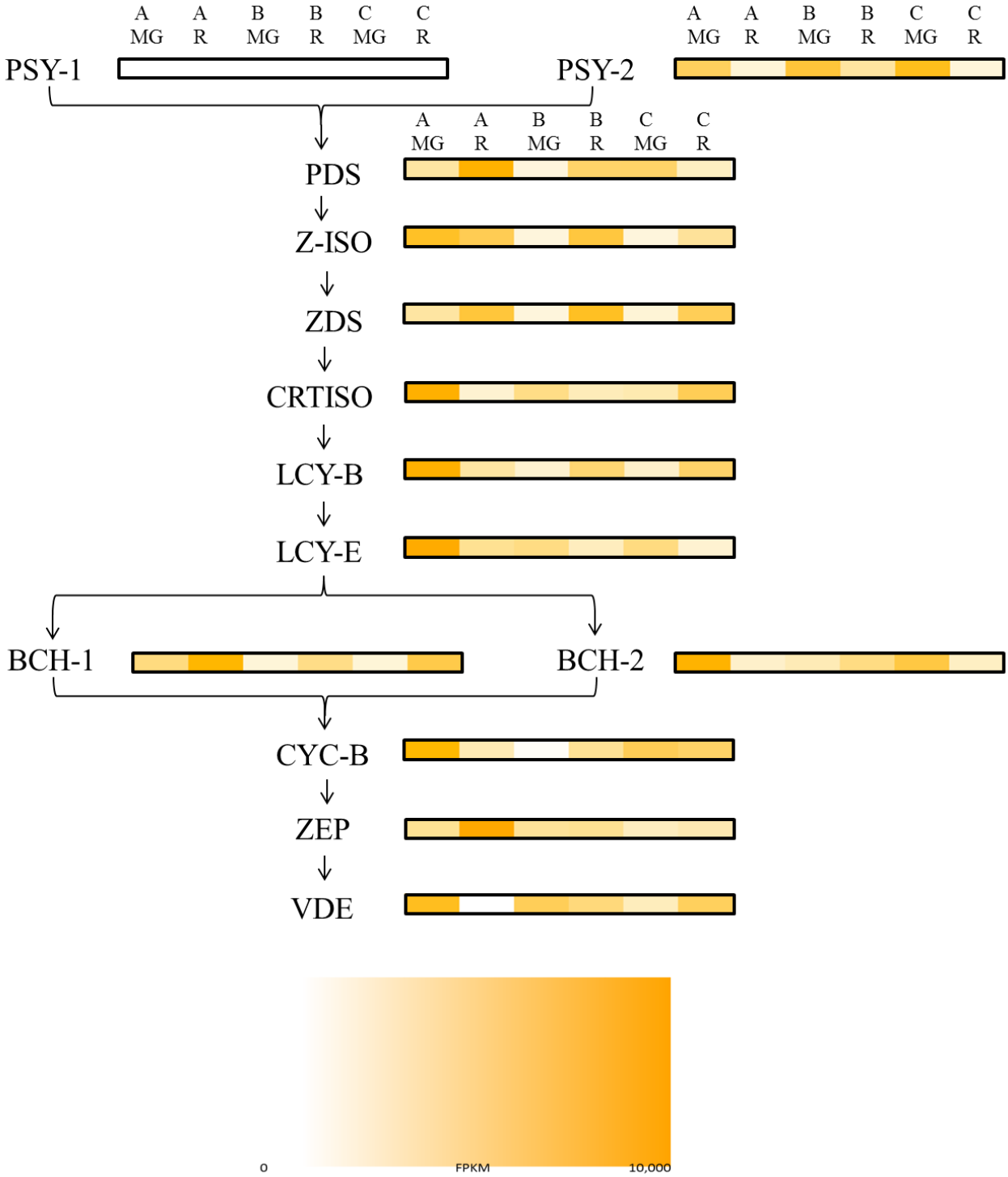


Figure 6.11. Expression levels of genes in the carotenoid biosynthesis pathway; FPKM values were determined by the Cuffdiff module of the Cufflinks package. A; 414, B: 255, C: 446. n=3 biological replicates

As colour is the fruit quality trait which the *c1* locus influences, it was essential to look at the expression levels of genes in the carotenoid biosynthesis pathway (Figure 6.11). All the lines in the panel are negative for *c2* (*Psy-1*), and correspondingly no *Psy-1* transcripts were detected in mature

green or ripe fruit. As would be expected from the model proposed in Chapters Three and Four (*c1* positive fruit which lack *c2* synthesise carotenoids in mature green fruit; these are then retained over the course of ripening, conferring the yellow colour on ripe fruit), expression levels of most genes are highest in mature green fruit of line 414. A notable exception to this is *Psy-2,* which, as discussed in Chapter Three is the phytoene synthase specifically expressed in vegetative tissue. RT-qPCR demonstrated that transcripts were at least twice as abundant in mature green than ripe fruit. Other exceptions are the desaturases *Pds* and *Zds*: more transcripts of both genes were found in ripe fruit of lines 414 (A) and 255 (B), however, slightly higher levels of *Pds* were observed in mature green fruit of line 446 (C) (Figure 6.11). For most of the genes selected, the lowest expression levels are in mature green and ripe fruit in line 446 (C) which is the palest of the two *c1* negative lines.

6.2.4.2 Photosynthetic metabolism

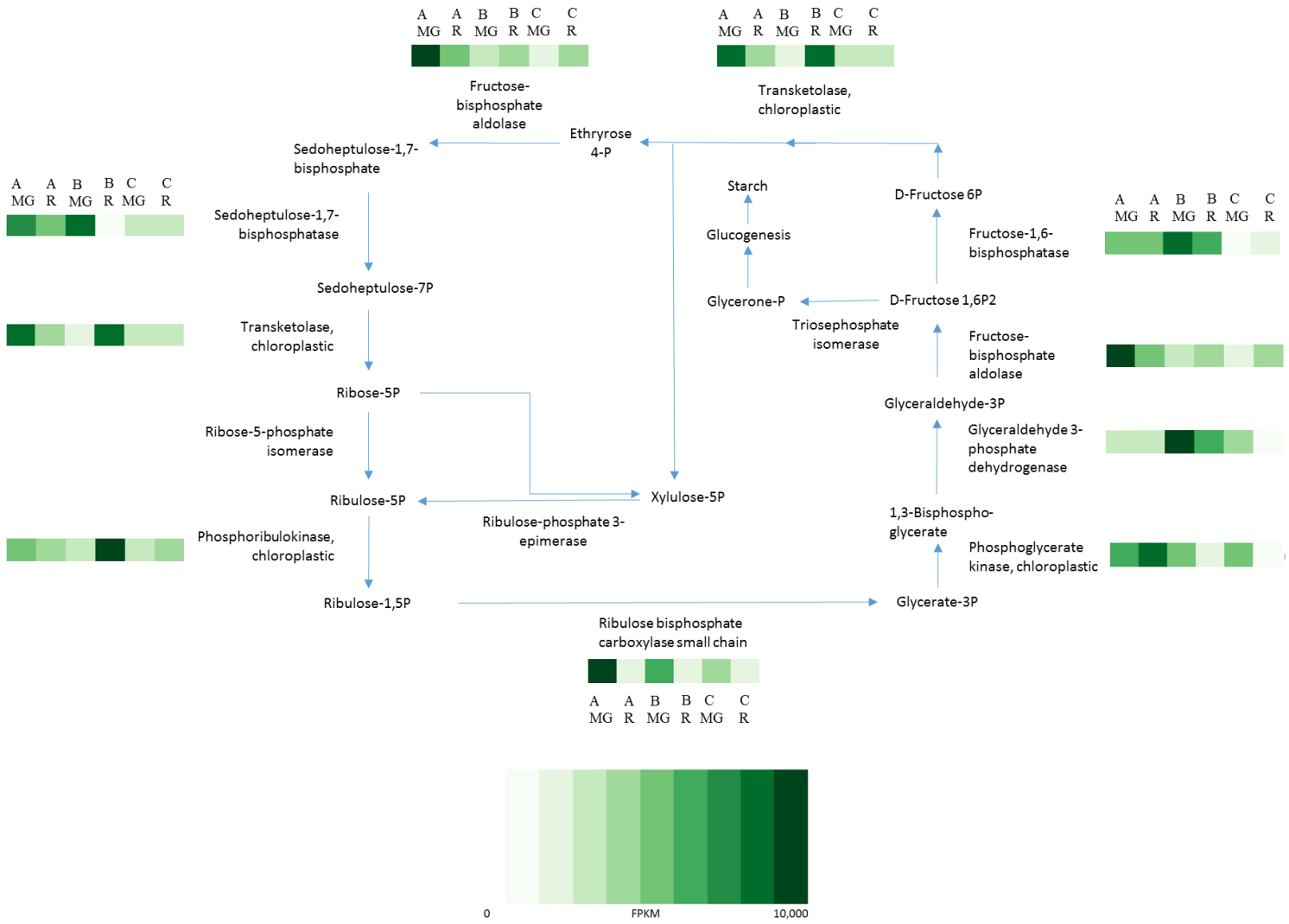


Figure 6.12. Expression levels of genes in the Calvin cycle; FPKM values were determined by the Cuffdiff module of the Cufflinks package. A; 414, B: 255, C: 446. n=3 biological replicates.

Looking into expression levels of genes involved in photosynthetic metabolism revealed that, for almost all genes encoding proteins in the Calvin cycle, expression was higher in mature green fruit for all three lines. Interestingly, the exceptions to this rule were the transketolase and phosphoribulokinase, both of which were more highly expressed in ripe fruit from line 255 than they were in mature green fruit from this line. Overall, this data supports the hypothesis that the most functional photosynthetic apparatus is in mature green fruit of the line possessing a functional copy of the *c1* locus, 414. This reinforces the conclusion that the *c1* locus mutation has a negative impact on chloroplast sub-structures in mature green fruit (Figure 6.12).

6.3 Discussion

6.3.1 *Capsicum* transcriptomics

Until recently, the only *Capsicum* transcriptomes published were assembled de novo, rather than aligned to a genome (Góngora-castillo et al., 2012). However in the last year *Capsicum* transcriptomes aligned to the genome have become available (Kim et al., 2014). *Capsicum* transcriptomes have been used to investigate a focussed question; for example, analysing dynamics of a particular transcription factor family, or identifying candidate genes associated with chilling tolerance (Jie Li et al., 2016; Wu, Cheng, Cui, Xu, & Liang, 2016). Transcriptomics in *Capsicum* has also been used to investigate how global gene expression changes in response to a particular condition, such as biotic stress from a virus (Gamage, Mcgrath, Persley, & Dietzgen, 2016). However, a comprehensive search of *Capsicum* transcriptome literature revealed no published transcriptomes comparing differently coloured accessions of sweet pepper, making the white and yellow pepper RNA-seq data presented here a valuable addition to the area.

6.3.2 Candidate genes selected for further analysis

Genes selected for further analysis were chosen for two reasons: if they were expressed in the accession with a functional copy of the c1 locus and not in those without; or if they were expressed in accessions without a functional copy of the c1 locus and not in the accession which did. As the nature of the mutation remains unknown, both options should be investigated. Although it is more likely for the mutation to result in a lack of transcripts for the associated gene, the possibility that it is a gain of function mutation cannot be ruled out. Differences of expression in mature green fruit were focussed on, due to the difference in plastid structure observed at this developmental stage.

6.3.2.1 Abscisic acid receptor PYL12-like

Abscisic acid receptor PYL12-like was one of three genes in the region of the *c1* locus found to be expressed in the accession with a functional copy of the locus and not expressed in either of the two accessions lacking *c1*. The gene encodes a receptor for abscisic acid (ABA), which is required for ABA-

mediated responses such as stomatal closure and germination inhibition. The protein inhibits the activity of group-A protein phosphatases type 2C (PP2Cs) in an ABA independent manner, but is more efficient when activated by ABA (Hao et al., 2011). ABA is involved in a huge range of processors in the plant, and the number of membrane proteins which can act as receptors is large. It is unclear how an absence of transcripts for this gene could relate to loss of colour in ripe fruit, however, as there are only three genes not expressed by mature green fruit in the region of the *c1* locus, further analysis with studies such as virus induced gene silencing (VIGS) could yield interesting results.

6.3.2.2 Fe(2+) transport protein 1-like

A second gene expressed in the accession with a functional copy of the *c1* locus and not in those without was that which encodes an Fe(2+) transport protein 1-like iron transporter. This high-affinity transporter plays a key role in the uptake of iron across the plasma membrane and acts as the principal regulator of iron homeostasis in planta. It also mediates the heavy metals uptake under iron-deficiency by its ability to transport cobalt, cadmium, manganese and/or zinc ions (Vert, Briat, & Curie, 2003). Sequence homology to other genes in the Arabidopsis Information Resource (TAIR) suggests that this protein has a role in the plant's response to ethylene, which could imply that pepper accessions without functional copies of the *c1* locus ripen differently to those with a functional copy. This could plausibly have a large impact on the colour of the mature fruit.

6.3.2.3 LYK2

The final gene found to have transcripts present accession 414 (A) and absent in accessions 255 (B) and 446 (C) encodes a probable receptor-like kinase (LYK2). As its biological functions are in cell surface receptor signalling and protein phosphorylation, implications of its non-expression could be far-reaching (uniprot.org/uniprot/Q9SJG2). Therefore, further analysis could provide information about whether absence of transcripts from this gene cause changes to ripe fruit colour phenotype.

6.3.2.4 Catf2

The gene *Capsicum* acyl-transferase 2 (*Catf-2*) displays the consensus motifs for acyl transferases: histidine,*,*,*, aspartic acid and aspartic acid, phenylalanine, glycine,*, glycine (Lang et al., 2006). Although the gene has not been well characterised itself, *Catf-1*, a close homologue, has been associated with pungency in *Capsicum* fruit. This functional characterisation of the gene is only relevant to pungent chilli pepper fruit; the homologue *Catf-2* cannot have a pungency-related role in sweet pepper.

6.3.2.5 NAC37

NAC37 is a member of the NAC gene family of transcription factors, characterised by their possession of a NAC DNA binding domain. The acronym NAC is derived from the names of the three initial genes in which the conserved domain was found: NAM (no apical meristem), ATAF-1 and -2 and CUC2 (cup-shaped cotyledon) (Aida, Ishida, Fukaki, Fujisawa, & Tasaka, 1997). The NAC family contains at least 105 members in *Arabidopsis*, and their functions are extremely diverse; they are involved in a wide range of processes including leaf development, fruit development and ripening, cell wall biosynthesis, and response to biotic and abiotic stress (Berger et al., 2009; Nuruzzaman et al., 2013; Ooka et al., 2003).

Although references to NAC transcription factors in pepper are sparse in the literature, NAC transcription factors have been linked to multiple types of stress response in closely related species such as tomato fruit. For example, several salt responsive NACs have been identified, as well as regulators of defence responses to fungal and bacterial pathogens (Wang et al., 2009; Yang, Deng, & Ouyang, 2011). NAC domain transcription factors have also been known to interact with viral proteins: there is evidence to suggest that *SlNAC1* is harnessed by the replication enhancer protein (REn) of tomato leaf curl virus, resulting in increased viral replication (Selth et al., 2005) They have also been linked to fruit quality traits in tomato: the important ripening regulator *nonripening* (*NOR*) belongs to the NAC family. The tomato gene *NOR* encodes a NAC domain transcription factor whose mutant possesses a similar phenotype to that of the *ripening inhibitor* (*rin*) mutant (Giovannoni, 2004, 2007). Mutant *rin* and *nor* fruit fail to ripen and to produce climacteric ethylene; while they are responsive to ethylene at the molecular level, they do not ripen in response to exogenous ethylene (Lincoln & Fischer, 1988). NOR was mapped to chromosome 10 in the tomato genome by positional cloning in 1995 (Giovannoni et al., 1995).

A small number of *Capsicum* NAC transcription factors have been characterised previously: *CaNAC1* expression was induced after incompatible interactions between viral and bacterial pathogens and chili pepper fruit (Oh, Lee, Yu, & Choi, 2005). In addition, the pepper transcription factor *CaNAC2* was characterised in 2012; as its expression was induced by salt, cold and abscisic acid (ABA) treatment, and inhibited by osmotic stress and salicylic acid (SA) treatment, it has been associated with response to abiotic stress (Guo et al., 2015).

NAC37 was selected for further analysis due to the fact that it is a member of such an influential transcription factor family, with links to the regulation of fruit ripening. However, the gene itself is not well characterised. Alignments to other members of the NAC family suggest that it has a role in cell wall maintenance.

6.3.2.6 ABCK

The 'acitivity of bc_1 complex kinase' (ABCK) protein family are one of the few to be found in plants in organelle-specific locations. The name of the family relates to their ancestral function in ubiquinone (UQ) biosynthesis: the first *ABCK* to be discovered (*ScCOQ8* in yeast) was found to be necessary for the redox activation of the mitochondrial bc_1 complex, which functions in respiration (Do, Hsu, Jonassen, Lee, & Clarke, 2001). Loss of functional *ScCOQ8* gene products led to an accumulation of the UQ precursor 3-hexaprenyl-4-hydroxybenzoic acid, which in turn destabilises the bc_1 complex(Bousquet, Dujardin, & Slonimski, 1991). For a decade, it was wrongly hypothesised that the ABCK protein ScCOQ8 took on a chaperone role, supressing a deleterious mutation in a cytochrome b translational activator, and assisting with cytochrome b folding (Bousquet et al., 1991; Brasseur, Tron, Slonimski, & Brivet-chevillotte, 1997). However, it has since been shown that this is instead carried out by a $tRNA^{Trp}$ gene; therefore the current view is that ScCOQ8 operates specifically on UQ biosynthesis, not acting as a chaperone (Do et al., 2001; Hsieh, Dinoso, & Clarke, 2004)

In plants, ABCKs are localised to mitochondria and chloroplasts (Lundquist, Davis, & Wijk, 2014). The gene family in Arabidopsis consists of seventeen members, proteins associated with seven of which have been located in the plastid: six in the plastoglobule (AtABC1K1 and 3-7), and one in the plastid inner envelope (AtABC1K8) (Austin et al., 2006; Lundquist et al., 2012; Ytterberg et al., 2006). Similar results were obtained from a study of protein localisations in maize leaf fractions (Majeran et al., 2012). Based on its functional homologue in yeast ABC1K13 is hypothesised to be localised in the mitochondria (Cardazzo, Hamel, Sakamoto, Wintz, & Dujardin, 1998).

Though they are found in organisms in each of the three domains of life, ABCKs have greatly proliferated in photosynthetic organisms Phylogenetic analysis focussing on 42 species of bacteria, archaea and eukaryotes have shown that the 247 ABC1K proteins identified to date can be divided into three major clades defined by evolutionary origin and subcellular location (plastid endosymbiosis, mitochondrial endosymbiosis, and ancestral) and an archaeal outgroup (Lundquist et al., 2014).

6.3.2.7 GDL62

GDL62 is a member of the GDSL family of esterase/lipases. The class of lipolytic enzymes were first identified by Upton and Buckley in 1995, and are so named due to their possession of a distinctive amino acid motif: glycine-aspartic acid-serine-leucine (GDSL), at or close to the N-terminus of the protein. Another characteristic of the family is their lack of the do-called 'nucleophile elbow', a feature possessed by other lipase classes. Rather than having the active site Ser in the middle of the amino acid sequence (as most esterase/lipases do), in GDSL esterases it is located close to the N-terminus of the protein (Akoh, Lee, Liaw, Huang, & Shaw, 2004; Upton & Buckley, 1995). The family

are known for their multifunctional properties; they possess broad substrate and region-specificity (Chepyshko et al., 2012). Another feature important for their classification is the fact that GDSL esterases have five blocks of conserved sequence (I-V). Enzymes of the family are further classified as SGNH hydrolases, due to the presence of a strictly conserved Ser-Gly-Asn-His motif in blocks I, II, III, and V (Brick et al., 1995; Upton & Buckley, 1995).

In plants, GDSL esterase/lipases are represented by at least 1100 genes found in twelve sequenced plant genomes. In Arabidopsis, the family consists of 108 members (Ling, 2008). Not all of these have been characterised, however, those which have are functional in a wide range of physiological processes including: growth and development, morphogenesis, biosynthesis of secondary metabolites, and the defence response.

Previous work has revealed that the family of enzymes are capable of acting upon Very Long Chain Fatty Acids (VLCFAs) in the sepal waxes of *Arabidopsis;* therefore it is possible that the homologue identified by this differential gene expression analysis acts in a similar way (Hong et al., 2017). This could account for some of the differences in VLCFA levels between lines possessing a functional copy of the *c1* locus and lines which do not.

6.3.3 Insights from differentially expressed genes

Mature green differentially expressed transcripts include several acyltransferase isoforms, as well as esterase/lipase isoforms.in concordance with GC-MS data from this selection of lines, fatty acid metabolism is one of the most striking differences between accessions with a functional copy of the *c1* locus and those without. Another set of differentially expressed genes more highly expressed in lines without a functional copy of *c1* are kinases, including plastid-specific kinase isoforms. This is interesting when seen in the light of the aberrant thylakoid structures in *c1*-lacking accessions: changing the phosphorylation state of thylakoid membrane proteins can change the width of the thylakoid lumen. It could be the case that increased expression of a kinase in the plastid changes the charge balance between the two sides of the lumen, thus forcing them further apart (as observed in Chapter Five). This could also contribute to the decreased number of thylakoids per granum in *c1*-negative mature green fruit: altered membrane phosphorylation states will affect membrane interactions, possibly making stacking energetically unfavourable.

6.4 Conclusion

- These are the first yellow and white bell pepper transcriptomes, and have the additional benefit of revealing differences between mature green and ripe fruit samples.
- Through narrowing down differentially expressed transcripts to the region of the genome associated with the *c1* locus, candidates for association with the locus have been identified.

- From the candidate genes selected, putative mechanisms by which the colour of the mature fruit could be changed have been proposed: membrane structure of thylakoids in mature green fruit is altered by charge imbalance caused by integral membrane protein phosphorylation (ABKC). Esterification is altered by increased expression of esterases/lipases or different acyl transferases (*GDL62* or *Catf2*).
- To further analyse the roles of these genes in mature fruit colour development, they will be silenced with VIGS to verify their impact on the phenotype. Another future study could be mining the RNAseq data presented here for sequence differences in genes of interest. Using this method, point mutations which could be related to the *c1* locus could be identified.

7 General Discussion

7.1 General Conclusions

7.1.1 Summary

The project aimed to assess the development of colour in sweet bell pepper (*Capsicum annuum*) from two different perspectives: orange and yellow fruit phenotypes in relation to the ripening-fruit copy of the enzyme catalysing the first step in carotenoid biosynthesis (Phytoene synthase-1 (PSY-1)), and yellow and white fruit phenotypes in relation to the *c1* locus whose identity is to date uncharacterised.

The diversity panel for *Psy-1* (the *c2* locus) consisted of one line with functional copies of all three loci controlling mature fruit colour (*y; Ccs, c2; Psy-1*, and *c1; unknown*) with a red ripe fruit phenotype, four lines which are positive for the *c2* and *c1* loci which produce orange fruit, ten lines which were positive for the *c1* locus only which produce yellow fruit, and one line which is negative for all three loci and produces white ripe fruit. The absence of *Psy-1* gene products in lines possessing the marker for deletion of the gene was confirmed. Pigment analysis with High Pressure Liquid Chromatography (HPLC) revealed that there was no increase in total carotenoid amount over the course of ripening in lines lacking *Psy-1*, leading to the conclusion that the yellow colour of these lines is due to retention of xanthophylls from the mature green stage. As the chromoplast is specialised for carotenoid synthesis and storage in ripe fruit, its structure in lines lacking the enzyme for ripening-related carotenoid biosynthesis was investigated. It was revealed that, although the same fibril protein and esterification-based storage mechanisms were employed, the chromoplast adapts to lower levels of pigment accumulation. The absence of ripening-related carotenoid biosynthesis has an effect on the broader metabolism of the fruit. To quantify this, metabolites were profiled with Gas Chromatography-Mass Spectrometry (GC-MS). As would be expected in a panel bred for differences in colour phenotype, the major differences were in fatty acids which were capable of esterifying xanthophylls.

The second diversity panel consisted of twelve lines, none of which possessed a functional copy of either the *y* or *c2* loci. Six were positive for the *c1* locus and produced yellow ripe fruit, and six were negative for all three loci and produced cream coloured or white ripe fruit. Carotenoid profiling revealed that the main pigments in yellow lines were violaxanthin and its derivatives, and lutein; while cream and white fruited lines contained only trace levels of lutein. Analysis of the broader fruit metabolome revealed significant differences in levels of Very Long Chain Fatty Acids (VLCFAs) between *c1* positive and negative lines, with the latter possessing the highest amounts. Investigation into mature green and ripe fruit plastids revealed that chloroplasts of lines lacking a functional copy of *c1* possessed atypical grana, containing only a few wide-lumen thylakoids per stack. This lead to the hypothesis that the gene associated with the *c1* locus is involved in control of plastid

compartment development; a mutation in this gene could result in plastids which are incapable of retaining pigments synthesised in green fruit over the course of ripening. To further investigate genes potentially associated with the locus transcriptome data from *c1* positive and negative lines at different stages of ripening were interrogated. Using information on the location of the marker associated with the locus on chromosome 2, several candidate genes were selected. These were: the kinase *ABC1K*, the esterase/lipase *GDL62*, the acyl-transferase *Catf2*, and the NAC transcription factor *NAC37.*

7.1.2 Aims and Objectives

The aims of the project can be divided into two sections: i) multi-level characterisation of the phenotype associated with the marker for *c2* (*Psy-1*) and ii) exploiting –omics approaches for selection of candidates for association with the *c1* locus. To achieve these aims, works towards the following objectives was carried out.

Work on the *Psy-1* diversity panel aimed to:

- Determine the carotenoid profiles of lines in a diversity panel for the marker associated with *c2*.
- Sequence *Psy-1* in lines of diverse phenotype to determine the nature of the mutation in genomic DNA.
- Confirm the absence of *Psy-1* transcripts from lines possessing the marker with quantitative real time PCR.
- Assess the impact of the absence of *Psy-1* gene products on transcript levels of other genes encoding enzymes in the carotenoid biosynthesis pathway.
- Profile fruit metabolites at different stages of ripening in lines with and without functional *Psy-1* gene products, using GC-MS.
- Assess the impact of the absence of *Psy-1* on the structural components of the organelles synthesising and storing carotenoids at different stages of fruit ripening, using microscopy, sub-chromoplast fractionation, western blots, and chromatography to characterise both pigment and protein content.

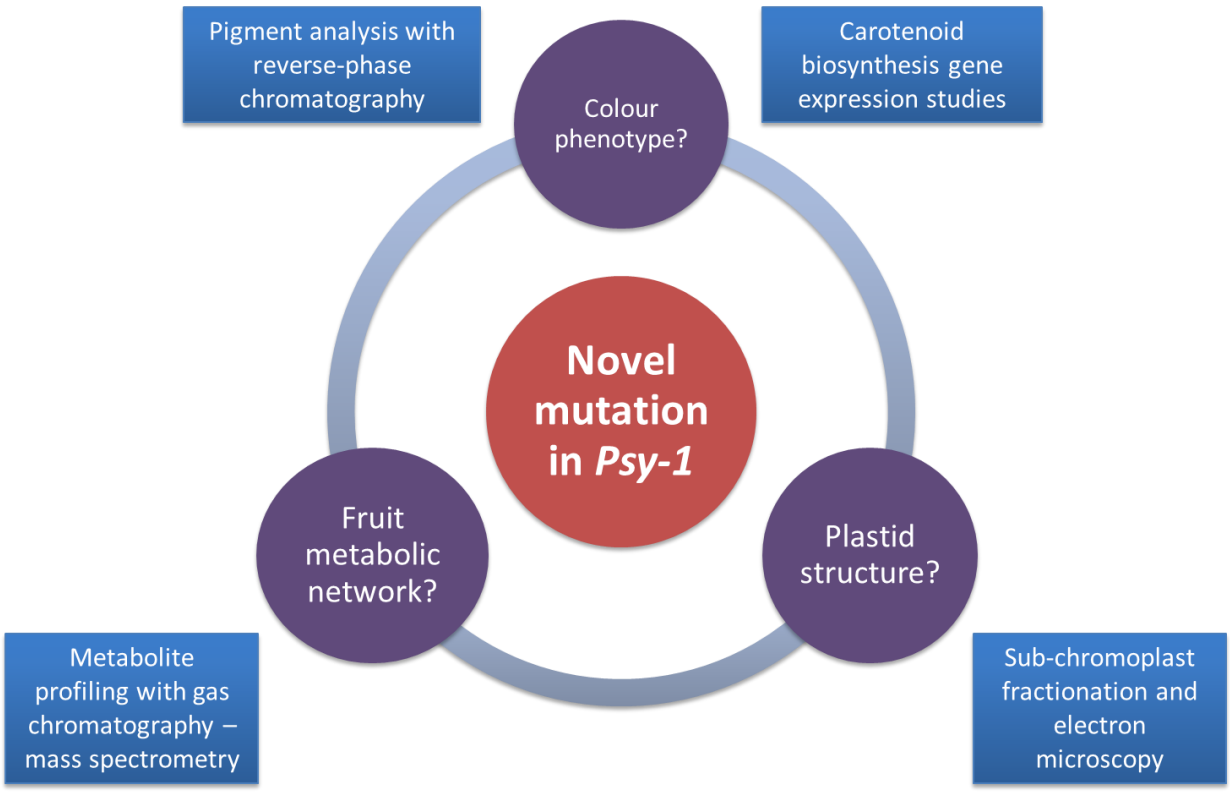


Figure 7.1. Outline of the main questions and experimental approaches used to evaluate the phenotypic impact of the marker for absence of the *c2* locus (Phytoene synthase-1 (*Psy-1*)).

Work on the *c1* diversity panel aimed to:

- Determine the pigment profiles from lines making up a second diversity panel for the possession of the *c1* locus, absence of which is associated with white ripe fruit.
- Employ GC-MS to profile fruit metabolites at different stages of fruit ripening in accessions with and without functional copies of the *c1* locus.
- Assess the impact of absence of the *c1* locus on the structure of plastids at different stages of fruit ripening, using microscopy, sub-chromoplast fractionation, chromatography and western blotting.
- Use RNA-seq to generate the first white and yellow bell pepper transcriptomes, and exploit this data to select potential gene candidates for association with the *c1* locus.

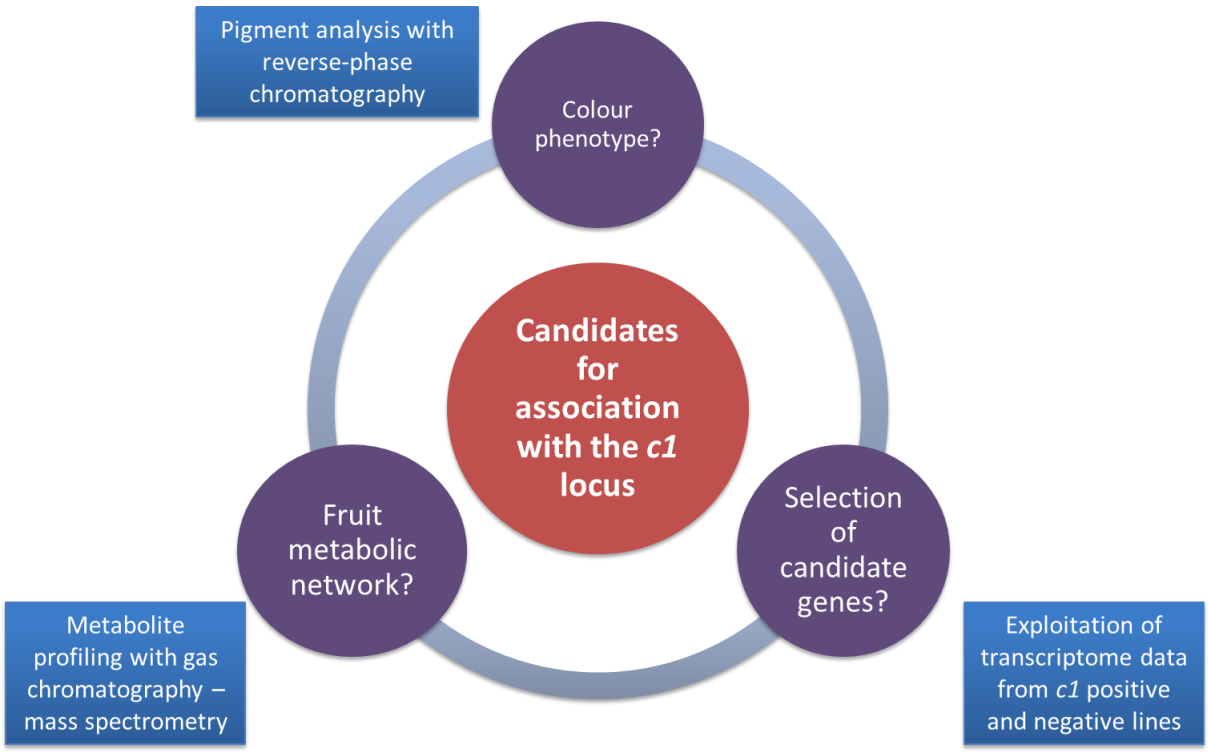


Figure 7.2. Outline of the questions and experimental approaches used to evaluate the impact of the *c1* locus on fruit phenotype, and development of candidate genes for further characterisation.

7.2 Carotenoid Accumulation

7.2.1 Benefits of pepper-specific xanthophyll accumulation

The *Capsicum annuum* accessions in *Psy-1* and *c1* diversity panels presented here were largely made up of orange and yellow fruit, which accumulate the high-value, nutritional pigments lutein and zeaxanthin. When compared to other commonly consumed fruits and vegetables in the UK, orange pepper cultivars contain the highest amount of zeaxanthin (Fraser & Bramley, 2004). The health benefits of these compounds have been well documented, and are particularly important in preventing age-related degeneration of the macula (Bernstein et al., 2017). Several large scale clinical trials have established the improvements in ocular health brought about by these pigments. The high concentration of lutein and zeaxanthin gives the fovea of the eye its yellow colour, resulting in the name '*macula lutea*' (Bernstein et al., 2017). As previously described, one of the most important functional properties of the carotenoids is their antioxidant capability. This is crucial for their role in the eye: potential for generation of Reactive Oxygen Species (ROS) here is high. The outer retina, especially the membranes of the outer segments of the photoreceptors, has high levels of polyunsaturated fatty acids which are susceptible to photo-oxidation (Cai, Nelson, Wu, Sternberg Jr, & Jones, 2000; Winkler, Boulton, Gottsch, & Sternberg, 1999). In the presence of ROS, lutein and zeaxanthin undergo oxidation and a series of transformations which protect the macula against oxidative damage. Lutein and zeaxanthin oxidation products have been identified in human serum,

monkey retinas, and the human eye (Bernstein et al., 2001; Khachik et al., 2002; Khachik, Bernstein, & Garland, 1997).

Clinical trials have shown the positive impact which carotenoid supplements have on the visual performance of participants. The Collaborative Optical Macular Pigment Assessment Study (COMPASS) was a randomized control trial involving 121 normal participants aged between 18 and 41, with the active group consuming 12mg lutein and 1mg zeaxanthin each day for 12 months (Nolan et al., 2011). Significant improvements of the antioxidant capability of the macula was observed in the group taking carotenoid supplements (Nolan et al., 2011). Another randomized, double-blind intervention trial also studied 120 normal recruits over the course of one year: in this case the active group took supplements of 20mg lutein each day (Yao et al., 2013). After measuring participants' visual acuity, visual performance and vision-related quality of life at time points of 1, 3, 6, and 12 months during the trial, results showed significant improvements in contrast sensitivity and glare disability as well as increases in vision-related quality of life (Yao et al., 2013).

Some orange and yellow pepper accessions in the two panels presented contained several hundred micrograms per gram dry weight (µg/g DW) of lutein (for example, RH 1 contained 577.98 µg/g DW). This is equal to approximately 58 µg/g fresh weight. A supermarket bought fresh bell pepper weighs on average 170g (Waitrose), meaning a consumer could hope to gain a total of 9 mg dietary lutein per fruit. Evidence from the clinical trials suggests this amount would contribute substantially to improvements in visual health Therefore these pepper accessions are highly nutritionally valuable.

7.2.2 Insights from profiling esterified carotenoids

Subjecting non-polar extracts from yellow, orange and red lines to High Pressure Liquid Chromatography (HPLC) revealed the distinctive pepper-specific chromatogram, featuring approximately fifty peaks eluting between ten and thirty-five minutes of the sixty minute run. Free carotenoids eluted in the first eighteen minutes of the run, the majority of these peaks were carotenoid mono- and di-esters. Combining mass spectrometry with liquid chromatography (LC-MS) enabled the identification of the specific fatty acids esterifying the hydroxyl groups of the carotenoids. The fact that pepper accumulates xanthophylls (with a hydroxyl group) as opposed to carotenes (for example, lycopene, which have no hydroxyl group), makes the storage mechanism of esterification possible for the species. Modifying carotenoids with the addition of fatty acid chains has nutritional implications for consumers of the fruit which contains them.

Due to the greatly increased complexity of chromatograms from esterified samples, most studies include a saponification step to remove the fatty acids before analysis. There are examples in the literature of focussed work on esters of a particular xanthophyll or class of xanthophylls, for example

Breithaupt and colleagues investigated esters of β-cryptoxanthin in orange juice, and Dugo and co-workers published a study of epoxycarotenoids in the same sample type (Breithaupt & Bamedi, 2001; Dugo, Giuffrida, Herrero, Donato, & Mondello, 2009). Carotenoid esters in red chilli pepper have been the subject of detailed analysis, with twenty-eight separate esterified carotenoids being identified by Schweiggert and colleagues (11 mono-esters and 17 di-esters) (Schweiggert et al., 2005). *Capsicum* carotenoid esters have also been used as biomarkers of fruit ripeness: Hornero-Mendez and Minguez-Mosquera measured un-saponified carotenoid composition of five *Capsicum* varieties over the course of ripening, and found that, independent of variety, all accessions achieved a similar balance between free, mono- and di- esterified carotenoids when fully ripe (Hornero-Mendez & Mınguez-Mosquera, 2000). Another conclusion of the study was that carotenoid esterification occurs independently of chlorophyll catabolism: one variety included in the investigation retained its chlorophyll as ripening progressed and reached the same free/esterified balance as the chlorophyll-losing varieties.

Esters of the nutritionally important carotenoids lutein and zeaxanthin were detected in yellow and orange accessions in the two panels studied in the present work. Does esterification of the compounds have a beneficial impact on their bioavailability or antioxidant properties? If so, is this dependent on the type of fatty acid? This could have implications for the nutritional value of fruit from the two panels presented in this work: the majority of carotenoids are esterified. Whether they are preferentially mono- or di- esterified depends on the compound, for example, more luteoxanthin and violaxanthin mono-esters were detected, whereas antheraxanthin, lutein, and zeaxanthin were preferentially di-esterified. Currently, it is understood that carotenoid absorption in human metabolism is inefficient: less than 30% of carotenoids are absorbed (Olson, 1994). There are three main stages to the process: release from the food matrix, incorporation into micelles, and uptake into the intestinal mucosa. These are influenced by several interacting factors; interestingly fat seems to exert the greatest effect. The hydroxyl groups of xanthophylls seem to provide an advantage: their polarity means that they are distributed on the surface of the emulsion, whereas carotenes are retained in the hydrophobic core. Therefore the former are more readily available for transfer from emulsions to micelles (During, Hussain, Morel, & Harrison, 2002). There is evidence to suggest that esterified carotenoids are more bioavailable than their un-esterified counterparts: a study in 1997 measured the lutein content in blood plasma over a time course from 0-408 hours after a dose of either free or esterified lutein. In the resulting concentration-time curves, esterified lutein had a higher total area. The increased liposolubility of the molecule is likely to be responsible for this (Herbst et.al., 1997). There is evidence suggesting that hydrolysis of esterified carotenoids occurs before absorption: an investigation into uptake of esterified β-cryptoxanthin found that the

same amount of free and esterified β-cryptoxanthin was absorbed by subjects, however, only the free carotenoid was found in plasma of subjects who received a dose of esterified compound. The enzymes responsible for this hydrolysis are yet to be discovered (Wingerath, Stahl, & Sies, 1995). It is possible that the addition of unsaturated fatty acids has a negative impact on the antioxidant properties of xanthophylls: in paprika samples it was found that unsaturated zeaxanthin esters degraded more quickly than the free carotenoid (Perez-Galvez & Minguez-Mosquera, 2002, 2004). However, if esterification has no negative effect on the bioavailability of the compound, and the fatty acids are hydrolysed from the carotenoid before absorption into the plasma, the lowering of antioxidant capability by esterification would not be a problem. Therefore, the carotenoid profiles of lines in the two panels presented in this thesis support the idea that yellow and orange *Capsicum* accessions are uniquely nutritionally valuable due to their high levels of free and esterified lutein and zeaxanthin.

7.3 Fatty Acid Accumulation

7.3.1 Relationship between esterifying fatty acids and carotenoid composition

Accessions from both the *c2* and *c1* diversity panels were subjected to global metabolite profiling with Gas Chromatography-Mass Spectrometry (GC-MS). One of the major outcomes from these experiments, for both panels, was the positive relationship between total carotenoid and C18≥ fatty acids. This is likely to be due to the fact that these fatty acids are capable of esterifying xanthophylls, and therefore would be preferentially accumulated in lines with high xanthophyll content. It has previously been described that xanthophylls are mono- or di- esterified with C12:0, C14:0, C16:0, C18:1, and C18:2 fatty acids (Biacs, Daood, Pavisa, & Hajdu, 1989). However, different fatty acid levels could be the linked to other differences between the lines. As membrane structure is dependent on fatty acid content, the membrane solubility of certain compounds in lines with slightly altered fatty acid chain length ratios could be changed.

7.3.2 Insights from Very Long Chain Fatty Acid (VLCFA) composition

One of the most striking differences between *c1* positive and negative lines was the higher accumulation of VLCFAs with acyl chain of lengths greater than twenty carbons (for example, docosanoic acid and hexcosanoic acid (C20:0, and C26:0 respectively)) in fruit lacking a functional copy of the *c1* locus. The resulting accumulation of VLCFAs is particularly interesting in the light of the altered membrane structures seen in TEM images of plastids from *c1* negative fruit. The fact that the thylakoid envelopes in *c1* negative lines could contain different ratios of fatty acids may contribute to the explanation of their aberrant shape and stacking behaviour. The accumulation of VLCFAs could also be related to a lack of xanthophyll esterification capability: fatty acids with chains longer than C18 are not observed as the chains added to the hyroxy groups of xanthophylls. If fatty acids with chain lengths longer than this are being accumulated by these lines, esterification during the course of ripening could be affected.

These have been previously detected in pepper (Conforti et al., 2007; Sora et al., 2015). They are normally associated with cuticular or epicuticular waxes; in *Arabidopsis*, knocking out a key gene in VLCFA biosynthesis led to embryo fatality, suggesting that the compounds play a crucial developmental role involving patterning the plant through polar auxin transport (Roudier et al., 2010). In fruit, VLCFAs are involved in water retention: it has been shown in tomato that β-ketoacyl-coenzyme A synthase (a key gene in VLCFA biosynthesis) knockout mutants experienced a 3-8 fold increase in water loss when fruit began to ripen (immature fruits remained unaffected) (Leide et al., 2007). This could imply that, with significantly less VLCFA, pepper fruits with functional copies of the *c1* locus are at risk of losing water at a greater rate.

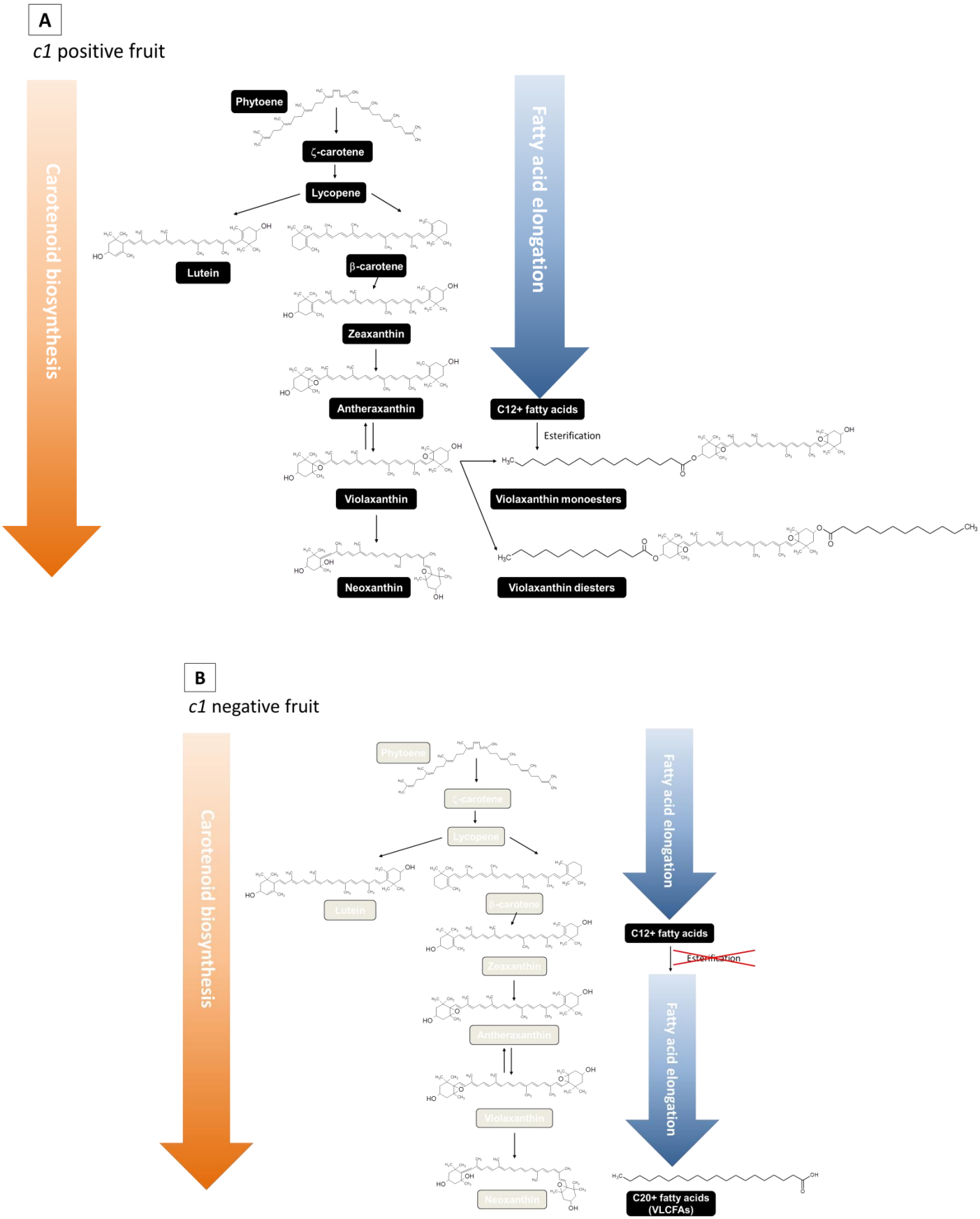


Figure 7.3. Outline of a possible explanation for the higher amounts of VLCFAs in fruit without a functional copy of the c1 locus. In the absence of carotenoids to esterify, fatty acid elongation continues resulting in fatty acids with lengths of C20+. A. Carotenoid biosynthesis and fatty acid elongation in *c1* positive fruit. B. Carotenoid biosynthesis and fatty acid elongation in *c1* negative fruit. Black textbox: compound accumulated. Grey textbox: compound not accumulated.

7.4 Plastid Structure and Pigment Storage

7.4.1 Chloroplast structure: carotenoid biosynthesis and storage

Accessions displaying aberrant thylakoid structures have the potential to become models for investigations into carotenoid retention and its associated membrane systems. This could be especially useful in the light of the lack of progress in stable transformation protocols for pepper: these lines could be utilised as null mutants for desired colour phenotypes. Detailed analysis of transcript levels and plastid electron microscopy images at many stages of ripening would give further clues about why pigment is not retained throughout ripening. High-throughput lipid analysis would also be crucial; comparing lipids in a detailed ripening series could inform about a potential missing factor or altered chain length ratio which could be held responsible for the sub-plastid structures. The technique electrospray ionisation (ESI) mass spectrometry has been used previously to profile lipids at sub-organelle levels: in 2016 it was used to generate lipid profiles of exosomes of colorectal cancer patients (Lydic, Townsend, Adda, Collins, & Reid, 2016; Welti & Wang, 2004). Lydic et al. aimed to uncover specific lipids which could be used as biomarkers for disease; the same could be applied to find specific chloroplast lipids associated with the white fruit colour phenotype.

These further investigations could also lead to application in a biotechnological context: in-depth knowledge of the mechanisms through which specific genes act could enable them to be used to modify chloroplast compartments in a targeted way. This could be carried out in non-pepper species; it would be interesting to observe whether alterations in homologous genes in related plants gave similar phenotypes. Furthermore, current transformation protocols in pepper would not facilitate stable modifications in targeted genes.

- Valuable information: what do lines not retaining pigment lack in the membrane? This must be something which the membrane possesses as the carotenoids are synthesised before the onset of ripening, and which they lose or do not synthesise when ripening begins.
- However, the background from which ripening begins is not typical: evidence from EM and RNAseq has shown this.
- How could this discovery be applied in a biotechnological context?
- If the functions of the genes causing the difference in plastid structure were dissected in detail, they could be used to modify compartments in targeted ways thus exploiting the biotechnological potential of the organelle.

7.4.2 Chromoplast structure: pepper-specific storage mechanisms

The unique carotenoid profile of pepper is made up of diverse xanthophylls and xanthophyll esters as well as the ketocarotenoids capsanthin and capsorubin in red pepper; these require unique storage mechanisms. To gain a comprehensive understanding of the production of colour phenotypes in ripe fruit, the structures used of pigment sequestration in diverse lines must be

studied. Employing a sub-chromoplast fractionation technique to do this provided a means of obtaining detailed information about pigments and proteins within defined sub-plastid structures. This is the first time the technique has been utilised in bell pepper, and adds important information to our understanding of the development of ripe colour phenotype (Nogueira et al., 2016).

Investigations into plastid structure presented in Chapters Four and Five of this thesis reveal important differences between lines possessing a functional copy of *Psy-1.* Broadly, accessions which synthesised ripening related carotenoids possessed more of the storage architecture and esterification-related fatty acids than lines which did not. As the gene is expressed at the onset of ripening, its deletion does not affect the chloroplast of mature green fruit. Protein and polar lipid structures known as fibrils are particularly important for chromoplast structure in pepper: they are hypothesised to be the final storage destination of carotenoids, and are thought to be a mechanism by which highly pigmented peppers facilitate carotenoid over-accumulation (Deruère et al., 1994b). A novel finding of the current work was that, in lines lacking *Psy-1* gene products, and therefore accumulating much lower levels of ripe fruit pigment, the amount of fibril protein decreases accordingly. This was observed in both western blot and electron microscopy results. The sub-plastid locations of carotenoid biosynthetic enzymes were also investigated; revealing their localisation in the membrane and fibril region, and in the stroma. This result for PSY-1 had been previously reported (Gómez-García & Ochoa-Alejo, 2013).

7.5 Future Directions

To assess the impact on fruit colour phenotype of the genes chosen for association with the *c1* locus, gene silencing techniques could be employed. For example, transiently silencing the genes with Virus Induced Gene Silencing (VIGS) would reveal the impact each had on fruit phenotypes at different stages of ripening. Carrying out electron microscopy studies on mature green fruit of lines transiently transformed with VIGS would determine the regulatory effect of the genes of interest on chloroplast compartment structure. Confirmation that the gene was responsible for the white ripe fruit phenotype would be obtained by transforming fruit which ripened yellow and were negative for the *y* and *c2* loci, and observing a white mature fruit.

Detailed fatty acid profiling of extracted plastids at precise time points throughout fruit ripening would provide results from which further hypotheses on the function of the fatty acid biosynthetic genes identified in relation to the chloroplast. Gene expression studies could confirm the impact of a lack of *ABC1K* gene products on their expression throughout ripening.

7.6 Project Outcomes

7.6.1 New molecular marker for absence of *Psy-1*

The *Psy-1* marker has been produced with a conventional breeding technique; there are advantages and disadvantages to this approach. The disadvantages are the potential off-target effects on unrelated genes in the genome, the slow speed of introduction into breeding populations, and the necessity of time consuming processes such as large-scale screening and back crossing. Correspondingly, the advantages are the relative low cost of the approach and its societal approval. The latter will ease translation of the resulting products to the market.

7.6.2 Economic implications of the new *Psy-1* absence marker

Peppers are expensive crops to grow: according to the Agricultural Marketing Resource Centre (AGMRC; http://www.agmrc.org) peppers grown in the United States are grown with traditional methods, using drip irrigation and mulch, and are harvested by hand approximately every week over a period of four to six weeks. In 2014, the U. S. grew 1,535 million pounds of bell pepper on over 46,000 acres. That year, the bell pepper crop was valued at $618.9 million. Although most of the crop is harvested at the mature green stage, growers receive a premium for selling higher-value speciality colour varieties. In 2015, one acre of bell pepper crop was given a gross value of $13,300 (AGMRC, 2015).

According to the Food and Agriculture Organisation's statistical unit (FAO Stat), ripe yellow and red peppers are valued at 2.5 times the price of a green bell pepper, but orange fruit obtain the highest prices, being worth 3 times the price of a green pepper. In the light of this information, it is clear that the new marker for the absence of *Psy-1* will be extremely useful in assisting breeders to bring fruit of the highest value to the marketplace.

7.6.3 Candidate genes for association with the *c1* locus

Insights from the transcriptomes of yellow and white bell peppers combined with information on the location of the molecular marker for the *c1* locus provided several potential genes for association with the locus. These were: *ABC1K*, the esterase/lipase *GDL62*, the acyl-transferase *Catf2*, and the NAC transcription factor *NAC37.*

As differences in both lipid composition and plastid compartments had been observed between *c1* positive and negative lines, it is interesting that several differentially expressed genes influencing both these factors were found in the region of chromosome 2 which the *c1* marker has been mapped to. As chloroplast compartment development could impact the retention of pigments as the fruit ripen, it is hypothesised that the plastid-specific kinase ABC1K would have the greatest impact on ripe fruit colour phenotype, and is therefore the protein encoded by the gene associated with the

c1 locus. Abnormal development of these compartments would affect the composition of the fatty acids they are composed of, and could explain differences in fatty acid biosynthetic gene expression. The hypothetical mechanism is outlined in Figure 7.4.

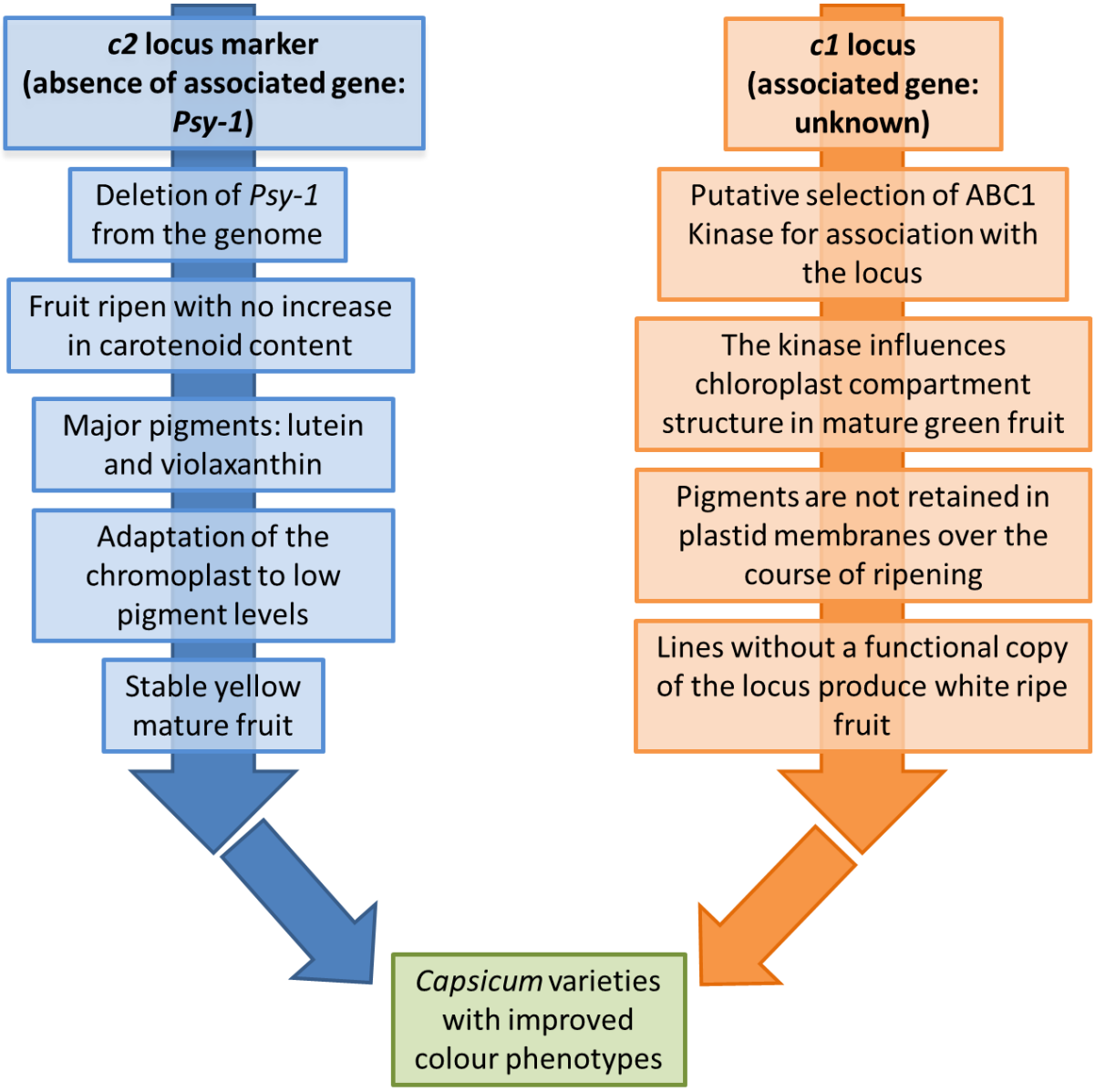


Figure 7.4. Summary of the mechanisms through which markers for the *c1* and *c2* loci can be used to produce high value fruit with stable colour phenotypes.

7.7 Overview

Objective	Outcome
i) 1. Determine the carotenoid profiles of lines in a diversity panel for the marker associated with *c2*.	Data produced is discussed in Chapter Three. Experiments revealed a lack of ripening associated carotenoid synthesis, giving fruit without *Psy-1* their yellow colour.
i) 2. Sequence *Psy-1* in lines of diverse phenotype to determine the nature of the mutation in genomic DNA.	Data produced is discussed in Chapter Three. Deletion of the gene from the genome in lines possessing the marker was confirmed.
i) 3. Confirm the absence of *Psy-1* transcripts from lines possessing the marker with quantitative real time PCR.	Data produced is discussed in Chapter Three. Absence of *Psy-1* gene products in lines possessing the marker was confirmed.
i) 4. Assess the impact of the absence of *Psy-1* gene products on transcript levels of other genes encoding enzymes in the carotenoid biosynthesis pathway.	Data produced is discussed in Chapter Three. A significant decrease in expression levels of *Pds* in lines possessing functional copies of *Psy-1* was observed.
i) 5. Profile fruit metabolites at different stages of ripening in lines with and without functional *Psy-1* gene products, using GC-MS.	Data produced is discussed in Chapter Four. Significant differences in long-chain fatty acid levels between Psy-1 positive and negative lines were observed and associated with carotenoid esterification.
i) 6. Assess the impact of the absence of *Psy-1* on the structural components of the organelles synthesising and storing carotenoids at different stages of fruit ripening, using microscopy, sub-chromoplast fractionation, western blots, and chromatography to characterise both pigment and protein content.	Data produced is discussed in Chapter Four. Chromoplasts in Psy-1 negative lines employ the same storage mechanisms, but adapt to decreased carotenoid levels.
ii) 1. Determine the pigment profiles from lines making up a second diversity panel for the possession of the *c1* locus, lack of which is associated with white ripe fruit.	Data produced is discussed in Chapter Five. Lines possessing the marker associated with the *c1* locus contained yellow and orange carotenoids, while *c1*-lacking lines contained only trace levels of lutein or no pigment in ripe fruit.
ii) 2. Employ GC-MS to profile fruit metabolites at different stages of fruit ripening in accessions with and without functional copies of the *c1* locus.	Data produced is discussed in Chapter Five. Lines without the c1 locus contained significantly more VLCFAs, which could be associated with the absence of xanthophylls for esterification.
ii) 3. Assess the impact of absence of the *c1* locus on the structure of plastids at different stages of fruit ripening, using electron microscopy.	Data produced is discussed in Chapter Five. Aberrant chloroplast structures in lines lacking a functional copy of the *c1* locus. Chromoplasts from white ripe fruit also adapt to their lack of carotenoid.
ii) 4. Use RNA-seq to generate the first white and yellow bell pepper transcriptomes, and exploit this data to select potential gene candidates for association with the *c1* locus.	Data produced is discussed in Chapter Six. Information on the location of the marker associated with *c1* was used to select candidates in the correct region of the genome.

Table 7.8. Summary of the experimental objectives of the project and their outcomes.

8 Bibliography

Aida, M., Ishida, T., Fukaki, H., Fujisawa, H., & Tasaka, M. (1997). Genes Involved in Organ Separation in Arabidopsis: An Analysis of the cup-shaped cotyledon Mutant. *The Plant Cell*, *9*(June), 841–857.

Akoh, C. C., Lee, G., Liaw, Y., Huang, T., & Shaw, J. (2004). Progress in Lipid Research GDSL family of serine esterases/lipases. *Progress in Lipid Research*, *43*, 534–552. http://doi.org/10.1016/j.plipres.2004.09.002

Alpha-Tocopherol Beta Carotene Cancer Prevention Study Group. (1994). The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. *The New England Journal of Medicine*, *330*(15), 1029–35.

Ampomah-Dwamena, C., Dejnoprat, S., Lewis, D., Sutherland, P., Volz, R. K., & Allan, A. C. (2012). Metabolic and gene expression apple (Malus x domestica) carotenogenesis. *Journal of Experimental Botany*, *63*(2), 695–709. http://doi.org/10.1093/jxb/err313

Andersson, M., Turesson, H., Nicolia, A., Falt, A., Samuelsson, M., & Hofvander, P. (2016). Efficient targeted multiallelic mutagenesis in tetraploid potato (Solanum tuberosum) by transient CRISPR-Cas9 expression in protoplasts. *Plant Cell Reports*, 1–12. http://doi.org/10.1007/s00299-016-2062-3

Andrade-Souza, V., Costa, M., Chen, C., Gmitter, F. J., & Costa, M. (2011). Physical location of the carotenoid biosynthesis genes Psy and β- Lcy in Capsicum annuum (Solanaceae) using heterologous probes from Citrus sinensis (Rutaceae). *Genetics and Molecular Research*, *10*(1), 404–409. http://doi.org/10.4238/vol10-1gmr1025

Andrews, J. (1993). DIFFUSION OF MESOAMERICAN FOOD COMPLEX TO SOUTHEASTERN EUROPE. *Geographical Review*, *83*(2), 194–204.

Angaman, D. M., Petrizzo, R., Hernández-gras, F., Romero-segura, C., Pateraki, I., & Boronat, A. (2012). Precursor uptake assays and metabolic analyses in isolated tomato fruit chromoplasts. *Plant Methods*, *8*, 1–10.

Auldridge, M. E., Block, A., Vogel, J. T., Dabney-smith, C., Mila, I., Bouzayen, M., … Klee, H. J. (2006). Characterization of three members of the Arabidopsis carotenoid cleavage dioxygenase family demonstrates the divergent roles of this multifunctional enzyme family. *The Plant Journal*, *45*, 982–993. http://doi.org/10.1111/j.1365-313X.2006.02666.x

Austin, J. R., Frost, E., Vidi, P.-A., Kessler, F., & Staehelin, L. A. (2006). Plastoglobules are lipoprotein subcompartments of the chloroplast that are permanently coupled to thylakoid membranes and contain biosynthetic enzymes. *The Plant Cell*, *18*(7), 1693–703. http://doi.org/10.1105/tpc.105.039859

Bach, L., & Faure, J. (2010). Role of very-long-chain fatty acids in plant development, when chain length does matter. *Comptes Rendus Biologies*, *333*, 361–370. http://doi.org/10.1016/j.crvi.2010.01.014

Bailey, R. L., Gahche, J. J., Lentino, C. V, Dwyer, J. T., Engel, J. S., Thomas, P. R., … Picciano, M. F. (2011). Dietary supplement use in the United States, 2003-2006. *Journal of Nutrition*, *141*(2), 261–266. http://doi.org/10.3945/jn.110.133025.participant

Baker, N. R. (2008). Chlorophyll Fluorescence: A Probe of Photosynthesis In Vivo. *Annual Review of Plant Biology*, *59*, 89–113. http://doi.org/10.1146/annurev.arplant.59.032607.092759

Barry, C. S., Aldridge, G. M., Herzog, G., Ma, Q., McQuinn, R. P., Hirschberg, J., & Giovannoni, J. J. (2012). Altered chloroplast development and delayed fruit ripening caused by mutations in a zinc metalloprotease at the lutescent2 locus of tomato. *Plant Physiology*, *159*(3), 1086–98. http://doi.org/10.1104/pp.112.197483

Barsan, C., Sanchez-bel, P., Rombaldi, C., Egea, I., Rossignol, M., Bouzayen, M., & Pech, J. (2010). Characteristics of the tomato chromoplast revealed by proteomic analysis. *Journal of Experimental Botany*, *61*(9), 2413–2431. http://doi.org/10.1093/jxb/erq070

Bartley, G. E., & Scolnik, P. A. (1993). cDNA Cloning , Expression during Development , and Genome Mapping of PSY2 , a Second Tomato Gene Encoding Phytoene Synthase. *Journal of Biological Chemistry*, *268*(34), 25718–25721.

Beaudoin, F., Wu, X., Li, F., Haslam, R. P., Markham, J. E., Zheng, H., … Kunst, L. (2009). Functional Characterization of the Arabidopsis beta-Ketoacyl-Coenzyme A Reductase Candidates of the Fatty Acid Elongase. *Plant Physiology*, *150*(July), 1174–1191. http://doi.org/10.1104/pp.109.137497

Berger, Y., Harpaz-saad, S., Brand, A., Melnik, H., Sirding, N., Alvarez, J. P., … Ori, N. (2009). The NAC-domain transcription factor GOBLET specifies leaflet boundaries in compound tomato leaves. *Development*, *136*, 823–832. http://doi.org/10.1242/dev.031625

Bernstein, P. S., Khachik, F., Carvalho, L., Muir, G. J., Zhao, D.-Y., & Katz, N. B. (2001). Identification and Quantitation of Carotenoids and their Metabolites in the Tissues of the Human Eye. *Experimental Eye Research*, *72*, 215–223. http://doi.org/10.1006/exer.2000.0954

Bernstein, P. S., Lia, B., Vachalia, P. P., Gorusupudia, A., Shyama, R., Henriksena, B. S., & Nolan, J. M. (2017). Lutein, Zeaxanthin, and meso-Zeaxanthin: The Basic and Clinical Science Underlying Carotenoid-based Nutritional Interventions against Ocular Disease. *Progress in Retina and Eye Research*, *50*, 34–66. http://doi.org/10.1016/j.preteyeres.2015.10.003.Lutein

Berset, D., & Pfander, H. (1985). Investigation of the Carotenoid Composition in the Petals of a Garden Hybrid Narcissus cvv “Golden Harvest.” *Helvetica Chimica Acta*, *68*, 1149–1154.

Beyer, P., Al-babili, S., Ye, X., Lucca, P., Schaub, P., Welsch, R., & Rice, G. (2002). Golden Rice: Introducing the beta-Carotene Biosynthesis Pathway into Rice Endosperm by Genetic Engineering to Defeat Vitamin A Deficiency. *The Journal of Nutrition*, *132*, 506–510.

Bhattacharjee, P., Das, R., Mandal, A., & Kundu, P. (2017). Functional characterization of tomato membrane-bound NAC transcription factors. *Plant Molecular Biology*, *93*(4), 511–532. http://doi.org/10.1007/s11103-016-0579-z

Biacs, P. A., Daood, H. G., Pavisa, A., & Hajdu, F. (1989). Studies on the Carotenoid Pigments of Paprika (Capsicum annuum L.var Sz-20). *Journal of Agricultural Food Chemistry*, *37*(2), 350–353.

Bian, W., Barsan, C., Egea, I., Purgatto, E., Chervin, C., Zouine, M., … Pech, J.-C. (2011). Metabolic and Molecular Events Occurring during Chromoplast Biogenesis Metabolic and Molecular Events Occurring during Chromoplast Biogenesis. *Journal of Botany*, (February 2015), 1–13. http://doi.org/10.1155/2011/289859

Bliss, F. A., Arulsekar, S., Foolad, M. R., Becerra, V., Gillen, A. M., Warburton, M. L., … Mydin, K. K. (2002). An expanded genetic linkage map of Prunus based on an interspecific cross between almond and peach. *Genome*, *45*, 520–529. http://doi.org/10.1139/G02-011

Borovsky, Y., Oren-Shamir, M., Ovadia, R., De Jong, W., & Paran, I. (2004). The A locus that controls anthocyanin accumulation in pepper encodes a MYB transcription factor homologous to Anthocyanin2 of Petunia. *Theoretical and Applied Genetics*, *109*, 23–29.

http://doi.org/10.1007/s00122-004-1625-9

Borovsky, Y., & Paran, I. (2008). Chlorophyll breakdown during pepper fruit ripening in the chlorophyll retainer mutation is impaired at the homolog of the senescence-inducible stay-green gene, 235–240. http://doi.org/10.1007/s00122-008-0768-5

Botte, C. Y., & Marechal, E. (2014). Plastids with or without galactoglycerolipids. *Trends in Plant Science*, *19*(2), 71–78. http://doi.org/10.1016/j.tplants.2013.10.004

Bousquet, I., Dujardin, G., & Slonimski, P. P. (1991). ABC1, a novel yeast nuclear gene has a dual function in mitochondria: it suppresses a cytochrome b mRNA translation defect and is essential for the electron transfer in the bc1 complex. *The EMBO Journal*, *10*(8), 2023–2031.

Bradford, M. M. (1976). A Rapid and Sensitive Method for the Quantitation Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Analytical Biochemistry*, *72*, 248–254.

Bramley, P., Teulieres, C., Blain, I., & Schuch, W. (1992). Biochemical characterization of transgenic tomato plants in which carotenoid synthesis has been inhibited through the expression of antisense RNA to pTOM5. *The Plant Journal*, *2*(3), 343–349.

Brand, A., Borovsky, Y., Hill, T., Rahman, K. A. A., Bellalou, A., Van Deynze, A., & Paran, I. (2014). CaGLK2 regulates natural variation of chlorophyll content and fruit color in pepper fruit. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *127*(10), 2139–48. http://doi.org/10.1007/s00122-014-2367-y

Brasseur, G., Tron, P., Slonimski, P. P., & Brivet-chevillotte, P. (1997). The nuclear ABC1 gene is essential for the correct conformation and functioning of the cytochrome bc, complex and the neighbouring complexes II and IV in the mitochondrial respiratory chain. *The European Journal of Biochemistry*, *246*, 103–111.

Bravo-Garcia, A., Yasumura, Y., & Langdale, J. A. (2009). Specialization of the Golden2-like regulatory pathway during land plant evolution. *New Phytologist*, *183*, 133–141.

Breithaupt, D. E., & Bamedi, A. (2001). Carotenoid Esters in Vegetables and Fruits: A Screening with Emphasis on Beta-Cryptoxanthin Esters. *Journal of Agricultural Food Chemistry*, *49*, 2064–2070.

Brick, D. J., Brumlik, M. J., Buckley, J. T., Cao, J., Davies, P. C., Misra, S., … Upton, C. (1995). A new family of lipolytic plant enzymes with members in rice, Arabidopsis and maize. *FEBS Letters*, *377*, 475–480.

Britton, G., Liaaen-Jensen, S., & Pfander, H. (2004). *Carotenoids Handbook*. (G. Britton, S. Liaaen-Jensen, & H. Pfander, Eds.) (1st ed.). Basel, Switzerland: Springer Basel AG.

Brooks, C., Nekrasov, V., Lippman, Z. B., & Eck, J. Van. (2014). Efficient Gene Editing in Tomato in the First Generation Using the Clustered Regularly Interspaced Short Palindromic Repeats/ CRISPR-Associated9 System. *Plant Physiology*, *166*(November), 1292–1297. http://doi.org/10.1104/pp.114.247577

Bujnicki, J., Dykstra, P., Fortunato, E., Heuer, R.-D., Keskitalo, C., Nurse, P., & Villani, C. (2017). *New techniques in Agricultural Biotechnology*.

Cai, J., Nelson, K. C., Wu, M., Sternberg Jr, P., & Jones, D. P. (2000). Oxidative Damage and Protection of the RPE. *Progress in Retinal and Eye Research*, *19*(2), 205–221.

Camara, B., Bardat, F., & Moneger, R. (1982). Sites of Biosynthesis of Carotenoids in Capsicum Chromoplast. *European Journal of Biochemistry*, *127*, 255–258.

Camara, B., & Dogbo, O. (1986). Demonstration and Solubilization of Lycopene Cyclase from

Capsicum Chromoplast Membranes. *Plant Physiology*, *80*, 172–174.

Campbell, R., Ducreux, L. J. M., Morris, W. L., Morris, J. A., Suttle, J. C., Ramsay, G., … Taylor, M. A. (2010). The Metabolic and Developmental Roles of Carotenoid Cleavage Dioxygenase4 from Potato. *Plant Physiology*, *154*, 656–664. http://doi.org/10.1104/pp.110.158733

Cardazzo, B., Hamel, P., Sakamoto, W., Wintz, H., & Dujardin, G. (1998). Isolation of an Arabidopsis thaliana cDNA by complementation of a yeast abc1 deletion mutant deficient in complex III respiratory activity. *Gene*, *221*, 117–125.

Cazzonelli, C. I., & Pogson, B. J. (2010). Source to sink: regulation of carotenoid biosynthesis in plants. *Trends in Plant Science*, *15*(5), 266–274. http://doi.org/10.1016/j.tplants.2010.02.003

Chen, Y., Li, F., & Wurtzel, E. T. (2010). Isolation and Characterization of the Z-ISO Gene Encoding a Missing Component of Carotenoid Biosynthesis in Plants. *Plant Physiology*, *153*(May), 66–79. http://doi.org/10.1104/pp.110.153916

Chepyshko, H., Lai, C., Huang, L., Liu, J., & Shaw, J. (2012). Multifunctionality and diversity of GDSL esterase/lipase gene family in rice (Oryza sativa L . japonica) genome: new insights from bioinformatics analysis. *BMC Genomics*, *13*, 309–328.

Ciska, M., Masuda, K., Moreno, S., & De, D. (2013). Lamin-like analogues in plants: the characterization of NMCP1 in Allium cepa. *Journal of Experimental Botany*, *64*(6), 1553–1564. http://doi.org/10.1093/jxb/ert020

Collard, B. C. Y., & Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society of London B.*, *363*, 557–572. http://doi.org/10.1098/rstb.2007.2170

Conesa, A., & Götz, S. (2008). Blast2GO: A Comprehensive Suite for Functional Analysis in Plant Genomics. *International Journal of Plant Genomics*, *2008*, 1–13. http://doi.org/10.1155/2008/619832

Conesa, A., Götz, S., García-gómez, J. M., Terol, J., Talón, M., Genómica, D., … Valencia, U. P. De. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*(18), 3674–3676. http://doi.org/10.1093/bioinformatics/bti610

Conforti, F., Statti, G. A., & Menichini, F. (2007). Food Chemistry Chemical and biological variability of hot pepper fruits (Capsicum annuum var . acuminatum L .) in relation to maturity stage. *Food Chemistry*, *102*, 1096–1104. http://doi.org/10.1016/j.foodchem.2006.06.047

Cosio, C., & Dunand, C. (2009). Specific functions of individual class III peroxidase genes. *Journal of Experimental Botany*, *60*(2), 391–408. http://doi.org/10.1093/jxb/ern318

Cuevas, H. E., Staub, J. E., Simon, P. W., & Zalapa, J. E. (2009). A consensus linkage map identifies genomic regions controlling fruit maturity and beta-carotene-associated flesh color in melon (Cucumis melo L.). *Theoretical and Applied Genetics*, *119*, 741–756. http://doi.org/10.1007/s00122-009-1085-3

Cuttriss, A. J., Chubb, A. C., Alawady, A., Grimm, B., & Pogson, B. J. (2007). Regulation of lutein biosynthesis and prolamellar body formation in Arabidopsis. *Functional Plant Biology*, *34*, 663–672.

Davies, H. B., Matthews, S., & Kirk, J. T. O. (1970). THE NATURE AND BIOSYNTHESIS OF THE CAROTENOIDS OF DIFFERENT COLOUR VARIETIES OF CAPSICUM ANNUUM. *Phytochemistry*, *9*(1955), 797–805.

Dellapenna, D., & Pogson, B. J. (2006). Vitamin Synthesis in Plants : Tocopherols and Carotenoids. *Annual Review of Plant Biology*, *57*(7), 11–38. http://doi.org/10.1146/annurev.arplant.56.032604.144301

Denic, V., & Weissman, J. S. (2007). A Molecular Caliper Mechanism for Determining Very Long-Chain Fatty Acid Length. *Cell*, *130*, 663–677. http://doi.org/10.1016/j.cell.2007.06.031

Depka, B., Jahns, P., & Trebst, A. (1998). Beta-Carotene to zeaxanthin conversion in the rapid turnover of the D1 protein of photosystem II. *FEBS Letters*, *424*, 267–270.

Deruère, J., Römer, S., d'Harlingue, A., Backhaus, R. A., Kuntz, M., & Camara, B. (1994a). Fibril assembly and carotenoid overaccumulation in chromoplasts: a model for supramolecular lipoprotein structures. *The Plant Cell*, *6*(1), 119–33. http://doi.org/10.1105/tpc.6.1.119

Deruère, J., Römer, S., d'Harlingue, a, Backhaus, R. a, Kuntz, M., & Camara, B. (1994b). Fibril assembly and carotenoid overaccumulation in chromoplasts: a model for supramolecular lipoprotein structures. *The Plant Cell*, *6*(1), 119–33. http://doi.org/10.1105/tpc.6.1.119

Dietrich, C. R., Perera, M. A. D. N., Yandeau-Nelson, M. D., Meeley, R. B., Nikolau, B. J., & Schnable, P. S. (2005). Characterization of two GL8 paralogs reveals that the 3-ketoacyl reductase component of fatty acid elongase is essential for maize (Zea mays L.) development. *The Plant Journal*, *42*, 844–861. http://doi.org/10.1111/j.1365-313X.2005.02418.x

Do, T. Q., Hsu, A. Y., Jonassen, T., Lee, P. T., & Clarke, C. F. (2001). A Defect in Coenzyme Q Biosynthesis Is Responsible for the Respiratory Deficiency in Saccharomyces cerevisiae abc1 Mutants. *The Journal of Biological Chemistry*, *276*(21), 18161–18168. http://doi.org/10.1074/jbc.M100952200

Dogbo, O., Bardat, F., Laferriere, A., Quennemet, J., Brangeon, J., & Camara, B. (1987). Metabolism of plastid terpenoids. I. biosynthesis of phytoene in plastid stroma isolated from higher plants. *Plant Science*, *49*, 89–101.

Dong, T., Hu, Z., Deng, L., Wang, Y., Zhu, M., Zhang, J., & Chen, G. (2013). A Tomato MADS-Box Transcription Factor, SlMADS1, Acts as a Negative Regulator of Fruit Ripening. *Plant Physiology*, *163*(October), 1026–1036. http://doi.org/10.1104/pp.113.224436

Dugo, P., Giuffrida, D., Herrero, M., Donato, P., & Mondello, L. (2009). Epoxycarotenoids esters analysis in intact orange juices using two-dimensional comprehensive liquid chromatography. *Journal of Separation Science*, *32*, 973–980. http://doi.org/10.1002/jssc.200800696

Dugo, P., Herrero, M., Giuffrida, D., Ragonese, C., Dugo, G., & Mondello, L. (2008). Original Paper Analysis of native carotenoid composition in orange juice using C 30 columns in tandem, 2151–2160. http://doi.org/10.1002/jssc.200800073

Dunn, T. M., Lynch, D. V., Michaelson, L. V., & Napier, J. A. (2004). A Post-genomic Approach to Understanding Sphingolipid Metabolism in Arabidopsis thaliana. *Annals of Botany*, *93*, 483–497. http://doi.org/10.1093/aob/mch071

During, A., Hussain, M. M., Morel, D. W., & Harrison, E. H. (2002). Carotenoid uptake and secretion by CaCo-2 cells: beta-carotene isomer selectivity and carotenoid interactions. *Journal of Lipid*, *43*, 1086–95. http://doi.org/10.1194/jlr.M200068-JLR200

Efrati, A., Eyal, Y., & Paran, I. (2005). Molecular mapping of the chlorophyll retainer (cl) mutation in pepper (Capsicum spp.) and screening for candidate genes using tomato ESTs homologous to structural genes of the chlorophyll catabolism pathway. *Genome*, *48*(2), 347–351. http://doi.org/10.1139/G04-119

Egea, I., Barsan, C., Bian, W., Purgatto, E., Latché, A., Chervin, C., … Pech, J. (2010). Chromoplast Differentiation: Current Status and Perspectives. *Plant Cell and Physiology*, *51*(10), 1601–1611. http://doi.org/10.1093/pcp/pcq136

Eugeni Piller, L., Abraham, M., Dormann, P., Besagni, C., & Kessler, F. (2012). Plastid lipid droplets at the crossroads of prenylquinone metabolism. *Journal of Experimental Botany*, *63*(4), 1609–1618. http://doi.org/10.1093/jxb/ers016

Falchi, R., Vendramin, E., Zanon, L., Scalabrin, S., Cipriani, G., Verde, I., … Morgante, M. (2013). Three distinct mutational mechanisms acting on a single gene underpin the origin of yellow flesh in peach. *The Plant Journal*, *76*, 175–187. http://doi.org/10.1111/tpj.12283

Ferreira, K. N., Iverson, T. M., Maghlaoui, K., Barber, J., & Iwata, S. (2004). Architecture of the Photosynthetic Oxygen-Evolving Center. *Science*, *303*(March), 1831–1836.

Fitter, D. W., Martin, D. J., Copley, M. J., Scotland, R. W., & Langdale, J. A. (2002). GLK gene pairs regulate chloroplast development in diverse plant species. *The Plant Journal*, *31*(6), 713–727.

Floss, D. S., & Walter, M. H. (2009). Role of carotenoid cleavage dioxygenase 1 (CCD1) in apocarotenoid biogenesis revisited. *Plant Signaling & Behavior*, *4*(3), 172–175. http://doi.org/10.4161/psb.4.3.7840

Fraser, P. D., & Bramley, P. M. (2004). The biosynthesis and nutritional uses of carotenoids. *Progress in Lipid Research*, *43*(3), 228–65. http://doi.org/10.1016/j.plipres.2003.10.002

Fraser, P. D., Enfissi, E. M. A., Halket, J. M., Truesdale, M. R., Yu, D., Gerrish, C., & Bramley, P. M. (2007). Manipulation of Phytoene Levels in Tomato Fruit: Effects on Isoprenoids, Plastids, and Intermediary Metabolism. *The Plant Cell*, *19*(October), 3194–3211. http://doi.org/10.1105/tpc.106.049817

Fraser, P. D., Truesdale, M. R., Bird, C. R., Schuch, W., & Bramley, P. M. (1994). Carotenoid Biosynthesis during Tomato Fruit Development (Evidence for Tissue-Specific Gene Expression). *Plant Physiology*, *105*(1), 405–413. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=159369&tool=pmcentrez&rendertype=abstract

Fray, R., & Grierson, D. (1993). Identification and genetic analysis of normal and mutant phytoene synthase genes of tomato by sequencing , complementation and co-suppression. *Plant Molecular Biology*, *22*, 589–602.

Gamage, S. M. K. W., Mcgrath, D. J., Persley, D. M., & Dietzgen, R. G. (2016). Transcriptome Analysis of Capsicum Chlorosis Virus-Induced Hypersensitive Resistance Response in Bell Capsicum. *PloS One*, *11*(7), 1–21. http://doi.org/10.1371/journal.pone.0159085

García-Limones, C., Schnäbele, K., Blanco-Portales, R., Bellido, M. L., Caballero, J. L., Schwab, W., & Muñoz-Blanco, J. (2008). Functional Characterization of FaCCD1: A Carotenoid Cleavage Dioxygenase from Strawberry Involved in Lutein Degradation during Fruit Ripening. *Journal of Agricultural Food Chemistry*, *56*, 9277–9285.

Gillissen, B., Bürkle, L., André, B., Kühn, C., Rentsch, D., Brandl, B., & Frommer, W. B. (2000). A New Family of High-Affinity Transporters for Adenine, Cytosine, and Purine Derivatives in Arabidopsis. *The Plant Cell*, *12*(February), 291–300.

Giovannoni, J. J. (2004). Genetic Regulation of Fruit Development and Ripening. *The Plant Cell*, *16*, 170–181. http://doi.org/10.1105/tpc.019158.Fruit

Giovannoni, J. J. (2007). Fruit ripening mutants yield insights into ripening control. *Current Opinion in*

Plant Biology, *10*, 283–289. http://doi.org/10.1016/j.pbi.2007.04.008

Giovannoni, J. J., Noensie, E. N., Ruezinsky, D. M., Lu, X., Tracy, S. L., Ganal, M. W., … Tanksley, S. D. (1995). Molecular genetic analysis of the ripening inhibitor and nonripening loci of tomato: a first step in genetic map-based cloning of fruit ripening genes. *Molecular and General Genetics*, *248*, 195–206.

Giuliano, G., & Diretto, G. (2007). Of chromoplasts and chaperones. *Trends in Plant Science*, *12*(12), 529–531.

Gómez-García, M. del R., & Ochoa-Alejo, N. (2013). Biochemistry and molecular biology of carotenoid biosynthesis in chili peppers (Capsicum spp.). *International Journal of Molecular Sciences*, *14*(9), 19025–53. http://doi.org/10.3390/ijms140919025

Gomez-Roldan, V., Fermas, S., Brewer, P. B., Puech-page, V., Pillot, J., Letisse, F., … Rochange, S. F. (2008). Strigolactone inhibition of shoot branching. *Nature*, *455*(September), 189–195. http://doi.org/10.1038/nature07271

Góngora-castillo, E., Fajardo-jaime, R., Fernández-cortes, A., Jofre-, A. E., Lozoya-gloria, E., Martínez, O., … Rivera-, R. (2012). The Capsicum transcriptome DB: a “hot” tool for genomic research. *Bioinformation*, *8*(1), 43–47.

Götz, S., Terol, J., Williams, T. D., Garcı, J. M., Nagaraj, S. H., Jose Nueda, M., … Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, *36*(10), 3420–3435. http://doi.org/10.1093/nar/gkn176

Gross, J., Eckhardt, G., Sonnen, H. D., & Lenz, F. (1979). Einfluß unterschiedlicher Stickstoffversorgung auf Flavedocarotinoide von Calamondinen (Citrus madurensis Lour .) / Effect of Different Nitrogen Supply on Flavedo Carotenoids of Calamondin Fruit Einfluß unterschiedlicher Stickstoffversorgung auf Flavedo-. *Gartenbeuwissenschaft*, *44*, 45–48.

Guo, W.-L., Wang, S.-B., Chen, R.-G., Chen, B.-H., Xiao-Du, H., Yin, Y.-X., … Zhang, Y.-Y. (2015). Characterization and expression profile of CaNAC2 pepper gene. *Frontiers in Plant Science*, *6*(September), 1–9. http://doi.org/10.3389/fpls.2015.00755

Guzman, I., Hamby, S., Romero, J., Bosland, P. W., & O’Connell, M. a. (2010). Variability of Carotenoid Biosynthesis in Orange Colored Capsicum spp. *Plant Science : An International Journal of Experimental Plant Biology*, *179*(1–2), 49–59. http://doi.org/10.1016/j.plantsci.2010.04.014

Ha, S.-H., Kim, J.-B., Park, J.-S., Lee, S.-W., & Cho, K.-J. (2007). A comparison of the carotenoid accumulation in Capsicum varieties that show different ripening colours: deletion of the capsanthin-capsorubin synthase gene is not a prerequisite for the formation of a yellow pepper. *Journal of Experimental Botany*, *58*(12), 3135–44. http://doi.org/10.1093/jxb/erm132

Hadley, C. W., Clinton, S. K., & Schwartz, S. J. (2003). Human Nutrition and Metabolism The Consumption of Processed Tomato Products Enhances Plasma Lycopene Concentrations in Association with a Reduced Lipoprotein Sensitivity to Oxidative Damage. *The Journal of Nutrition*, *133*, 727–732.

Hao, Q., Yin, P., Li, W., Wang, L., Yan, C., Lin, Z., … Yan, N. (2011). The Molecular Basis of ABA-Independent Inhibition of PP2Cs by a Subclass of PYL Proteins. *Molecular Cell*, *42*(5), 662–672. http://doi.org/10.1016/j.molcel.2011.05.011

Heiser, C. B., & Pickersgill, B. (1969). Names for the Cultivated Capsicum Species (Solanaceae). *Taxon*, *18*(3), 277–283.

Hieber, A. D., Bugos, R. C., & Yamamoto, H. . (2000). Plant lipocalins: Violaxanthin de-epoxidase and zeaxanthin epoxidase. *Biochimica et Biophysica Acta*, *1482*, 84–91. http://doi.org/10.1016/S0167-4838(00)00141-2

Hirakawa, H. I., Shirasawa, K. E., Miyatake, K. O. J. I., Nunome, T. S., Negoro, S. A., Ohyama, A. K. I. O., … Fukuoka, H. I. (2014). Draft Genome Sequence of Eggplant (Solanum melongena L.): the Representative Solanum Species Indigenous to the Old World. *DNA Research*, *21*(September), 649–660.

Holzl, G., & Dormann, P. (2007). Structure and function of glycoglycerolipids in plants and bacteria. *Progress in Lipid Research*, *46*, 225–243. http://doi.org/10.1016/j.plipres.2007.05.001

Hong, L., Brown, J., Segerson, N. A., Rose, J. K. C., & Adrienne, H. K. (2017). CUTIN SYNTHASE2 maintains progressively developing cuticular ridges in Arabidopsis sepals. *Molecular Plant*. http://doi.org/10.1016/j.molp.2017.01.002

Hornero-Mendez, D., & Mınguez-Mosquera, I. (2000). Xanthophyll Esterification Accompanying Carotenoid Overaccumulation in Chromoplast of Capsicum annuum Ripening Fruits Is a Constitutive Process and Useful for Ripeness Index. *Journal of Agricultural Food Chemistry*, *48*, 1617–1622.

Horton, P., & Ruban, A. (2005). Molecular design of the photosystem II light-harvesting antenna: Photosynthesis and photoprotection. *Journal of Experimental Botany*, *56*, 365–373. http://doi.org/10.1093/jxb/eri023

Howitt, C. A., Cavanagh, C. R., Bowerman, A. F., Cazzonelli, C., Rampling, L., Mimica, J. L., & Pogson, B. J. (2009). Alternative splicing, activation of cryptic exons and amino acid substitutions in carotenoid biosynthetic genes are associated with lutein accumulation in wheat endosperm. *Functional and Integrative Genomics*, *9*(June), 363–376. http://doi.org/10.1007/s10142-009-0121-3

Hsieh, E. J., Dinoso, J. B., & Clarke, C. F. (2004). A tRNA(TRP) gene mediates the suppression of cbs2-223 previously attributed to ABC1/COQ8. *Biochemical and Biophysical Research Communications*, *317*, 648–653. http://doi.org/10.1016/j.bbrc.2004.03.096

Huff, A. (1983). Nutritional Control of Regreening and Degreening in Citrus Peel Segments. *Plant Physiology*, *73*, 243–249.

Hugueney, P., Badillo, A., Chen, H.-C., Klein, A., Hirschberg, J., Camara, B., & Kuntz, M. (1995). Beta-cyclase cDNA - expression theoughout fruit development. *The Plant Journal*, *8*(3), 417–424.

Hugueney, P., Romer, S., Kuntz, M., & Camara, B. (1992). Characterization and molecular cloning of a flavoprotein catalyzing the synthesis of phytofluene and zeta-carotene in Capsicum chromoplasts. *European Journal of Biochemistry*, *209*, 399–407.

Huh, J. H., Kang, B. C., Nahm, S. H., Kim, S., Ha, K. S., Lee, M. H., & Kim, B. D. (2001). A candidate gene approach identified phytoene synthase as the locus for mature fruit color in red pepper (Capsicum spp.). *TAG Theoretical and Applied Genetics*, *102*(4), 524–530. http://doi.org/10.1007/s001220051677

Hurtado-hemandez, H., & Smith, P. G. (1985). Inheritance of mature fruit color hi Capsicum armuum L ., (June), 211–213.

Hurtado-Hernandez, H., & Smith, P. G. (1985). Inheritance of mature fruit color in Capsicum annuum L. *Journal of Heredity* , *76*(3), 211–213. Retrieved from http://jhered.oxfordjournals.org/content/76/3/211.abstract

Iglesias, D. J., Tadeo, F. R., Legaz, F., Primo-millo, E., & Talon, M. (2001). In vivo sucrose stimulation of colour change in citrus fruit epicarps: Interactions between nutritional and hormonal signals. *Phyiologia Plantarum*, *112*, 244–250.

Ito, Y., Nishizawa-yokoi, A., Endo, M., Mikami, M., & Toki, S. (2015). CRISPR/Cas9-mediated mutagenesis of the RIN locus that regulates tomato fruit ripening. *Biochemical and Biophysical Research Communications*, 1–30. http://doi.org/10.1016/j.bbrc.2015.09.117

Jarvis, P., Dormann, P., Peto, C. A., Lutes, J., Benning, C., & Chory, J. (2000). Galactolipid deficiency and abnormal chloroplast development in the Arabidopsis MGD synthase 1 mutant. *PNAS*, *97*(14), 8175–8179.

Jenkins, M. (1926). A SECOND GENE PRODUCING GOLDEN PLANT COLOR IN MAIZE. *The American Naturalist*, *LX*, 484–488.

Joubès, J., Raffaele, S., Bourdenx, B., Garcia, C., Laroche-Traineau, J., Moreau, P., … Lessire, R. (2008). The VLCFA elongase gene family in Arabidopsis thaliana: Phylogenetic analysis, 3D modelling and expression profiling. *Plant Molecular Biology*, *67*, 547–66. http://doi.org/10.1007/s11103-008-9339-z

Kahlau, S., & Bock, R. (2008). Plastid Transcriptomics and Translatomics of Tomato Fruit Development and Chloroplast-to-Chromoplast Differentiation: Chromoplast Gene Expression Largely Serves the Production of a Single Protein. *The Plant Cell*, *20*, 856–874. http://doi.org/10.1105/tpc.107.055202

Karrer, P., & Jucker, E. (1945). Partialsynthesen des Flavoxanthins, Chrysanthemaxanthins, Antheraxanthins, Violaxanthins Mutatoxanthins und Auroxanthins. *Helvetica Chimica Acta*, *28*(1), 300–315.

Khachik, F., Bernstein, P. S., & Garland, D. L. (1997). Identification of Lutein and Zeaxanthin Oxidation Products in Human and Monkey Retinas. *Investigative Opthalmology and Visual Science*, *38*(9), 1802–1811.

Khachik, F., Moura, F. F. De, Zhao, D., Aebischer, C., Bernstein, P. S., & Esults, R. R. (2002). Transformations of Selected Carotenoids in Plasma, Liver, and Ocular Tissues of Humans and in Nonprimate Animal Models. *Investigative Opthalmology and Visual Science*, *43*(11), 3383–3392.

Kim, S., Park, M., Yeom, S.-I., Kim, Y.-M., Lee, J. M., Lee, H.-A., … Choi, D. (2014). Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. *Nature Genetics*, (October 2013). http://doi.org/10.1038/ng.2877

Kleinig, H., & Liedvogel, B. (1980). Fatty Acid Synthesis by Isolated Chromoplasts from the Daffodil. Energy Sources and Distribution Patterns of the Acids. *Planta*, *169*, 166–169.

Klie, S., Osorio, S., Tohge, T., Drincovich, M. F., Fait, A., Giovannoni, J. J., … Nikoloski, Z. (2014). Conserved Changes in the Dynamics of Metabolic Processes during Fruit Development and Ripening. *Plant Physiology*, *164*(January), 55–68. http://doi.org/10.1104/pp.113.226142

Kobayashi, K., Fujii, S., Sasaki, D., Baba, S., Ohta, H., Masuda, T., & Wada, H. (2014). Plant Systems Biology Transcriptional regulation of thylakoid galactolipid biosynthesis coordinated with chlorophyll biosynthesis during the development of chloroplasts in Arabidopsis Article type : Received on : Accepted on : Citation : Transcriptional r. *Frontiers in Plant Science*, *5*(272). http://doi.org/10.3389/fpls.2014.00272

Kobayashi, K., Kondo, M., Fukuda, H., Nishimura, M., & Ohta, H. (2007). Galactolipid synthesis in chloroplast inner envelope is essential for proper thylakoid biogenesis, photosynthesis, and

embryogenesis. *PNAS*, *104*(43), 17216–17221.

Kobayashi, K., Narise, T., Sonoike, K., Hashimoto, H., Sato, N., Kondo, M., & Nishim-, M. (2013). Role of galactolipid biosynthesis in coordinated development of photosynthetic complexes and thylakoid membranes during chloroplast biogenesis in Arabidopsis. *The Plant Journal*, *73*, 250–261. http://doi.org/10.1111/tpj.12028

Kohno, M., Momoi, M., Oo, M. L., Paik, J., Lee, Y., Venkataraman, K., … Hla, T. (2006). Intracellular Role for Sphingosine Kinase 1 in Intestinal Adenoma Cell Proliferation. *Molecular and Cellular Biology*, *26*(19), 7211–7223. http://doi.org/10.1128/MCB.02341-05

Kovtun, Y., Chiu, W., Tena, G., & Sheen, J. (1999). Functional analysis of oxidative stress-activated mitogen-activated protein kinase cascade in plants. *PNAS*, *97*(6), 2940–2945.

Kraft, K. H., Brown, C. H., Nabhan, G. P., Luedeling, E., Jesús, J. De, & Ruiz, L. (2014). Multiple lines of evidence for the origin of domesticated chili pepper, Capsicum annuum, in Mexico. *PNAS*, *111*(17), 6165–6170. http://doi.org/10.1073/pnas.1308933111

Kunst, L., Taylor, D. C., & Underhill, E. W. (1992). Fatty acid elongation in developing seeds of Arabidopsis thaliana. *Plant Physiology and Biochemistry*, *30*(4), 425–434.

Lang, Y., Yanagawa, S., Sasanuma, T., & Sasakuma, T. (2006). A Gene Encoding a Putative Acyl-transferase Involved in Pungency of Capsicum. *Breeding Science*, *56*, 55–62.

Langdale, J. A., & Kidner, C. A. (1994). bundle sheath defective , a mutation that disrupts cellular differentiation in maize leaves. *Development*, *681*, 673–681.

Leide, J., Hildebrandt, U., Reussing, K., Riederer, M., & Vogg, G. (2007). The Developmental Pattern of Tomato Fruit Wax Accumulation and Its Impact on Cuticular Transpiration Barrier Properties : Effects of a Deficiency in a beta -Ketoacyl-Coenzyme A Synthase (LeCER6). *Plant Physiology*, *144*(July), 1667–1679. http://doi.org/10.1104/pp.107.099481

Li, F., Murillo, C., & Wurtzel, E. T. (2007). Maize Y9 Encodes a Product Essential for 15-cis-zeta-Carotene Isomerization. *Plant Physiology*, *144*(June), 1181–1189. http://doi.org/10.1104/pp.107.098996

Li, F., Vallabhaneni, R., & Wurtzel, E. T. (2008). PSY3, a New Member of the Phytoene Synthase Gene Family Conserved in the Poaceae and Regulator of Abiotic Stress-Induced Root Carotenogenesis. *Plant Physiology*, *146*(March), 1333–1345. http://doi.org/10.1104/pp.107.111120

Li, J., Aach, J., Norville, J. E., Mccormack, M., Bush, J., Church, G. M., & Sheen, J. (2014). Multiplex and homologous recombination-mediated plant genome editing via guide RNA/Cas9. *Nature Biotechnology*, *31*(8), 688–691. http://doi.org/10.1038/nbt.2654.Multiplex

Li, J., Yang, P., Kang, J., Gan, Y., Yu, J., & Calderón-urrea, A. (2016). Transcriptome Analysis of Pepper (Capsicum annuum) Revealed a Role of 24-Epibrassinolide in Response to Chilling. *Frontiers in Plant Science*, *7*(2181), 1–17. http://doi.org/10.3389/fpls.2016.01281

Li, L., Yang, Y., Xu, Q., Owsiany, K., Welsch, R., Chitchumroonchokchaie, C., … Thannhauser, T. W. (2012). The Or Gene Enhances Carotenoid Accumulation and Stability During Post-Harvest Storage of Potato Tubers. *Molecular Plant*, *5*(2). http://doi.org/10.1093/mp/ssr099

Li, L., & Yuan, H. (2013). Chromoplast biogenesis and carotenoid accumulation. *Archives of Biochemistry and Biophysics*, *539*(2), 102–9. http://doi.org/10.1016/j.abb.2013.07.002

Li, Z., Wang, S., Gui, X.-L., Chang, X.-B., & Gong, Z.-H. (2013). A further analysis of the relationship between yellow ripe-fruit color and the capsanthin-capsorubin synthase gene in pepper

(Capsicum sp.) indicated a new mutant variant in C. annuum and a tandem repeat structure in promoter region. *PloS One*, *8*(4), e61996. http://doi.org/10.1371/journal.pone.0061996

Liang, Z., Chen, K., Li, T., Zhang, Y., Wang, Y., Zhao, Q., … Gao, C. (2017). Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. *Nature Communications*, *8*, 1–5. http://doi.org/10.1038/ncomms14261

Lincoln, J. E., & Fischer, R. L. (1988). Regulation of Gene Expression by Ethylene in Wild-Type and rin Tomato (Lycopersicon esculentum) Fruit. *Plant Physiology*, *88*, 370–374.

Lindqvist, A., & Andersson, S. (2002). Biochemical properties of purified recombinant human beta-carotene 15,15'-monooxygenase. *The Journal of Biological Chemistry*, *277*(26), 23942–8. http://doi.org/10.1074/jbc.M202756200

Ling, H. (2008). Sequence Analysis of GDSL Lipase Gene Family in Arabidopsis thaliana. *Pakistan Journal of Biological Sciences*, *11*(5), 763–767.

Liu, L., Shao, Z., Zhang, M., & Wang, Q. (2015). Regulation of Carotenoid Metabolism in Tomato. *Molecular Plant*, *8*(1), 28–39. http://doi.org/10.1016/j.molp.2014.11.006

Livingstone, K. D., Lackney, V. K., Blauth, J. R., Wijk, R. Van, & Jahn, M. K. (1999). Genome Mapping in Capsicum and the Evolution of Genome Structure in the Solanaceae. *Genetics*, *152*, 1183–1202.

Long, A., Benz, B. F., Donahue, D. J., & Jull, A. J. T. (1989). FIRST DIRECT AMS DATES ON EARLY MAIZE FROM TEHUACAN, MEXICO. *Radiocarbon*, *31*(3), 1035–1040.

Lu, S., Van Eck, J., Zhou, X., Lopez, A. B., O'Halloran, D. M., Cosman, K. M., … Li, L. (2006). The cauliflower Or gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of beta-carotene accumulation. *The Plant Cell*, *18*(12), 3594–605. http://doi.org/10.1105/tpc.106.046417

Lundquist, P. K., Davis, J. I., & Wijk, K. J. Van. (2014). ABC1K atypical kinases in plants; filling the organellar kinase void. *Trends in Plant Science*, *17*(9), 546–555. http://doi.org/10.1016/j.tplants.2012.05.010.ABC1K

Lundquist, P. K., Poliakov, A., Bhuiyan, N. H., Zybailov, B., Sun, Q., & Wijk, K. J. Van. (2012). The Functional Network of the Arabidopsis Plastoglobule Proteome Based on Quantitative Proteomics and Genome-Wide Coexpression Analysis. *Plant Physiology*, *158*, 1172–1192. http://doi.org/10.1104/pp.111.193144

Lydic, T. A., Townsend, S., Adda, C. G., Collins, C., & Reid, G. E. (2016). Rapid and Comprehensive "Shotgun" Lipidome Profiling of Colorectal Cancer Cell Derived Exosomes. *Methods*, 1–31.

Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., … Wijk, K. J. Van. (2012). Nucleoid-Enriched Proteomes in Developing Plastids and Chloroplasts from Maize Leaves: A New Conceptual Framework for Nucleoid Functions. *Plant Physiology*, *158*, 156–189. http://doi.org/10.1104/pp.111.188474

Maoka, T. (2009). Recent progress in structural studies of carotenoids in animals and plants. *Archives of Biochemistry and Biophysics*, *483*(2), 191–195. http://doi.org/10.1016/j.abb.2008.10.019

Marz, U. (2015). *The Global Market for Carotenoids*.

Matsufuji, H., Ishikawa, K., Nunomura, O., Chino, M., & Takeda, M. (2007). Anti-oxidant content of different coloured sweet peppers, white, green, yellow, orange and red (Capsicum annuum L .). *International Journal of Food Science and Technology*, *42*, 1482–1488. http://doi.org/10.1111/j.1365-2621.2006.01368.x

Matus, Z., Baranyai, M., Toth, G., & Szabolcs, J. (1981). Identification of Oxo, Epoxy and Some cis-Carotenoids in High- Performance Liquid Chromatography. *Chromatographia*, *14*(6), 337–340.

Matus, Z., Deli, J., & Szabolcs, J. (1991). Carotenoid Composition of Yellow Pepper during Ripening : Isolation of Beta-Cryptoxanthin 5,6-Epoxide. *Journal of Agricultural Food Chemistry*, *39*, 1907–1914.

Mercadante, A. Z., & Rodriguez-Amaya, D. B. (1998). Effects of Ripening, Cultivar Differences, and Processing on the Carotenoid Composition of Mango. *Journal of Agricultural Food Chemistry*, *46*, 128–130.

Michaelson, L. V, Haslam, R., Bellec, Y., Gissot, L., Marion, J., Domergue, F., … Faure, J. (2008). The very-long-chain hydroxy fatty acyl-CoA dehydratase PASTICCINO2 is essential and limiting for plant development. *PNAS*, *105*(38), 14727–14731.

Millar, A. A., Clemens, S., Zachgo, S., Giblin, E. M., Taylor, D. C., & Kunst, L. (1999). CUT1, an Arabidopsis Gene Required for Cuticular Wax Biosynthesis and Pollen Fertility, Encodes a Very-Long-Chain Fatty Acid Condensing Enzyme. *The Plant Cell*, *11*(May), 825–838.

Millar, A. A., & Kunst, L. (1997). Very-long-chain fatty acid biosynthesis is controlled through the expression and specificity of the condensing enzyme. *The Plant Journal*, *12*(March), 121–131.

Millar, A. A., Wrischer, M., & Kunst, L. (1998). Accumulation of very-long-chain fatty acids in membrane glycerolipids is associated with dramatic alterations in plant morphology. *The Plant Cell*, *10*(11), 1889–1902. http://doi.org/10.1105/tpc.10.11.1889

Mizusawa, N., & Wada, H. (2012). The role of lipids in photosystem II. *BBA - Bioenergetics*, *1817*(1), 194–208. http://doi.org/10.1016/j.bbabio.2011.04.008

Moran, T. V, & Walker, J. C. (1993). Molecular cloning of two novel protein kinase genes from Arabidopsis thaliana. *Biochimica et Biophysica Acta*, *1216*, 9–14.

Mundy, J., Olsen, A. N., Mundy, J., & Skriver, K. (2002). Peptomics, identification of novel cationic Arabidopsis peptides with conserved sequence motifs. *In Silico Biology*, *2*(4), 441–451.

National Institutes of Health. (2014). *Vitamin A: Fact sheet for consumers*. Retrieved from https://ods.od.nih.gov/factsheets/VitaminA-HealthProfessional/

Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J. D. G., & Kamoun, S. (2013). Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. *Nature Biotechnology*, *31*(8), 691–693. http://doi.org/10.1038/nbt.2654

Nelson, N., & Ben-shem, A. (2004). The Complex Architecture of Oxygenic Photosynthesis. *Nature Reviews. Molecular Cell Biology*, *5*(December), 1–13. http://doi.org/10.1038/nrm1525

Nguyen, C. V, Vrebalov, J. T., Gapper, N. E., Zheng, Y., Zhong, S., Fei, Z., & Giovannoni, J. J. (2014). Tomato GOLDEN2-LIKE transcription factors reveal molecular gradients that function during fruit development and ripening. *The Plant Cell*, *26*(2), 585–601. http://doi.org/10.1105/tpc.113.118794

Nogueira, M., Berry, H., Nohl, R., Klompmaker, M., Holden, A., & Paul, D. (2016). Subchromoplast Fractionation Protocol for Different Solanaceae Fruit Species. *Bio-Protocol*, *6*(13), 1–14.

Nogueira, M., Mora, L., Enfissi, E. M. a, Bramley, P. M., & Fraser, P. D. (2013). Subchromoplast sequestration of carotenoids affects regulatory mechanisms in tomato lines expressing different carotenoid gene combinations. *The Plant Cell*, *25*(11), 4560–79. http://doi.org/10.1105/tpc.113.116210

Nolan, J. M., Loughman, J., Akkali, M. C., Stack, J., Scanlon, G., Davison, P., & Beatty, S. (2011). The impact of macular pigment augmentation on visual performance in normal subjects: COMPASS. *Vision Research*, *51*(5), 459–469. http://doi.org/10.1016/j.visres.2010.12.016

Nunn, N., & Qian, N. (2010). The Columbian Exchange: A History of Disease, Food, and Ideas. *Journal of Economic Perspectives*, *24*(2), 163–188. http://doi.org/10.1257/jep.24.2.163

Nuruzzaman, M., Sharoni, A. M., & Kikuchi, S. (2013). Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. *Frontiers in Microbiology*, *4*(September), 1–16. http://doi.org/10.3389/fmicb.2013.00248

Oh, S.-K., Lee, S., Yu, S. H., & Choi, D. (2005). Expression of a novel NAC domain-containing transcription factor (CaNAC1) is preferentially associated with incompatible interactions between chili pepper and pathogens. *Planta*, *222*, 876–887. http://doi.org/10.1007/s00425-005-0030-1

Ohmiya, A., Kishimoto, S., Aida, R., Yoshioka, S., & Sumitomo, K. (2006). Carotenoid Cleavage Dioxygenase (CmCCD4a) Contributes to White Color Formation in Chrysanthemum Petals. *Plant Physiology*, *142*, 1193–1201. http://doi.org/10.1104/pp.106.087130

Olsen, A. N., Ernst, H. A., Leggio, L. Lo, & Skriver, K. (2005). NAC transcription factors: structurally distinct, functionally diverse. *Trends in Plant Science*, *10*(2), 79–87. http://doi.org/10.1016/j.tplants.2004.12.010

Olson, J. A. (1994). Absorption, transport, and metabolism of carotenoids in humans. *Pure and Applied Chemistry*, *66*(5), 1011–1016.

Ooka, H., Satoh, K., Doi, K., Nagata, T., Otomo, Y., Murakami, K., … Kikuchi, S. (2003). Comprehensive Analysis of NAC Family Genes in Oryza sativa and Arabidopsis thaliana. *DNA Research*, *247*, 239–247.

Osorio, S., Alba, R., Nikoloski, Z., Kochevenko, A., Fernie, A. R., Giovannoni, J. J., … Germany, S. O. (2012). Integrative Comparative Analyses of Transcript and Metabolite Pro fi les from Pepper and Tomato Ripening and Development Stages Uncovers Species-Speci fi c Patterns of Network Regulatory Behavior [W][OA]. *Plant Physiology*, *159*(August), 1713–1729. http://doi.org/10.1104/pp.112.199711

Paolillo, D. J., Garvin, D. F., & Parthasarathy, M. V. (2004). The chromoplasts of Or mutants of cauliflower (Brassica oleracea L. var. botrytis). *Protoplasma*, *224*, 245–253. http://doi.org/10.1007/s00709-004-0059-1

Paran, I., & Knaap, E. Van Der. (2007). Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. *Journal of Experimental Botany*, *58*(14), 3841–3852. http://doi.org/10.1093/jxb/erm257

Pathway, B., Bartley, G. E., Viitanen, P. V, Bacot, K., & Scolniks, P. A. (1992). A Tomato Gene Expressed during Fruit Ripening Encodes an Enzyme of the Carotenoid Biosynthesis Pathway. *Journal of Biological Chemistry*, *8*, 5036–5039.

Paul, S., Gable, K., Cahoon, E., Jaworski, J., Napier, J. A., & Dunn, T. M. (2006). Members of the Arabidopsis FAE1-like 3-Ketoacyl-CoA Synthase Gene Family Substitute for the Elop Proteins of Saccharomyces cerevisiae. *The Journal of Biological Chemistry*, *281*(14), 9018–9029. http://doi.org/10.1074/jbc.M507723200

Peng, X., Zhao, Y., Cao, J., Zhang, W., Jiang, H., Li, X., … Zhu, S. (2012). CCCH-Type Zinc Finger Family in Maize: Genome-Wide Identification, Classification and Expression Profiling under Abscisic Acid and Drought Treatments. *PloS One*, *7*(7), 1–15.

http://doi.org/10.1371/journal.pone.0040120

Perez-fons, L., Wells, T., Corol, D. I., Ward, J. L., Gerrish, C., Beale, M. H., … Fraser, P. D. (2014). A genome-wide metabolomic resource for tomato fruit from Solanum pennellii. *Nature Scientific Reports*, *4*(3859), 1–8. http://doi.org/10.1038/srep03859

Perez-Galvez, A., & Minguez-Mosquera, I. (2002). Degradation of non-esterified and esterified xanthophylls by free radicals. *Biochimica et Biophysica Acta*, *1569*, 31–34.

Perez-Galvez, A., & Minguez-Mosquera, I. (2004). Degradation, under Non-Oxygen-Mediated Autooxidation, of Carotenoid Profile Present in Paprika Oleoresins with Lipid Substrates of Different Fatty Acid Composition. *Journal of Agricultural Food Chemistry*, *52*, 632–637.

Peterson, B. A., Haak, D. C., Nishimura, M. T., Teixeira, P. J. P. L., James, R., Dangl, J. L., & Nimchuk, Z. L. (2016). Genome-Wide Assessment of Efficiency and Specificity in CRISPR/Cas9 Mediated Multiple Site Targeting in Arabidopsis. *PloS One*, 1–11. http://doi.org/10.1371/journal.pone.0162169

Pettus, B. J., Chalfant, C. E., & Hannun, Y. A. (2002). Ceramide in apoptosis: an overview and current perspectives. *Biochimica et Biophysica Acta*, *1585*, 114–125.

Pickersgill, B. (1971). Relationships Between Weedy and Cultivated Forms in Some Species of Chili Peppers (Genus Capsicum). *Evolution*, *25*(4), 683–691.

Pogson, B. J., & Rissler, H. M. (2000). Genetic manipulation of carotenoid biosynthesis and photoprotection. *Philosophical Transactions of the Royal Society of London B.*, *355*, 1395–1403. http://doi.org/10.1098/rstb.2000.0701

Pogson, B., Mcdonald, K. A., Truong, M., Britton, G., & Dellapenna, D. (1996). Arabidopsis Camtenoid Mutants Demonstrate That Lutein s Not Essential for Photosynthesis in Higher Plants. *The Plant Cell*, *8*(September), 1627–1639.

Pott, I., Breithauptb, D. E., & Carle, R. (2003). Detection of unusual carotenoid esters in fresh mango (Mangifera indica L. cv. “Kent”). *Phytochemistry*, *64*, 825–829. http://doi.org/10.1016/S0031-9422(03)00466-7

Powell, A. L. T., Nguyen, C. V., Hill, T., Cheng, K. L., Figueroa-Balderas, R., Aktas, H., … Bennett, A. B. (2012). Uniform ripening Encodes a Golden2 Like Transcription Factor Regulating Tomato Fruit Chloroplast Development. *Science (New York, N.Y.)*, (June), 1711–1716.

Qin, C., Yu, C., Shen, Y., Fang, X., Chen, L., Min, J., & Cheng, J. (2014). Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum domestication and specialization. *PNAS*, *111*(14), 5135–5140. http://doi.org/10.1073/pnas.1400975111

Rao, G. U., & Paran, I. (2003). Polygalacturonase : a candidate gene for the soft flesh and deciduous fruit mutation in Capsicum. *Plant Molecular Biology*, *51*, 135–141.

Ray, J., Moureau, P., Bird, C., Bird, A., Grierson, D., Maunders, M., … Schuch, W. (1992). Cloning and characterization of a gene involved in phytoene synthesis from tomato. *Plant Molecular Biology*, *19*, 401–404.

Rinaldi, R., Deynze, A. Van, Portis, E., Rotino, G. L., Toppino, L., Hill, T., … Lanteri, S. (2016). New Insights on Eggplant/Tomato/Pepper Synteny and Identification of Eggplant and Pepper Orthologous QTL. *Frontiers in Plant Science*, *7*(1031). http://doi.org/10.3389/fpls.2016.01031

Rissanen, T., Voutilainen, S., Nyyssonen, K., & Salonen, J. (2002). Lycopene, Atherosclerosis, and Coronary Heart Disease. *Experimental Biology and Medicine*, *227*(15), 900–907.

Rockholm, D. C., & Yamamoto, H. Y. (1996). Violaxanthin De-Epoxidase. *Plant Physiology*, *110*, 697–703.

Rodriguez-Uribe, L., Guzman, I., Rajapakse, W., Richins, R. D., & O'Connell, M. a. (2012). Carotenoid accumulation in orange-pigmented Capsicum annuum fruit, regulated at multiple levels. *Journal of Experimental Botany*, *63*(1), 517–26. http://doi.org/10.1093/jxb/err302

Rodriguez-Villalon, A., Gas, E., & Rodrıuez-Concepcion, M. (2009). Colors in the dark: A model for the regulation of carotenoid biosynthesis in etioplasts. *Plant Signaling & Behavior*, *4*(10), 965–967. http://doi.org/10.1111/j.1365-313X.2009.03966.x.arotenoids

Rodrıguez-Villalon, A., Gas, E., & Rodrıuez-Concepcion, M. (2009). Phytoene synthase activity controls the biosynthesis of carotenoids and the supply of their metabolic precursors in dark-grown Arabidopsis seedlings. *The Plant Journal*, *60*, 424–435. http://doi.org/10.1111/j.1365-313X.2009.03966.x

Rohmer, M., Louis, U., Cnrs, P., Bel, I. Le, Pascal, B., & Cedex, S. (1999). The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Natural Product Reports*, *16*(5), 565–574.

Romer, S., Hugueney, P., Bouvier, F., Camara, B., & Kuntz, M. (1993). Expression of the Genes Encoding the Early Carotenoid Biosynthetic Enzymes in Capsicum. *Biochemical and Biophysical Research Communications*, *196*(3), 1414–1421.

Ronen, G., Carmel-goren, L., Zamir, D., & Hirschberg, J. (2000). An alternative pathway to ᴺᴸ -carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato.

Rossel, J. B., Wilson, I. W., & Pogson, B. J. (2002). Global Changes in Gene Expression in Response to High Light in Arabidopsis. *Genome Analysis*, *130*(November), 1109–1120. http://doi.org/10.1104/pp.005595.Under

Rossini, L., Cribb, L., Martin, D. J., & Langdale, J. A. (2001). The Maize Golden2 Gene Defines a Novel Class of Transcriptional Regulators in Plants. *The Plant Cell*, *13*(May), 1231–1244.

Rottet, S., Besagni, C., & Kessler, F. (2015). The role of plastoglobules in thylakoid lipid remodeling during plant development. *BBA - Bioenergetics*, *1847*(9), 889–899. http://doi.org/10.1016/j.bbabio.2015.02.002

Rottet, S., Devillers, J., Glauser, G., Douet, V., Besagni, C., & Kessler, F. (2016). Identification of Plastoglobules as a Site of Carotenoid Cleavage. *Frontiers in Plant Science*, *7*(December), 1–11. http://doi.org/10.3389/fpls.2016.01855

Roudier, F., Gissot, L., Beaudoin, F., Haslam, R., Michaelson, L., Marion, J., … Faure, J. (2010). Very-Long-Chain Fatty Acids Are Involved in Polar Auxin Transport and Developmental Patterning in Arabidopsis. *The Plant Cell*, *22*(February), 364–375. http://doi.org/10.1105/tpc.109.071209

Saini, R. K., & Keum, Y. S. (2016). GC–MS and HPLC–DAD analysis of fatty acids and tocopherols in sweet peppers (Capsicum annuum L.). *Journal of Food Measurement and Characterization*, *10*(3), 685–689. http://doi.org/10.1007/s11694-016-9352-x

Särkinen, T., Bohs, L., Olmstead, R. G., & Knapp, S. (2013). A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. *BMC Evolutionary Biology*, *13*, 214–229.

Sarpras, M., Gaur, R., Sharma, V., Chhapekar, S. S., Das, J., Kumar, A., … Ramchiary, N. (2016). Comparative Analysis of Fruit Metabolites and Pungency Candidate Genes Expression between

Bhut Jolokia and Other Capsicum Species. *PloS One*, *11*(12), 1–19. http://doi.org/10.1371/journal.pone.0167791

Schilmiller, A. L., Gilgallon, K., Ghosh, B., Jones, A. D., & Last, R. L. (2016). Acylsugar Acylhydrolases: Carboxylesterase-Catalyzed Hydrolysis of Acylsugars in Tomato Trichomes. *Plant Physiology*, *170*(March), 1331–1344. http://doi.org/10.1104/pp.15.01348

Schneiter, R., Hitomi, M., Ivessa, A. S., Fasch, E., Kohlwein, S. D., & Tartakoff, A. M. (1996). A Yeast Acetyl Coenzyme A Carboxylase Mutant Links Very-Long-Chain Fatty Acid Synthesis to the Structure and Function of the Nuclear Membrane-Pore Complex. *Molecular and Cellular Biology*, *16*(12), 7161–7172.

Schneiter, R., Ugger, B. B. R., Amann, C. M., Prestwich, G. D., Epand, R. F., Wieland, F. T., & Epand, R. M. (2004). Identification and biophysical characterization of a very-long-chain-fatty- acid-substituted phosphatidylinositol in yeast subcellular membranes. *Biochemical Journal*, *381*, 941–949.

Schwartz, S. H., Tan, B. C., Gage, D. A., Zeevaart, J. A. D., & Mccarty, D. R. (1997). Specific Oxidative Cleavage of Carotenoids by VP14 of Maize. *Science*, *276*, 1872–1874.

Schweiggert, U., Kammerer, D. R., Carle, R., & Schieber, A. (2005). Characterization of carotenoids and carotenoid esters in red pepper pods (Capsicum annuum L.) by high-performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry. *Rapid Communications in Mass Spectrometry : RCM*, *19*(18), 2617–28. http://doi.org/10.1002/rcm.2104

Seddon, J. M., Ajani, U. A., Sperduto, R. D., Hiller, R., Blair, N., Burton, T. C., … Miller, D. T. (1994). Dietary Carotenoids , Vitamins A , C , and E , and Advanced Age-Related Macular Degeneration. *Journal of the American Medical Association*, *272*, 1413–1420.

Selth, L. A., Dogra, S. C., Rasheed, M. S., Healy, H., Randles, J. W., & Rezaian, M. A. (2005). A NAC Domain Protein Interacts with Tomato leaf curl virus Replication Accessory Protein and Enhances Viral Replication. *The Plant Cell*, *17*(January), 311–325. http://doi.org/10.1105/tpc.104.027235.the

Shumskaya, M., Bradbury, L. M. T., Monaco, R. R., & Wurtzel, E. T. (2012). Plastid localization of the key carotenoid enzyme phytoene synthase is altered by isozyme, allelic variation, and activity. *The Plant Cell*, *24*(9), 3725–41. http://doi.org/10.1105/tpc.112.104174

Simkin, A. J., Bramley, P. M., Fraser, P. D., & Kuntz, M. (2007). Fibrillin influence on plastid ultrastructure and pigment content in tomato fruit. *Phytochemistry*, *68*, 1545–1556. http://doi.org/10.1016/j.phytochem.2007.03.014

Smith, B. D. (1997). RECONSIDERING THE OCAMPO CAVES AND THE ERA OF INCIPIENT CULTIVATION IN MESOAMERICA. *Latin American Antiquity*, *8*(4), 342–383.

Smith, S. H. (2014). In the shadow of a pepper-centric historiography: Understanding the global diffusion of capsicums in the sixteenth and seventeenth centuries. *Journal of Ethnopharmacology*, *167*, 64–77. http://doi.org/10.1016/j.jep.2014.10.048

Sora, G. T. de S., Henrique, A., Souza, P., Antônio, A., Zielinski, F., Haminiuk, C. W. I., … Peralta, R. M. (2015). Fatty acid composition of Capsicum genus peppers. *Ciência E Agrotecnologia*, *39*(4), 372–380.

Sperling, P., Franke, S., Lüthje, S., & Heinz, E. (2005). Are glucocerebrosides the predominant sphingolipids in plant plasma membranes? *Plant Physiology and Biochemistry*, *43*, 1031–1038. http://doi.org/10.1016/j.plaphy.2005.10.004

Spicher, L., & Kessler, F. (2015). Unexpected roles of plastoglobules (plastid lipid droplets) in vitamin K1 and E metabolism. *Current Opinion in Plant Biology*, *25*, 123–129. http://doi.org/10.1016/j.pbi.2015.05.005

Spurr, A. R., & Harris, W. M. (1968). ULTRASTRUCTURE OF CHLOROPLASTS AND CHROMOPLASTS IN CAPSICUM ANNUUM I . THYLAKOID MEMBRANE CHANGES DURING FRUIT RIPENING. *American Journal of Botany*, *55*(10), 1210–1224.

Stahl, W., & Sies, H. (2003). Antioxidant activity of carotenoids. *Molecular Aspects of Medicine*, *24*, 345–351. http://doi.org/10.1016/S0098-2997(03)00030-X

Stange, C. (2016). *Carotenoids in Nature*. (C. Stange & J. R. Harris, Eds.). Springer.

Stewart, C. J., Kang, B., Liu, K., Mazourek, M., Moore, S. L., Yoo, E. Y., … Jahn, M. M. (2005). The Pun1 gene for pungency in pepper encodes a putative acyltransferase. *The Plant Journal*, *42*, 675–688. http://doi.org/10.1111/j.1365-313X.2005.02410.x

Subagio, A., Morita, N., & Sawada, S. (1996). Carotenoids and Their Esters in Banana Peel. *J. Nutr. Sci. Vitaminol.*, *42*, 553–566.

Tabata, S. (2000). Sequence and analysis of chromosome 5 of the plant Arabidopsis thaliana. *Nature*, *408*(December), 823–826.

Tanaka, Y., & Ohmiya, A. (2008). Seeing is believing: engineering anthocyanin and carotenoid biosynthetic pathways. *Current Opinion in Biotechnology*, *19*, 190–197. http://doi.org/10.1016/j.copbio.2008.02.015

Tao, L., & Cheng, Q. (2004). Novel β-carotene ketolases from non-photosynthetic bacteria for canthaxanthin synthesis. *Molecular Genetics and Genomics*, *272*(5), 530–537. http://doi.org/10.1007/s00438-004-1083-8

Telef, N., Stammitti-Bert, L., Mortain-Bertrand, A., Maucourt, M., Carde, J. P., Rolin, D., & Gallusci, P. (2006). Sucrose deficiency delays lycopene accumulation in tomato fruit pericarp discs. *Plant Molecular Biology*, *62*, 453–469. http://doi.org/10.1007/s11103-006-9033-y

Telfer, A. (2002). What is beta-carotene doing in the photosystem II reaction centre? *Philosophical Transactions of the Royal Society of London*, *357*, 1431–1440. http://doi.org/10.1098/rstb.2002.1139

Tewksbury, J. J., & Nabhan, G. P. (2001). Directed deterrence by capsaicin in chillies. *Nature*, *412*, 403–404.

The Potato Genome Consortium. (2011). Genome sequence and analysis of the tuber crop potato. *Nature*, *475*, 189–197. http://doi.org/10.1038/nature10158

The Tomato Genome Consortium. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, *485*, 635–641. http://doi.org/10.1038/nature11119

Thorup, T. A., Tanyolac, B., Livingstone, K. D., Popovsky, S., Paran, I., & Jahn, M. (2000). Candidate gene analysis of organ pigmentation loci in the Solanaceae. *PNAS*, *97*(21), 11192–11197.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., … Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, *7*, 562–578. http://doi.org/10.1038/nprot.2012.016

Umehara, M., Hanada, A., Yoshida, S., Akiyama, K., Arite, T., Takeda-kamiya, N., … Yamaguchi, S. (2008). Inhibition of shoot branching by new terpenoid plant hormones. *Nature*, *455*, 195–201. http://doi.org/10.1038/nature07272

Upton, C., & Buckley, J. T. (1995). A new family of lipolytic enzymes? *Trends in Biochemical Science*, *20*(5), 178–179.

Vallabhaneni, R., Bradbury, L. M. T., & Wurtzel, E. T. (2010). The carotenoid dioxygenase gene family in maize, sorghum, and rice. *Archives of Biochemistry and Biophysics*, *504*(1), 104–111. http://doi.org/10.1016/j.abb.2010.07.019

Vert, G. a, Briat, J.-F., & Curie, C. (2003). Dual regulation of the Arabidopsis high-affinity root iron uptake system by local and long-distance signals. *Plant Physiology*, *132*(2), 796–804. http://doi.org/10.1104/pp.102.016089

Vizzotto, M., Cisneros-zevallos, L., & Byrne, D. H. (2007). Large Variation Found in the Phytochemical and Antioxidant Activity of Peach and Plum Germplasm. *Journal of the American Society for Horticultural Science*, *132*(May 2006), 334–340.

Wahyuni, Y., Ballester, A., Sudarmonowati, E., Bino, R. J., & Bovy, A. G. (2011). Metabolite biodiversity in pepper (Capsicum) fruits of thirty-two diverse accessions : Variation in health-related compounds and implications for breeding. *Phytochemistry*, *72*(11–12), 1358–1370. http://doi.org/10.1016/j.phytochem.2011.03.016

Walsh, B. M., & Hoot, S. B. (2012). Phylogenetic Relationships of Capsicum (Solanaceae) Using DNA Sequences from Two Noncoding Regions: The Chloroplast atpB - rbcL Spacer Region and Nuclear waxy Introns. *International Journal of Plant Sciences*, *162*(6), 1409–1418.

Walter, M. H., & Strack, D. (2011). Carotenoids and their cleavage products: Biosynthesis and functions. *Natural Product Reports*, *28*, 663–692. http://doi.org/10.1039/c0np00036a

Wang, X., Basnayake, B. M. V. S., Zhang, H., Li, G., Li, W., Virk, N., … Song, F. (2009). The Arabidopsis ATAF1, a NAC Transcription Factor, is a Negative Regulator of Defense Responses Against Necrotrophic Fungal and Bacterial Pathogens. *Molecular Plant-Microbe Interactions*, *22*(10), 1227–1238.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., & Qiu, J. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature Biotechnology*, (July), 1–6. http://doi.org/10.1038/nbt.2969

Wang, Y., Yang, Y., Fei, Z., Yuan, H., Fish, T., Thannhauser, T. W., … Li, L. (2013). Proteomic analysis of chromoplasts from six crop species reveals insights into chromoplast function and development. *Journal of Experimental Botany*, *64*(4), 949–961. http://doi.org/10.1093/jxb/ers375

Weinhold, A., & Baldwin, I. T. (2011). Trichome-derived O -acyl sugars are a fi rst meal for caterpillars that tags them for predation. *PNAS*, *108*(19), 7855–7859. http://doi.org/10.1073/pnas.1101306108

Welti, R., & Wang, X. (2004). Lipid species profiling: a high-throughput approach to identify lipid compositional changes and determine the function of genes involved in lipid metabolism and signaling. *Current Opinion in Plant Biology*, *7*, 337–344. http://doi.org/10.1016/j.pbi.2004.03.011

Whitaker, B. D. (1986). Fatty-acid composition of polar lipids in fruit and leaf chloroplasts of “16:3”- and “18:3”-plant species. *Planta*, *196*, 313–319.

Wingerath, T., Stahl, W., & Sies, H. (1995). Beta-Cryptoxanthin Selectively Increases in Human Chylomicrons upon Ingestion of Tangerine Concentrate Rich in Beta-Cryptoxanthin Esters. *Archives of Biochemistry and Biophysics*, *324*(2), 385–390.

Winkler, B. S., Boulton, M. E., Gottsch, J. D., & Sternberg, P. (1999). Oxidative damage and age-related macular degeneration. *Molecular Vision*, *5*(32).

Wu, Z., Cheng, J., Cui, J., Xu, X., & Liang, G. (2016). Genome-Wide Identification and Expression Profile of Dof Transcription Factor Gene Family in Pepper (Capsicum annuum L .). *Frontiers in Plant Science*, *7*(April), 1–15. http://doi.org/10.3389/fpls.2016.00574

Yamamoto, H. Y. (2006). Functional roles of the major chloroplast lipids in the violaxanthin cycle. *Planta*, *224*, 719–724. http://doi.org/10.1007/s00425-006-0257-5

Yang, R., Deng, C., & Ouyang, B. (2011). Molecular analysis of two salt-responsive NAC-family genes and their expression analysis in tomato. *Molecular Biology Reports*, *38*, 857–863. http://doi.org/10.1007/s11033-010-0177-0

Yao, Y., Qiu, Q., Wu, X., Cai, Z., Xu, S., & Liang, X. (2013). Lutein supplementation improves visual performance in Chinese drivers: 1-year randomized, double-blind, placebo-controlled study. *Nutrition*, *29*, 958–964. http://doi.org/10.1016/j.nut.2012.10.017

Yokoyama, R., & Nishitani, K. (2006). Identification and characterization of Arabidopsis thaliana genes involved in xylem secondary cell walls. *Journal of Plant Research*, *119*, 189–194.

Ytterberg, A. J., Peltier, J.-B., & van Wijk, K. J. (2006). Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. *Plant Physiology*, *140*(3), 984–97. http://doi.org/10.1104/pp.105.076083

Yu, B., Lydiate, D. J., Young, L. W., Schafer, U. A., & Hannoufa, A. (2008). Enhancing the carotenoid content of Brassica napus seeds by downregulating lycopene epsilon cyclase. *Transgenic Research*, *17*, 573–585. http://doi.org/10.1007/s11248-007-9131-x

Zhang, Q., Meng, Y., Zhang, L., Chen, J., & Zhu, D. (2009). RNF13: a novel RING-type ubiquitin ligase over-expressed in pancreatic cancer. *Cell Research*, *19*(3), 348–357. http://doi.org/10.1038/cr.2008.285

Zheng, H., Rowland, O., & Kunst, L. (2005). Disruptions of the Arabidopsis Enoyl-CoA Reductase Gene Reveal an Essential Role for Very-Long-Chain Fatty Acid Synthesis in Cell Expansion during Plant Morphogenesis. *The Plant Cell*, *17*(May), 1467–1481. http://doi.org/10.1105/tpc.104.030155.at

Zhigila, D. A., Abdulrahaman, A. A., Kolawole, O. S., & Oladele, F. A. (2014). Fruit Morphology as Taxonomic Features in Five Varieties of Capsicum annuum L . Solanaceae. *Journal of Botany*, *2014*, 1–6.

Zhou, X., Mcquinn, R., Fei, Z., Wolters, A. A., Eck, J. V. A. N., Brown, C., … Li, L. (2011). Regulatory control of high levels of carotenoid accumulation in potato tubers. *Plant, Cell and Environment*, *34*, 1020–1030. http://doi.org/10.1111/j.1365-3040.2011.02301.x

Zouni, A., Witt, H.-T., Kern, J., Fromme, P., Krauss, N., Saenger, W., & Orth, P. (2001). Crystal structure of photosystem II from Synechococcus elongatus at 3.8 A resolution. *Nature*, *409*, 739–743.

9 Appendix

9.1 Gas Chromatography-Mass Spectrometry (GC-MS)

9.1.1 Table of Identified Compounds and Relative Amounts in c1 positive and negative Lines

Compound	**Average (Y)**	**Std. dev. (Y)**	**Average (W)**	**Std. dev. (W)**	**Fold change**
Tocopherol-y	0.316646	0.184447	0.0516522	0.0300631	0.163123
Tocopherol-ç	0.316646	0.184447	0.0516522	0.0300631	0.163123
Dodecanoic acid	0.159665	0.0777804	0.0465669	0.0353049	0.291654
à-Tocopherol (vitamin E)	13.7641	3.532	4.04233	1.98552	0.293685
Hexadecanoic acid	10.3633	0.868405	12.1872	2.02778	1.176
Docosanoic acid	0.915267	0.374621	1.41232	0.330253	1.54307
1,4-Butanediamine	0.0225155	0.00751097	0.0357954	0.0125305	1.58982
Tetracosanoic acid	0.773989	0.372229	1.47343	0.374207	1.90369
4-Hydroxyanthraquinone	0.100134	0.0450443	0.192701	0.0296472	1.92443
Stigmasterol	1.29164	0.443278	2.60041	0.326614	2.01326
a-Amyrin	0.704139	0.496337	1.42307	0.574489	2.02101
Octadecadienoic acid, cis 9,12	4.86346	1.34112	9.8392	3.6246	2.02309
Heptadecanoic acid	0.054542	0.0180776	0.119287	0.0571907	2.1870

	8				4
Hexacosanoic acid	0.0971376	0.144802	0.316158	0.0618721	3.25475
Benzoic acid	0.0207282	0.00567824	0.0740726	0.0451003	3.57351
Unknown - 26.5836 [NPS] (N,O-Bis(trimethylsilyl)-L-phenylalanine) (ions 218+192)	0.270209	0.150672	0.0872106	0.0942547	0.322753
N,O-Bis(trimethylsilyl)-L-phenylalanine	0.270209	0.150672	0.0911002	0.0905134	0.337148
Phosphoric acid, 2-[bis(trimethylsilyl)amino]ethyl bis(trimethylsilyl) ester	0.731511	0.226241	0.470348	0.246027	0.642981
L-Isoleucine, N-(trimethylsilyl)-, trimethylsilyl ester	0.020199	0.00616831	0.0303471	0.0150056	1.50241
N,O-Bis-(trimethylsilyl)isoleucine	0.020199	0.00616831	0.0303471	0.0150056	1.50241
Linolenic acid, trimethylsilyl ester	0.923899	0.60593	2.17005	2.10756	2.3488
á-Sitosterol trimethylsilyl ether	19.901	3.65775	23.376	3.16088	1.17462
Unknown (Hexacosanoic acid (1TMS)) (453)	0.691555	0.483808	0.291802	0.0826896	0.42195
Butanedioic acid, bis(trimethylsilyl) ester	0.0237029	0.00911022	0.0863233	0.126732	3.64189
1H-Indole-3-ethanamine, N,N,1-tris(trimethylsilyl)-	0.0575777	0.0519151	0.0200398	0.00361646	0.348047
L-Proline, 1-(trimethylsilyl)-, trimethylsilyl ester	0.430022	0.41755	0.128007	0.115478	0.297674
Unknown - 16.4675 [P+NP] contamination	0.671924	0.270281	0.425739	0.33398	0.633613
Galactose oxime hexakis(trimethylsilyl)	5.90732	6.06946	1.60719	1.17807	0.272067
Unknown - 34.5246 d-Glucose, 2,3,4,5,6-pentakis-O-(trimethylsilyl)-, o-methyloxyme, (1Z)-iso 2 (319+205+160)	5.882	6.08516	1.60719	1.17807	0.273238

Hexacosanoic acid, 2-[(trimethylsilyl)oxy]-, trimethylsilyl ester	0.433794	0.360277	0.170917	0.171601	0.394004
Tetradecanoic acid, trimethylsilyl ester	0.968831	0.263582	0.761017	0.226596	0.7855
Unknown - 18.7878 (L-Serine 3TMS) (204)	0.0242597	0.00576328	0.0200398	0.00361646	0.826051
Pentane, 3,3-dimethyl-	0.0390127	0.028256	0.0200398	0.00361646	0.513673
Piperonyl butoxide	1.25974	1.0675	0.491068	0.659488	0.389817
1,4-Benzenedicarboxylic acid, bis(trimethylsilyl) ester	0.020199	0.00616831	0.0322537	0.0263257	1.5968
Sucrose (8TMS) / ?-D-Glucopyranoside, 1,3,4,6-tetrakis-O-(trimethylsilyl)-?-D-fructofuranosyl 2,3,4,6-tetrakis-O-(trimethylsilyl)	0.926051	0.741026	0.428807	0.244842	0.463049
à-D-Glucopyranoside, 1,3,4,6-tetrakis-O-(trimethylsilyl)-á-D-fructofuranosyl 2,3,4,6-tetrakis-O-(trimethylsilyl)-	0.920399	0.736658	0.428807	0.244842	0.465893
Phosphoric acid, bis(trimethylsilyl) 2,3-bis[(trimethylsilyl)oxy]propyl ester	0.892086	0.299627	1.21866	0.614749	1.36608
Standard - 30.5735 [P] (Glycerol-3-Phosphate peak 2 (4TMS)) (299+357)	0.910021	0.278357	1.21866	0.614749	1.33916
Eicosanoic acid, trimethylsilyl ester	1.05207	0.461148	0.75096	0.193181	0.713794
Unknown - 46.1008 [NPS])Eicosanoic acid (1TMS)) (369)	1.05207	0.461148	0.75096	0.193181	0.713794
Standard - 26.4499 unknown	0.0987537	0.129942	0.0200398	0.00361646	0.202927
Unknown	8.99557	8.46373	3.79312	3.35437	0.421665
á-Amyrin trimethylsilyl ether	1.98223	0.521525	2.32167	0.422175	1.17124
1,2,3-Propanetricarboxylic acid, 2-[(trimethylsilyl)oxy]-, tris(trimethylsilyl) ester	0.118532	0.17266	0.0200398	0.00361646	0.169066
L-Aspartic acid (3TMS) - 23.8666 [P] (232)	0.023560	0.0057523	0.0200398	0.0036164	0.8505

	8	4		6	55
N,N,o'-Tris-(trimethylsilyl)aspartic acid	0.0235608	0.00575234	0.0200398	0.00361646	0.850555
D-Fructose, 1,3,4,5,6-pentakis-O-(trimethylsilyl)-, O-methyloxime	9.02469	8.44278	4.16004	3.05423	0.460962
Unknown	9.02469	8.44278	4.16004	3.05423	0.460962
Nonadecane	0.622404	0.251001	0.478022	0.104843	0.768024
Campesterol tms	7.41088	1.3743	8.25	0.96738	1.11323
Standard - 16.1972 [P] (Glycerol 3TMS) (205)	0.494738	0.579619	0.872758	0.600406	1.76408
Trimethylsilyl ether of glycerol	0.494738	0.579619	0.872758	0.600406	1.76408
Unknown - 16.5908 (Isoleucine) (158)	0.0261503	0.0077617	0.0320642	0.0120502	1.22615
L-Cysteine (3TMS) - 24.7316 [P] (ion 220+218 - not 219)	0.223508	0.386038	0.0200398	0.00361646	0.0896601
Unknown - 16.7390 [P] (L-Proline (2TMS)) (142)	0.020199	0.00616831	0.0237604	0.00467486	1.17632
Octadecanoic acid (1TMS) - 42.2655 [NP] (341)	7.15211	0.572013	7.62626	1.15237	1.0663
Octadecanoic acid, trimethylsilyl ester	7.15211	0.572013	7.62626	1.15237	1.0663
Propanoic acid, 2-[(trimethylsilyl)oxy]-, trimethylsilyl ester	0.5757	0.396738	0.366243	0.219119	0.63617
Myo-Inositol, 1,2,3,4,5,6-hexakis-O-(trimethylsilyl)-	0.0566622	0.0741237	0.0200398	0.00361646	0.353671
Unknown - 15.7362 [P] (Ethanolamine (3TMS)) (174)	0.0267358	0.0122325	0.0344718	0.0151244	1.28935
N,O-Bis-(trimethylsilyl)valine	0.0277631	0.0161987	0.0200398	0.00361646	0.721813
GABA (3TMS) - 24.0623 [P] (174,304,147)	0.025649	0.0118418	0.0200398	0.0036164	0.7812

	5			6	92
Citric acid (4TMS) - 32.234 [P] (363+347+375)	0.101241	0.175743	0.0200398	0.00361646	0.19794
L-Valine, N-(trimethylsilyl)-, trimethylsilyl ester	0.028709	0.0187573	0.0200398	0.00361646	0.698031
2-Monopalmitin (2TMS) - 48.3986 [NP] (313)	1.45682	0.035221	1.50478	0.151137	1.03292
2-Monopalmitin trimethylsilyl ether	1.45682	0.035221	1.50478	0.151137	1.03292
Lactic acid (2TMS) - 9.4338 [P] (117)	0.502951	0.315389	0.366243	0.219119	0.728188
Butanoic acid, 4-[bis(trimethylsilyl)amino]-, trimethylsilyl ester	0.0229974	0.00751085	0.0200398	0.00361646	0.871391
Mannonic acid, 2,3,4,6-tetrakis-O-(trimethylsilyl)-, lactone	0.0321984	0.0325928	0.0200398	0.00361646	0.622383
Standard - 27.7515 [P] (unknown)	0.0689064	0.137511	0.0200398	0.00361646	0.290826
Phytol (1TMS) - 40.4894 [NPS] (143+123)	0.0300359	0.0281133	0.0200398	0.00361646	0.667193
Sedoheptulose, o-methyloxime, hexakis-O-(trimethylsilyl)-	1.08021	2.99629	0.0200398	0.00361646	0.0185517
Unknown - 21.7339 [NP] contamination 307+350	0.534951	0.214994	0.458916	0.015134	0.857866
Docosanoic acid, trimethylsilyl ester	0.190213	0.251579	0.0952387	0.14686	0.500696
Standard - 29.1814 [Xylitol 5TMS) (217+129)	0.351013	0.939454	0.0200398	0.00361646	0.0570912
Unknown - 11.7874	1.45328	0.624225	1.2252	0.277958	0.84306
Arabinofuranose, 1,2,3,5-tetrakis-O-(trimethylsilyl)-	0.0671882	0.0449014	0.0473778	0.0560961	0.705151
3,8-Dioxa-2,9-disiladecane, 2,2,9,9-tetramethyl-	0.053826	0.09369	0.0226062	0.0075441	0.4199

	8			6	79
2-Monostearin trimethylsilyl ether	1.95112	0.0688221	1.98441	0.111562	1.01706
Unknown - 23.8057	0.0222303	0.00628625	0.0200398	0.00361646	0.901462
Unknown - 35.3840	0.0222972	0.00713769	0.0200398	0.00361646	0.898755
Serine tritms	0.0218606	0.00590546	0.0200398	0.00361646	0.916705
2-Piperidinecarboxylic acid, 1-(trimethylsilyl)-, trimethylsilyl ester	0.0220541	0.00697499	0.0200398	0.00361646	0.908662
L-Proline, 5-oxo-1-(trimethylsilyl)-, trimethylsilyl ester	0.0220541	0.00697499	0.0200398	0.00361646	0.908662
Standard - 23.7741 [P] (L-Proline, 5-oxo-1-(trimethylsilyl)-, trimethylsilyl ester) pyroglutamic acid 2tms (156)	0.0220541	0.00697499	0.0200398	0.00361646	0.908662
Bis(trimethylsilyl)monostearin	87.1821	2.90761	88.5386	6.21822	1.01556
Monostearin (2TMS) Octadecanoic acid, 2,3-bis[(trimethylsilyl)oxy]propyl ester (399)	87.1821	2.90761	88.5386	6.21822	1.01556
Unknown - 10.5132	0.148102	0.0527464	0.163374	0.0311051	1.10312
D-Myo-Inositol, 1,2,4,5,6-pentakis-O-(trimethylsilyl)-, bis(trimethylsilyl) phosphate	1.6915	0.477442	1.8442	0.70008	1.09027
Myo-Inositol, 1,3,4,5,6-pentakis-O-(trimethylsilyl)-, bis(trimethylsilyl) phosphate	1.6915	0.477442	1.8442	0.70008	1.09027
Unknown - 45.4536 [P] (D-Myo-Inositol, 1,2,4,5,6-pentakis-O-(trimethylsilyl)-, bis(trimethylsilyl) phosphate) (318+315+387)	1.6915	0.477442	1.8442	0.70008	1.09027
Unknown - 62.9128 [NPS] (Triacontanoic acid (1TMS)) (509+524)	0.307016	0.215578	0.253936	0.147635	0.82711
1,4-Butanediamine, N,N,N',N'-tetrakis(trimethylsilyl)-	0.0523016	0.0729452	0.0360391	0.0334873	0.689064
L-Valine (2TMS) - 14.1272 [P] (144)	0.027763	0.0161987	0.0314568	0.0163718	1.1330

	1				5
Lauric acid, 2,3-bis(trimethylsiloxy)propyl ester	0.039289	0.0283014	0.0332967	0.0272669	0.84748
(E,E)-7,11,15-Trimethyl-3-methylene-hexadeca-1,6,10,14-tetraene	0.0790113	0.1656	0.0505754	0.0628561	0.640104
1,6,10-Dodecatriene, 7,11-dimethyl-3-methylene-, (Z)-	0.0790113	0.1656	0.0505754	0.0628561	0.640104
RI=2977.9, 56.3584 [NP] fattty acid (193)	5.08698	0.97156	5.24514	1.07491	1.03109
Trimethyl(2,6 ditert.-butylphenoxy)silane	0.499858	0.126403	0.485072	0.0899037	0.97042
9-Octadecenoic acid, 2-[(trimethylsilyl)oxy]-1-[[(trimethylsilyl)oxy]methyl]ethyl ester (218)	0.0801208	0.168736	0.0943424	0.145068	1.1775
Cyclooctasiloxane, hexadecamethyl-	0.0414857	0.0253138	0.0392038	0.0401588	0.944997
Silanol, trimethyl-, phosphate (3:1)	17.6956	2.70599	17.5191	4.27536	0.990022
Unknown - 15.9984	17.6956	2.70599	17.5191	4.27536	0.990022
Benzene, 1,3-bis(1,1-dimethylethyl)-	0.18901	0.117827	0.193663	0.0646953	1.02462
Standard - 53.8311 [P] (Narigenin peak 2?)	0.0979402	0.210303	0.105184	0.0979452	1.07396
1-Monopalmitin (2TMS) - 49.2047 [NP] (371+459)	80.4447	2.471	80.5549	6.17206	1.00137
Hexadecanoic acid, 2,3-bis[(trimethylsilyl)oxy]propyl ester	80.4447	2.471	80.5549	6.17206	1.00137
Internal Standard D-27 Myristic acid 32.4833 min	36.6482	9.86769	36.6546	5.1327	1.00018

9.1.2 Table of Identified Compounds and Relative Amounts in *Psy-1* Positive and Negative Lines

Psy-1* vs no *Psy-1						
Var ID	**Probability**	**Average (Plus)**	**Average (Null)**	**Std. dev. (Plus)**	**Std. dev. (Null)**	**Fold change**
1H-Indole-3-ethanamine, N,N,1-tris(trimethylsilyl)-	0.0370875	0.00043699	0.012253	0.000173117	0.0178203	28.0396
Azelaic acid, bis(trimethylsilyl) ester	0.0119719	0.00043699	0.00513443	0.000173117	0.00577635	11.7495
4-Hydroxyanthraquinone-2-carboxylic acid, di-TMS	0.0339378	0.000479392	0.00157518	0.000192349	0.00161674	3.28579
D-Myo-Inositol, 1,2,4,5,6-pentakis-O-(trimethylsilyl)-, bis(trimethylsilyl) phosphate	0.00808521	0.0702955	0.0400084	0.0413806	0.0194403	0.569146
Unknown - 45.4536 [P] (D-Myo-Inositol, 1,2,4,5,6-pentakis-O-(trimethylsilyl)-, bis(trimethylsilyl) phosphate) (318+315+387)	0.00808521	0.0702955	0.0400084	0.0413806	0.0194403	0.569146
a-Tocopherol (1TMS) - 58.5124 [NP] (502)	0.0185058	0.728952	0.386683	0.317204	0.392615	0.530465
à-Tocopherol (vitamin E), trimethysilyl derivative	0.0185058	0.728952	0.386683	0.317204	0.392615	0.530465
Docosanoic acid, trimethylsilyl ester	0.0031504	0.0780923	0.0369662	0.0498911	0.0230656	0.473366
(E,E)-7,11,15-Trimethyl-3-methylene-hexadeca-1,6,10,14-tetraene	0.0387048	0.0111119	0.00447881	0.00957905	0.00748498	0.403066
GABA (3TMS) - 24.0623 [P] (174,304,147)	0.0142065	0.00406103	0.00162088	0.00371179	0.00161947	0.399132
Unknown - 35.3840 (L-Ascorbic acid, 2,3,5,6-tetrakis-O-(trimethylsilyl)-) (332+374)	0.0380423	0.0043747	0.00172965	0.0045526	0.00239222	0.395375
Butanoic acid, 4-[bis(trimethylsilyl)amino]-, trimethylsilyl ester	0.0137465	0.00406103	0.00160278	0.00371179	0.00163165	0.394672
Tetradecanoic acid (1TMS) - 32.9840 [NP] (285)	4.30155E-05	0.0947178	0.0358686	0.0503727	0.0192553	0.378689
Tetradecanoic acid, trimethylsilyl ester	4.30155E-05	0.0947178	0.0358686	0.0503727	0.0192553	0.378689

Unknown - 46.1008 [NPS])Eicosanoic acid (1TMS)) (369)	1.23928E-05	0.107867	0.0388215	0.0509493	0.0243724	0.359901
Standard - 23.7741 [P] (L-Proline, 5-oxo-1-(trimethylsilyl)-, trimethylsilyl ester) pyroglutamic acid 2tms (156)	0.0092517	0.00223865	0.000782839	0.0023209	0.000521508	0.349692
L-Proline, 1-(trimethylsilyl)-, trimethylsilyl ester	0.0130549	0.00520913	0.00181544	0.00467017	0.00264564	0.348512
Phosphoric acid, bis(trimethylsilyl) 2,3-bis[(trimethylsilyl)oxy]propyl ester	0.000821112	0.0976121	0.0315274	0.0727214	0.0278318	0.322986
Standard - 30.5735 [P] (Glycerol-3-Phosphate peak 2 (4TMS)) (299+357)	0.000821112	0.0976121	0.0315274	0.0727214	0.0278318	0.322986
Unknown - 27.5137 [NPS] (Dodecanoic acid (1TMS)) (257+117)	0.000370152	0.0217988	0.00693743	0.0152206	0.00572657	0.318248
Dodecanoic acid, trimethylsilyl ester	0.000360296	0.0217988	0.00683996	0.0152206	0.00583228	0.313777
L-Ascorbic acid, 2,3,5,6-tetrakis-O-(trimethylsilyl)-	0.0098084	0.00514623	0.00158097	0.00494057	0.00242703	0.307208
RI=2722.0, 51.9331 [NP] 9-Octadecenoic acid, 2-[(trimethylsilyl)oxy]-1-[[(trimethylsilyl)oxy]methyl]ethyl ester (218)	0.000561335	0.0887464	0.0262747	0.044903	0.0427693	0.296065
L-Proline, 5-oxo-1-(trimethylsilyl)-, trimethylsilyl ester	0.0038833	0.00223865	0.000622524	0.0023209	0.000441433	0.27808
RI=1709.5, 29.5161 min [NP] Phosphoric acid, 2-(trimethylsiloxy)-1-[(trimethylsiloxy)methyl]ethyl bis(trimethylsilyl) ester (299)	0.00517915	0.00763916	0.00198878	0.00775699	0.00281494	0.26034
RI=1726.2, 29.9664 [NP] Phosphoric acid, 2-(trimethylsiloxy)-1-[(trimethylsiloxy)methyl]ethyl bis(trimethylsilyl) ester / Glycero-2-phosphate (4TMS) (299)	0.00517915	0.00763916	0.00198878	0.00775699	0.00281494	0.26034
Phosphoric acid, 2-(trimethylsiloxy)-1-[(trimethylsiloxy)methyl]ethyl bis(trimethylsilyl) ester	0.0400032	0.0058746	0.00148433	0.00870173	0.00272383	0.25267
Phytol (1TMS) - 40.4894 [NPS] (143+123)	2.50677E-05	0.00245961	0.000437258	0.00187952	0.000153855	0.177776
Lauric acid, 2,3-bis(trimethylsiloxy)propyl ester	0.00574242	0.0114973	0.00180255	0.0143623	0.00344505	0.156779
RI=3168.9, 59.4247 [NP] Hexacosanoic acid, 2-[(trimethylsilyl)oxy]-,	0.00512869	0.0204346	0.00239618	0.0271297	0.00428589	0.11726

trimethylsilyl ester (439)						1
Tocopherol-ç-tms-derivative (high mass adjustment=100%)	0.00404303	0.0933617	0.00892885	0.125504	0.00935032	0.095637
RI=2958.3, 56.0345 min [NP] Tocopherol-y-tms-derivative (488)	0.00403941	0.0933617	0.00891832	0.125504	0.00936047	0.095524
1-Monooleoylglycerol trimethylsilyl ether	0.021375	0.0208367	0.000926854	0.0380249	0.00222352	0.044482
L-Aspartic acid, N-(trimethylsilyl)-, bis(trimethylsilyl) ester	0.0541197	0.00211285	0.000703081	0.00305687	0.000826999	0.332765
Hexadecanoic acid (1TMS) - 37.8220 [NP] (313)	0.0550707	1.08773	0.623393	0.628974	0.623004	0.573115
Hexadecanoic acid, trimethylsilyl ester	0.0550707	1.08773	0.623393	0.628974	0.623004	0.573115
Sedoheptulose, o-methyloxime, hexakis-O-(trimethylsilyl)-	0.056682	0.00135966	0.000606462	0.00148354	0.000680428	0.446039
Stigmasterol (1TMS) - 60.3160 [NPS] (394+484)	0.0569874	0.0380924	0.0604709	0.00970442	0.0365604	1.58748
Standard - 27.5198 [P] (Lyxose isomer 1 MEOX 4TMS) (217+277+307)	0.0578709	0.552757	0.302583	0.441375	0.27719	0.547407
L-Aspartic acid (3TMS) - 23.8666 [P] (232)	0.0582246	0.00215484	0.000774136	0.00303502	0.000847005	0.359255
N,N,o'-Tris-(trimethylsilyl)aspartic acid	0.0582246	0.00215484	0.000774136	0.00303502	0.000847005	0.359255
Phosphoric acid, 2-[bis(trimethylsilyl)amino]ethyl bis(trimethylsilyl) ester	0.0653796	0.0201303	0.012727	0.0141554	0.00787652	0.632233
RI=1761.8, 30.9280 min [NP] Phosphoric acid, 2-[bis(trimethylsilyl)amino]ethyl bis(trimethylsilyl) ester (299)	0.0700485	0.0201303	0.0129691	0.0141554	0.00756052	0.644256
D-Fructose, 1,3,4,5,6-pentakis-O-(trimethylsilyl)-, O-methyloxime	0.0775842	0.552757	0.323529	0.441375	0.270095	0.585302
Standard - 27.9270 [P] (Lyxose isomer 2 (MEOX 4TMS)) (217+227+307)	0.0775842	0.552757	0.323529	0.441375	0.270095	0.585302
Unknown - 33.4797 [P] (D-Fructose, 1,3,4,5,6-pentakis-O-(trimethylsilyl)-, O-methyloxime iso 1) (217+307+364)	0.0775842	0.552757	0.323529	0.441375	0.270095	0.585302

Unknown - 33.7659 D-Fructose, 1,3,4,5,6-pentakis-O-(trimethylsilyl)-, O-methyloxime iso 2 (217+307+364)	0.0775842	0.552757	0.323529	0.441375	0.270095	0.585302
Standard - 16.1972 [P] (Glycerol 3TMS) (205)	0.0803405	0.0860968	0.0512926	0.0688429	0.0404283	0.595756
RI=3325.1, 61.8173 (b-Sitosterol trimethylsilyl ether pepper)	0.0833868	0.617157	0.825967	0.0967831	0.377491	1.33834
á-Sitosterol trimethylsilyl ether	0.0833868	0.617157	0.825967	0.0967831	0.377491	1.33834
2-Monopalmitin (2TMS) - 48.3986 [NP] (313)	0.0874335	0.133551	0.0759618	0.120913	0.0647608	0.568787
2-Monopalmitin trimethylsilyl ether	0.0874335	0.133551	0.0759618	0.120913	0.0647608	0.568787
Unknown - 64.4143 [NPS] (Dotriacontanol (1TMS)) (523)	0.0874469	0.000673028	0.000502187	0.000303313	0.000235056	0.74616
Benzene, 1,3-bis(1,1-dimethylethyl)-	0.0932517	0.000566315	0.00112922	0.000521426	0.00100283	1.99398
Bis(trimethylsilyl)monostearin	0.109858	4.15251	2.91031	2.71369	1.57241	0.700857
Monostearin (2TMS) Octadecanoic acid, 2,3-bis[(trimethylsilyl)oxy]propyl ester (399)	0.109858	4.15251	2.91031	2.71369	1.57241	0.700857
Trimethylsilyl ether of glycerol	0.115037	0.0816741	0.0481731	0.0737123	0.043547	0.589822
Myristic acid, 2,3-bis(trimethylsiloxy)propyl ester	0.119515	0.0779392	0.0499693	0.0626903	0.0365167	0.641132
Unknown - 45.3104 [NP] (Myristic acid, 2,3-bis(trimethylsiloxy)propyl ester) (343)	0.119515	0.0779392	0.0499693	0.0626903	0.0365167	0.641132
RI=3539.9, Unknown 65.0845 min 006 MG 2 1 NP	0.12476	0.0900624	0.0658582	0.0425631	0.0404673	0.731251
Mannonic acid, 2,3,4,6-tetrakis-O-(trimethylsilyl)-, lactone	0.1261	0.00280581	0.0013081	0.00352749	0.00189446	0.466211
RI=3343.0, 62.0888 (b-Amyrin 1TMS pepper) (218+203)	0.128227	0.00428493	0.000746209	0.0104674	0.000791477	0.174147
Lactic acid (2TMS) - 9.4338 [P] (117)	0.129324	0.0165134	0.0223471	0.0115538	0.00920537	1.35327
Propanoic acid, 2-[(trimethylsilyl)oxy]-, trimethylsilyl ester	0.129324	0.0165134	0.0223471	0.0115538	0.00920537	1.35327

Myo-Inositol, 1,2,3,4,5,6-hexakis-O-(trimethylsilyl)-	0.131811	0.00751794	0.00318616	0.0109266	0.00499588	0.423808
Decanoic acid, trimethylsilyl ester	0.131896	0.00123564	0.000774564	0.00104866	0.000640053	0.626855
á-Amyrin trimethylsilyl ether	0.14149	0.032719	0.0180713	0.0246499	0.0267399	0.552319
Unknown - 34.5246 d-Glucose, 2,3,4,5,6-pentakis-O-(trimethylsilyl)-, o-methyloxyme, (1Z)- iso 2 (319+205+160)	0.15066	0.0181614	0.000589665	0.0554346	0.00068572	0.032468
Silanol, trimethyl-, phosphate (3:1)	0.156183	0.81896	0.648091	0.369202	0.285087	0.791358
Unknown - 15.9984 (Phosphate) (299)	0.156183	0.81896	0.648091	0.369202	0.285087	0.791358
1-Monopalmitin (2TMS) - 49.2047 [NP] (371+459)	0.161035	3.70343	2.75146	2.28332	1.46446	0.742948
Hexadecanoic acid, 2,3-bis[(trimethylsilyl)oxy]propyl ester	0.161035	3.70343	2.75146	2.28332	1.46446	0.742948
Unknown - 15.7362 [P] (Ethanolamine (3TMS)) (174)	0.165135	0.000914491	0.00157213	0.000578813	0.00146496	1.71913
RI=2194.5, 41.4416 min [NP] a-Linolenic acid, trimethylsilyl ester (335+75)	0.171818	0.00939827	0.000437258	0.0297892	0.000153855	0.046525
RI=2977.9, 56.3584 [NP] fattty acid (193)	0.180552	0.144895	0.00361391	0.47912	0.014552	0.024942
2-Monostearin - 51.9413 [NP] (218+341)	0.190353	0.172177	0.106294	0.182306	0.0977905	0.617353
3,8-Dioxa-2,9-disiladecane, 2,2,9,9-tetramethyl-	0.192646	0.00292089	0.000534967	0.00830724	0.000444654	0.183152
Galactose oxime hexakis(trimethylsilyl)	0.195066	0.371715	0.199242	0.528959	0.208514	0.536008
9,12-Octadecadienoic acid (Z,Z)-, trimethylsilyl ester	0.203636	0.00043699	0.0169504	0.000173117	0.0418121	38.7891
Glycine, N,N-bis(trimethylsilyl)-, trimethylsilyl ester	0.203685	0.0339289	0.00288256	0.110415	0.00922883	0.084959

1,2,3-Propanetricarboxylic acid, 2-[(trimethylsilyl)oxy]-, tris(trimethylsilyl) ester	0.205136	0.0236655	0.0120027	0.0333775	0.0179158	0.507181
Citric acid (4TMS) - 32.234 [P] (363+347+375)	0.205136	0.0236655	0.0120027	0.0333775	0.0179158	0.507181
1,4-Butanediamine, N,N,N',N'-tetrakis(trimethylsilyl)-	0.20753	0.00147837	0.00837298	0.00167228	0.0175723	5.66367
RI=2202.7, 41.6227 min 006 R 3 1 NP	0.213656	0.137994	0.0715749	0.0956522	0.158167	0.518679
N,O,O-Tris(trimethylsilyl)-L-threonine	0.215293	0.00131267	0.000762313	0.001527	0.000938507	0.580735
1,4-Benzenedicarboxylic acid, bis(trimethylsilyl) ester	0.221669	0.0040234	0.0092961	0.00306397	0.0137322	2.31051
2,3,4-Trihydroxybutyric acid tetrakis(trimethylsilyl) deriv.	0.227809	0.00043699	0.000557257	0.000173117	0.000297212	1.27522
L-Threonine (3TMS) - 19.6641 [P] (218)	0.233026	0.00131267	0.000784894	0.001527	0.000932727	0.597937
Linolenic acid, trimethylsilyl ester	0.236272	0.00043699	0.00559373	0.000173117	0.01404	12.8006
2-Monostearin trimethylsilyl ether	0.237342	0.166543	0.106294	0.18694	0.0977905	0.638237
Heptadecanoic acid, glycerine-(1)-monoester, bis-O-trimethylsilyl-	0.23919	0.00638859	0.00357059	0.00778984	0.00541153	0.5589
Unknown - 16.9371 (Glycine coeluting, ion 174)	0.247991	0.000605847	0.000437258	0.000629356	0.000153855	0.72173
Standard - 31.7308 [P] (Fructose isomer 1 (5TMS) (437+217)	0.248888	0.00627197	0.0292055	0.0086138	0.0638876	4.65651
Standard - 31.9771 [P] (Fructose isomer 2 (5TMS)) (437+217)	0.248888	0.00627197	0.0292055	0.0086138	0.0638876	4.65651
Standard - 27.7515 [P] (unknown?) lyxose something (204+217+191)	0.25789	0.003866	0.0113976	0.00478887	0.0212206	2.94817
Unknown - 16.4675 [P+NP] contamination	0.263978	0.04708	0.0377591	0.0304409	0.0162055	0.802021
L-Isoleucine, N-(trimethylsilyl)-, trimethylsilyl ester	0.265122	0.000743384	0.00127026	0.000606794	0.00146521	1.70875
Octadecanoic acid (1TMS) - 42.2655 [NP] (341)	0.266212	0.636006	0.444273	0.509949	0.424375	0.698537
Octadecanoic acid, trimethylsilyl ester	0.266212	0.636006	0.444273	0.509949	0.424375	0.69853

						7
Unknown - 15.1261 [P] (Benzoic acid trimethylsilyl ester) reagent (179+135)	0.27562	0.00251502	0.00148257	0.00349946	0.00179916	0.589487
N,O-Bis-(trimethylsilyl)isoleucine	0.278604	0.000743384	0.00125539	0.000606794	0.00146533	1.68875
Octadecadienoic acid, cis 9,12 (1TMS) - 41.4953 [NP] (337)	0.307784	0.00043699	0.0121809	0.000173117	0.037247	27.8746
Standard - 53.8311 [P] (Narigenin peak 2?)	0.308079	0.00043699	0.000553866	0.000173117	0.000350134	1.26746
RI=2613.0, 49.9371 min [NP] Docosanoic acid, trimethylsilyl ester (397)	0.309304	0.00043699	0.00319519	0.000173117	0.0087751	7.31183
2-Piperidinecarboxylic acid, 1-(trimethylsilyl)-, trimethylsilyl ester	0.316857	0.00066576	0.000486233	0.000774787	0.000191482	0.730343
L-Tyrosine (3TMS) - 35.1283 [P] L-Tyrosine, N,O-bis(trimethylsilyl)-, trimethylsilyl ester (218+280)	0.326707	0.00043699	0.000545023	0.000173117	0.000334867	1.24722
Standard - 26.4499 [P] (sugar - need to check) (204+191+101+103)	0.337093	0.00503789	0.0113976	0.00444411	0.0212206	2.26238
MSTFA - 17.7999 [NPS] same as 17.8092NP	0.339642	0.00320837	0.00375655	0.00127816	0.00162451	1.17086
Unknown - 16.7390 [P] (L-Proline (2TMS)) (142)	0.339689	0.00043699	0.000510949	0.000173117	0.000218939	1.16925
RI=3238.3, 60.4968 (Campesterol tms pepper) (ions 343+382)	0.350815	0.245213	0.286812	0.0391216	0.141753	1.16964
RI=1349.9, 18.7247 min [NP] Nonanoic acid, trimethylsilyl ester (215)	0.355161	0.000497514	0.000857342	0.000265397	0.00124675	1.72325
Unknown - 18.7878 (L-Serine 3TMS) (204)	0.36316	0.000675148	0.000547945	0.000477825	0.000302185	0.811592
Benzoic acid trimethylsilyl ester	0.367668	0.00232075	0.00145641	0.0035765	0.00180941	0.627561
Pyridine, 3-trimethylsiloxy - 8.6699 [P] contamination same as 8.6781	0.377307	0.00245105	0.0019861	0.00144851	0.00136589	0.810307
Hexacosanoic acid, 2-[(trimethylsilyl)oxy]-, trimethylsilyl ester	0.387403	0.0130287	0.00759678	0.0252598	0.00981195	0.58308
Malic acid (3TMS) - 22.9733 [P] (233)	0.396706	0.00164037	0.00110414	0.00209553	0.00141982	0.673105
Heptadecanoic acid (1TMS) - 40.0664 [NP] (327)	0.398898	0.00563946	0.00419967	0.00337879	0.00499408	0.74469

						4
Benzaldehyde, 4-methyl-, O-methyloxime	0.409728	0.00538788	0.00309041	0.0085847	0.00670202	0.573586
Standard - 29.1814 [Xylitol 5TMS) (217+129)	0.469574	0.00043699	0.000582544	0.000173117	0.000642352	1.33308
Standard - 29.6663 [P] (Ribitol (5TMS)) (217+129)	0.469574	0.00043699	0.000582544	0.000173117	0.000642352	1.33308
RI=2960.4 Unknown 56.0693 min 006 MG 2 1 NP	0.478089	0.00043699	0.00361391	0.000173117	0.014552	8.27002
Nonanoic acid, trimethylsilyl ester	0.505874	0.00160902	0.00228438	0.00287143	0.00260196	1.41974
L-Lysine (4TMS) - 34.7410 [P] (174)	0.509202	0.00043699	0.000503138	0.000173117	0.00030195	1.15137
Serine tritms	0.534707	0.00164518	0.00114194	0.00253389	0.00193423	0.694112
Unknown - 34.1195 d-Glucose, 2,3,4,5,6-pentakis-O-(trimethylsilyl)-, o-methyloxyme, (1E)- iso 1 (319+205+160)	0.555196	0.371715	0.282407	0.528959	0.32044	0.759738
RI=3383.9, 62.7117 (a-Amyrin trimethylsilyl ether pepper) (218+189)	0.572369	0.00836344	0.0127464	0.0121282	0.0237662	1.52406
Unknown - 21.7339 [NP] contamination 307+350	0.576828	0.0285009	0.0254707	0.010334	0.0160926	0.89368
L-Valine (2TMS) - 14.1272 [P] (144)	0.63277	0.000979461	0.00121247	0.00100396	0.00142073	1.23789
L-Asparagine, N,N2-bis(trimethylsilyl)-, trimethylsilyl ester	0.652911	0.00125825	0.000917904	0.00277567	0.00149203	0.729509
L-Valine, N-(trimethylsilyl)-, trimethylsilyl ester	0.675822	0.000950588	0.00115325	0.000973438	0.00142132	1.2132
Standard - 38.3135 [P] (trans-Ferulic acid (2TMS))	0.691039	0.00161796	0.00127122	0.00276669	0.00206309	0.785694
Standard - 38.7522 [P] (Cinnamic acid, 4-methoxy-3-(trimethylsiloxy)-, trimethylsilyl ester/ Ferulic acid MEOX TMS) (338+323+249)	0.691039	0.00161796	0.00127122	0.00276669	0.00206309	0.785694
Unknown - 11.7874 [P+NP] contamination?	0.698004	0.0596203	0.0647076	0.0201305	0.0402905	1.08533
Unknown - 10.5132 [NP] (Undecane)	0.712153	0.00549788	0.00426331	0.0128478	0.00603483	0.77544

						7
N,O-Bis-(trimethylsilyl)valine	0.713493	0.000975909	0.00115325	0.000957704	0.00142132	1.18172
Standard - 42.3233 [P] (trans-Sinapinic acid TMS/ Cinnamic acid, 3,5-dimethoxy-4-(trimethylsiloxy)-, trimethylsilyl ester) (368+353+338)	0.738403	0.000961814	0.000823782	0.00128858	0.000992717	0.856488
Unknown - 62.9128 [NPS] (Triacontanoic acid (1TMS)) (509+524)	0.761043	0.00981866	0.0111006	0.0120863	0.0107651	1.13056
3-Chloro-1,2-propanediol, di(trimethylsilyl) ether	0.762791	0.00175904	0.00147593	0.00232261	0.0025803	0.839056
N,O-Bis(trimethylsilyl)-L-phenylalanine	0.807287	0.00537715	0.00591588	0.00421534	0.006558	1.10019
Unknown - 26.5836 [NPS] (N,O-Bis(trimethylsilyl)-L-phenylalanine) (ions 218+192)	0.807287	0.00537715	0.00591588	0.00421534	0.006558	1.10019
n-Pentadecanoic acid, trimethylsilyl ester	0.825248	0.00198217	0.00209209	0.00134904	0.00131414	1.05545
Unknown - L-Alanine (2TMS) - 10.2859 [P] (116)	0.827099	0.000545194	0.000583268	0.000438618	0.000476581	1.06984
Butanedioic acid, bis(trimethylsilyl) ester	0.841271	0.00251389	0.00228367	0.00305947	0.00306324	0.90842
Sucrose (8TMS) / ?-D-Glucopyranoside, 1,3,4,6-tetrakis-O-(trimethylsilyl)-?-D-fructofuranosyl 2,3,4,6-tetrakis-O-(trimethylsilyl)- (361+437+451)	0.862709	0.0443176	0.0522691	0.0542693	0.145026	1.17942
à-D-Glucopyranoside, 1,3,4,6-tetrakis-O-(trimethylsilyl)-á-D-fructofuranosyl 2,3,4,6-tetrakis-O-(trimethylsilyl)-	0.862709	0.0443176	0.0522691	0.0542693	0.145026	1.17942
RI=1796.5, 31.8635 min [NP] shikimic acid? (255)	0.865577	0.000464736	0.000454231	0.000174336	0.00016059	0.977397
L-Glutamine (3TMS) - 30.9868 [P] (156)	0.878008	0.000563865	0.000536766	0.000496457	0.000456679	0.95194
Nonadecane	0.881099	0.0129816	0.0133247	0.00615804	0.00608676	1.02643
1,2-Ethandimine, N,N'-ditrifluoroacetyl-	0.895816	0.000708945	0.000754022	0.000846374	0.000950462	1.06358
Phosphoric acid, bis(trimethylsilyl)monomethyl ester	0.938908	0.000714064	0.000729039	0.000600792	0.000475423	1.02097
Hexane, 2,5-dimethyl-2,5-bis(trimethylsilyloxy)-	0.943257	0.00237845	0.00227051	0.00340603	0.00432336	0.954615

Arabinofuranose, 1,2,3,5-tetrakis-O-(trimethylsilyl)-	0.974542	0.000788863	0.000774468	0.00118022	0.00121249	0.981753
Unknown - 29.7065 [P] (1,4-Butanediamine, N,N,N',N'-tetrakis(trimethylsilyl)-) (174+200)	0.98933	0.000464293	0.000465331	0.000173091	0.000221804	1.00224
Unknown - 56.8069 min [NPS] (Hexacosanoic acid (1TMS)) (453)	0.994147	0.0154144	0.0153671	0.017932	0.0167667	0.996934
Unknown - 16.5908 (Isoleucine) (158)	0.994798	0.000469876	0.000470407	0.000190844	0.000229108	1.00113
RI=1896.9, 34.4435 [NP] L-Lysine, N2,N6,N6-tris(trimethylsilyl)-, trimethylsilyl ester (174)	0.996444	0.00043699	0.000437258	0.000173117	0.000153855	1.00061
RI=1768.6, 31.1090 [NP [1,4-Benzenedicarboxylic acid, bis(trimethylsilyl) ester / Terephthalic acid trimethylsilyl ester (295)	0.998028	0.0017483	0.00174548	0.00205425	0.00343163	0.998385

9.2 Carotenoid content: Morocco grown samples

Accession	Violaxanthin	Standard Error	Antheraxanthin	Standard Error	Neoxanthin	Standard Error	Lutein	Standard Error	β-carotene	Standard Error	Zeaxanthin	Standard Error	Phytoene	Standard Error
RH 4	235.74	25.00	36.67	3.45	83.56	3.23	168.17	20.62	0.00		19.42	1.50	278.11	3.90
RH 3	88.14	22.66	14.39	2.25	54.58	4.87	103.77	11.76	0.00	0.00	10.33	0.38	184.13	22.76
RH 2	42.99	8.42	12.77	1.82	49.36	3.87	86.61	14.02	154.05	22.51	26.68	5.35	519.90	103.63
RH 0	15.79	2.69	70.44	15.81	12.08	1.33	15.48	6.23	0.00	0.00	34.21	4.72	31.53	10.20
RH 1	110.47	24.74	22.57	4.86	35.55	2.55	82.13	14.05	313.89	0.80	37.28	1.07	626.65	9.38
RH 4 (set 2)	240.61	54.02	53.18	26.71	76.90	15.52	221.07	50.67	0.00	0.00	32.05	4.48	281.38	23.63
RH 15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

9.3 Plastome/Genome Ratio

A PCR based study comparing levels of *Rubsico large subunit (RbcL)* (plastid encoded) and *Phytoene desaturase (Pds)* (genome encoded) was used to estimate per-cell plastid number and compare between accessions of different *Psy-1* genotypes.

Accession	*RbcL/Pds*
RH 1	4.78
RH 2	2.91
RH 10	9.47
RH 15	3.62

9.4 *Psy-1* Genomic DNA

Lowercase letters denote intron sequences, uppercase denote exon sequences. Highlighted sections represent primer sequences; primer pairs are highlighted in the same colour.

GGAGC

AATAAAAGAGGAGAAAAACCAACAAACTACTTGTAATTCTTGAAGAATACAAAGCAGACA

TTGTTGAAAAGTGGCCATTTTTGCAAGgtttgtttcacttcaattcttgttatcgaattc

ataactttagttcaaaccatgtaacatgttaacttaaaaatattataaaatttagaattc

ataaactttaaattttgactcgttatgtttattgatatacaaagttgcaatatttaacat

tctttgtaactttctattgctattgtagtgttcttgttgctattctaggggcggagctag

agtgtcagttacgcattcagttgaattcagtaactttagttcaaaccccgtgatgtgtta

acttagaaatattagagttcagaattcatgaacatcaaattttgacacgttacatttctt

gatatacaaagttgcagtctttgacaatctgtgtggctttttattgctattgtagtgttc

ttgttgctatacatgggcggagctagagtgttagttacgggtttagttgaattcagtaac

tttagttcagaccccgtatcgtgttaacttaaaaaatattagaattctgaacccataaag

ttcaaattttgactcgattcgacttcttgatgtataaagtttcaatctttaacattcttt

gtaactttctattgttctggtagtgtgtgtgttggtctactgggcggagctagagtatca

gttacggatttagttgaatttagtagcttagcttaaaccctgtaacatagtaacttaaaa

atattgtaaaatttagaactcataaacttcaaattttgactcattacatttcttgatata

caaagttgcaatctttgacattctgtgtggctttttattgctatggttgtgcgcttgttg

ctatacatgggcggatctagagtgttagttacgggtttagttgaattcagtaactttagt

tcaaaccccgtaatgtgtttacttagaaatattagaatttagaactcataaatttcaaat

tttggttcgttacatttcttgatatataaaattgcaatctttgacattctctgtggctct

ttattgctatggtagtgtgcttgttgctatactaggggtggagctagagtgtcaattatg
ggtttggttgaattcagtaactttagttcaaaccccataacgtgttaaacttagaaatat
tagattttagaactcataaacttcaaattttgactcgttacattttcttgatatacgaag
ttgcaatctttgacattctgtgtaactttctattgctattgtagtgttgttgttgctata
ctaggggcagagctacagtgtgagttatgtattcagttgaattcagtagctttggttcaa
accccataacatagtaacttaaaaatgttaaaaatttagaactcataaacttcaaatgtt
gactcgttacatttcttgatatacaaagttgcaatctttgacattctttgcggcttttta
ttgctatggttgtgtgcttgttgctatacatgggcggagctagagtgttagttacgggtt
tagttgaatttagtaggtttagttcagaccccgtaatttgtaacttaaaaaacattagaa
ttctgaacccataaagttcaaattttgactcggttcgacttctagatgtataaagtttca
atctttaacattctttgtaacttttattgctatagtaatgttcttttgctatactagg
ggcggaactagagtgtaagtcgaattcagtaactttagttcaaaccatgtaacctagtat
cttaaaaatattaaaattcggaaccccataaacttcaaattttagctccgactatacaag
aatagaatttgagagtggcctaaaatttagtgtccattgtttagtggagaaccaattatc
aaggtttgtttcccttcacttcttgatatacaagttgcaatgcttaacattccctgtaac
tttctattccactggtagtgtgcttgttgctataccggggcggagctagagtgttagtta
cgggtttggtcgactttagtagggcagcccggtcactacagctcccgctatgcgcagggt
ggcgggaaggggcggaccacaagagtcttttaacatttttgcaaggagctgtttccacgc
ttgaaccggggacctcctagtcactagaattcagtagctttagttcagaatccacaaact
tcaaattctagctccgcctactagtgactataaaataatagaaaatgagcacttgcctta
tgaatatagcttctacgtgtaccaaaattagaaagtgaggtgcttattataatctagttg
actaaatatagaaagatcccattacctccagaaagtgtgattccactttgtgctttcaat
agtgtaaaatagtttctcaaacatcctttctttgtgccattggtaggtaagattgctgtt
ttgtcttgttttggagatgtttttaaagttaaaattgtttatactcaggaagtctataa
tgctggttacgagttcatgtgaactcagtagttttcgctaagcaatccactaatatctat
aaatatttgaccgtgaactccgcttgtatcatttcattaacttgaggtcgctataggagc
cgataaacttccaacttctgaatcagtttttgtttatactcaagtatgatgtttggttta
tctcatattgcagAAGCCAAGAAATAGGTTATTTCTTTGTTTGATAGTGGAAGTATACTC
TAGTGGGAATCTACTAGGAGTTACTTATTTTCTATAAAGAAGACAAAAACCTTGGAGTTG
CTTTAGACAACCAAGGTTTTTCTTGTTCAGAATGTCTGTTGCCTTGTTATGGGTTGTTTC
TCCTTGTGACGTCTCAAACGGGACAGGATTCTTGGTATCCGTTCGTGAGGGAAACCGGAT

TTTTGATTCGTCGGGGCGTAGGAATTTGGCGTGCAATGAGAGAATCAAGAGAGGAGGTGG
AAAACAAAGGTGGAGTTTTGGTTCTTGCTTGGGAGGAGCACAAACTGGAAGTGGACGGAA
ATTTTCTGTACGTTCTGCTATCGTGGCTACTCCGGCTGGAGAAATGACGATGTCATCAGA
ACGGATGGTATATGATGTGGTTTTGAGGCAGGCAGCCTTGGTGAAGAGACAGCTGAGATC
GACCGATGAGTTAGATGTGAAGAAGGATATACCTATTCCGGGGACTTTGGGCTTGTTGAG
TGAAGCATATGATAGGTGTAGTGAAGTATGTGCAGAGTACGCAAAGACGTTTTACTTAGg
ttagcttctttatctatccgttcgtttaccaatactatatatgcctgtttatgtatctga
tgaagacaaaatttgatctttgtttggttattcagGAACGATGCTAATGACTCCGGAGAG
atctttatctttgtttgattcagagttctacgtcgctttagctgaatgctttccgcatgg
AAGAAAGGCTATCTGGGCAATATACGgtgaggattttagccatgttgtatcagatacaca
actaaacacatatgattaatcggagacgaagaaaaagaactaagatttgagtttgagggt
caccagtgatacgtaaaatctgagtttcactagcttgaaggcatcgtaagaagttttatt
gtcatatcaagcttaatctttacaatgattgctagaggttgcggttattgcatatataaa
cgaacttcgtgagttgaatagtttctcattatctgctgtttaaacagTATGGTGCAGGAG
AACAGACGAACTTGTTGATGGTCCGAATGCATCACACATTACTCCGGCGGCCTTAGATAG
GTGGGAAGACAGGCTAGAAGATGTTTTCAGTGGACGGCCATTTGACATGCTCGATGCTGC
TTTGTCCGACACAGTTTCCAAATTTCCAGTTGATATTCAGgttagtcttccaattctatg
attactcgtattctttcatcctcacagaaaatgtctttagttggatcagttcctctcgtc
ctagattctacttgtactcggttttctgtttcacagttttgcgaactcttgtttttcac
cttttggtgttgtcttcactacactttttggcttcgtttctttggaacagatacattat
tcattcacataatttgatttgctttcttcaacaacgtaaattttctaatctttcgatata
tgttgcagCCATTCAGAGATATGATTGAAGGAATGCGTATGGACTTGAGGAAGTCAAGAT
ACAGAAACTTTGACGAACTATACCTATATTGTTATTACGTTGCTGGTACGGTTGGGTTGA
TGAGTGTTCCAATTATGGGCATCGCACCTGAATCAAAGGCAACAACGGAGAGCGTATATA
ATGCTGCTTTGGCTTTGGGGATCGCAAATCAGCTGACCAACATACTTAGAGATGTTGGAG
AAGAgtaagtacaacgatacattttatgcacatcaaaagattcgctacacctgtctactc
gacgcttatcacttaaaaaagattcgatttagttatctttcaggccccaaaattttcatg
tcgtggaattctttccccctcgacatatagactctgacaagccacgttcaatcgctttcc
ttttttccttcattgaagtaccataatttttgctcttaaggcttccttttttcttatatgt
tgctttcgctgcagTGCCAGAAGAGGAAGAGTCTATTTGCCTCAAGATGAATTAGCACAG
GCAGGTCTATCCGACGAAGACATATTTGCTGGAAGAGTGACCGATAAATGGAGAATCTTC

ATGAAGAAACAAATTCAGAGGGCAAGAAAGTTCTTTGACGAGGCAGAGAAAGGAGTGACC
GAATTGAGCGCAGCTAGTAGATGGCCTgtaagcattcataaacactcttagttatatgaa
tatgattccccttccccgtttcgagtcaggtgtccaggcgccttgactgttgtgcgctag
acattactagcgctggccgaacactgccataccggcgtcaccatcctaggcatggtgggc
acagagggtgatgggtggctctgataccactgttagctaataaggtgggagagcccgagg
ctataaataacccccgccattactttgtagtaccaatgtgagactttttgcaacggatca
taacatatttcttggtcgatgaagttgagacgaaagggttttgaagttttaacaatgttt
actgggtgtcgtcatttgctaacccatccggagtctaggaattggtcatgttactctgaa
cggactagaattctaaccttatatcaggacctacagacgaatgtttaaccttgcatttcc
gaatatttgctctaacatgacattgtgggatcctgaaatgctttggattggatcaaacaa
aaacattcaatctgacaacatacttaatgttagaattacaatattccatagctcgcatct
ttaatattctcttatttattacttgtatctacgaattttggagaataacctcacccggtg
ctcggatttttttcagGTGTTGGCATCTCTGCTGTTGTACCGCAGGATACTGGACGAGAT
CGAAGCCAATGACTACAACAACTTCACAAAGAGAGCTTATGTGAGCAAACCAAAGAAGTT
GATTGCATTACCTATTGCATATGCAAAATCTCTTGTGCCTTCTACAAGAACATGAAATCA
GGATTTTATATAAATCAAGGCCAATGAAGCCAATATACATTTAGAAGAAAAAAAACAAGT
GTTTATAAAGTAGAATTATTGAAGGGGAGGCTTGGAGTAACTGGTAAAGTTGTTGTCATG
TGACTGGGAAGTCACGGGTTCAAGCCTTGGAAACAGCCTCTGGCAGAAATGCAAGGTAAG
GTTGCGTACAATATACCGTTAAGGTGGGGTCCTTCCCAGTACACCGCGCATAGCGATAGA
TTTAGTGCACCGGGTCGCCTTTTTTCTAAAGTAGAATAATTGTTGTATTCATGTCAATGT
ATATCATCAAAATTAGGTGGTAGTAAAATCCAATGTAACAATCTCCCACTTTCACCAGTT
ATTCACTCCGCACTTGGCTActcagtgtcttgctatgggcacgaaaactggttcgccgaa
ttaaaacacactaaatgaatggataaaacttatagcatcgagtatttatgccaaaccaat
aaatgatatacatgtgctttttatatacacatttatcattaggcaatggttaggtgctat
aggggcggctcaagcatataagtggcctaaggccaaaatcgaacggaggccttaattttt
tttaataaagtctattttaaaaattttataatccttaataatataatcgatgcattaatc
ttttccgagacatagtactttagatatattaattttttttaataatttttatcttcatat
aaaaaagttatttttcatcaaataatattattatctttctcaaaaaaaaaaaattcatta
ttgctagaaaatgaggccctcaaatttgggggccttaggtagttgccttttttgtaatga
agtcgagccgtccctggtcctatcacccccctaatctattttttcatgtatttattacccc
tccaaatgctgagctggtaaagaaggtgagaacacccccctctgggcgcgttggtttagtc

ctaacaactccaaaagtaagattttagtcttccacctgtataatacacatcattaattga
ttttttgaagaaataaatcatgaaaaatgactctctctctctcttcaacggtcatcttcc
tcttctttatatcactcttttctttgtccacctactgccttcatctttctccatttttt
tattttatttctttaacctacgttttagcttttaatcccaatttctacttgaattcttgg
agttttgttgtgtaatttggttcaactcttaatttaggatttgtaatttgttgattgatt
ttatcggaattcatgaaatgacaaatttgaaattgaataagctttgtatggatgaatcag
acccaatgctcaatgctgttgtgagatgcaaacatggtattttgcttgacttacaaattt

9.5 RNAseq Figures

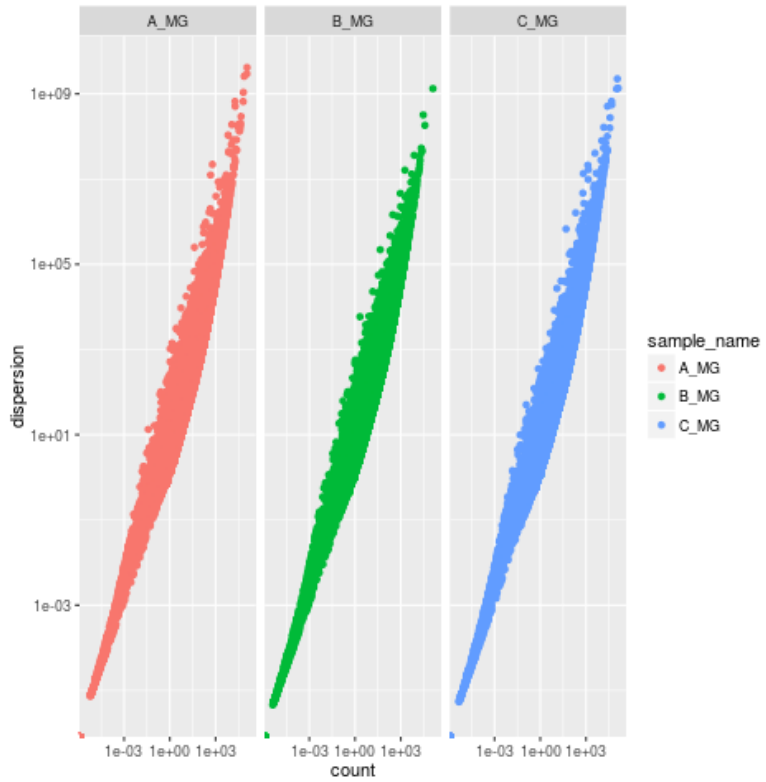


Figure 9.1. Dispersion plot showing the dispersion of counts from mature green samples. The plot was generated with the R package CummeRbund. (n=3 biological replicates)

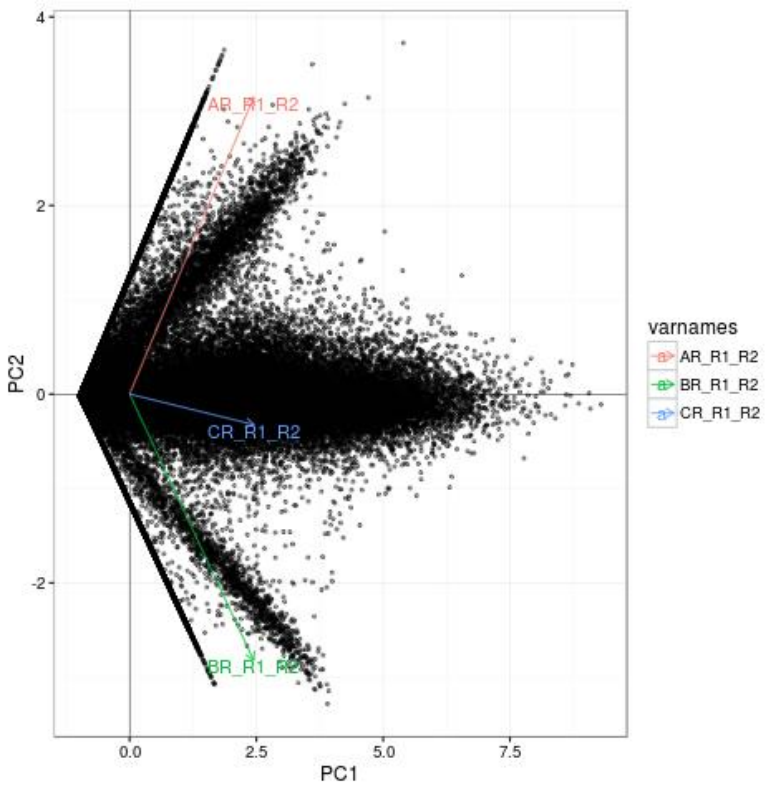


Figure 9.2. PCA showing separation of samples based on transcript expression levels. The plot was generated with the R package CummeRbund (n=3 biological replicates).

Using PCA to investigate the how transcripts from ripe fruit samples group together showed a broadly similar trend to the mature green samples: the *c1* positive line (A: 414) pulls transcripts to the top of the plot, whereas the *c1* negative lines (B: 255; and C: 446) group in the lower half of the plot. However, the *c1* negative lines group less closely in the ripe fruit samples (Figure 9.2).

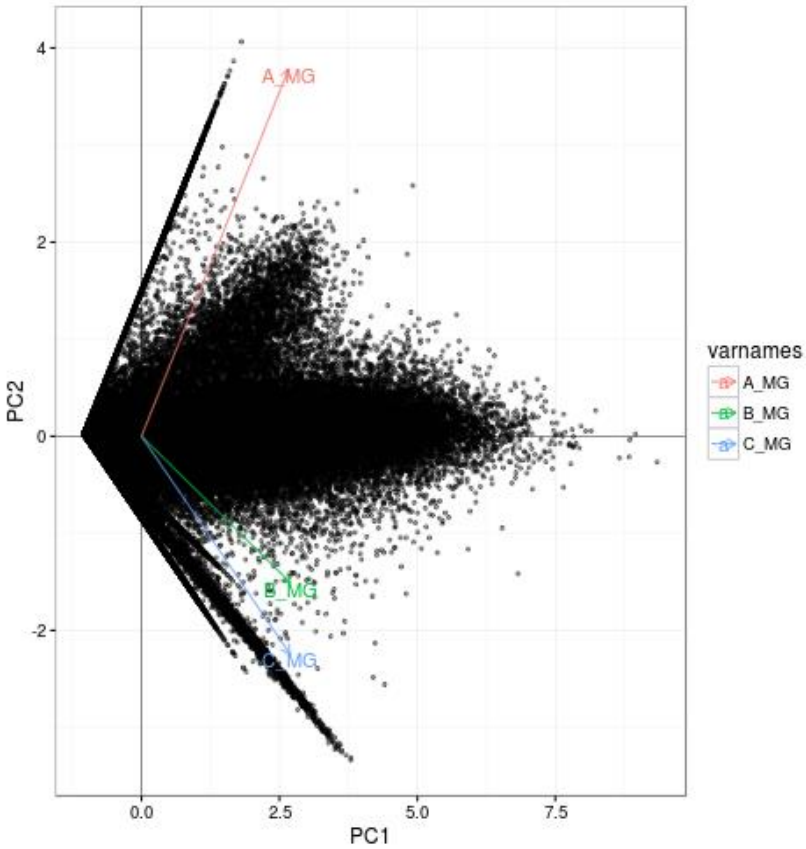


Figure 9.3 PCA showing separation of samples based on transcript expression levels. The plot was generated with the R package CummeRbund. (n=3 biological replicates)

Principal Components Analysis (PCA) was used to investigate how the samples could be separated based on levels of the transcripts detected. When expression levels of all transcripts were plotted with this technique it was clear that the two lines lacking a functional copy of the *c1* locus (B: 255; and C: 446) clustered together away from the line which possessed a functional copy of the locus (A: 414) (Figure 9.3).

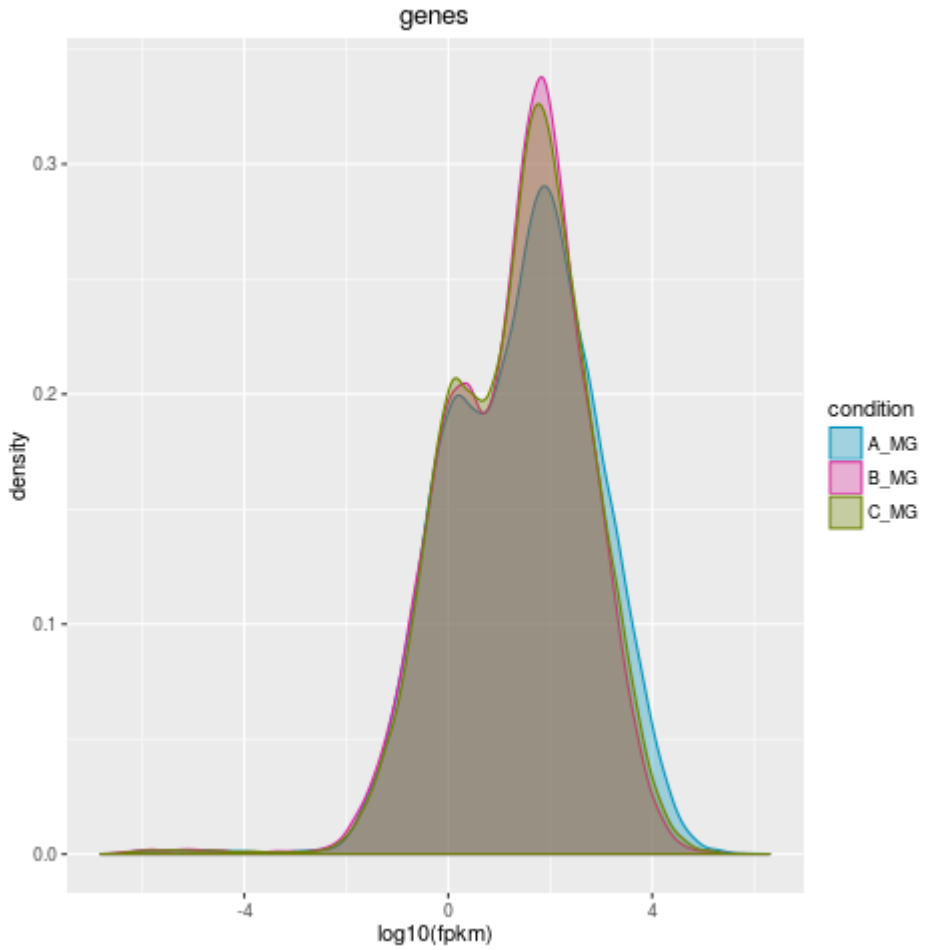


Figure 9.4. Density plot for mature green fruit samples. The plot was generated with the R package CummeRbund (n=3 biological replicates).

The density plot for the three sets of mature green samples showed that transcript expression levels were distributed similarly across densities for all conditions, with the highest frequency expression level being approximately 3.8 log10(fpkm) (Figure 9.4).