Identification and characterisation of phenolic based quantitative trait loci from wild tomato relatives

Daniel Viner Rickett

This thesis is submitted for the degree of Doctor of Philosophy at Royal Holloway, University of London, in September 2012

Declaration of authorship

I, Daniel Rickett, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed:

Date:

Abstract

Quality traits associated with consumer health are an important aspect of modern plant breeding programmes. Of particular interest are phenolic compounds, a group that includes phenylpropanoids, flavonoids and anthocyanins. Previous studies have shown an inverse correlation between the incidence of chronic disease states and the intake of fruits and vegetables rich in phenolics. Molecular linkage mapping populations that utilise wild tomato relatives in traditional breeding strategies exploit wild germplasm for the identification of novel quantitative trait loci (QTL) without the use of genetic modification.

In this study, the *Solanum neorickii* backcross inbred line population was screened for novel phenolic profiles. Marker data available within the EU Sol consortium were used to identify chromosomes 5 and 10 as possible QTL-containing regions for high rutin and *p*-coumaric acid phenotypes, and accessions neo-111 and neo-123 were selected as candidate lines, respectively. Accession 3939 was incorporated from the *Solanum habrochaites* near isogenic line population because of its similarities with neo-111. Each accession was characterised for metabolomic and physiological properties throughout fruit development and ripening. Results indicated multiple metabolic pathways were affected, including increases in isoprenoid intermediates; fruit ripening time was increased; and fruit size was altered. Antioxidant capacity of polar extracts increased, especially in fruit tissues accumulating phenolics. Transcriptomic analysis throughout fruit development identified 186 differentially expressed transcription factor (TF) genes. The promoter regions of two of these TF genes, a MYB-related in chromosome 5 and a LIM family in chromosome 10, were sequenced for differences between the wild relative and *Solanum lycopersicum* parents. These differences provide candidate QTL identities for the previously characterised phenotypes, and offer the potential for future utilisation in plant breeding programmes for high phenolic commercial lines.

Table of contents

Table of figures

Table of tables

Abbreviations

3GT	flavonoid-3-O-glucosyltransferase
4CL	4-coumaroyl:CoA-lyase
ABTS	2,2’-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
$ABTS^{\cdot+}$	ABTS radical cation
AMDIS	Automated Mass Spectral Deconvolution and Identification System
ASTM	American Society for Testing and Materials
Av.	average
BIL	backcross inbred line
Br	breaker
C	Celsius
C3H	*p*-coumaroyl ester 3-hydroxylase
C4H	cinnamate-4-hydroxylase
CC	core collection
cDNA	complementary DNA
CGA	chlorogenic acid
CHI	chalcone isomerase
CHS	chalcone synthase
COSII	conserved orthologs set II
D	Dalton
DAD	diode array detector
DIMS	direct infusion mass spectrometry
DNA	deoxyribonucleic acid
Dpa	days post anthesis
Dpb	days post breaker
Dph	days post harvest
DW	dry weight
EBG	early biosynthetic gene
EDTA	ethylenediaminetetraacetic acid
F3’5’H	flavonoid-3’5’-hydroxylase
F3’H	flavonoid-3’-hydroxylase
F3H	flavanone-3-hydroxylase
FLS	flavonol synthase

G	multiples of gravity
G	gram
GC	gas chromatography
GM	genetic manipulation
H	hour
HCT	cinnamoyl-CoA shikimate/quinate transferase
Hp	high pigment
HPLC	high performance liquid chromatography
HQT	hydroxycinnamoyl-CoA quinate transferase
HUJI	Hebrew University of Jerusalem, Israel
IAA	isoamyl-alcohol
IL	introgression line
K3OR	kaempferol-3-*O*-rutinoside
Kb	kilo-base pair
L	litre
LBG	late biosynthetic gene
LC	liquid chromatography
LOD	limit of detection
LOI	limit of identification
LOQ	limit of quantification
M	metre
MG	mature green
Min	minute
Mol	moles
MPI-MP	Max-Planck-Institut für Molekulare Pflanzenphysiologie, Germany
MS	mass spectrometry
MSTFA	*N*-methy-*N*-(trimethylsily) trifluoroacetamide
MTBE	methyl *tert*-butyl ether
N	number of replicates
N.D.	not detected
NIL	near isogenic line
NIST	National Institute of Standards and Technology
PAL	phenylalanine ammonia lyase
PC	principal component

PCA	principal component analysis
PCR	polymerase chain reaction
PDA	photo diode array
QC	quality control
qRT-PCR	quantitative real-time PCR
QTL	quantitative trait locus/loci
R	ripe
RFLP	restriction fragment length polymorphism
RHUL	Royal Holloway, University of London
RIL	recombinant inbred line
RNA	ribonucleic acid
RT	flavonoid-3-O-rhamnosyltransferase
S	second
sect.	section
SEM	standard error of the mean
SGN	Sol Genomics Network
T	turning
TEAC	Trolox equivalent antioxidant capacity
TF	transcription factor
Tris	tris(hydroxymethyl)aminomethane
Trolox	6-hydroxy-2,5,7,8-tetramethychroman-2-carboxylic acid
Tt	*transparent testa* Arabidopsis mutant line, or gene (*TT*)
UNK	unknown compound
UNKnp	unknown compound from non-polar extraction
UNKp	unknown compound from polar extraction
UPLC	ultra performance liquid chromatography
UV	ultra-violet
v/v	volume per volume
W	Watt
w/v	weight per volume

Acknowledgements

I thank my supervisors at Royal Holloway, Dr Paul Fraser and Prof. Peter Bramley, for their encouragement, support and humour throughout this project, and my supervisor at Syngenta, Dr Charlie Baxter, for his consultation, advice and guidance.

This project would not have been possible without funding from an EPSRC Syngenta CASE award and facilities provided by Royal Holloway and the Bramley/Fraser group. Additional resources were provided by Syngenta and the EU Sol consortium.

This project has benefited from the many people who have willingly provided their time or skill set in either collaborative work or advice. I am grateful to Dr Johannes Rohrmann and Dr Alisdair Fernie at MPI-MP for the TF platform analysis, and to Dr Shannon McDonald, Dr Nan Zhou and Dr Brandon Hurr at Syngenta for microarray analysis, promoter motif analysis, and consultation on post-harvest storage. I thank Mark Levy for enthusiastic horticultural support and knowledge. I am grateful to those at Royal Holloway who have helped me with learning new techniques and provided invaluable advice, and want to especially mention Dr Genny Enfissi, Dr Laura Perez-Fons, Dr Tom Wells, Chris Gerrish, Dr Matthew Jones, and Dr Dana Heldt. I thank too all members of the Bramley/Fraser group, past and present, who have made it a unique and enjoyable place to work.

To my parents, Valerie and Alan Rickett, and my brother Chris, I thank you for your encouragement, foundation and for instilling the need to ask why.

Above all I thank Nicole. For your belief in me, your time, your unconditional support, and for saying ‘yes’.

1 Introduction

1.1 Aims and objectives

The aim of this study was to screen fruit from the *S. neorickii* BIL population for accessions with stable perturbations in phenolic profiles associated with health-promoting properties; fully characterise biochemical and physiological changes within these accessions by comparison with the elite *S. lycopersicum* background; and identify quantitative trait loci (QTL)-containing introgressed regions and QTLs thought to be responsible for the perturbations in phenolic profile. This was achieved by fulfilling the following objectives:

01. Screen *S. neorickii* BIL population cultivated in a field environment for fruit phenolic profiles and identify stable lines by comparison with metabolite data available within the EU Sol consortium

02. Confirm stable phenotypes when reproduced under glasshouse conditions

03. Nominate candidate introgressed regions as possible sources of metabolite QTL using marker data available within the EU Sol consortium

04. Fully characterise perturbations within the phenolic pathway throughout fruit development stages and in specific tissue types

05. Assess the post-harvest properties of accessions containing the QTL

06. Assess *in vitro* the antioxidant characteristics of accessions containing the QTL

07. Utilise in-house metabolomic platforms to identify coordinated metabolic changes to elucidate the molecular basis of the QTL

08. Identify specific polymorphisms within the QTL in order to nominate a genetic basis for the phenotypic changes

1.2 Tomato as a valuable foodstuff

1.2.1 Tomato industry production

Solanum lycopersicum (tomato; formerly called *Lycopersicon esculentum*) is a domesticated crop plant grown for its red coloured fruit. It is within the same genus as crops *S. tuberosum* (potato) and *S. melongena* (aubergine, or eggplant). *Solanum* is a member of the Solanaceae family, which comprises more than 3,000 species (Grandillo et al., 2011) and includes the crop plants *Capsicum annuum* (bell and chilli peppers), *Physalis peruviana* (cape gooseberry), *P. philadelphica* (tomatillo), and *Nicotiana tabacum* (tobacco). Solanaceae also encompass non-edible nightshade species, such as *Atropa belladonna* (deadly nightshade), and ornamental species of horticultural importance from the *Petunia* and *Nicotiana* genera.

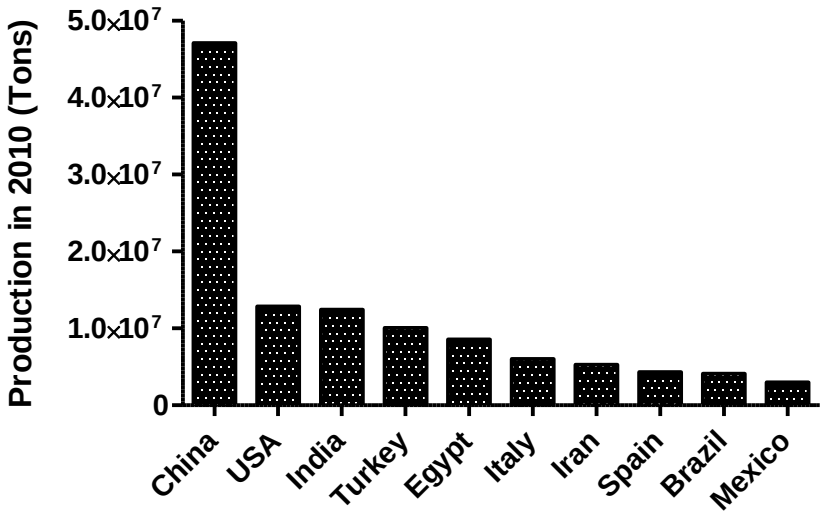

Figure 1 Tomato production for ten highest tomato producing nations in 2010

Data taken from FAOSTAT database (2012) http://faostat.fao.org/

Different morphologies of 'elite' tomato cultivars are available that include cherry, cocktail and larger plum or round cultivars used as fresh produce, as well as processing line cultivars used to make tomato products such as sauces, ketchups, pastes, juices and chopped tomato foodstuffs (Labate et al., 2007). Global production of processing and fresh fruit tomato crops combined has increased steadily by 291 % between 1961 and 2002 to more than 100 million tons per year (http://www.fas.usda.gov), and in 2009 this increased beyond 150 million tons. The majority of processing line fruits is used to make paste and sauces, and export of tomato

products from the USA alone exceeded $232 million in 2002. In 2010, China was the largest producer of tomatoes, with 47.1 million tons (Figure 1; http://faostat.fao.org/). USA, India and Turkey each produced more than 10 million tons in 2010. The UK produced 89.3 thousand tons in 2010 (http://faostat.fao.org/).

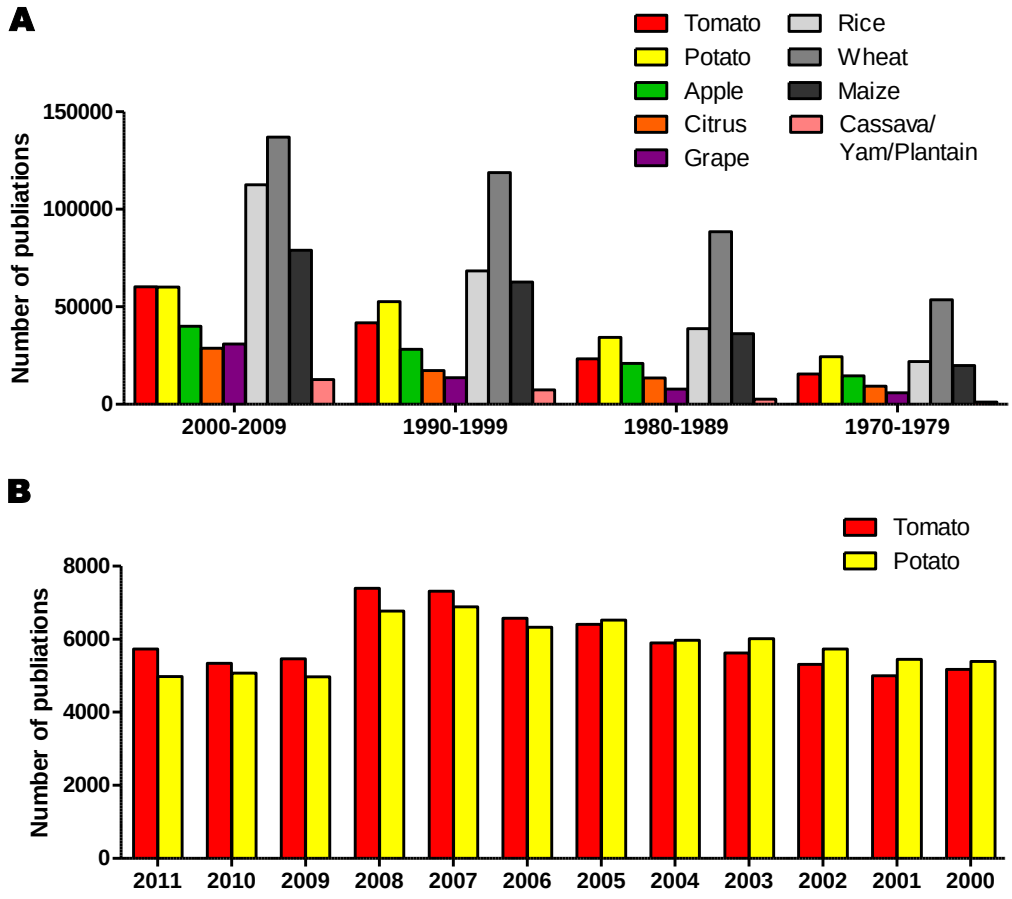

Figure 2 Research output from important food crops

Number of publications listed on Web of Knowledge (http://apps.webofknowledge.com) in each year as a result of literature search with 'topic' key words indicated. Cassava/yam/plantain represents sum of three searches combined.

In addition to an increase in global tomato production, research output from tomato has increased (Figure 2). While research into cereals continues to dominate in number of publications, the number of tomato publications rivals that of potato (Figure 2B), a staple food source in developed countries, and far exceeds that of crops used as staple foods in the developing world, such as cassava, yam, and plantain. Tomato research also exceeds that of other commercially important fruits (Figure 2A). Other than commercial value tomato possesses further qualities that allow it to attract this

level of research. Firstly, tomato products are a large component of Western diets; therefore, any nutritional improvement has the potential to benefit large sections of the population (Labate et al., 2007). Secondly, tomato is a model system for fleshy fruit and its development, and is closely related to other non-fruit crops such as potato. Any research outputs, therefore, have the potential for wider impact in crop science. The high status of tomato as a model crop is reflected in the recent publication by Sato and colleagues (2012) of the tomato genome.

1.2.2 Fruit quality traits in tomato

Fruit quality traits are the parameters by which the fruits are valued during cultivation, preparation and consumption as either fresh produce or processing line tomato products. Tomato fruit quality traits can be categorised into three classes: organoleptic, agroeconomic, and health-related. Organoleptic fruit quality traits are predominantly led by the consumer, and include fruit colour, taste, aroma, sugar/acid balance, crispness, mealiness, firmness, size and shape. Efforts have been made to understand the genetic basis of organoleptic traits in tomato with the view to commercial exploitation (Ashrafi et al., 2012; Carli et al., 2011; Causse et al., 2001; Saliba-Colombani et al., 2001); however, it has also been recognised that consumer preference can vary greatly (Causse et al., 2010).

In contrast to organoleptic traits, agroeconomic fruit quality traits are those led by the producer. These traits include yield, production or ripening time, resistance to diseases, and tolerance to post-harvest conditions. In addition there are important traits encompassing the whole plant, for example defence against herbivore attack and tolerance to abiotic stress such as drought, light, temperature, and soil pH or salinity. Again, the commercial benefit of improving these traits is apparent and there is much research in these areas (Carrera et al., 2012; de Castro et al., 2010; Gur et al., 2011).

Health-promoting quality traits are defined by levels of specific metabolites or compound classes that when consumed are believed to promote health or are associated with disease prevention. Foodstuffs that have been manipulated to improve health-promoting quality traits and therefore contain enhanced levels of specific metabolites are sometimes labelled ‘biofortified’ or ‘functional foods’, and have been extensively reviewed (Tucker, 2003; Mayer et al., 2008).

It should be recognised that fruit quality traits are not necessarily restricted to one category. For example, the carotenoids lycopene and β-carotene are antioxidants that have been associated with the prevention of some disease states when consumed by humans (reviewed by Fraser and Bramley, 2004), but also contribute much of the red and orange pigment, respectively, in tomato fruit. Perturbations to levels of these compounds can therefore affect both health-promoting and organoleptic quality traits (for example Apel and Bock, 2009).

1.2.3 Domestication of tomato

It is believed that tomato was domesticated from ancestral species in the Americas, likely southern Mexico or Peru, at a time predating the Spanish conquest of Mexico in 1512 (Labate et al., 2007). Domestication altered the appearance of tomato over generations for desirable organoleptic and agroeconomic quality traits by selectively breeding to achieve the large red fruit we now recognise as *S. lycopersicum*.

Evidence that demonstrates the extent of these changes can be seen from so-called 'wild relatives', which have not undergone domestication (Labate et al., 2007). The species *S. lycopersicum* is part of the *Solanum* genus section called *Lycopersicon*. Within *Lycopersicon* section are 12 other species of wild relatives of tomato (Table 1), and approximately 75,000 accessions exist in seedbanks throughout more than 120 countries (Grandillo et al., 2011). Wild tomato relatives are usually small and berry-like and, with the exception of *S. cheesmaniae*, *S. galapagense* and *S. pimpinellifolium*, exist in various shades of green, even when ripe.

In relatively recent years, breeding within domesticated cultivars has also allowed selection of plants based on their nutritional and health-promoting quality traits; however, the original domestication has unintentionally resulted in a narrowing of the gene pool. Vavilov in 1940 notably highlighted this as a concern (reviewed by Tanksley and McCouch, 1997), which resulted in establishing seed stocks of wild relative varieties, known as 'exotics'. It is thought that while the exotic phenotypes within these germplasm banks are undesirable, exploitation of the 'masked' genetic information could be beneficial to our current agricultural gene pools. As a result it is suggested that when aiming to improve phenotypic traits in agricultural crops, such as

tomato, with a 'non-GM' approach, focus be shifted from phenotype-led to genetic-region-led searching (Tanksley and McCouch, 1997).

Table 1 Species within *Solanum* sect. *Lycopersicon*

Current *Solanum* and former *Lycopersicon* genera names for 13 species in *Solanum* section *Lycopersicon*. Table adapted from (Grandillo et al., 2011) and (Peralta et al., 2008).

Species under *Solanum* genus	Formerly known as species under *Lycopersicon* genus	Fruit morphology when ripe
S. lycopersicum	*L. esculentum*	Red, large by comparison
S. arcanum	Part of *L. peruvianum*	Green, dark green stripes
S. cheesmaniae	*L. cheesmaniae*	Green-yellow, orange, 0.5 to 1.5 cm
S. chilense	*L. chilense*	White-green, green, purple stripes
S. chmielewskii	*L. chmielewskii*	Green, dark green stripes, 1.0 to 1.5 cm
S. corneliomulleri	*L. glandulosum* or part of *L. peruvianum*	Green, dark green, purple, striped
S. galapagense	Part of *L. cheesmaniae*	Yellow, orange, 0.5 to 1.0 cm
S. habrochaites	*L. hirsutum*	Green, dark green stripes, hairy
S. huaylasense	Part of *L. peruvianum*	Green, purple-green, dark green stripes
S. neorickii	*L. parviflorum*	Green, dark green stripes
S. pennellii	*L. pennellii*	Green
S. peruvianum	*L. peruvianum*	Green, white-green, purple
S. pimpinellifolium	*L. pimpinellifolium*	Red, 0.5 to 1.0 cm

1.3 Phenolic compounds in tomato: a health-promoting trait

1.3.1 Introduction to phenolic compounds

The biosynthesis of phenolic compounds encompasses many diverse structures that, among others, include phenylpropanoids, flavonoids and anthocyanins (Figure 3). In the literature this collective series of reactions is therefore often referred to as the phenylpropanoid, flavonoid or anthocyanin pathway; however, here it is labelled phenolic biosynthesis due to the defining phenolic ring structures of each compound within the pathway. The shikimic acid pathway is situated at the start of phenolic biosynthesis, where the eponymous phenolic ring structure is formed. The resulting phenylalanine molecule leads into phenylpropanoid metabolism. This pathway, comprising largely colourless compounds, builds upon the phenolic ring of phenylalanine to feed many other pathways, such as those of lignin and flavonoid biosynthesis. The flavonoid pathway comprises the synthesis of both flavonoids and isoflavonoids, which range in colour from yellow through to orange and red. Flavonoids in turn lead to the biosynthesis of anthocyanins, which range in

pigmentation from purple to blue, and brown coloured proanthocyanidins (condensed tannins). The biosynthesis of each subsection of this pathway has been extensively reviewed (Gonzali et al., 2009; Rausher, 2008; Shirley, 1996; Vogt, 2010; Winkel, 2008; Winkel-Shirley, 2001a; Winkel-Shirley, 2001b; Yu and Jez, 2008; Zhao and Moghadasian, 2008) .

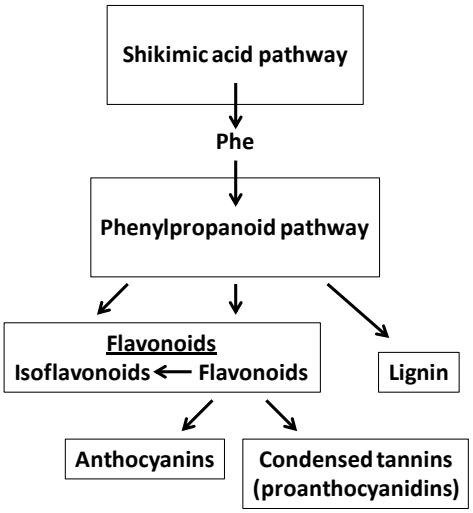

Figure 3 Biosynthesis of phenolic compounds

Schematic representing relationship between some of the divisions of phenolic compound biosynthesis. Phe, phenylalanine.

1.3.2 In vivo function of phenolics in plants

Phenolic compounds comprise a major group of plant secondary metabolites. Their biosynthesis and structural diversity is based on their defining phenolic ring structures. Sequential changes to this structure consequentially give rise to an array of functions *in planta* including allelopathy; molecular signalling; protection against UV radiation; defence against microbes, insects and mammalian herbivores; tissue pigmentation; male fertility; and mechanical stability (reviewed by Chopra et al., 2008; Peer and Murphy, 2008; Shirley, 1996; Vogt, 2010; Weisshaar and Jenkins, 1998; Winkel-Shirley, 2001a) .

Pigmentation is one of the more commonly recognised functions. Phenolic compounds contribute to a range of pigmented plant tissues including flowers, fruit, leaves and seeds. While many compounds from early biosynthesis are colourless (do not absorb wavelengths within the visible range), they are still detectable to some pollinating insects. Further progression along the biosynthetic pathway features

compounds with an increased number of conjugated double bonds, and consequently later flavonoid compounds absorb wavelengths within the visible range (Shirley, 1996). Both flavonoid pigments and complexes of anthocyanin pigments with flavonol co-pigments colour flower and fruit tissues in order to attract pollen and seed dispersers (Winkel-Shirley, 2001a). Chalcone-naringenin is a yellow flavonoid found in the peel of tomato fruits, thereby contributing to the overall fruit colour. Fruit lacking this pigment appear pink (Ballester et al., 2010) due to the red flesh seen through the colourless skin. Proanthocyanidins (condensed tannins) are flavonoid polymers that are brown in colour and contribute to seed coat (testa) colour as well as robustness. Arabidopsis mutants from the *transparent testa* (*tt*) class produce seeds with various shades of light brown to yellow colouring due to the reduced levels or absence of proanthocyanidins (reviewed by Chopra et al., 2008; Shirley, 1996).

Another function of phenolics that is extensively researched is the protection against harmful UV-B radiation. Many phenylpropanoid and flavonoid compounds have absorption maxima in the UV-B range of 280 to 320 nm; furthermore, flavonols have two absorption maxima within this range (Shirley, 1996). Many phenolic compounds are found in high abundance in leaf and pollen epidermis, ideally located for this function, and plants have been shown to induce flavonoid synthesis upon UV-B radiation exposure (Bieza and Lois, 2001). Several studies support this function, such as UV-B sensitivity in flavonoid-deficient Arabidopsis (Li et al., 1993; Lois and Buchanan, 1994) and high flavonoid levels in a UV-B tolerant Arabidopsis mutant (Bieza and Lois, 2001). Naringenin, rutin, and flavonoid extracts from apple have been shown to prevent UV-B induced DNA damage *in vitro* (Kootstra, 1994), and tolerance to UV-B radiation has been linked with increased levels of the phenylpropanoid chlorogenic acid, total phenolics, and specific phenolic profiles in tomato (Cle et al., 2008).

1.3.3 Health-promoting properties of phenolic compounds in human diets

Foodstuffs possessing health-related quality traits, sometimes labelled ‘functional foods’ or ‘nutraceuticals’, contain levels of particular compounds that, when consumed, are believed to contribute to promoting health or preventing disease. Although the range of these compounds encompasses a wide variety of metabolite classes, flavonoids are considered the major active nutraceutical in plants (Lin and

Weng, 2008). Diets rich in flavonoids, fruits and vegetables are protective against diseases such as cardiovascular disease (CVD) and some cancers (Ness and Powles, 1997; Vauzour et al., 2010).

Reactive oxygen species (ROS) are known to damage DNA and contribute to cellular aging and the oxidation of low-density lipoproteins; they are also believed to play a role in coronary heart disease and carcinogenesis (Lin and Weng, 2008). Flavonoids such as quercetin and kaempferol have been shown to possess antioxidant activity *in vitro* (Dwyer, 1995). It is believed flavonoids can exert their protective effect, therefore, by the transfer of free radical electrons, as well as the inhibition of oxidases, and the activation of antioxidant enzymes (Lin and Weng, 2008).

Direct links between dietary intake of plant derived phenolic compounds have been made with prevention or reversal of age-related diseases (Joseph et al., 1999); reduced occurrence of coronary heart disease, CVD, and associated mortality (Mink et al., 2007); and the contribution to reducing occurrence of stroke and thrombosis by inhibiting platelet aggregation (reviewed by Nijveldt et al., 2001). Seeram and colleagues (Seeram et al., 2004) used human oral, colon and prostate tumour cell lines to demonstrate an inverse correlation between proliferation and exposure to flavonoids extracted in polyphenolic fractions of cranberries. It is hypothesised that systematic increases in dietary flavonoids of interest could reduce the occurrence of all cancers by 7-31% and of deaths due to coronary heart disease by 30-40% (Hertog et al., 1993).

1.3.4 Biosynthesis of phenolics in plants

Phenylalanine is an essential amino acid comprising an aromatic ring and three-carbon side chain. The shikimic acid pathway exists in plants, fungi and bacteria, but not animals, and metabolises erythrose 4-phosphate and phosphoenol-pyruvate through a series of intermediates including shikimate to synthesise the eponymous phenolic ring and resulting phenylalanine structure (Figure 4). In a reaction that bridges primary and secondary plant metabolism, phenylalanine ammonia lyase (PAL) then catalyses the production of cinnamic acid. Not only is this the most important regulatory step of phenylpropanoid metabolism (Yu and Jez, 2008), but since its discovery in 1961 PAL is one of the most characterised enzymes of the pathway (Koukol and Conn, 1961; MacDonald and D'Cunha, 2007).

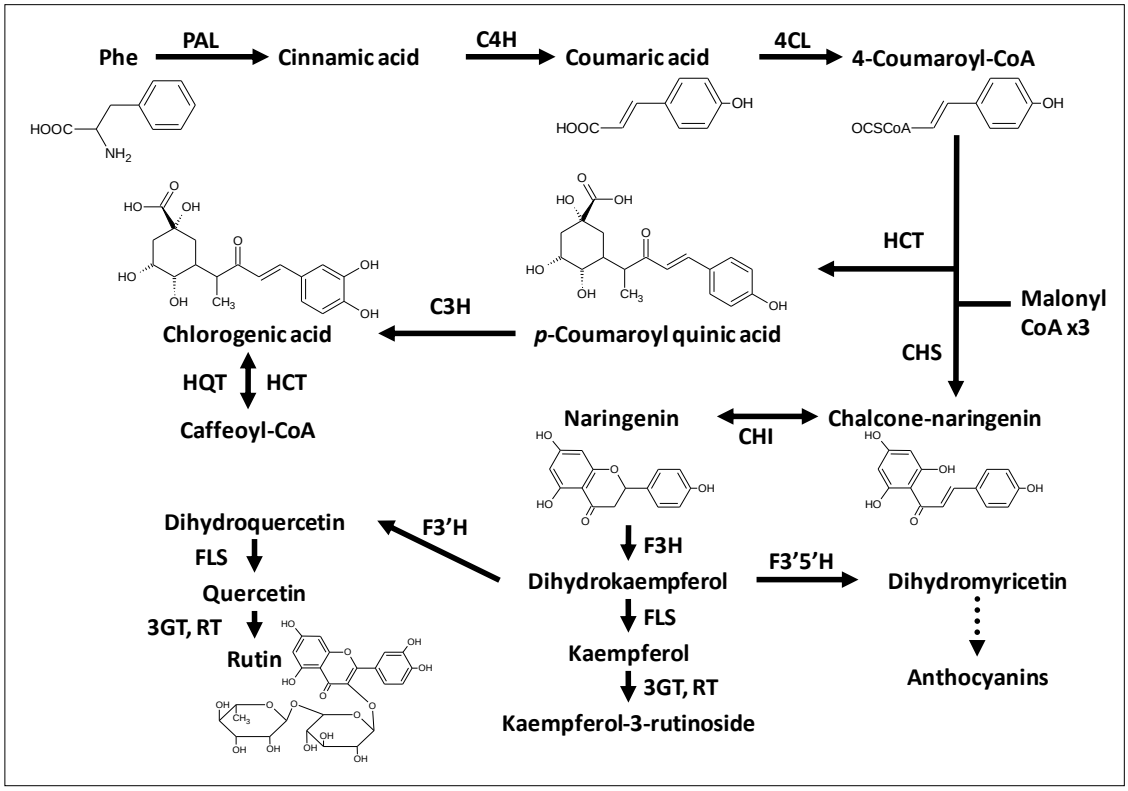

Figure 4 Schematic of phenolic biosynthesis

Overview of central phenylpropanoid pathway and flavonoid intermediates in tomato, illustrating some of the major catalytic steps and intermediate structures. Phe, phenylalanine; PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumaroyl:CoA-lyase; HCT, cinnamoyl-CoA shikimate/quinate transferase; C3H, p-coumaroyl ester 3-hydroxylase; HQT, hydroxycinnamoyl-CoA quinate transferase; CHS, chalcone synthase; CHI, chalcone isomerise; F3H, flavanone-3-hydroxylase; FLS, flavonol synthase; F3'5'H, flavonoid-3'5'-hydroxylase; 3GT, flavonoid-3-O-glucosyltransferase; RT, flavonoid-3-O-rhamnosyltransferase; F3'H, flavonoid-3'-hydroxylase.

Cinnamic acid is hydroxylated by cinnamate-4-hydroxylase (C4H, a cytochrome P450 enzyme) to form 4-coumaric acid (more commonly called *para*- or *p*-coumaric acid). Alternatively, in some monocots *p*-coumaric acid may be produced in a single step reaction from tyrosine by tyrosine ammonia lyase (TAL) (Rosler et al., 1997). *p*-coumaric acid may feed into a series of reactions towards lignin biosynthesis via intermediates such as ferulic acid (Weisshaar and Jenkins, 1998; Zhao and Moghadasian, 2008). Alternatively, *p*-coumaric acid may be metabolised in a series of reactions towards caffeic acid and subsequently chlorogenic acid (Weisshaar and Jenkins, 1998). However, should *p*-coumaric acid continue along the central phenylpropanoid pathway towards flavonoid biosynthesis it will be catalysed by 4-coumaroyl:CoA-lyase (4CL) to result in 4-coumaroyl-CoA. The flux may once again lead away from the central phenylpropanoid pathway and feed into coumarin

biosynthesis (Yu and Jez, 2008). However, should the flux continue towards flavonoid metabolism chalcone synthase (CHS) catalyses the final step that bridges phenylpropanoid and flavonoid metabolisms. CHS is an important regulatory step, and since the first descriptions of the CHS enzyme and *CHS* gene from parsley the complex mechanism of CHS is now better understood (Kreuzaler and Hahlbrock, 1975; Kreuzaler et al., 1983; Yu and Jez, 2008). Molecules of malonyl-CoA derived from fatty acid biosynthesis each extend 4-coumaroyl-CoA by two carbons in sequential condensation reactions. The limited size of the active site of CHS will not permit more than three stepwise reactions, before cyclisation forms a chalcone ring from the newly added carbons to form chalcone-naringenin. A heterocyclic ring (the flavonoid C ring) is formed by chalcone isomerase (CHI) in a reaction that catalyses chalcone-naringenin into the flavanone called naringenin. Flavanones are a subgroup of the wider categorised flavonoids, defined by their diphenylpropane 6C-3C-6C backbone. There are believed to exist more than 6,400 known flavonoid compound structures (Winkel, 2008), and the biosynthetic network forms a 'grid' structure comprising many compound classes. Some flavonoid classes are more specific to individual plant species, such as the isoflavonoids found in legume species, and the phlobaphenes common to some cereals (Winkel-Shirley, 2001b). Some classes are more commonly accumulated in specific plant tissues, such as proanthocyanidins in seed coats. On the other hand, some flavonoid classes, such as the flavonols and anthocyanins, are found throughout the higher plant kingdom and in many tissue types (Shirley, 1996).

More than 70 flavonoids have been found in tomato (Moco et al., 2006). In addition to naringenin, the flavonols quercetin and kaempferol, together with their glycosides, are common components of tomato fruit. Naringenin is catalysed into the first of the dihydroxyflavonols, dihydrokaempferol, by flavanone-3-hydroxylase (F3H). Dihydrokaempferol can be catalysed by flavonoid-3'-hydroxylase (F3'H) into an alternative dihyrdroflavonol called dihydroquercetin. Each of the dihyroflavonols, dihydrokaempferol and dihydroquercetin, can be catalysed by flavonol synthase (FLS) to form respective flavonols, kaempferol and quercetin. Flavonol glycosides can be formed by sequential reactions catalysed by flavonoid-3-*O*-glucosyltransferase (3GT) and flavonoid-3-*O*-rhamnosyltransferase (RT). The resulting compounds are kaempferol-3-*O*-rutinoside (K3OR) and quercetin-3-*O*-rutinoside (rutin) (Ballester et al., 2010; Winkel, 2008; Winkel-Shirley, 2001a).

1.3.5 Regulation of phenolic biosynthesis

As previously detailed in section 1.3.2, phenolic compounds are believed to function *in planta* as protection against harmful UV-B radiation (Bieza and Lois, 2001), and plants exposed to UV-B are shown to exhibit elevated levels of phenolics (Ryan et al., 2001). Phenolic profiles have also been observed to change under other abiotic or environmental stresses such as limited water (Sanchez-Rodriguez et al., 2012), nitrogen deficiency (Bonguebartelsman and Phillips, 1995), and temperature change (Gautier et al., 2008). Lovdal and colleagues (2010) demonstrated not only that nitrogen depletion, temperature decrease, and increased light intensity each resulted in an increase in phenolic content and biosynthetic gene expression in tomato leaves, but that these abiotic stresses exhibited synergistic effects on phenolic biosynthesis when in combination. Perturbations to phenolic profiles are also observed under pathogenic (biotic) or herbivore (mechanical) stresses (Atkinson et al., 2011; Conceicao et al., 2006; Mellway et al., 2009).

Plant stress, whether derived from biotic or abiotic stimuli, consequently affects the regulation of phenolic biosynthesis. These phenolic changes due to stress often affect groups of compounds, or the phenolic profile, rather than isolated phenolic intermediates, or target compounds (Quattrocchio et al., 2008). This is both due to the structure of biosynthetic enzymes and the mechanism of regulation. Stafford first suggested in 1975 that biosynthetic enzymes in phenolic biosynthesis may form multi-enzyme complexes (reviewed by Winkel-Shirley, 2001a). Direct associations have been found between CHS, CHI, F3H and DFR in Arabidopsis (Burbulis and Winkel-Shirley, 1999). It is proposed that expression of biosynthetic enzymes in branches of the phenolic 'biosynthetic grid' pathway can therefore be regulated independently; however, that the regulation of some sections of this grid is still unknown (Quattrocchio et al., 2008). Dicotyledons, such as tomato, possess two clusters of co-regulated biosynthetic genes known as the early biosynthetic genes (EBGs) and the late biosynthetic genes (LBGs) (Mol et al., 1998). Although a few examples of the regulation of phenolic biosynthesis are due to post-translational modification (for example Pairoba and Walbot, 2003), most are controlled by the action of transcription factors (TFs) (reviewed by Quattrocchio et al., 2008).

The regulation of proanthocyanidins (condensed tannins) has been extensively researched in Arabidopsis due to the many *tt* (*transparent testa*) mutants exhibiting

reduced levels or absence of these compounds in the seed coat (testa). Some of the genes underlying *tt* mutant phenotypes have been assigned to TF families (Quattrocchio et al., 2008). These genes include *TT1*, a WIP family TF (a new family containing a zinc-finger domain) (see review by Quattrocchio et al., 2008); *TT2*, an R2R3-MYB family TF (a subgroup of MYB family that contains an R2R3 domain and is named after the myeloblastosis gene) (Martin and PazAres, 1997; Stracke et al., 2001); *TT8*, a bHLH family TF (named after the conserved basic helix-loop-helix domain) (Toledo-Ortiz et al., 2003); *TT16*, a MADS-box family TF (a conserved domain named from an acronym of four of the original members of the TF family, MCM1 in *Saccharomyces cerevisiae*, AGAMOUS in Arabidopsis, DEFICIENS in *Antirrhinum majus*, SRF in human) (see review by Quattrocchio et al., 2008); and *TTG2*, WRKY family TF (named after the conserved amino acid sequence tryptophan, arginine, lysine, tyrosine) (Rushton et al., 2010). Regulation of transcription of a subset of LBGs of proanthcyanidin biosynthesis has been shown by a complex of TF proteins (Figure 5A) including the R2R3-MYB TT2 and the bHLH TT8 in Arabidopsis (Quattrocchio et al., 2008).

Likewise, transcription of anthocyanin biosynthetic genes has been shown to be regulated by TF complexes. AN1 and JAF13 are two bHLH proteins known to regulate anthocyanin biosynthesis in petunia (Spelt et al., 2000). These bHLH TFs can form hetero- or homo-dimers and construct a complex with one of two R2R3-MYB TFs AN2 or AN4 to differentially affect expression of anthocyanin LBGs (Quattrocchio et al., 1999; Spelt et al., 2000). While many TF are positive regulators of groups of phenolic biosynthetic steps, examples exist of negative regulators. Strawberry *FaMYB1* and Arabidopsis *AtMYB4* (Figure 5C) are two R2R3-MYB negative regulators of anthocyanin biosynthesis that inhibit transcription of biosynthetic genes by forming likewise TF complexes.

Relatively less is known about transcriptional control of flavonols, or flavonoids in general. Flavonol production has been shown to be altered as a result of the manipulation of TFs targeted to anthocyanin biosynthesis and LBG expression. For example, increased flavonol accumulation was seen in Arabidopsis as a result of the overexpression of the endogenous R2R3-MYB TF *PAP1* (Tohge et al., 2005). In tomato fruit, increased expression of EBGs was observed and an accumulation of flavonols, including rutin, was found in tomato fruit skin as a result of the simultaneous expression of the anthocyanin biosynthesis regulatory genes from *Antirrhinum majus* (snapdragon) *Del* (a bHLH family TF) and *Ros1* (a MYB-related family TF) (Butelli et

al., 2008a). Likewise, Bovy and colleagues (2002) used maize anthocyanin TF regulators LC and C1 (bHLH and MYB) to induce flavonol accumulation in tomato fruit flesh. Negative regulators have demonstrated similar effects. Aharoni and colleagues (2001) overexpressed strawberry *FaMYB1* R2R3-MYB family TF in tobacco, which resulted in decreased expression of LBGs and decreased accumulation of anthocyanin and flavonol compounds in flowers.

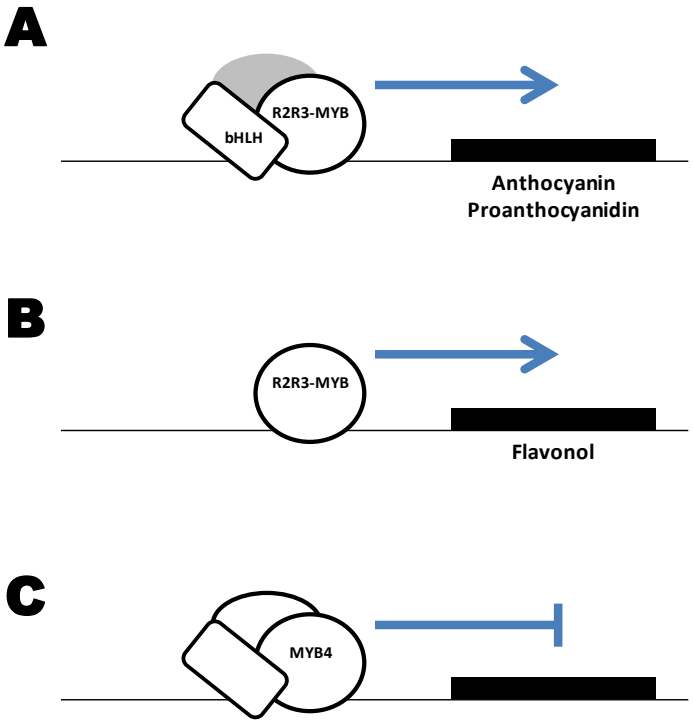

Figure 5 Schematic overview of regulation of phenolic biosynthetic genes by TF complexes

(A) Regulatory complex comprising R2R3-MYB family TF and bHLH family TF components together with third unannotated regulatory component, such as WDR, in grey. Complex regulates transcription (blue arrow) of anthocyanin and anthocyanidin biosynthetic genes (black box). (B) Regulation of flavonol biosynthetic genes by R2R3-MYB family TF in no known complex. (C) Negative regulation (inhibition, blue line) of biosynthetic genes by TF complex containing AtMYB4 as an example of MYB family TF. Schematics based on (Gonzali et al., 2009; Quattrocchio et al., 2008).

The Arabidopsis TF *AtMYB12*, however, is noted for its difference from these previously highlighted examples of R2R3-MYB TFs (Mehrtens et al., 2005). Not only has AtMYB12 been shown to induce flavonol accumulation in Arabidopsis by the regulation of CHS, CHI, F3H, and FLS without affecting anthocyanin LBGs, but AtMYB12 appears to also function independently from any known co-activator, bHLH TF family or otherwise (Figure 5B).

While these many examples illustrate the effects of artificially perturbing expression of TF regulators on subsections of phenolic biosynthetic gene expression and subsequent accumulations of phenolic compounds, they do not address how endogenous expression of TFs is itself controlled. Promoter regions of these TF genes possess TF binding sites, and together with feedback mechanisms from accumulations of phenolic compound intermediates, likely contribute to the control of TF expression (Quattrocchio et al., 2008; Winkel-Shirley, 2001a). The consequence of this is a cascade of TF regulation. For example, R2R3-MYB TFs AN2 and AN4 in petunia are believed to regulate expression of the bHLH TF *AN1* (Spelt et al., 2000), and the R2R3-MYB TF TT2 in Arabidopsis is believed to regulate expression of the bHLH TF *TT8* (reviewed by Quattrocchio et al., 2008).

1.3.6 Potential application in tomato for manipulation of phenolics

While levels of phenolic compounds in domesticated tomato *S. lycopersicum* are less abundant than is desirable for health-promoting benefits, tomato maintains a functional pathway that can be manipulated to optimise phenolic compound production. Many of the genes in phenylpropanoid biosynthesis, for example, are highly conserved among all higher plants (Yu and Jez, 2008) allowing the potential application of exogenous biosynthetic genes from related plant species to assist with this optimisation. Furthermore, much about the mechanisms of regulating flux dynamics among flavonoid pathway branches is still unknown (Winkle, 2008). Tomato wild relative germplasm possess as yet unexploited genetic information that could regulate expression of these conserved pathways.

1.4 Genetic manipulation of phenolic biosynthesis in tomato

Although phenolic compounds encompass a wide range of structures, some groups of which are specific to individual plant species or families, the central pathway is largely conserved among all plants (Tanaka and Ohmiya, 2008). An increasing understanding of branches and individual steps of phenolic biosynthesis and its regulation has allowed genetic manipulation (GM) of these biosynthetic steps in tomato, as well as other plant crops. The transfer of biosynthetic and regulatory genes between species has been performed in order to optimise accumulation of phenolic intermediates.

Several reviews summarise recent advances in tomato (Davies, 2007; Sevenier et al., 2002; Tanaka and Ohmiya, 2008).

1.4.1 Manipulation of biosynthetic genes

By manipulating individual steps in the series of inter-connected pathways, or 'biosynthetic grid', a better understanding of the overall flux can be gained (Quattrocchio et al., 2008). Niggeweg and colleagues (2004) used both gene silencing and overexpression in tomato to perturb the levels of hydroxycinnamoyl-CoA quinate:hydroxycinnamoyl transferase (HQT), an enzyme involved in chlorogenic acid synthesis. The resulting decrease in chlorogenic acid in response to *HQT* silencing demonstrated that HQT is in fact integral to chlorogenic acid accumulation (Niggeweg et al., 2004). Overexpression successfully increased chlorogenic acid accumulation in tomato fruit. Likewise, the overexpression of petunia *CHI* in tomato resulted in an increase of flavonol intermediate accumulation to such an extent that whole fruit levels were comparable to onion (Muir et al., 2001). Flavonol content in tomato fruit peel, which predominantly comprised rutin, was shown to have increased by up to 78-fold as a result of the genetic manipulation. Furthermore, overexpression of *CHS* and *FLS* in combination resulted in synergistic increases in fruit flesh flavonol content (Verhoeyen et al., 2002).

The overexpression of biosynthetic genes can be used to manipulate tomato into accumulating intermediates not normally detected in tomato fruit. Stilbenes, such as resveratrol, are common components of grape (Brewer, 2011). Isoflavones are a subgroup of flavonoids produced by legumes, but less common in other plant species (Yu and Jez, 2008). The production of both, however, was stimulated in tomato fruit (Schijlen et al., 2006; Shih et al., 2008). Constitutive expression of grape stilbene synthase (*STS*) in tomato under the cauliflower mosaic 35S promoter (*CaMV 35S*) produced high levels of stilbenes such as resveratrol in tomato fruit peel and flesh (Schijlen et al., 2006). Constitutive expression of soybean isoflavone synthase (*GmIFS2*) under CaMV 35S resulted in high levels of genistin (an isoflavone) in vegetative tissue and moderate levels in fruit peel without any drastic effects on existing profiles of fruit phenolics (Shih et al., 2008). Schijlen and colleagues (2006) formed constructs containing two biosynthetic steps with similar aims to engineer novel tomato phenolic profiles. Petunia chalcone synthase (*CHS1*) and alfalfa chalcone reductase

(*CHR*) when constitutively expressed in tomato produced high levels of deoxychalcones, such as butein, in fruit peel (Schijlen et al., 2006). Petunia chalcone isomerase (*CHI*) and gerbera flavone synthase (*FNS-II*) were constitutively expressed in tomato and resulted in significant accumulation of flavones, such as luteolin, in fruit peel (Schijlen et al., 2006).

Together these examples demonstrate that manipulation of individual biosynthetic steps can result in targeted effects on tomato fruit phenolic profiles. Tomato can also be manipulated to synthesise novel compounds in a targeted manner. However, wider impact on phenolic profiles can be obtained by the manipulation of regulatory elements, as described in the following section.

1.4.2 Manipulation of pathway by transcriptional regulation

Section 1.3.5 provided an overview of the role regulatory elements play in controlling the synthesis of phenolic compounds, especially with regards to key TF families such as MYB and bHLH. Genetic manipulation of these regulatory elements, therefore, is a rich resource for manipulation of phenolic biosynthesis in tomato fruit.

Several examples exist of the manipulation of endogenous regulatory genes in tomato affecting phenolic fruit profiles. Constitutive expression of the endogenous TF *AN1* (an R2R3-MYB) in tomato fruit resulted in up-regulation of EBGs and LBGs from the anthocyanin pathway (Mathews et al., 2003). Plants displayed vegetative tissue and tomato fruit with purple pigmentation as a result of anthocyanin accumulation. Giliberto and colleagues (2005) increased accumulation of flavonoids in tomato fruit and anthocyanins in vegetative tissue by the overexpression of endogenous *CRY2*, a photoreceptor regulatory element. Conversely, down-regulation of the tomato regulatory element *DET1* by RNA interference resulted in accumulation of flavonoid compounds (Davuluri et al., 2005), and phenolic compounds such as chalcone-naringenin, chlorogenic acid and quercetin derivatives were later shown to exhibit qualitative differences in ripe fruit (Enfissi et al., 2010).

The combined manipulation of two TF genes from maize was used in tomato. Both were expressed in tomato in a fruit-specific manner (Bovy et al., 2002). The result of *LC* (a bHLH) and *C1* (a R2R3-MYB) expression was an accumulation of leaf

anthocyanins, as well as an increase in fruit flesh flavonols such as kaempferol and fruit flesh flavanones such as naringenin by up to 20-fold.

1.5 Mapping populations utilising wild relative germplasm

Although the GM strategies highlighted in section 1.4 are successful in optimising phenolic profiles in tomato, application of these products as foodstuffs is restricted by a lack of public acceptance of GM technology. New products are additionally answerable to stringent parameters before becoming available commercially (Kok and Kuiper, 2003).

It has long been recognised that domestication has resulted in a genetic bottleneck for some traits (Tanksley and McCouch, 1997). It is estimated that cultivated tomato varieties contain as little as 5 % of the genetic variation found in wild species (Miller and Tanksley, 1990). A study by Garcia-Martinez and colleagues (2006), which aimed to characterise phenotypically distinct domesticated tomato cultivars, was unable to do so based on genetic polymorphisms alone. Wild relatives of tomato, on the other hand, possess vast genetic variation, especially self incompatible species. There is greater genetic variation within a single accession of some self incompatible wild relative species than in all accessions of some self compatible species (Miller and Tanksley, 1990). Therefore, wild relative germplasm offers the potential for the exploitation of genomes with the aim of improving quality traits in domesticated crops (Grandillo et al., 2011; Tanksley and McCouch, 1997).

1.5.1 Overview of wild germplasm derived mapping populations

The first reported use of hybridising wild relative germplasm for the improvement of a desired trait was by Bohn and Tucker in 1940 (reviewed by Grandillo et al., 2011). The widespread interest of utilising wild relative germplasm for specifically improving health-promoting quality traits in tomato is demonstrated by reports in national media (for example http://www.dailymail.co.uk/sciencetech/article-2094331), popular science literature (Jones, 2001; Levin, 2008), and scientific review (Grandillo et al., 2011; Labate et al., 2007; Lippman et al., 2007; Zamir, 2001). Willits and colleagues (2005) demonstrated for the first time an approach to improving flavonoid profile in tomato by hybridising a *S. pennellii* accession with a domesticated

S. lycopersicum elite Syngenta cultivar. Resulting F_1 progeny were shown to accumulate quercetin mono- and diglycosides in fruit flesh and peel at higher levels than either parent (Willits et al., 2005). The development of F_2 and backcross (BC) populations for the identification of desired traits is commonly used in plant mapping (Grandillo et al., 2011; Zamir, 2001). Sometimes referred to as 'opening' the population, these allow greater genetic variation than F_1 hybrids. The resulting segregation is relatively easy to obtain, and less time consuming than more advanced breeding programs (Asins, 2002). However, F_2 and BC are limited because once a population has been 'opened' or segregated from the F_1, these genotypes cannot be repeatedly grown in several environments as can be done with populations using double haploidy (DH; although not with tomato, personal communication C. Baxter), and introgression line (IL; such as recombinant inbred lines, RILs) technologies (Asins, 2002).

The development of introgression mapping populations begins with initial hybrid crosses between elite tomato cultivars and wild relative species. This is followed by subsequent crossing strategies to produce populations in which each progeny genotype possesses an elite cultivar genome background with relatively small genomic regions substituted from the wild relative genome (Grandillo et al., 2011). The regions of wild relative genome substituted into the elite cultivar background differ between individual genotypes of the population in an attempt to achieve complete coverage of wild relative genome substitution by the collective population.

Molecular markers are utilised to distinguish wild relative from elite cultivar tomato genome sequences. Examples of molecular markers used for this purpose are restriction fragment length polymorphism (RFLP); conserved ortholog set II (COSII); simple sequence repeats (SSR); and amplified fragment length polymorphism (AFLP) (reviewed by Grandillo et al., 2011; Labate et al., 2007). These molecular markers can be used in marker assisted breeding. In the same way that Jones suggested linkage between tomato genes for dwarfness and fruit shape in 1917 (reviewed by Grandillo et al., 2011), molecular markers can be found today that are linked with traits of interest. As a result, allelic variation of these markers can be used to indirectly select for the trait.

By distinguishing genomic regions from either wild relative or elite cultivar origin, and by establishing differences in quantitative traits between all genotypes in a population, links can be made between genomic regions and quantitative traits. The

effects that these substituted regions have on the quantitative traits can be measured, and quantitative trait loci (QTL) can by identified (Grandillo et al., 2011; Labate et al., 2007). As mentioned in section 1.2.1, domesticated tomato, *S. lycopersicum*, is a member of the *Lycopersicon* section of *Solanum* genus. The *Lycopersicon* section contains 12 other wild relative members (detailed in section 1.2.3, Table 1) and is divided into four groups (Table 2) based on recent classification (Peralta et al., 2008).

Many of the species within *Solanum* sect. *Lycopersicon* and allied sect. *Lycopersicoides* and sect. *Juglandifolia* have been hybridised (Grandillo et al., 2011). Those initial crosses between wild relative accessions within *Solanum* sect. *Lycopersicon* species and elite *S. lycopersicum* cultivars have been used to produce mapping populations (such as introgression lines) and for subsequent QTL analysis.

Table 2 Classification of 13 species in *Solanum* sect. *Lycopersicon* into groups

Hierarchy of classification of four groups within section are defined by Peralta and colleagues (2008)

Genus	***Solanum***
Section	***Lycopersicon***
Group	**Lycopersicon**
Species	*S. lycopersicum*
	S. pimpinellifolium
	S. cheesmaniae
	S. galapagense
	Neolycopersicon
	S. pennellii
	Eriopersicon
	S. habrochaites
	S. huaylasense
	S. corneliomulleri
	S. peruvianum
	S. chilense
	Arcanum
	S. arcanum
	S. chmielewskii
	S. neorickii

1.5.2 Identification of QTL using mapping populations

Mapping populations (whereby breeding strategies have introgressed tomato wild relative genomes into elite tomato cultivars) that utilise molecular markers have been produced for most of the wild species within *Solanum* sect. *Lycopersicon* (Table 2). These have been reviewed by Grandillo and colleagues (2011). Populations using *S. pennellii* accessions remain one of the most heavily researched, especially when used

in the production of introgression line (IL) populations (Eshed et al., 1992; Eshed and Zamir, 1994; Grandillo et al., 2011). Genotypes from *S. pennellii* IL populations have been utilised to identify QTL for traits as diverse as fruit yield, aroma, colour, shape and size; leaf morphology; disease resistance; and tolerance to salt or drought (Grandillo et al., 2011). Eshed and colleagues (1996) and Eshed and Zamir (1995) used a *S. pennellii* IL population in the processing line background M82 to identify QTL for fruit yield. Three QTL for sugar yield from *S. pennellii* ILs were identified as *Brix9-2-5*, a fruit-specific *Brix9-2-5*, and a shoot-specific *PW9-2-5* (Fridman et al., 2002; Fridman et al., 2004).

Advanced backcross (AB) QTL analysis studies have been used to identify a total of 11 QTL contributing to fruit colour in the *S. habrochaites* and *S. peruvianum* AB populations (Bernacchi et al., 1998b; Fulton et al., 1997). Further to this, 222 QTL for 15 traits associated with fruit flavour, including fruit sugar levels, organic acid levels and pH, were identified from the four AB populations of *S. neorickii*, *S. habrochaites*, *S. peruvianum*, and *S. pimpinellifolium* (Fulton et al., 2002). Paran and colleagues (1997) used the *S. galapagense* recombinant inbred line (RIL) population for the identification of QTL for plant morphological traits such as height, mass, and leaf morphology. Kabelka and colleagues (2002) used the *S. habrochaites* backcross inbred line (BIL) population to map a QTL for resistance to the bacterial pathogen *Clavibacter michiganensis* subspecies *michiganensis* (*Cmm*).

QTL analyses using mapping populations have not been limited to traits for organoleptic and agroeconomic quality. A total of 59 QTL were identified for seven traits including fruit lycopene content using the *S. pimpinellifolium* backcross (BC) populations BC_1 and BC_1S_1 (Chen and Foolad, 1999). Rousseaux and colleagues (2005) conducted a QTL analysis on *S. pennellii* ILs identifying QTL for improved nutritional quality including antioxidant capacity, and contents of ascorbic acid and total phenolics in fruit. Three of these ILs (IL7-3, IL10-1, and IL12-4) were later characterised in plant tissue types for phenolic profile (Minutolo et al., 2012).

In this study accessions from the *S. neorickii* BIL population (Fulton et al., 2000; Grandillo et al., 2011) and one accession from the *S. habrochaites* near isogenic line (NIL) population (Monforte and Tanksley, 2000) are used.

1.5.3 S. neorickii BIL and S. habrochaites NIL populations

The *S. neorickii* (formerly called *L. parviflorum*) (Rick et al., 1976) mapping population was first developed as an AB population (Fulton et al., 2000). Relative to mapping populations with other *Solanum* sect. *Lycopersicon* species, *S. neorickii* has been exploited far less by geneticists (Grandillo et al., 2011). The parental lines of the *S. neorickii* AB population used for the original hybrid cross comprise *S. neorickii* wild relative accession LA2133 and the processing line domesticated cultivar *S. lycopersicum* TA209 (also known as E6203; Figure 6).

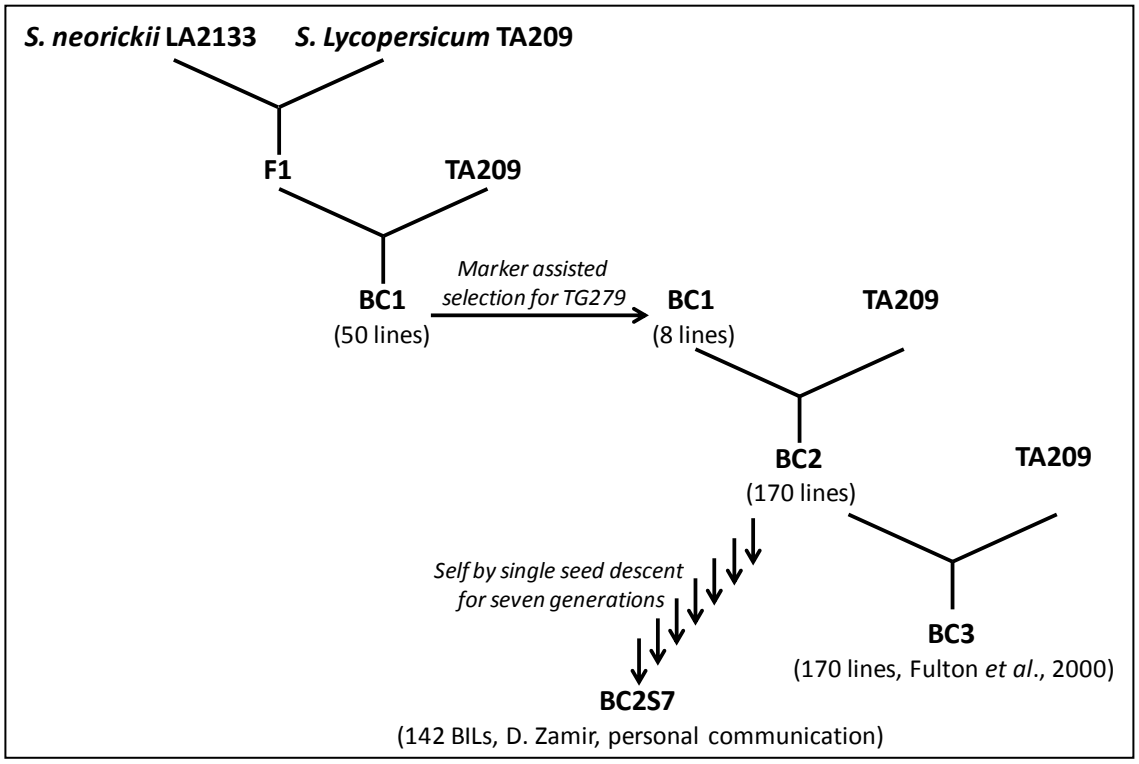

Figure 6 Schematic illustrating the production of *S. neorickii* BIL population

Process from initial cross between LA2133 and TA209 to the resulting BIL population is shown. BC, backcross; S, self.

A subsequent backcross of these produced 50 BC_1 plants (Fulton et al., 2000). Marker assisted selection for the RFLP marker TG279 located on chromosome 6 was used to select BC_1 lines homozygous for the *S. lycopersicum* allele, ensuring determinate growth was maintained by all subsequent lines (a trait desired for future field grown plants). Eight BC_1 lines were selected for further backcrossing with TA209 to produce a BC_2 population of 170 lines. These were backcrossed once more for the

final BC_3 population reported by Fulton and colleagues (2000) (Figure 6). The BC_2 AB population was used to generate the BIL population used in this study (Grandillo et al., 2011; D. Zamir, personal communication). Single seed descent on the BC_2 population for seven generations produced the BC_2S_7 BIL population of 142 BILs (Figure 6).

The *S. habrochaites* (formerly *L. hirsutum*) mapping population used in this study is a NIL population (Monforte and Tanksley, 2000). Similar to *S. neorickii* BIL population, the *S. habrochaites* NIL population was developed from an AB population derived from the hybridisation between *S. habrochaites* wild relative accession LA1777 and the same processing line domesticated cultivar *S. lycopersicum* TA209 (Bernacchi et al., 1998b). RFLP marker assisted selection was used to screen the AB population at the BC_2 and BC_3S_1 stages for the presence of *S. habrochaites* introgressions that were homozygous for expected QTL-containing regions expected to affect favourable traits (Bernacchi et al., 1998a). The resulting population comprised 35 NILs that formed part of an introgression library of 99 accessions (Bernacchi et al., 1998a; Bernacchi et al., 1998b).

The genetic variation present within the *S. neorickii* BIL population and *S. habrochaites* NIL population that is derived from the respective wild accession genomes was used in this project to fulfil objective 01 (section 1.1). This genetic variation was the cause of perturbations in phenolic profiles observed between genotypes within the *S. neorickii* BIL population and reported by T Wells (personal communication) with regards to *S. habrochaites* NIL population. As stated in section 1.1, these lines, together with metabolite and marker data available within the EU Sol consortium, were used to confirm a reproducible phenotype and identify possible QTL-containing introgressed regions (objectives 02 and 03). Once identified, selected lines from these populations were used to fully characterise these perturbations in phenolic profiles and to assess post-harvest characteristics and antioxidant properties (objectives 04, 05, and 06) with the potential for future application as a health-promoting trait. The in-depth study of these selected lines was continued by utilising available metabolomic and transcriptomic platforms (objectives 07 and 08) that would elucidate the molecular basis of the QTL with the possible application for optimising production of specific phenolic compounds in future germplasm.

2 Materials and methods

2.1 Materials

2.1.1 Chemical reagents

Unless otherwise stated, chemical reagents were purchased from Sigma-Aldrich Company Ltd. (Dorset, UK), and solvents were purchased from Fisher Scientific UK Ltd. (Leicestershire, UK). MS (mass spectrometry) grade water and molecular biology reagent grade water were purchased from Sigma-Aldrich Company Ltd. (Dorset, UK).

Methanol, acetonitrile, methyl *tert*-butyl ether (MTBE), hexane, and ethyl acetate were used at high performance liquid chromatography (HPLC) grade. Ethanol, chloroform and isopropanol were used at laboratory reagent grade. Formic acid was used at MS grade. Water was used, where indicated, at HPLC, MS and molecular biology reagent grade.

Chalcone naringenin and kaempferol standards were purchased from Apin Chemicals Ltd. (Oxon, UK). Naringenin, kaempferol-3-O-rutinoside, rutin, quercetin, isoferulic acid, and β-carotene standards were purchased from Extrasynthèse (Genay, France). Salicylic acid and all other phenolic standards were purchased from Sigma-Aldrich Company Ltd. (Dorset, UK). All other isoprenoid standards were purchased from CaroteNature GmbH (Lupsingen, Switzerland).

2.1.2 Biological material

Seed for *Solanum lycopersicum* (tomato) domesticated processing cultivars TA209 (alternatively called E6203) and M82 were from stocks kept at Royal Holloway, University of London (RHUL). Seed for the tomato wild relatives *Solanum neorickii* (Rick et al., 1976) accession number LA2133 and *Solanum habrochaites* (reviewed by Peralta et al., 2008) accession number LA1777 were obtained from the Tomato Genetics Resource Centre (TGRC), University of California, Davis. Seed for selected lines of the molecular linkage mapping population *S. neorickii* backcross inbred lines (BILs) (Fulton et al., 2000; Grandillo et al., 2011) were provided by D. Zamir at the Hebrew

University of Jerusalem, Israel (HUJI) and C. Baxter at Syngenta (Jealott's Hill International Research Centre, UK). Seed for line 3939 from *S. habrochaites* near isogenic line (NIL) molecular linkage mapping population (Monforte and Tanksley, 2000) were provided by G. Seymour at University of Nottingham, UK. Tomato fruit for field grown *S. neorickii* BILs and accessions from *S. lycopersicum* core collection (CC) colour mutant population 2009 (HUJI) were provided by D. Zamir (HUJI).

2.1.3 Primers

Primers, purchased from Eurofins MWG Operon (London, UK), are shown in Table 3.

Table 3 Primer sequences used

Gene	Primers	Reference
AJ224356	forward 5'-CGAAGCAAGCGTGAACAAAT-3' reverse 5'-TGCGGAGATTAGGATGGACA-3'	(Rohrmann et al., 2011)
TF1 promoter	forward 5'-GTTTTCAACAAGGGTTTGATGG-3' reverse 5'- GTATAACTGGAACTTGAAAGCC-3'	This work
TF2 promoter	forward 5'-TGCGATTAAAAGCCTTCTCC-3' reverse 5'-AGCTGACAAAAGCTCAACTG-3'	This work
TF3 promoter	forward 5'- GAAAACTCAGTTATTTATAGAAAGG-3' reverse 5'- GAAGATCCATAACAATTTCTTCC-3'	This work
TF4 promoter	forward 5'-CAAATGGACACATTGTTCACC-3' reverse 5'-GTTTCCGTACCTTCTTGCTAG-3'	This work
TF5 promoter	forward 5'- CTTGACCTAAGTGTAGAGTCC-3' reverse 5' GCAACGGACGTAGCAAAGAT-3'	This work

2.1.4 Chromedia

Analysis by high performance liquid chromatography with diode array detector (HPLC-DAD) used an Agilent 1100 Series (Agilent Technologies UK Ltd., Wokingham, UK) that comprised G1313A ALS autosampler, G1322A degasser, G1311A QuatPump, and G1315B DAD units. The stationary phase consisted of a 5 μm reverse phase C18 column (4.6 x 300 mm, 16 % carbon loading) and guard column (4.6 x 25 mm) (Hichrom Ltd., Reading, UK).

Ultra performance liquid chromatography with photodiode array detection (UPLC-PDA) was conducted with a Waters Acquity system (Waters Ltd.,

Hertfordshire, UK) comprising sample manager, binary solvent manager, and PDA eλ (extended wavelength) detector. Separation used an Acquity UPLC BEH C18 column (1.7 µm, 2.1 x 100 mm) stationary phase with a Waters VanGuard precolumn (2.1 mm x 5 mm) (Waters Ltd., Hertfordshire, UK).

For gas chromatography mass spectrometry (GC-MS) analyses, an Agilent 7890A gas chromatograph with an online 5975C mass spectrometer (Agilent Technologies UK Ltd., Wokingham, UK) was used. Separation was achieved using a DB-5MS with guard column (40 m x 250 µm x 0.25 µm column, including a 10 m guard column, and temperature limit of 350 °C) (J&W Scientific, Folsom, California, USA).

2.2 Methods

2.2.1 Experimental design

The experimental approach is shown in Figure 7.

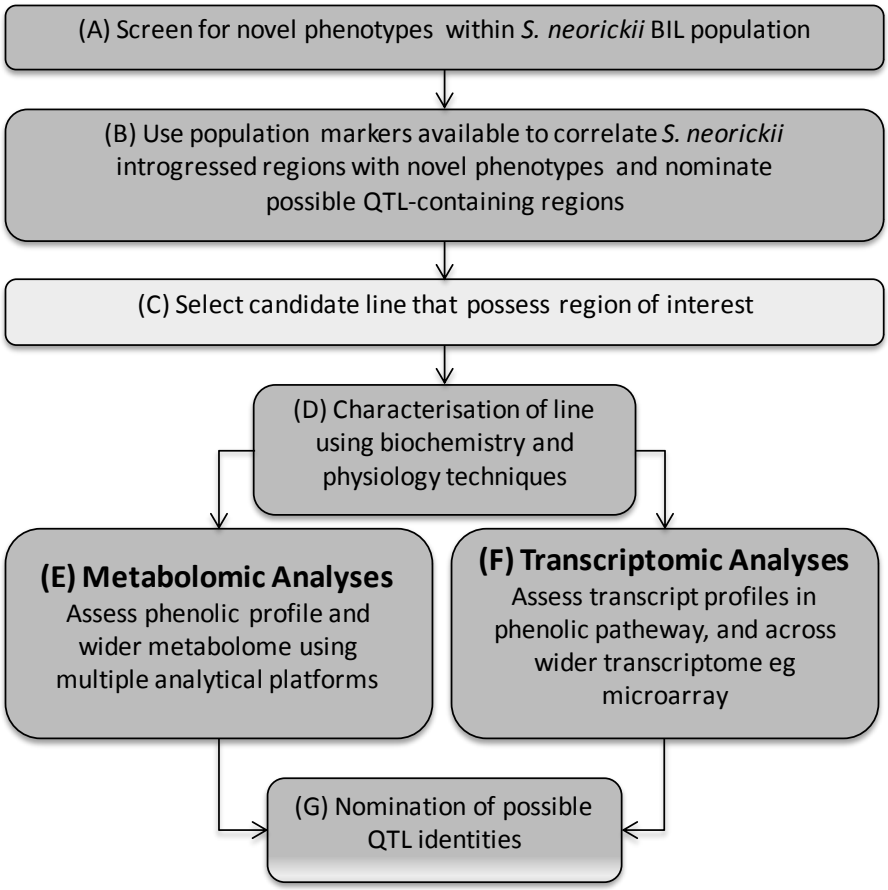

Figure 7 Experimental design

Schematic workflow representing intended approach to identify QTL. (A) crop 2 of *S. neorickii* BIL population is screened for phenolic profile by HPLC-DAD and novel phenotypes are identified. (B) available marker data provided by the EU Sol consortium is used to correlate *S. neorickii* introgressed regions with novel phenotypes of interest. (C) a candidate line possessing both the novel phenotype and the correlating introgressed region is selected for (D) characterisation of phenolic pathway throughout fruit development and assessment of fruit quality traits; (E) metabolomic and (F) transcriptomic analyses. (G) data is integrated to identify possible nominated QTL identity.

2.2.2 Cultivation and generation of tomato plants and sample material

Tomato field crops (Table 4) were grown in Akko, Israel, and provided by D. Zamir (HUJI). Routinely, four plants per line were sampled from randomised plots, and

three fruit from each plant were harvested and pooled. Harvested fruits were transported to UK intact, where they were halved, frozen immediately with liquid nitrogen, and stored at -80 °C. Samples were freeze dried commercially (European Freeze Dry Ltd, Preston, UK) for 72 hours. Once freeze dried, tissue was stored at -20 °C. Tissue was homogenised until a fine powder using a TissueLyser (Qiagen, Crawley, UK) for 5 min at a frequency of 30 times s^{-1}, and stored at -20 °C until analysis by HPLC or UPLC.

S. neorickii crop 2 was cultivated two years following crop 1. Where possible, seed was collected from crop 1 in 2006 and used the following year (2007) for an intermediate crop (not available for this study) cultivated chronologically between crops 1 and 2. Where seed was not available plants were grown from seeds stocks used to cultivate crop 1. This same process was repeated between the intermediate crop in 2007 and those cultivated for crop 2 (in 2008). As a result, the majority of BILs in crop 2 are an additional two selfed generations from that of crop 1. It is likely that some BILs, however, are replicated plants grown in different seasons. To the best of available knowledge, there is no way to determine for which BILs this is the case (personal communication, D. Zamir).

Table 4 Summary of fruit crops grown at Hebrew University of Jerusalem, Israel

Crop names are in reference to this work only. Dates of harvest are approximate.

Crop name	Date of harvest	Analysis
S. neorickii BIL crop 1	July 2006	E. Enfissi (unpublished)
S. neorickii BIL crop 2	July 2008	This work
S. neorickii BIL crop 3 [a]	July 2009	None
S. neorickii BIL backcross lines[a]	July 2009	None
Core collection colour mutant population	July 2009	This work

[a] indicates that only specific lines were selected for analysis from these crops

Tomato plants grown at RHUL were cultivated under glasshouse conditions with supplementary lighting of 110 µmol m^{-2} s^{-1} (H. Berry, personal communication) provided by 400 W Son-T high pressure sodium bulbs (Osram Ltd., Berkshire, UK). A light/dark cycle of 16/8 h was maintained with temperatures controlled at 20-25 °C

during light and 15-18 °C during dark. Plants were cultivated in Levington M3 high nutrient pot and bedding compost (The Scotts Company LLC, Ohio, USA). Typically five plants per line were grown in randomised block positions.

Leaf material was harvested from three separate leaves and pooled for each plant. Five plants were sampled, and each plant represented a biological replicate. Flowers were harvested from five plants and pooled in order to obtain adequate material for analysis. A minimum of 20 flowers were sampled per genotype. Leaf and flower tissues were frozen in liquid nitrogen, and then stored at -80 °C.

Figure 8 Fruit development stages selected for analysis

Typical TA209 fruit morphology for four developmental stages defined here (from left to right) mature green (MG), breaker (Br), turning (T), and ripe (R). Images not shown to scale.

Fruits were harvested at four stages to represent fruit development, and are shown in Figure 8. In order of ripening these were, mature green (MG) stage, where fruits were full size, but remained green in colour; breaker (Br) stage, where colour other than green (usually yellow or orange) was first detected; turning (T) stage, where no more MG green colour was visible, fruits were normally orange but not necessarily uniform in colour, and fruits had not yet developed their final ripe colour; and ripe (R) stage, where fruits were uniformly their final colour (red for parental genotype line TA209) and began to lose firmness. To represent these developmental stages, three fruit were harvested and pooled from each plant. Three to five plants were sampled for developmental stages, representing three to five biological replicates. Fruit samples were halved or diced depending on their intended use, and then frozen and stored as described above. Ten stages were harvested to represent a ripening series over periods of time following anthesis and breaker stage. Anthesis was defined as flowers that were fully open with all petals at or surpassing 90° from the floral axis (Figure 9). Breaker stage was defined as above. This series comprised 14 dpa (days post anthesis), 21 dpa,

28 dpa, 35 dpa, 0 dpb (days post breaker, therefore breaker stage), 3 dpb, 6 dpb, 9 dpb, 12 dpb, and 16 dpb. Due to a limited number of available fruit for each plant, three single fruits were harvested for each stage from three of the five plants, and each single fruit represented biological replication. Fruit samples were halved, and then frozen and stored as described above. Additional ripe fruit were harvested to assess different compositions of tissue types. Ripe fruit were defined as above. Three fruit were harvested and pooled for each plant. Three of the five plants were sampled, which represented three biological replicates. These fruit were blanched in water at 90-100 °C for 5 s to facilitate removal of skin. Fruits were segmented and jelly samples were removed, which comprised juice, placenta and seeds. The remaining fruit flesh represented pericarp samples. Tissue samples were frozen and stored as described above.

Material intended for metabolite profiling was freeze dried using a Lyovac GT2 freeze dryer lyophilzer (Leybold-Heraeus GmbH, Hanau, Germany) for 72 hours and homogenised into a fine powder using a TissueLyser as described above, then stored at -20 °C until analysis. Material intended for RNA or DNA extraction was stored at -80 °C, in diced form, and ground under liquid nitrogen immediately before use. Grinding was achieved by pestle and mortar, followed by homogenisation using a 6750 Freezer/Mill (Spex CertiPrep Ltd., New Jersey, USA) for one cycle of 1.0 min and at a rate of 10 repeats s^{-1}, and stored at -80 °C.

Figure 9 Criteria used to determine anthesis

Images illustrate flowers at anthesis, as defined here. (A) single flower of *S. neorickii* LA2133 at anthesis, present with senescing flowers on truss. (B) three flowers of *S. neorickii* BIL neo-111 at anthesis. Hand-tied labels are shown that were used to record the date of anthesis and therefore determine dpa. Images are not shown to scale.

At all stages of preparation, samples were processed in batches that contained background parental line control samples (TA209 in all cases except core collection colour mutants, which had no equivalent control) and quality control samples (M82 in all cases).

Seeds were collected to maintain stocks and to be re-grown for future analysis of fruit from subsequent generations. Fruit, separate to those used for analysis, were harvested for the sole purpose of seed collection. Fruit placenta, or jelly, containing seeds were removed and bathed in 16 % (v/v) hydrochloric acid for 20 min, before being rinsed in water to remove jelly, placenta and residual fruit pericarp. Seeds were allowed to dry on absorbent paper, and then stored at room temperature.

2.2.3 Targeted analytical procedures for the determination of plant material composition

2.2.3.1 Extraction and analysis of phenolic compounds by HPLC-DAD

Freeze-dried tomato material, homogenised into a fine powder (20 mg) was weighed directly into screw capped Pyrex tubes in triplicate, to create three technical replicates. Phenolic compounds were extracted as described previously (Davuluri et al., 2005). To each sample, methanol (1 ml) was added, containing 20 μg ml^{-1} salicylic acid internal standard. Samples were incubated for 1 h at 80 °C, before cooling on ice for 20 min. Centrifugation at 2,500 *g* for 5 min pelleted the debris, and the methanol supernatants were removed with a pipette. Particulates were removed by passing extracts from a syringe (1 ml) through 0.2 μm nylon filters (Chromacol Ltd., Hertfordshire, UK). Samples were either stored at 5 °C and analysed within one day of extraction or frozen at -20 °C for a maximum of four days before analysis.

Compounds were quantified from methanol extracts following the procedure of (Melendez-Martinez et al., 2010) with slight modification. Extracts (1 ml) were placed into sample vials (1.5 ml; Agilent Technologies UK Ltd., Wokingham, UK), and analysed using HPLC-DAD (described in Materials section 2.1.4). Sample volumes of 25 μl were injected and a mobile phase flow rate of 1 ml min^{-1} was used. The mobile phase consisted of (A) water (HPLC grade) containing 2 % (v/v) methanol and 0.1 % (v/v) 6 M hydrochloric acid, which when combined was 0.2 μm filtered, and (B)

acetonitrile. A linear gradient of 5 to 40 % solvent B for 20 min followed by 40 to 60 % solvent B for 15 min was used. A conditioning phase of 5 min returned the column to starting conditions. Elution from the column was monitored continuously by the on-line unit, which scanned from 210 to 500 nm. Chromatographic components were integrated using Chemstation software version A.10.02 (Agilent Technologies UK Ltd., Wokingham, UK). The phenolic components were identified using an in-house library verified by authenticated standards, and quantification was carried out by comparison to the internal standard (salicylic acid).

2.2.3.2 Extraction and analysis of phenolic compounds by UPLC-PDA

Phenolic compounds were extracted as described in section 2.2.3.1. From 1 ml methanol extraction containing phenolic compounds, 200 µl was taken to dryness by rotary evaporation (EZ-2plus personal evaporator, GeneVac Ltd., Ipswich, UK; low boiling point setting and lamp off) then resuspended in methanol (50 µl). Samples were placed into sample vials containing vial inserts (1.5 ml and 250 µl respective volumes; Agilent Technologies UK Ltd., Wokingham, UK). A method for the separation of phenolic compounds using UPLC-PDA (Materials section 2.1.4) was developed as part of this work, and is discussed in section 4.2. Sample Manager was maintained at 8 °C, and column temperature was 40 °C. The partial loop mode was used to inject 5 µl of sample, and separation was achieved using 0.4 ml min^{-1} flow rate. Mobile phase was (A) water (MS grade) comprising 2 % (v/v) methanol and 0.1 % (v/v) formic acid and (B) acetonitrile. A gradient for 9 min was conducted as follows: 10 % solvent B at initial time and held isocratic until 0.5 min, linear gradient 10 to 30 % B until 5.0 min, isocratic until 6.0 min, 30 to 100 % B using Waters gradient curve shape 10 until 6.5 min, isocratic until 7.0 min, a linear gradient to return to initial conditions of 10 % B until 7.5 min, and an isocratic column reconditioning stage until 9.0 min. Data were collected using on-line PDA detection over a wavelength range 210 to 600 nm with resolution 1.2 nm and sampling rate 20 points s^{-1}. Peaks were integrated using Empower 2 software (Waters Ltd., Hertfordshire, UK), identified by spectral profiles, verified by authenticated standards, and quantified using a salicylic acid internal standard.

2.2.3.3 Extraction and analysis of isoprenoid compounds by UPLC-PDA

Freeze-dried, homogenous fine powdered tomato (10 mg) was weighed into microcentrifuge tubes in triplicate to represent three technical replicates. Sequentially, methanol (250 µl) and chloroform (500 µl) were added and vortexed. Samples remained on ice, in darkness, for 20 min. To the samples, 100 mM Tris-HCl buffer (250 µl, pH7.5) in water (HPLC grade) was added and mixed by vortex. Centrifugation at 13,500 *g* for 5 min separated the non-polar chloroform hypophase from the polar aqueous epiphase. The hypophase containing isoprenoid extract was transferred to a new microcentrifuge tube. Additional chloroform (500 µl) was added to the remaining aqueous phase, and a second extraction by vortex and centrifugation was conducted as described above. Both chloroform extracts were combined and taken to dryness using a rotary evaporator (as stated above), and stored at -20 °C until analysis.

Where stated section 4.3.2, an alternative extraction with MTBE was performed in place of chloroform. This procedure was identical to that described above, except equal volumes of MTBE were used in place of chloroform and the resulting extracts were therefore collected in the epiphase when centrifuged (see section 4.3.2, Figure 21).

A method for the separation of isoprenoids was provided by P. Fraser (personal communication) and the method assessed and validated as described in section 4.3 and section 4.4. Dry isoprenoid extract was resuspended in ethyl acetate (70 µl). The Waters Acquity UPLC system was used, as detailed in section 2.1.4. Samples were stored at 8 °C until injection, and column temperature was 30 °C. A partial loop mode was used to inject 5 µl sample. Mobile phase solvents were (A) 1:1 (v/v) methanol/water (HPLC grade), and (B) 3:1 (v/v) acetonitrile/ethyl acetate. A gradient for 8 min and flow rate of 0.6 ml min^{-1} was conducted as follows: 70 % solvent B at initial time and held isocratic until 0.5 min, linear gradient 70 to 99.9 % B until 5.0 min, isocratic until 6.0 min, a waters gradient curve shape 9 to recondition the column to 70 % B until 8.0 min. Data were collected over a wavelength range 250 to 600 nm, with resolution 1.2 nm and sampling rate 20 points s^{-1}. Peaks were integrated using Empower 2 software (Waters Ltd., Hertfordshire, UK), identified and quantified respectively by spectral profiles and calibration curves, each verified by co-chromatography and identical spectral properties with authenticated standards.

2.2.4 Non-targeted analytical procedures for the determination of plant material composition

2.2.4.1 Extraction and analysis of polar compounds by GC-MS

Polar compounds were extracted and analysed by GC-MS as described by (Enfissi et al., 2010) with some modification. Homogenised freeze dried powder (10 mg) was weighed into microcentrifuge tubes, using three to five biological replicates. Extraction buffer (1 ml) was added, comprising 0.04 mg ml^{-1} ribitol internal standard in 4:1 (v/v) methanol/acidified water (0.21 M HCl; HPLC grade). Samples were vortexed then agitated by repeatedly inverting for 1 h. Samples were then pelleted for 5 min by centrifugation at 20,000 *g* and an aliquot (20 µl) was dried by rotary evaporator, as described previously. Dried samples were derivatised into methoxymated and silylated forms, as described by (Halket et al., 2005), by first resuspending by vortex in pyridine (30 µl) containing 20 mg ml^{-1} methoxylamine hydrochloride and incubating at 40 °C for 1 h, and second by adding *N*-methy-*N*-(trimethylsily) trifluoroacetamide (MSTFA, 80 µl) and re-incubating at 40 °C for 1 h. Derivatised samples were left at room temperature for up to 24 h before being subjected to analysis by GC-MS using chromedia described in section 2.1.4. A volume of 1 µl was injected in splitless mode, and a second injection (1 µl) was subsequently used for a split mode (1 in 10) for quantification of high abundant sugars in ripe and turning fruit samples. A gradient for 67.5 min comprised 5 min at initial temperature of 70 °C, followed by a linear increase to 320 °C at a rate of 4 °C min^{-1}. The oven equilibration time between samples was 0.5 min. Helium was employed as the carrier gas and the flow rate was set at 0.5 ml min^{-1}. The MS inlet source was set to 280 °C and the mass selective detectors (MSD) transfer line was set to 250 °C. Following a 9 min solvent delay, the MS performed in full scan acquisition mode from 15 to 800 D. Data were processed using the Automated Mass Spectral Deconvolution and Identification System (AMDIS) software version 2.69. Chromatographic components were identified using the automated deconvolution function, and a mixture of known *n*-alkane calibration standards ranging from C_8 to C_{32} was used for retention index external calibration. Identities were assigned to components using the National Institute of Standards and Technology (NIST) version 2.0 library database, and therein an in-house MS library was constructed. Where available, entries were verified by authenticated standards. This purpose-built MS library was used to identify components in all remaining samples. AMDIS was then

used to determine peak areas. From these areas compounds were quantified relative to the ribitol internal standard. For compounds with multiple chromatographic peaks as a consequence of the derivatisation steps (such as aspartic acid) the sum areas of these peaks were used for quantification. Significantly abundant components with no known identity were assigned 'unknown' identities by a 'UNKp' prefix.

2.2.4.2 Extraction and analysis of non-polar compounds by GC-MS

Non-polar compounds were extracted as described by Jones and colleagues (in preparation) with some modification. Homogenised freeze dried powder (20 mg) was weighed into microcentrifuge tubes using three to five biological replicates, and to this 4:1 (v/v) methanol/water (250 µl, HPLC grade) was added. Samples were spiked with hexane (5 µl) containing 0.5 mg ml^{-1} nonadecanoic acid as an internal standard, and vortexed. Chloroform (500 µl) was added, vortexed, and incubated on ice, in the dark, for 20 min. Phases were separated by the addition of water (250 µl, HPLC grade), vortexing, and centrifugation at 20,000 *g* for 5 min. The non-polar chloroform hypophase was removed to a new glass vial, leaving the polar aqueous epiphase. A second chloroform extraction was conducted on the remaining epiphase as before and combined with the first. The resulting chloroform extract (1 ml) was taken to dryness by rotary evaporator, as described above. Samples were saponified by the addition of 4:1 (v/v) methanol/water (250 µl, HPLC grade) containing 6 % (w/v) potassium hydroxide and incubation for 1 h at 55 °C. Sequentially, chloroform (500 µl) and water (250 µl, HPLC grade) were added, vortexed, and centrifuged at 20,000 *g* for 5 min. The non-polar chloroform hypophase was removed and taken to dryness by rotary evaporator. Extracts were resuspended in chloroform (100 µl), from which an aliquot (20 µl) was dried by rotary evaporator. As described in section 2.2.4.1 samples were derivatised, and analysed by GC-MS using the splitless mode only. Similarly, a library was constructed of non-polar components using NIST identity matches, and significantly abundant components of 'unknown' identities were assigned 'UNKnp' prefixes. Quantification was conducted as described by section 2.2.4.1.

2.2.5 Analysis of antioxidant capacity

Antioxidant capacity was determined using the Trolox equivalent antioxidant capacity (TEAC) assay as described by (Re et al., 1999). In brief, a freshly prepared

stock solution of 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) radical cations ($ABTS^{\cdot+}$) was prepared by diluting ABTS diammonium salt and potassium persulphate in water to respective final concentrations of 7 mM and 2.45 mM. The stock solution was allowed to react for 16 h in darkness, at room temperature. $ABTS^{\cdot+}$ working solution was formed by diluting $ABTS^{\cdot+}$ stock solution in ethanol to a stable absorbance reading at 734 nm of 0.70 ± 0.02, using a bench top spectrophotometer (Novaspec Plus visible spectrophotometer, Amersham Biosciences, Buckinghamshire, UK). Phenolic extract in methanol (10 µl) was added to a cuvette containing 1 ml $ABTS^{\cdot+}$ working solution, and the decrease in absorbance after 6 min was recorded. Values were compared to an equivalent standard curve, whereby 10 µl Trolox (6-hydroxy-2,5,7,8-tetramethychroman-2-carboxylic acid) at working concentrations of 0.25 to 1.75 mM (final concentrations 2.5 to 17.5 µM) were likewise added to $ABTS^{\cdot+}$ working solution causing an absorbance decrease ranging 20-80 %. Therefore, for each sample assayed, Trolox equivalent decreases in absorbance were calculated from the standard curve gradient. Assays were replicated in triplicate and antioxidant activity was shown to be dose responsive. Trolox standard curve was repeated daily with fresh reagents. Spectrophotometer was blanked with ethanol. Control assays using 10 µl methanol, ethanol, and 20 µg ml^{-1} salicylic acid internal standard in methanol exhibited no significant decrease in absorbance. Additionally, there were no significant differences between phenolic extracts replicated with and without the presence of salicylic acid internal standard. It was concluded firstly, that neither solvent nor internal standard contributes to antioxidant activity and secondly, that any possible interaction between phenolic extract constituents and internal standard does not affect antioxidant activity.

2.2.6 Assessment of tomato plant and fruit parameters

2.2.6.1 Fruit firmness

Firmness was assessed using an Analog HP-FFF mechanical fruit firmness tester durometer, with an exchangeable 0.25 cm^2 test anvil (Qualitest International Inc., New York, USA), using the American Society for Testing and Materials (ASTM) International D2240 (standard for hardness) type A scale ranging 0 to 100. Single

measurements were taken between the fruit shoulder and equator of three separate fruit and mean averages were calculated.

2.2.6.2 Fruit mass and yield, seed mass and number

Average fruit masses were calculated from six individual whole fruit measurements (± 0.01 g), taken directly after harvesting. Fruit yields were estimated by observation. Average seed masses and yields were calculated from six fruit per plant and a minimum of three (up to five) plants per genotype. For each fruit, all seeds were removed and bathed in 16 % (v/v) hydrochloric acid for 20 min, before being rinsed in water to remove jelly, placenta and residual fruit pericarp. Once dried, seed mass for each fruit was estimated from the combined mass of ten random seeds per fruit (± 0.1 mg). Total seed number was estimated using this estimated seed mass per fruit and the total mass of all seeds per fruit.

2.2.6.3 Water content of fruit

Water content was calculated based on the loss of mass of whole fruits before and after freeze drying. The difference in mass was represented as a percentage of initial mass.

2.2.6.4 Post-harvest properties

For each genotype, six fruit were harvested at 6 dpb and their masses recorded. Whole fruit were stored by resting on blotting paper at least 8 cm apart for 10 days post harvest (dph) at 18.0 °C (± 1.0 °C) in a well ventilated room and away from direct sunlight. The mass of each fruit was recorded at 2 day intervals. In addition, at 10 dph, fruit firmness was recorded (see section 2.2.6.1), water content was estimated (see section 2.2.6.3), and fruits were frozen and homogenised (see section 2.2.2) for targeted metabolite analysis (see section 2.2.3.2).

2.2.7 Analyses of gene expression

2.2.7.1 Extraction of DNA and RNA using kits

In accordance with manufacturer's instructions, RNA and DNA were extracted using RNeasy Plant Mini Kit and DNeasy Plant Mini Kit (Qiagen, Crawley, UK), respectively.

2.2.7.2 Extraction of RNA and digestion of DNA for the purpose of transcription factor platform and microarray analyses

For each genotype, three biological replicates were prepared, each replicate comprising three pooled fruit from a single plant. Fruit at MG, Br and T developmental stages were deseeded and then homogenised under liquid nitrogen, as described in Section 2.2.2.

RNA extraction was conducted as described by (Bugos et al., 1995), with modifications detailed here. Extraction buffer was prepared in water (molecular biology reagent grade), comprising 10 % (v/v) Tris-NaOH (pH 9.0, 1 M), 4 % (v/v) sodium chloride (5 M), 3 % (v/v) EDTA (pH 8.0, 0.5 M), and 0.5 % (w/v) *N*-lauroyl sarcosine sodium salt. 2-Mercaptoethanol was added to the extraction buffer immediately prior to use, to achieve a concentration of 8 μl ml^{-1}. Extraction buffer (4 ml), containing 2-mercaptoethanol, was measured into 50 ml centrifuge tubes and frozen homogenised fresh tissue (-80 °C; approximately 1.0 g) was added. Sequentially, phenol (4 ml), 4 % (v/v) isoamyl-alcohol (IAA) in chloroform (1.6 ml), and 3 M sodium acetate (1.3 ml) was added and mixed well. Samples were incubated on ice for 15 min. Centrifugation at 2,000 *g* for 10 min at 4 °C separated the phases. The aqueous hypophase was removed to a new centrifuge tube, and to this one volume of phenol containing 2% (v/v) IAA and 48 % (v/v) chloroform was added. Samples were centrifuged and the aqueous hypophase was transferred to a new centrifuge tube as before. One volume of isopropanol was added, and samples were incubated at -80 °C for 20 min. RNA was pelleted by centrifugation at 9,500 *g* for 10 min at 4 °C, and the pellet was washed in 80 % (v/v) chilled ethanol in molecular biology reagent grade water (1 ml). Pellets were centrifuged as before, and pellets were dried by leaving the centrifuge tubes inverted for approximately 10 min. RNA pellets were resuspended in molecular biology reagent grade water (1 ml) and transferred to microcentrifuge tubes. Insoluble material was

pelleted by centrifugation at 2,300 *g* for 3 min at 4 °C, and samples were transferred to new microcentrifuge tubes. Chilled 8 M lithium chloride (0.5 ml, -20 °C) was added, and RNA precipitate was allowed to form during overnight incubation at 4 °C. RNA was pelleted by centrifugation at 4,000 *g* for 10 min at 4 °C and the pellet was washed in 80 % (v/v) chilled ethanol in molecular biology reagent grade water (1 ml). RNA was pelleted again by centrifugation at 9,500 *g* for 10 min at 4 °C. Pellets were left to dry for approximately 1 h by inverting microcentrifuge tubes. RNA was resuspended in molecular biology reagent grade water (50 µl) and its concentration was estimated with a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA). RNA extracts were diluted to approximately 300 ng μl^{-1}.

RNA extracts (50 µl) intended for transcription factor (TF) platform analysis were digested with DNase I using Ambion TURBO DNA-free kit (Applied Biosystems, Life Technologies Ltd., Paisley, UK) in accordance with manufacturer's instructions. Samples intended for microarray analysis were subjected to DNase I digestion and clean-up with RNeasy Plant Mini Kit (Qiagen, Crawley, UK), according to the manufacturer's instructions. Irrespective of the DNA digestion method used, all samples were assessed for the presence of contaminating genomic DNA by amplification of AJ224356 intron sequence (Table 3), using illustra PuReTaq Ready-To-Go PCR beads (GE Healthcare Life Sciences, Buckinghamshire, UK), and visualisation by gel electrophoresis with ethidium bromide on 2 % (w/v) agarose gels. Final RNA concentrations, following DNA digestion, were assessed with a Nanodrop ND-1000 spectrophotometer. RNA integrity was assessed by gel electrophoresis with ethidium bromide on 1 % (w/v) agarose gels using 5 µl of 30 ng μl^{-1} RNA extract.

2.2.7.3 Synthesis of cDNA for analysis by transcription factor platform

RNA (2 µg, see section 2.2.7.2) was used to synthesise cDNA using SuperScript III Reverse Transcriptase kit (Invitrogen, Life Technologies Ltd., Paisley, UK) and Oligo (dT) 15 Primer (Promega, Wisconsin, USA), in accordance with the manufacturer's instructions.

2.2.7.4 Transcription factor platform analysis

Synthesised cDNA (see section 2.2.7.3) was transported to Max-Planck-Institut für Molekulare Pflanzenphysiologie (MPI-MP, Potsdam-Golm, Germany) on dry ice.

TF platform assays were prepared by J. Rohrmann (MPI-MP), as detailed by (Caldana et al., 2007), and stored at 4 °C in 384-well plates. qRT-PCR reactions were conducted, as described by (Rohrmann et al., 2011), using an ABI PRISM 7900 HT sequence detection system (Applied Biosystems, Life Technologies GmbH, Darmstadt, Germany) for three biological replicates per genotype, using primer sets for 1077 TF genes. Data was analysed using SDS software version 2.3 (Applied Biosystems, Life Technologies GmbH, Darmstadt, Germany). For each individual reaction, quality assurance comprised manual verification of melt curve shape, and by this method samples were excluded from further analysis where non-specific amplification was detected by irregular melt curve shapes. Additionally, each PCR reaction was manually verified using LinRegPCR version 11.0 (Ruijter et al., 2009) to identify the exponential phase, and this was corrected for samples that had not been identified by automated functions. Data were returned to J. Rohrmann (MPI-MP) for normalising and processing using R programming language, as described by (Rohrmann et al., 2011). Data were retrieved as 40-ΔC_t values. These were converted to expressions relative to TA209 and t-test values were calculated to identify TFs that had significantly ($p \leq 0.05$) altered expression.

2.2.7.5 Preparation of sample for microarray analysis

An RNA slurry was used to prevent degradation of RNA during transportation to USA for microarray analysis. To each RNA extract (40 µl) containing 30 to 60 µg RNA (see section 2.2.7.2), molecular biology reagent grade water (244 µl), 5 M ammonium acetate (16 µl), and ethanol (700 µl) were added sequentially. Samples were transported on dry ice to E. Bondo and S. McDonald (Syngenta Biotechnology Inc., North Carolina, USA), for analysis.

2.2.7.6 Cloning and sequencing of transcription factor promoter regions

Promoter regions were defined as approximately 1 kb flanking upstream of TF gene and primers were designed to amplify these regions (see Table 3) from the International Tomato Annotation Group (ITAG) Release 2.31 genome sequence, available at Sol Genomics Network (SGN) (Bombarely et al., 2011). Sequences were amplified from *S. lycopersicum* TA209, *S. neorickii*, and *S. habrochaites* genomic DNA using Novagen KOD Hot Start DNA Polymerase kit (Merck Biosciences Ltd., Nottingham, UK), according to the manufacturer's instructions. PCR products were verified as single products of correct band size by gel electrophoresis using GelRed on 1

% (w/v) agarose gels. DNA was purified using Promega Wizard SV Gel and PCR Clean-Up System (Promega UK, Sounthampton, UK), following the manufacturer's instructions for DNA purification by centrifugation. PCR products were cloned using Zero Blunt PCR Cloning Kit (Invitrogen, Life Technologies Ltd., Paisley, UK) according to manufacturer's instructions for ligation into pCR-Blunt and transformation into One Shot TOP10 *Escherichia coli* (Invitrogen, Life Technologies Ltd., Paisley, UK). Colonies were grown on LB-agar at 37 °C for approximately 16 h with kanamycin (50 μg ml^{-1}). Positive colonies were verified by colony PCR (illustra PuReTaq Ready-To-Go PCR beads, GE Healthcare Life Sciences, Buckinghamshire, UK), and cultured in liquid LB medium containing kanamycin (50 μg ml^{-1}). DNA was extracted using Promega Wizard Plus SV Minipreps DNA Purification System (Promega UK, Sounthampton, UK), according to manufacturer's instructions. Sequences were obtained from forward and reverse reads of products from duplicate colonies using Eurofins MWG Operon (Ebersberg, Germany).

3 Screening of natural variation for novel phenolic profiles

3.1 Introduction

In order to identify possible QTL for desirable traits, large segregating populations must first be screened for intra-population natural variation. This has been shown many times, for example Fulton and colleagues (2002) identified QTL for mapping populations of *S. neorickii*, *S. habrochaites*, *S. peruvianum*, and *S. pimpinellifolium*. The sizes of mapping populations used for screening can vary widely, and larger populations with a greater number of recombination events can provide more precision for QTL identification. This range in population size can be illustrated by *S. galapagense* (UC204B x LA0483) F_2 population (Paterson et al., 1991) comprising 350 genotypes, and *S. habrochaites* (Moneymaker x LYC4) IL population (Finkers et al., 2007a) comprising only 30 genotypes. The time and resources required for routine analytical procedures applied to the screening of larger-scale populations, although necessary, can therefore provide a bottleneck in experimental design.

In this chapter, work is described on the genotypes from *S. neorickii* BIL population (Fulton et al., 2000; Grandillo et al., 2011), which were screened by targeted HPLC-DAD analysis for levels of phenolic compounds in ripe fruit. Consistency of phenotype was assessed by comparison with previous unpublished data (E. Enfissi). These results, together with the available marker data, were used to select potential QTL-containing genome regions that suggest a possible association with, and may therefore affect, the resulting changes in phenolic profiles observed. Candidate BILs were chosen to represent these regions; a genotype from the *S. habrochaites* NIL population (Monforte and Tanksley, 2000) was additionally selected for comparison; and the phenotypes of three lines were confirmed when grown in the UK under glasshouse conditions.

3.2 Screening of *S. neorickii* BIL population for novel phenolic profiles

3.2.1 Analysis of S. neorickii BIL population phenolic profiles

Genotypes from the *S. neorickii* BIL population comprising crop 2 were grown in 2008 in the field at HUJI and supplied by D. Zamir (see section 2.2.2, Table 4). Samples were analysed by HPLC-DAD for a phenolic profile consisting of rutin, *p*-coumaric acid, chlorogenic acid, chalcone-naringenin and naringenin levels (Figure 10). Results were compared with data from crop 1 (provided by E. Enfissi, RHUL, a replication of crop 2 that was grown and analysed previously) for consistency of phenotype, defined in terms of direction of change (increase or decrease in levels relative to background TA209), of reproducibility of levels ≥ 2 fold, and of trend across the whole population. All phenolic compounds assessed showed a greater abundance in crop 1 than crop 2. This was shown by quantitative values of TA209, and was most evident for chalcone-naringenin and chlorogenic acid levels.

Relative rutin levels ranged from 23.38 ± 1.25 fold to levels below detection in crop 2, and from 90.52 ± 0.43 to 0.43 ± 0.01 fold in crop 1 (Figure 10A). The trend in relative rutin levels between both crops showed little consistency. Many BILs exhibited a relative decrease in crop 2, such as neo-88 with 0.69 ± 0.02 fold and neo-101 with 0.34 ± 0.01 fold, but showed a large relative increase in crop 1 (55.59 ± 0.48 and 3.45 ± 0.06 fold, respectively). However, the reverse (an increase in crop 2 and a decrease in crop 1) was less common and exhibited less extreme fold changes. Only three examples (neo-12, -32 and -42) displayed this trend, with increases in crop 2 that did not exceed 1.5 fold (1.03 ± 0.65, 1.18 ± 0.14, and 1.31 ± 0.87 fold, respectively) and decreases in crop 1 that did not exceed 2 fold (0.97 ± 0.01, 0.91 ± 0.02, and 0.69 ± 0.02 fold, respectively). Comparisons between crops 1 and 2 displayed some consistency with BILs exhibiting greater fold increases. 15 BILs showed an increase of ≥ 2 fold in crop 2. 11 of these BILs also displayed a ≥ 2 fold increase in crop 1, two BILs showed increases of 1.97 ± 0.01 fold (neo-53) and 1.44 ± 0.03 (neo-77), and two BILs had no available sample for crop 1. None of the BILs possessing ≥ 2 fold increase in rutin for crop 2 displayed a decrease in crop 1.

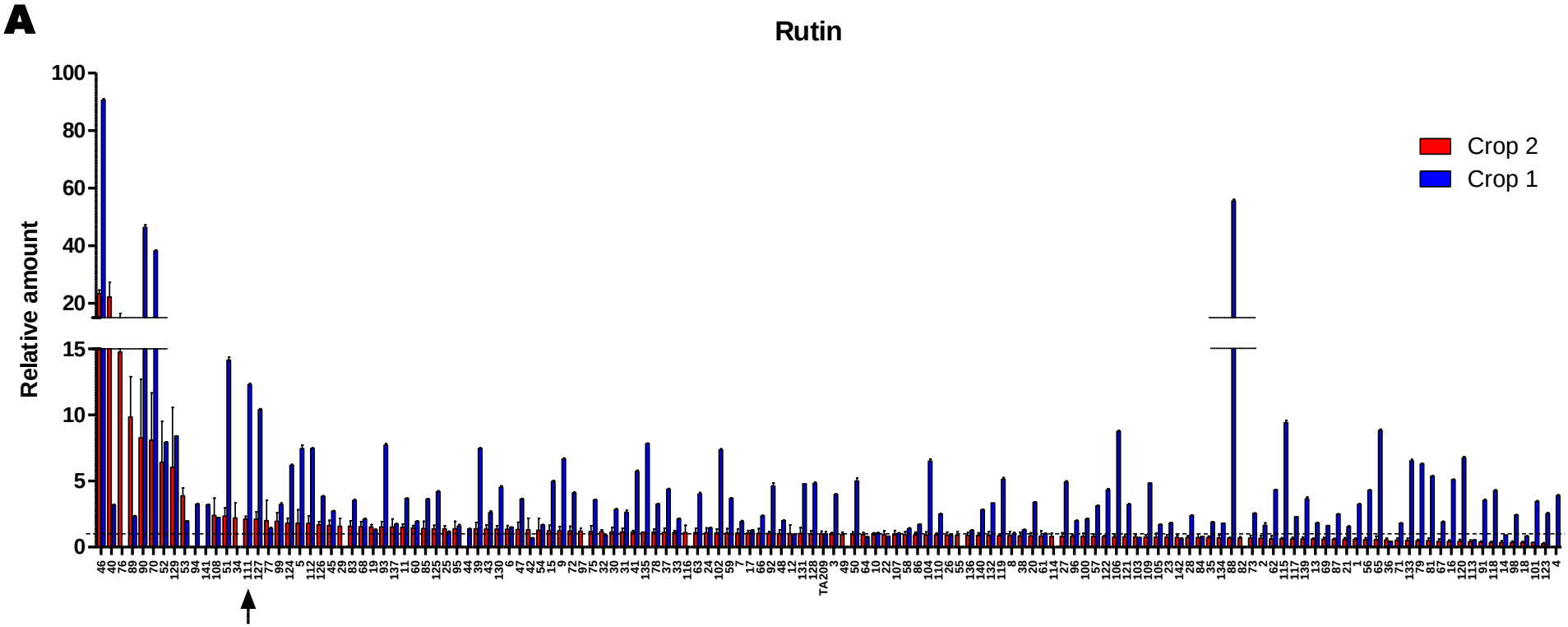

Figure 10 Relative amounts of phenolic compounds in crop 1 and crop 2 of the *S. neorickii* BIL population grown in the field.

Average compound levels are provided as amount relative to TA209. Sample names indicate neo accession numbers from *S. neorickii* BIL population. Error bars represent standard error of the mean (SEM; n=4, where available; biological replicates for crop 2, technical replicates for crop 1). Zero values indicate compound was not detected by HPLC-DAD, except where samples were not available for analysis: crop 1 neo-29, -34, -49, -55, -76, -82, -97, -114, -116; crop 2 neo-44, -94, -141. Data for crop 1 are provided by E. Enfissi. Broken horizontal line represents TA209 levels. (A) Rutin. TA209 levels are 332.29 ± 54.94 μg g^{-1} DW (crop 1) and 243.86 ± 48.01 μg g^{-1} DW (crop 2). Arrow indicates neo-111.

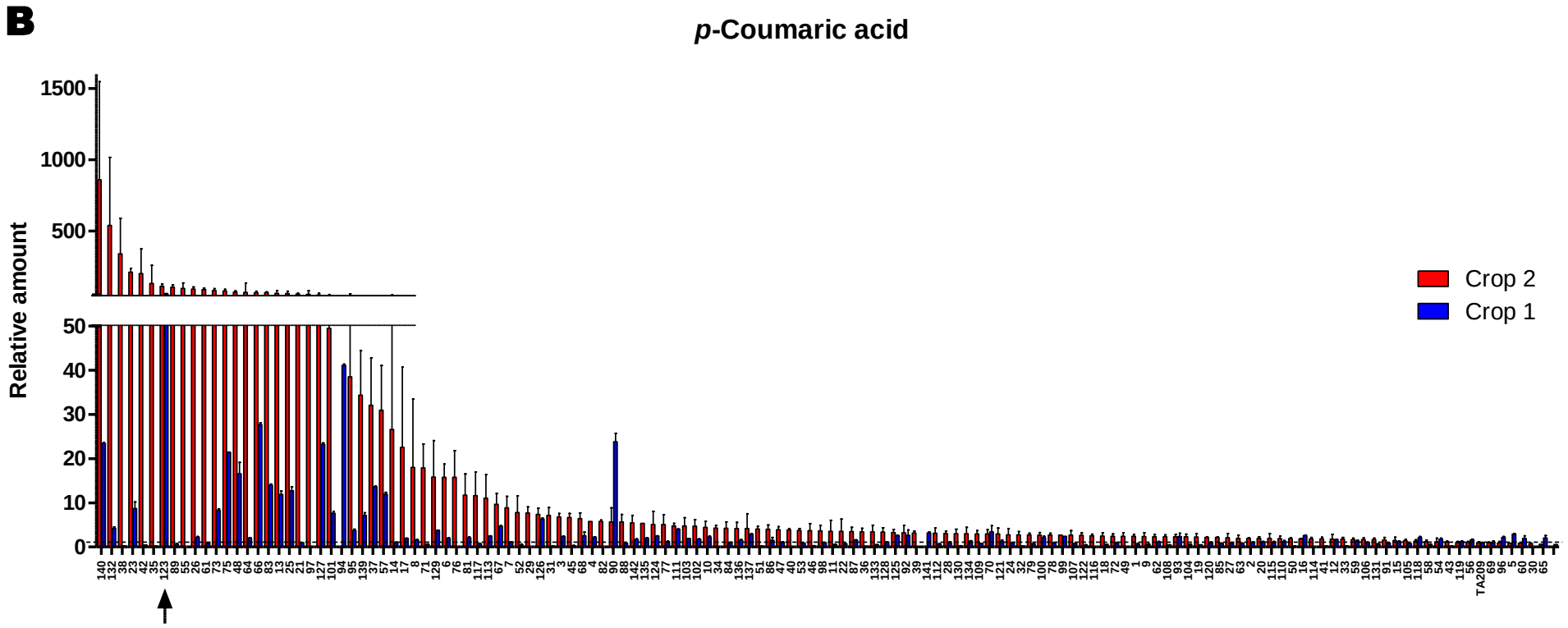

Figure 10 continued.

(B) *p*-Coumaric acid. Levels include derivatives of unknown structure. TA209 levels are 803.50 ± 28.01 μg g^{-1} DW (crop 1) and 52.88 ± 7.17 μg g^{-1} DW (crop 2). Arrow indicates neo-123. (C) Chlorogenic acid. Levels for TA209 are 337.03 ± 38.70 (crop 1) and 198.68 ± 17.56 μg g^{-1} DW (crop 2). (D) Chalcone-naringenin. TA209 levels are 177.50 ± 23.53 μg g^{-1} DW (crop 1) and 35.86 ± 14.76 μg g^{-1} DW (crop 2). (E) Naringenin crop 2 data (crop 1 unavailable). TA209 level is 46.02 ± 8.63 μg g^{-1} DW.

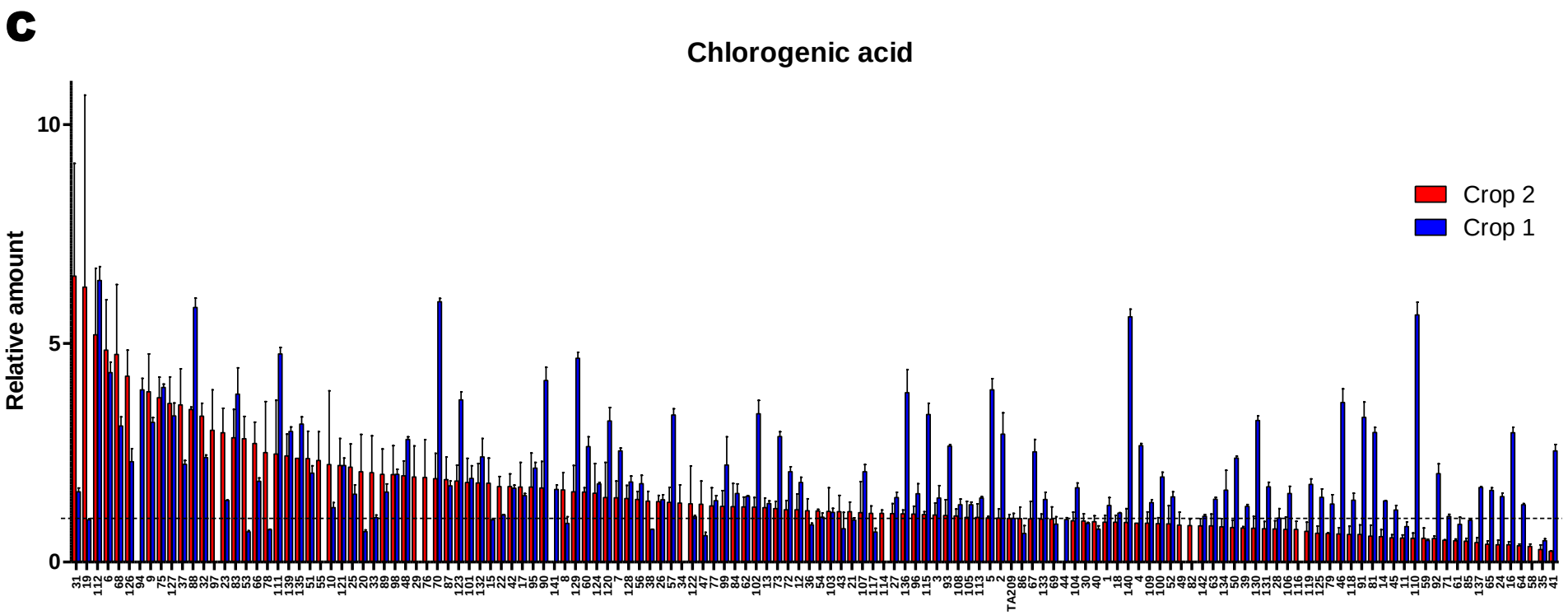

Figure 10 continued.

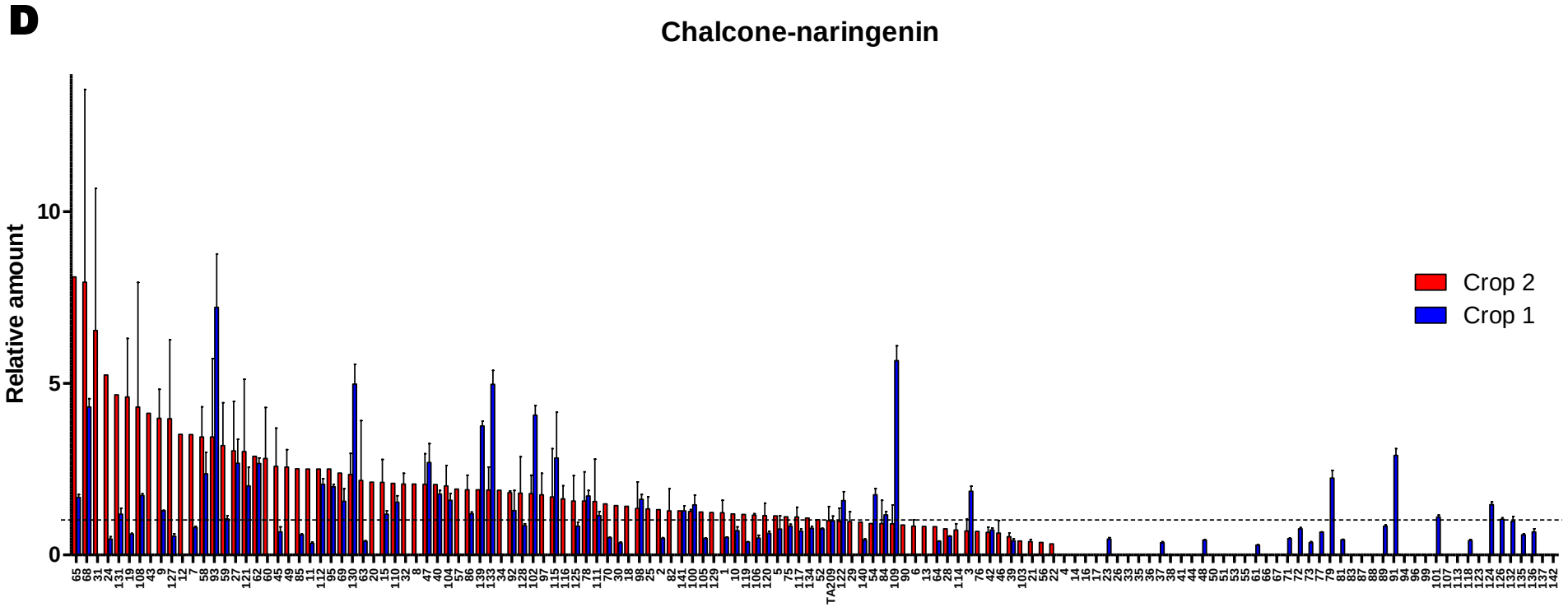

Figure 10 continued.

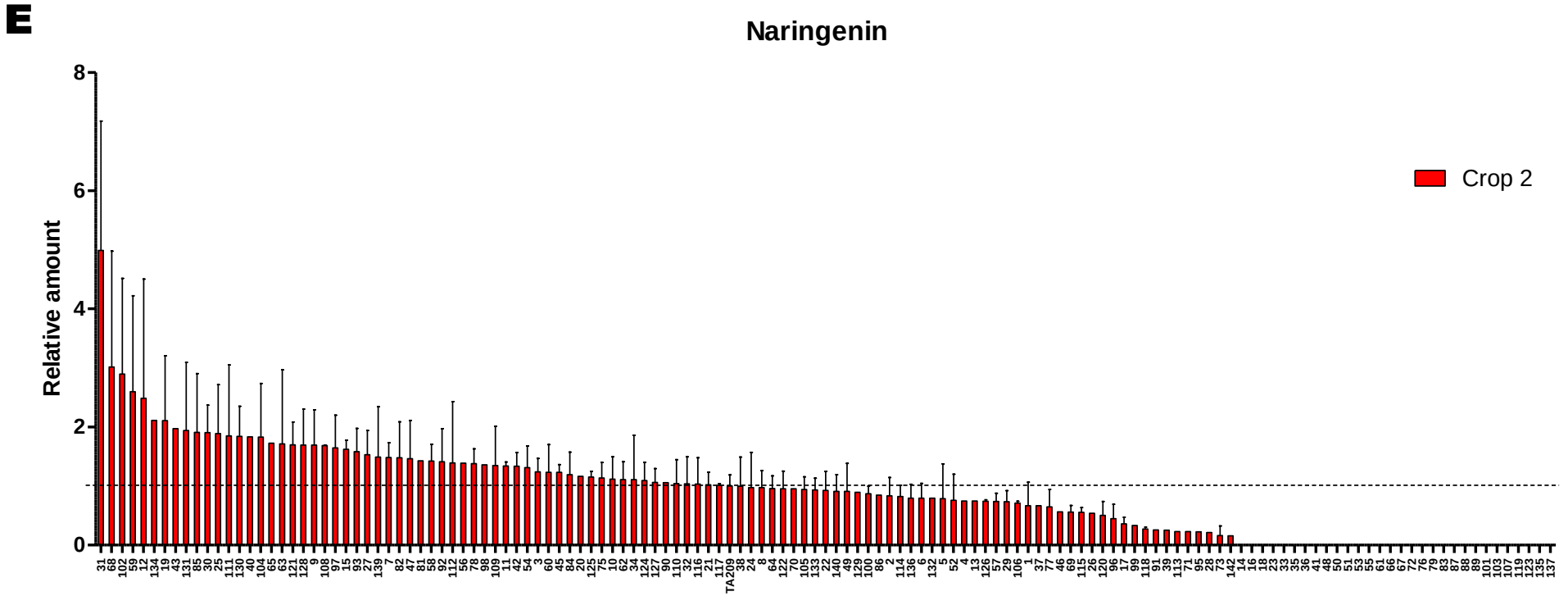

Figure 10 continued.

Relative levels of *p*-coumaric acid displayed a greater range (Figure 10B): from 860.22 ± 688.28 to 0.62 ± 0.01 fold in crop 2 and from 59.30 ± 1.16 to non-detectable levels in crop 1. Some lines showed consistency of phenotype for both crops while others did not conform. In crop 2, *p*-coumaric acid levels relative to TA209 increased for 130 of 136 BILs. By comparison, only 78 of 130 BILs showed a relative increase in crop 1. Although crop 2 exhibited lower absolute levels than crop 1, relative changes were generally higher. BILs with the highest *p*-coumaric acid levels relative to TA209 showed large variation between biological replicates (neo-140, 860.22 ± 688.28 fold; neo-132, 540.03 ± 476.87 fold; neo-38, 339.79 ± 249.74 fold). In crop 2 there were 29 BILs showing ≥ 20 fold *p*-coumaric acid levels. While six of these showed a decrease in crop 1, and two had no available sample for crop 1, 21 of these 29 BILs showed an increase in crop 1, and six of which also showed increases of ≥ 20 fold.

Likewise, the trend in chlorogenic acid levels between crop 1 and 2 (Figure 10C) showed some consistency in BILs containing ≥ 2 fold increase in crop 2 (18 out of 30 BILs also showed ≥ 2 fold increase in crop 1). However, large fold increases in crop 1 displayed throughout the gradient of crop 2 distort any visible whole population association. Examples included neo-5, -140 and -110, which show near TA209 levels and decreases in crop 2 (1.01 ± 0.04, 0.90 ±0.32, and 0.55 ± 0.12 fold, respectively), but large increases in crop 1 (3.94 ± 0.26, 5.61 ± 0.17, and 5.65 ± 0.29 fold, respectively). Chlorogenic acid levels showed similar ranges in both crops: from 6.53 ± 2.58 to 0.25 ± 0.01 fold in crop 2, and from 6.43 ± 0.32 to 0.49 ± 0.04 fold in crop 1.

Consistency in levels of chalcone-naringenin (Figure 10D) was far less than that of rutin, *p*-coumaric acid, and chlorogenic acid. Fold increases ranged up to 8.10 ± 0.09 in crop 2 and 7.21 ± 1.55 in crop 1. Fold decreases for both crops exceeded levels of detection. There are 36 BILs that showed a ≥ 2 fold increase in chalcone-naringenin in crop 2; however, only nine of these also showed increases of ≥ 2 fold in crop 1. In addition, levels for chalcone-naringenin were not detected for a large proportion of BILs (45 out of 136 BILs for crop 2, and 46 out of 130 BILs for crop 1). Relative naringenin levels ranged from 4.99 ± 2.19 fold to non-detectable levels (Figure 10E). Levels in crop 2 were comparable to those of chalcone-naringenin. Of the 59 BILs that displayed an average increase in relative naringenin levels, 49 also showed increases in chalcone-naringenin levels. There were seven BILs showing increases in naringenin ≥ 2 fold. All of these also showed increases in chalcone-naringenin in crop 2, and five of these showed increases ≥ 2 fold.

In an attempt to estimate the relationship between metabolite levels in crops 1 and 2, comparisons of relative levels of rutin, *p*-coumaric acid, chlorogenic acid and chalcone-naringenin were plotted (Figure 11) and Pearsons correlations were calculated. In support of the observations detailed in Figure 10, rutin levels showed the greatest correlation between in crops 1 and 2 when expressed as relative amounts (Figure 11A; $r = 0.6029$). In comparison to other metabolites, a high level of correlation was also observed when rutin levels were expressed as log values of relative amounts (Figure 11B; $r = 0.4918$). Both rutin correlations showed the highest levels of significance ($p \leq 0.0001$) among this dataset.

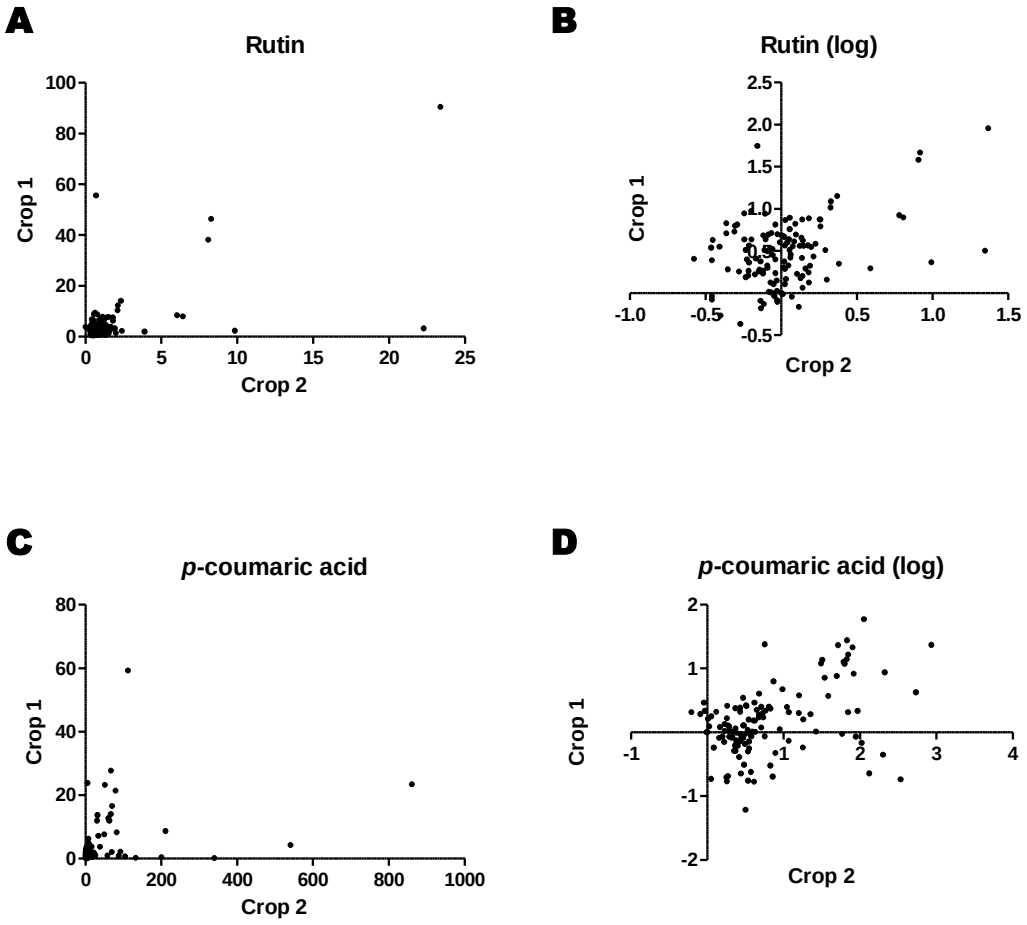

Figure 11 Comparison of relative metabolite levels in crop 1 and crop 2 of the *S. neorickii* BIL population grown in the field.

Relative amounts shown in (A), (C), (E) and (G) are expressed as log values as indicated in (B), (D), (F) and (H). Missing values and levels below limits of detection are represented by a relative amount of zero.

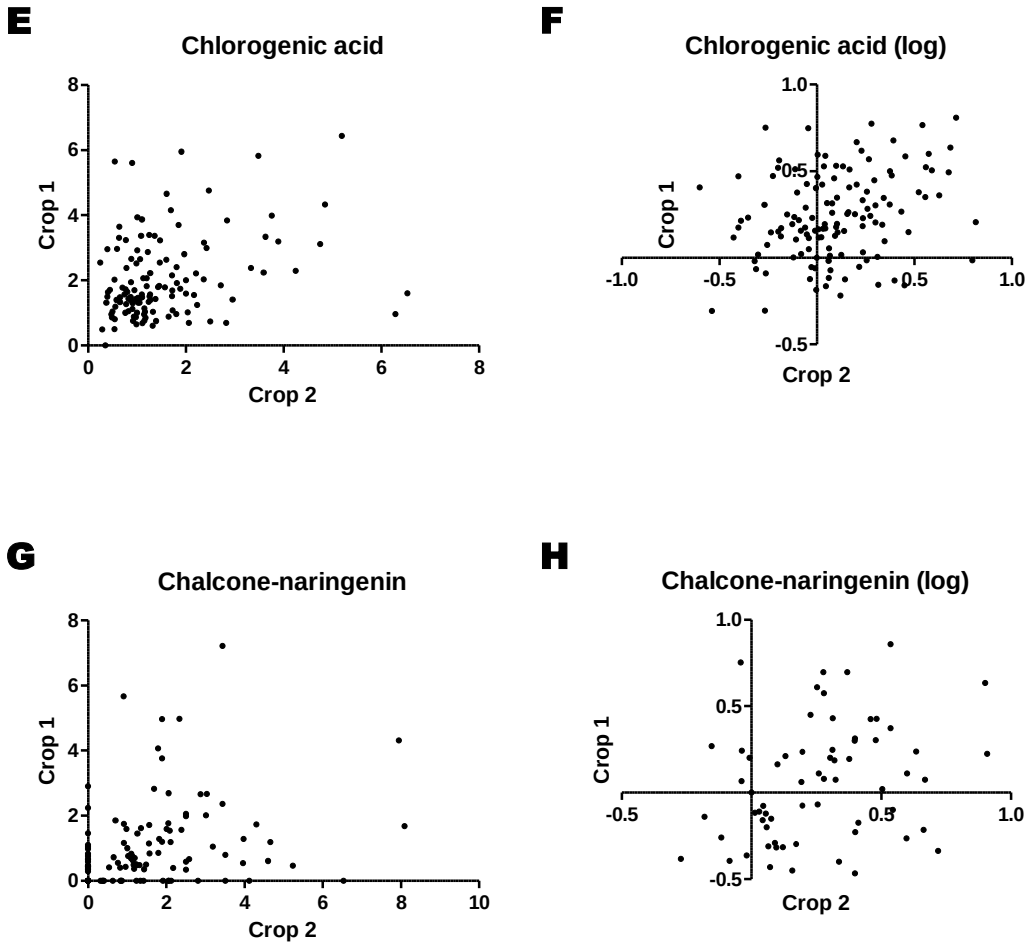

Figure 11 continued.

Surprisingly, levels of chalcone-naringenin and *p*-coumaric acid showed similar levels of correlation (Figure 11C and G), both ranking second to the correlation of rutin levels (*p*-coumaric acid $r = 0.3275$, $p \leq 0.0005$; chalcone-naringenin $r = 0.3532$, $p \leq 0.0001$). By comparison, the log values of *p*-coumaric acid levels (Figure 11D) showed greater correlation than any other log expressed values, including that of rutin (*p*-coumaric acid $r = 0.4524$, $p \leq 0.0001$; chalcone-naringenin $r = 0.2992$, $p \leq 0.05$, which is the least significant correlation, Figure 11H).

In agreement with the trends of relative chlorogenic acid levels in crop 1 and 2 observed in Figure 10, relative levels of chlorogenic acid also showed the least correlation with some of the least significance (Figure 11E, $r = 0.2995$, $p \leq 0.001$; Figure 11F, $r = 0.2697$, $p \leq 0.005$).

Due to lack knowledge as to whether or not each genotype within crop 2 was identical to its counterpart in crop 1 (via the assumed two seed generations across two subsequent growing seasons), heritability calculations were deemed difficult, if not impossible, to calculate. Further to this, it was suspected that for some genotypes plants in crop 1 and crop 2 were from common seed pools; that is to say, for these genotypes crop 1 and crop 2 plants were simply biological replicates grown in two different seasons. As mentioned previously, a further complication to these data lie in the fact that metabolite data for crop 1 each represent a pooled homogenate of biological replicates that were analysed using four technical replicates. In contrast, metabolite data for crop 2 was generated from between one and four biological replicates that were not pooled before metabolite analysis. Due to the limitations of these data, relative metabolite levels for crop 2 were preferentially used to suggest future relationships between fruit phenolic profile and *S. neorickii* BIL inserts.

3.2.2 Suggested relationship between phenolic content profile and S. neorickii inserts in BIL population

The *S. neorickii* BIL population has been mapped with both restriction fragment length polymorphism (RFLP) markers (Figure 12) and conserved orthologs set II (COSII) markers (Figure 13) (EU Sol, personal communication). These data were available within the EU Sol consortium (http://www.eu-sol.net), and have been represented here schematically. Each of the 142 *S. neorickii* BILs contained between zero and 18 RFLP markers for *S. neorickii* from a possible total of 54 markers. According to these markers, 69.7 % of BILs contained *S. neorickii* inserts in more than one chromosome; however, 9.2 % contain no RFLP markers for *S. neorickii*. Genomic regions were present as both heterozygous and homozygous inserts for *S. neorickii*. By comparison, data on 114 COSII markers existed for *S. neorickii* BILs (Figure 13). Some chromosome inserts identified by RFLP were seemingly not present according to COSII markers. Likewise, the reverse is true, and some BILs contained inserts according to COSII markers where previously no insert was detected by RFLP markers. A greater proportion of inserts shown by COSII markers are homozygous than with RFLP.

During the early stages of this project, RFLP were the only available markers; therefore, these were used for the identification of possible relationships between phenolic content seen in section 3.2.1 and genome regions from *S. neorickii*. These possible associations were suggested by crude and non-quantitative observations since little information was available about the RFLP marker data and how well they represented the genotypes made available during this phenotype screening. The lineage (including the number of generations) between the population that was genotyped for the RFLP marker data and those genotypes used for phenotype screening is still unknown. It was determined, therefore, that any possible observed relationship between phenolic content and genome regions from *S. neorickii* would be at best speculative and should not be confused with statistically validated QTL-mapping. As a result, with the information made available, an association was made between RFLP markers showing *S. neorickii* genomic sequences and BILs containing consistently high rutin levels. When the schematics of BIL genomes using RFLP markers were reordered according to rutin levels relative to TA209 in crop 2 (as shown by Figure 14A), a cluster was identified on chromosome 5. There were 15 BILs that possessed rutin levels ≥ 2 fold relative to TA209 in crop 2 (Figure 14B, labelled neo-46 to -77), 11 of which showed inserts from *S. neorickii* according to RFLP markers. An additional four BILs were shown within the cluster exhibiting increases in relative rutin levels ranging from 1.55 to 1.80 fold increases and chromosome 5 inserts. Further to this, 12 BILs exhibited ≥ 2 fold increases in rutin in both crops 1 and 2, and the cluster remains visible in chromosome 5 (Figure 14C). The majority of the inserts within the cluster (in both Figure 14B and Figure 14C) were heterozygous for *S. neorickii*.

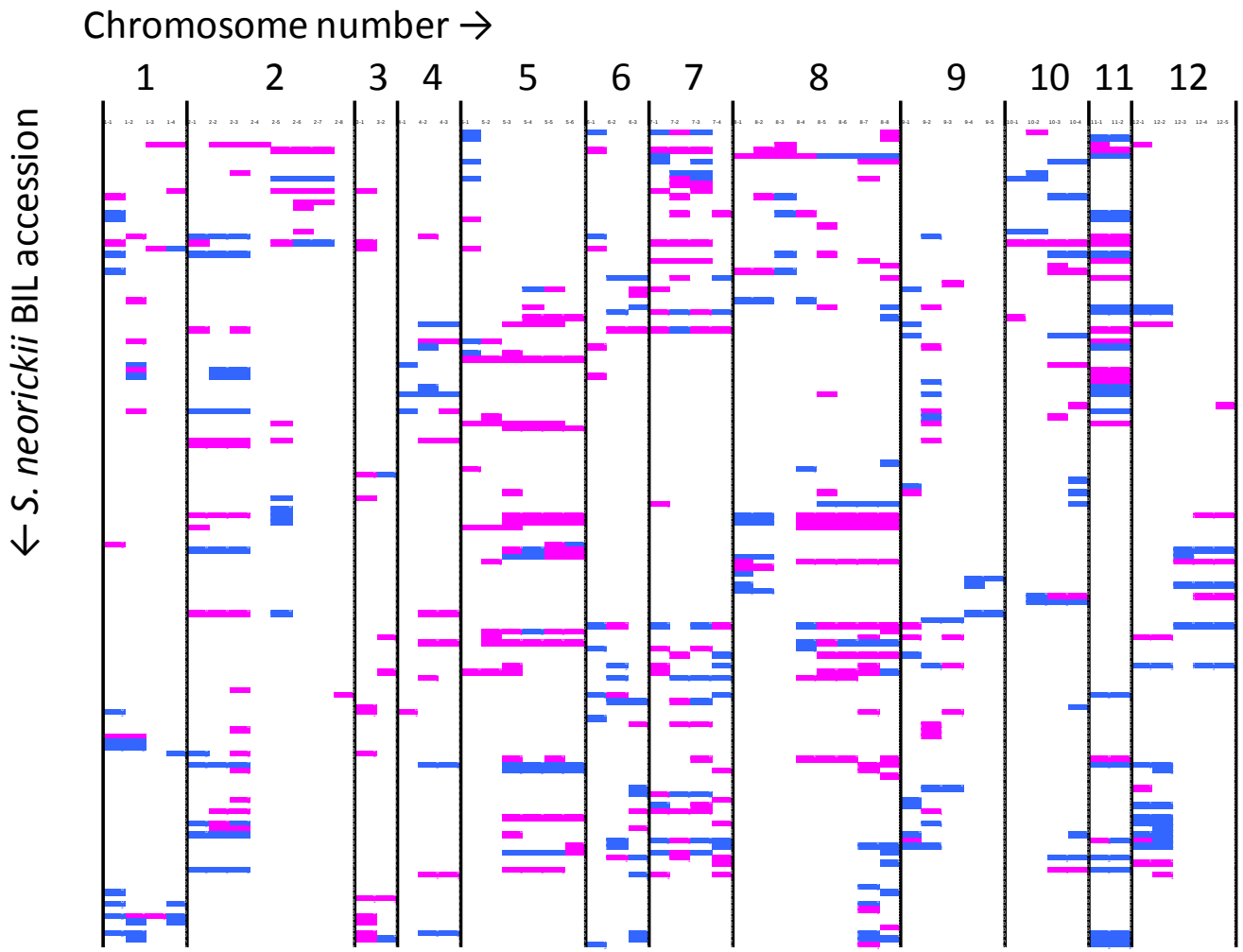

Figure 12 Schematic representation of inserts from *S. neorickii* for all BIL population accessions according to RFLP markers.

Chromosomes 1 to 12 are separated by thick lines and labelled at the head of each column. RFLP markers are located within each chromosome. Rows indicate *S. neorickii* BIL genomes from neo-1 to neo-142. For any single row, white background represents TA209 genome. Coloured horizontal bars represent inserts from *S. neorickii*, where a pink bar indicates heterozygous insertion, and blue bar represents homozygous insertion. Original data taken from http://www.eu-sol.net, available within the EU Sol consortium. Chromosome length and insert positions are not shown to scale.

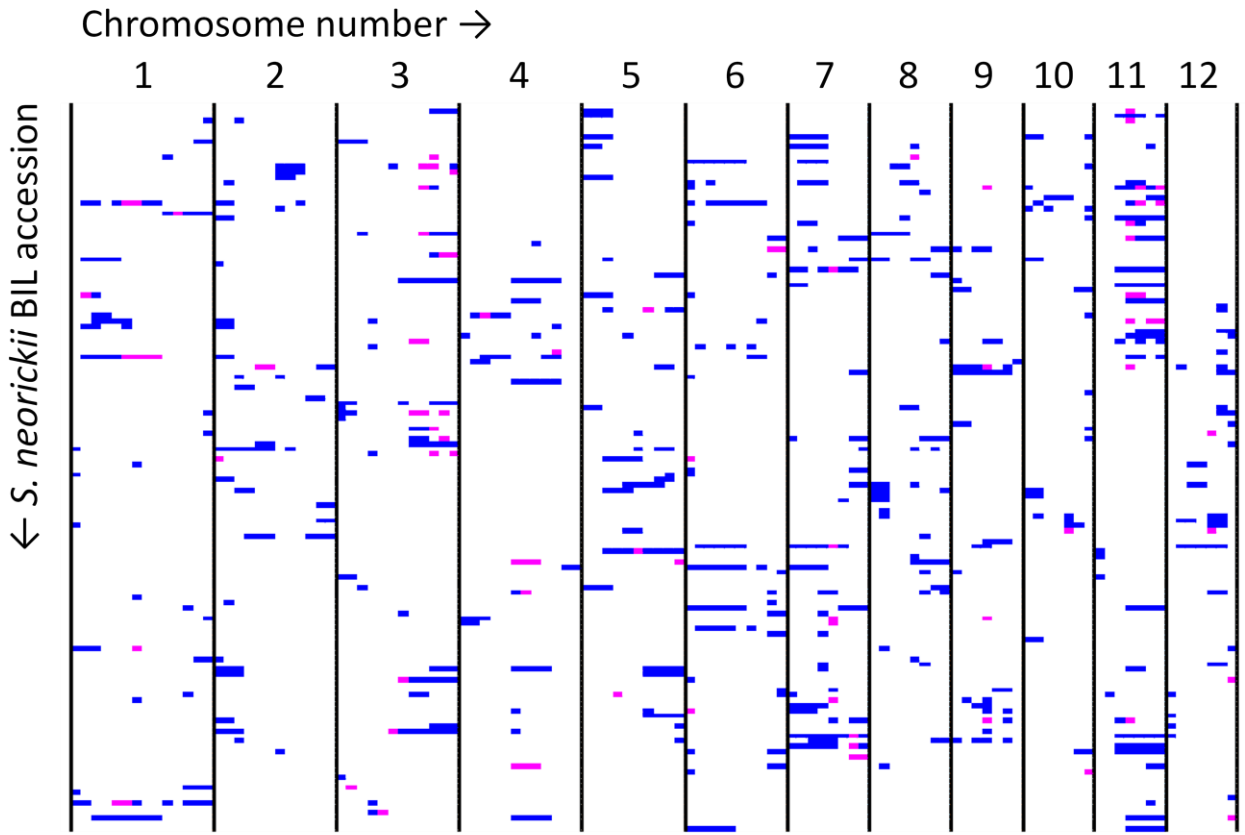

Figure 13 Schematic representation of inserts from *S. neorickii* for all BIL population accessions according to COSII markers.

Chromosomes 1 to 12 are separated by thick lines and labelled at the head of each column. COSII markers are located within each chromosome. Rows indicate *S. neorickii* BIL genomes from neo-1 to neo-142. For any single row, white background represents TA209 genome. Coloured horizontal bars represent inserts from *S. neorickii*, where a pink bar indicates heterozygous insertion, and blue bar represents homozygous insertion. Original data taken from http://www.eu-sol.net, available within the EU Sol consortium. Chromosome length and insert positions are not shown to scale.

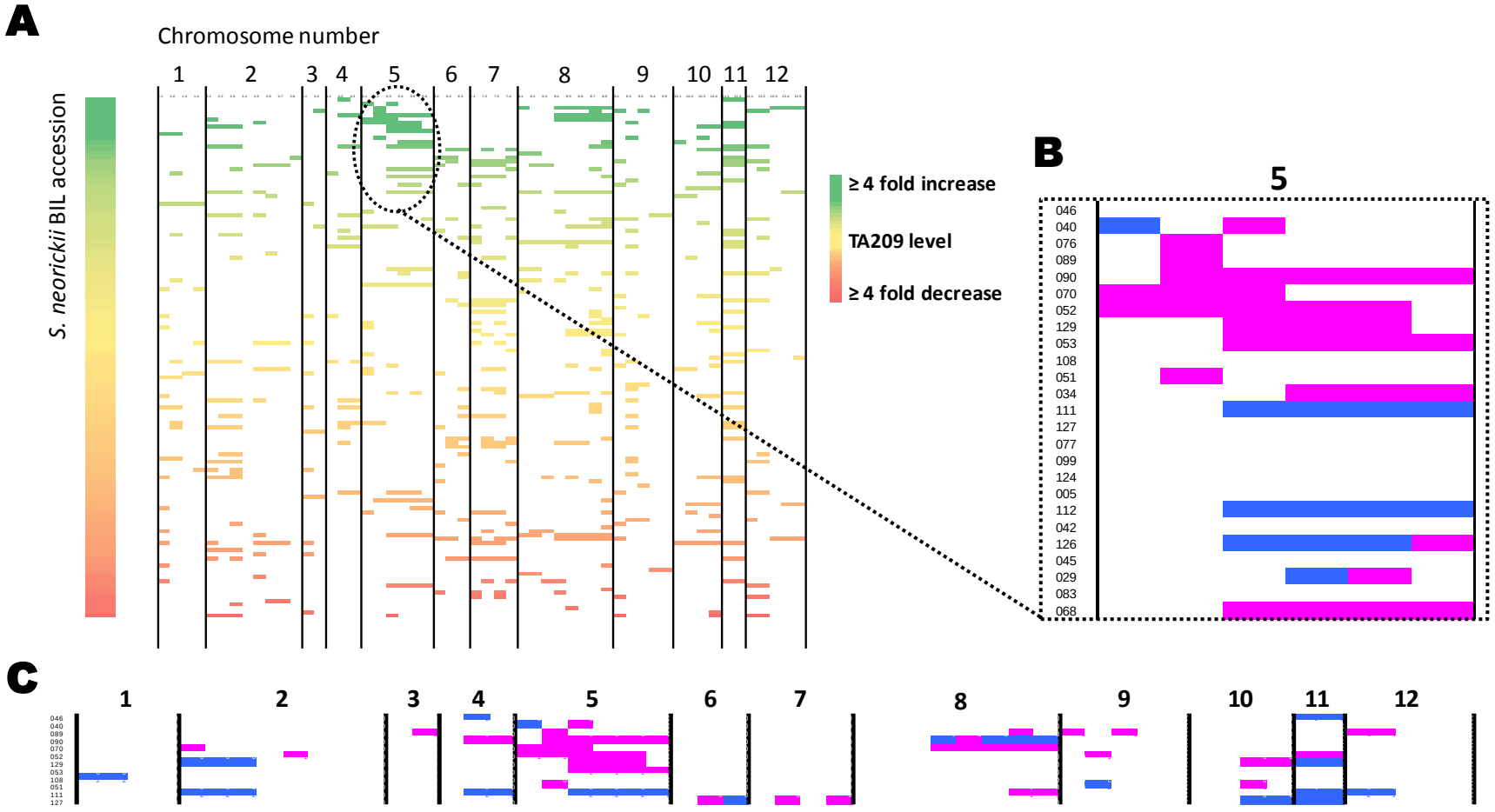

Figure 14 Suggested relationship between high rutin phenotype and *S. neorickii* insert regions

(Legend overleaf)

Figure 14 (legend, continued from previous page)

(A) Schematic representation of *S. neorickii* BIL population genomes shown in Figure 12 reorganised in descending order of rutin levels relative to TA209 in crop 2. Cluster of chromosome 5 inserts correlating with rutin levels ≥ 4 fold relative to TA209 in crop 2 are identified by broken ring. (B) *S. neorickii* BIL accession numbers for accessions within cluster and (C) schematic for RFLP inserts for accessions possessing ≥ 2 fold increases in rutin in both crops 1 and 2. Rows are BILs and columns represent chromosome number. Zygosity of inserts according to RFLP markers from *S. neorickii* are represented by pink (heterozygous) and blue (homozygous) bars. White background represents TA209 genomic background. Chromosome and marker lengths are not shown to scale.

BILs were similarly organised in descending order of *p*-coumaric acid levels according to crop 2 (Figure 15A) in order to similarly make possible associations between high relative metabolite levels and *S. neorickii* chromosomal inserts shown by RFLP markers. Of the 36 BILs possessing ≥ 10 fold increases in *p*-coumaric acid levels relative TA209, 14 showed *S. neorickii* genome sequence for one of both of the latter two RFLP markers (TG241 and CT95) on chromosome 10. This is shown by the cluster in Figure 15A, and detailed in Figure 15B, in which all BILs except neo-113, -67 and -7 contained *p*-coumaric acid levels ≥ 10 fold relative to TA209. Of the 15 insert regions in this cluster, ten were shown to be homozygous for *S. neorickii*. There were 17 BILs that contain consistently high increases (≥ 5 fold) in *p*-coumaric acid levels for both crops 1 and 2 (Figure 15C). Nine of these also displayed *S. neorickii* genomic insert for one or both TG241 and CT95, and eight of these nine were homozygous for *S. neorickii*.

Figure 15 (legend; figure overleaf)

(A) Schematic representation of *S. neorickii* BIL population genomes shown in Figure 12 reorganised in descending order of *p*-coumaric acid levels relative to TA209 in crop 2. Cluster of chromosome 10 inserts correlating with *p*-coumaric acid levels ≥ 10 fold relative to TA209 in crop 2 are identified by broken ring. (B) *S. neorickii* BIL accession numbers for accessions within cluster and (C) schematic for RFLP inserts for accessions possessing ≥ 5 fold increases in *p*-coumaric acid in both crops 1 and 2. Rows are BILs and columns represent chromosome number. Zygosity of inserts according to RFLP markers from *S. neorickii* are represented by pink (heterozygous) and blue (homozygous) bars. White background represents TA209 genomic background. Chromosome and marker lengths are not shown to scale.

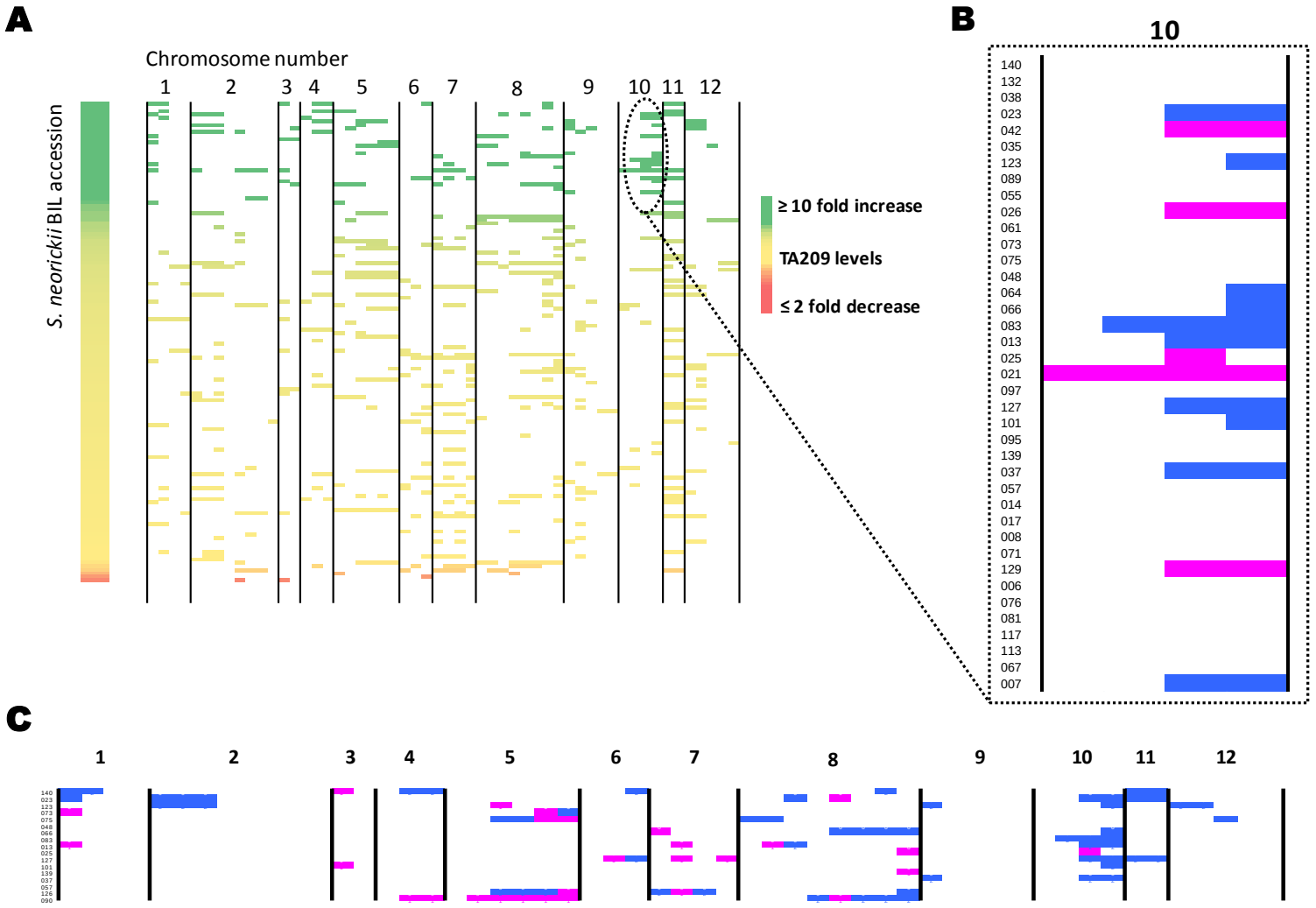

Figure 15 Suggested relationship between high *p*-coumaric acid phenotype and *S. neorickii* insert regions (legend on previous page)

3.2.3 Selection of candidate lines

Candidate lines were selected to represent the chromosome cluster regions discussed in section 3.2.2, and these are summarised in Figure 16. Neo-111 was selected to represent a high rutin phenotype and *S. neorickii* insert in chromosome 5 according to RFLP markers. In addition to possessing an increase in both crops 1 and 2, neo-111 displayed sequences homozygous for *S. neorickii* within chromosome 5 with both RFLP and COSII markers. Levels of other assessed phenolics were also increased (Figure 16). Similarities were identified between neo-111 and an genotype 3939 from *S. habrochaites* NIL population (Monforte and Tanksley, 2000). T. Wells (RHUL, unpublished) reported increases in rutin relative to TA209 in NIL 3939, which also possesses a *S. habrochaites* insertion in chromosome 5 according to RFLP markers. It was decided, therefore, to include NIL 3939 in future analyses to compare with *S. neorickii* BIL neo-111. Neo-123 was selected to represent a high *p*-coumaric acid phenotype and chromosome 10 insert from *S. neorickii*. This candidate was chosen, firstly because it displayed a consistently high increase ≥ 50 fold relative to TA209 in crops 1 and 2, and secondly because it possesses RFLP markers for *S. neorickii* as a homozygous insert in chromosome 10. Unfortunately, when COSII marker data was later made available, no insert was shown to be detected by COSII markers.

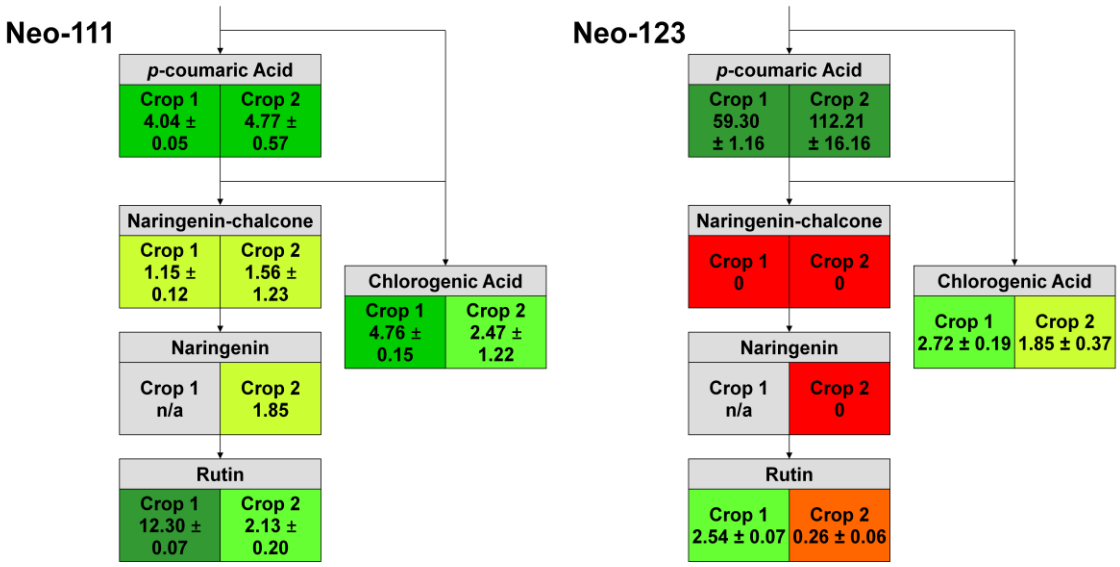

Figure 16 Summary of levels of phenolic compounds in candidate lines neo-111 and neo-123

Values are relative to TA209 ± standard error of the mean. Colour indicates degree of increased (green) or decreased (red) fold change in crops 1 and 2. n/a indicates no data available. Crop 1 data provided by E. Enfissi (RHUL).

3.2.4 Confirmation of novel phenotype under different growth conditions

Fruit was cultivated in a glasshouse environment in the UK in order to ascertain whether the phenotype was reproducible under different growth conditions. Fruit at 7 and 14 days post breaker (dpb) were harvested in January 2010. *S. habrochaites* NIL 3939 and *S. neorickii* BILs neo-111 and -123 maintained their respective high rutin and high *p*-coumaric acid phenotypes seen previously when grown under glass at both 7 and 14 dpb (Figure 17A-C) relative to TA209 levels (Figure 17D).

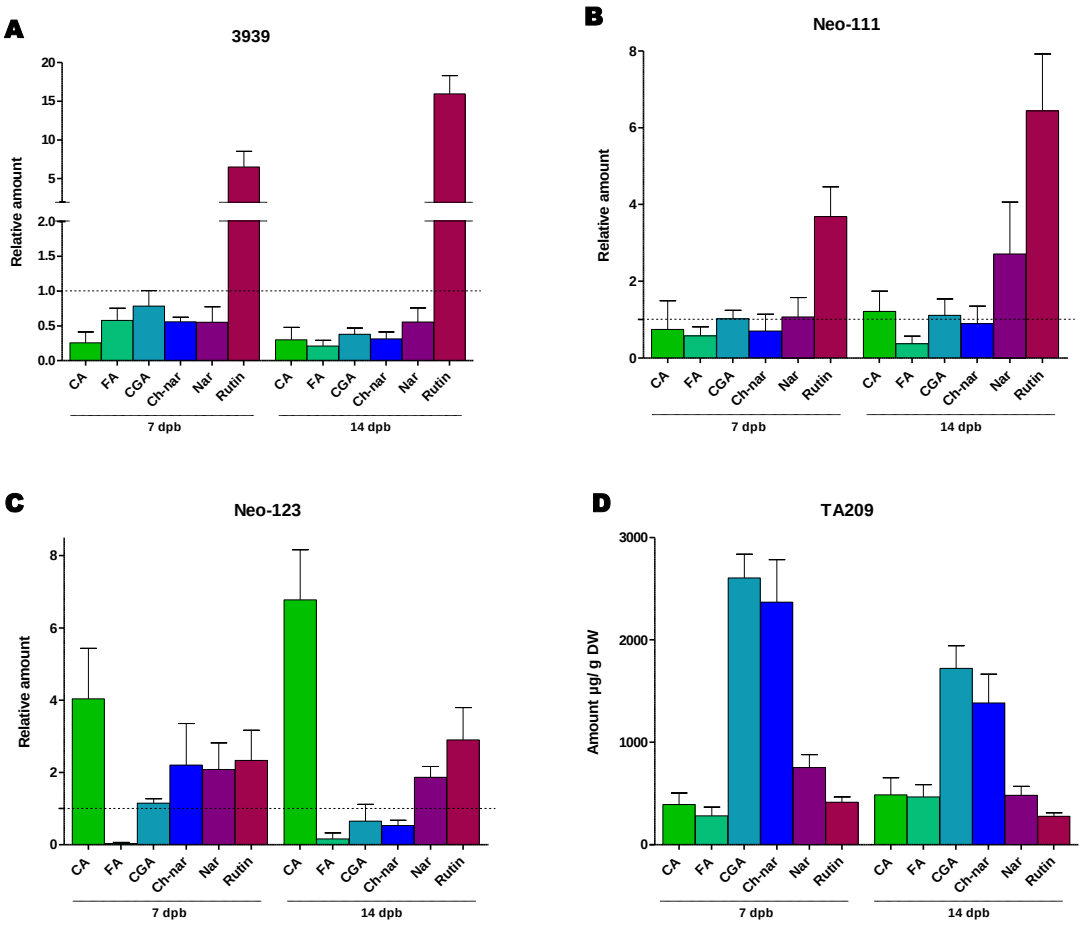

Figure 17 Confirmation of phenolic profile for selected lines cultivated under glass

Average phenolic levels of selected lines (A) *S. habrochaites* NIL 3939 and (B and C) *S. neorickii* BILs neo-111 and -123, respectively, relative to (D) TA209 absolute values at 7 and 14 days post breaker (dpb) for fruit grown under glass at RHUL and analysed by HPLC-DAD. Error bars represent SEM (n=4 for 3939; n=3 for neo-111 and -123; n=7 for TA209, 7 dpb; n=4 for TA209, 14 dpb). CA represents *p*-coumaric acid and derivatives of unknown structure; FA, ferulic acid; CGA represents chlorogenic acid and unknown compounds of similar UV spectral shape; Ch-nar, chalcone-naringenin; Nar, naringenin.

NIL 3939 showed greater increases in rutin than neo-111, and displayed relative decreases in levels of all other phenolics (Figure 17A). No previous data were available to compare reproducibility in levels of other phenolics. Although the high rutin phenotype of neo-111 was sustained (Figure 17B), levels of other phenolics were similar to levels of TA209, except for ferulic acid, which decreased, and naringenin, which increased at 14 dpa only. This profile is different from that seen in the field in crops 1 and 2 (Figure 16), where all assessed metabolites increased. Neo-123 displayed high levels of *p*-coumaric acid relative to TA209 (Figure 17C), but these fold changes were not as high as seen previously. Increases were observed for the flavonoids naringenin, chalcone-naringenin (at 7 dpb only), and rutin. These changes are contrary to those seen in crops 1 and 2, except for rutin levels in crop 1, which also increased. Ferulic acid relative levels decreased, and chlorogenic acid levels showed little change from TA209 compared with the increased observed in crops 1 and 2. Levels of all phenolics from glasshouse grown TA209 fruit (Figure 17D) are higher than those seen in the field (Figure 10) in crops 1 and 2, with the exception of *p*-coumaric acid levels for crop 1.

3.3 Discussion

The *S. neorickii* BIL population (Fulton et al., 2000; Grandillo et al., 2011) comprising 142 genotypes was screened for levels of phenolics with the aim of identifying and characterising genomic regions from *S. neorickii* influencing levels of health-promoting phenolic profiles in ripe tomato fruit. *S. neorickii* BIL population was chosen for this purpose due in part to the low level of previous exploitation of the wild relative compared with other *Solanum* sect. *Lycopersicon* species, such as *S. pennellii* and *S. habrochaites*, as a result of the relatively recent discovery of *S. neorickii* (Rick et al., 1976). However, *S. neorickii* BIL population was also chosen because of the large variation of phenolic profiles observed during a previous screen (Crop 1) in 2006 (E. Enfissi, unpublished).

The screened phenolic profile comprised the phenylpropanoids *p*-coumaric acid and chlorogenic acid, and the flavonoids naringenin, chalcone-naringenin, and rutin. These compounds represent some of the major phenolic components of ripe domesticated tomato fruit (Martinez-Valverde et al., 2002; Slimestad et al., 2008). TA209 levels of chlorogenic acid and chalcone-naringenin in at least two of the three

environments were within the range of previously published levels for ripe tomato cultivars (Long et al., 2006; Melendez-Martinez et al., 2010). Qualitative levels of chlorogenic acid, chalcone-naringenin, rutin, *p*-coumaric acid, and ferulic acid were within ranges of previously published results (published as fresh weight levels) allowing for inter-cultivar variation (Bovy et al., 2002; Martinez-Valverde et al., 2002; Slimestad et al., 2008).

Levels of phenolic compounds were in higher abundance in Crop 1 (E. Enfissi, unpublished) compared with Crop 2. This can be observed from reported levels of *p*-coumaric acid, chlorogenic acid, chalcone-naringenin and rutin in the *S. neorickii* background line TA209 (Figure 10). Levels of some phenolics in TA209 were higher still when BIL plants were reproduced under glasshouse environments in the UK (Figure 17). This was observed for levels of chlorogenic acid, chalcone-naringenin, *p*-coumaric acid (compared with Crop 2 only), and naringenin (compared with Crop 2 only), but not for rutin. Reproducibility of levels of phenolics between Crops 1 and 2 was low, both in terms of directionality (fold increase or decrease relative to TA209) and in terms of population trend (comparison of ranking order of BILs between Crops 1 and 2 shown in Figure 10). This was especially observed for levels of rutin, chlorogenic acid, and chalcone-naringenin. Reproducibility of directionality for levels of *p*-coumaric acid fared better; however, this is likely due to low basal levels in TA209 resulting in the majority of BILs (94 % in Crop 2) exhibiting a fold increase.

As detailed above, in section 3.2, some individual BILs did show consistent levels of phenolics relative to TA209 across Crops 1 and 2. 73 % of BILs that showed a fold increase of two or more in rutin levels in Crop 2 likewise showed such as increase in Crop 1.

Phenolic biosynthesis is known to be affected by changes in environments and abiotic stress, such as increased light intensity and drought, which result in an increased accumulation of compound intermediates in tomato fruit (Lovdal et al., 2010; Sanchez-Rodriguez et al., 2012). Likewise, the phenotypic effects of some QTL have been shown to be altered by environmental conditions. Patterson and colleagues (1991) showed that not all effects observed as a result of identified QTL alleles from *L. cheesmanii* (former name) hybrids were likewise observed under replicated trials in alternative environmental conditions. Environment has been shown to affect QTL for antioxidants, especially water soluble phenolics, when reproduced over more than one

season in the field (Rousseaux et al., 2005). In addition, Rousseaux and colleagues (2005) demonstrated that within their study, the effects of most QTL identified from field grown crops were not observed when plants were reproduced in the glasshouse. Finally, Goodfellow (2008) also observed a low level of reproducibility in secondary metabolite fruit profiles when *S. pennellii* ILs were grown in multiple environments.

The environmental conditions in the field during cultivation of Crop 1 (in 2006) were reported to have likely caused greater plant stress due to higher light intensity, temperature and drought than during the cultivation of Crop 2 (in 2008) (D. Zamir, personal communication). The change in environmental conditions between 2006 and 2008, and likewise between a field and a glasshouse environment, could explain the observed increases in levels of phenolic compounds in Crop 1 compared with Crop 2, and in the glasshouse lines compared with field grown lines.

The RFLP and COSII marker data presented in Figure 12 and Figure 13 illustrate the complexity of introgressed regions of *S. neorickii* genome in *S. neorickii* BIL population compared with populations such as *S. pennellii* IL or *S. habrochaites* NIL where accessions contain a single chromosomal region homozygous for the wild relative sequence (Eshed and Zamir, 1995; Monforte and Tanksley, 2000). Firstly, the prevalence of heterozygous regions from *S. neorickii*, especially according to RFLP markers, causes uncertainty with regards to the exact genotype in successive generations of BILs. An allele that according to RFLP or COSII markers may be heterozygous for *S. neorickii* in Crop 1 (2006) may have segregated away from the genotype by Crop 2 (in 2008, which is two generations later via an unanalysed intermediate crop in 2007). Secondly, BILs that possess more than one introgressed chromosome region in two separate genomic locations might additionally affect two QTL for the same complex trait (such as accumulation of phenolic compounds). These QTL may in fact have cumulative, opposing or synergistic effects on the regulation of phenolic biosynthesis and the accumulation of compound intermediates. The complex genotypes of *S. neorickii* BIL population may therefore have not only contributed to non-reproducible phenotypes in subsequent generations, such as Crops 1 and 2, but may also explain large variations between biological replicates of individual BILs, such as the large SEM values for high *p*-coumaric acid lines in Crop 2 (that were not seen in Crop 1 due to pooling of biological replicates; E. Enfissi, personal communication).

This chapter hypothesised a possible association between an introgressed region from *S. neorickii* in chromosome 5 and a high rutin phenotype. A similarly tentative association between an introgressed region in chromosome 10 and a high *p*-coumaric acid phenotype was highlighted. Candidate BILs, neo-111 and -123, were identified to represent these associated regions, respectively. Due to the previously mentioned complexities of the introgressed regions defined by RFLP and COSII markers, any associations made between reproducible BIL phenotypes and common introgressed regions therein can merely be speculative. The limitations of the marker data are exemplified by the fact that 9 % of the BILs contain no *S. neorickii* introgressed region according to RFLP marker data, and there are no COSII marker data available for 4 % of the BILs. There is disagreement between RFLP and COSII datasets; that is to say some introgressed regions are present in one dataset but missing from another, which could be due to low marker coverage or different levels of polymorphism. This disagreement between RFLP and COSII markers occurred for the chromosome 10 insert in neo-123, which was the key parameter for which it was chosen as a candidate BIL. With hindsight, neo-123 would not have been selected to represent a chromosome 10 introgressed region had COSII data been made available when neo-123 was selected. This information does not, however, preclude the presence of a chromosome 10 introgressed region in neo-123. It is plausible that either neo-123 contains an introgressed region in chromosome 10 not defined by the COSII markers or that the genotypes of the experimental plants grown in Crops 1 and 2 differ from those used to anchor the COSII markers by the presence of a chromosome 10 introgressed region (S. Grandillo, personal communication).

A previous study on *S. habrochaites* NIL population (Monforte and Tanksley, 2000) identified a genotype, 3939, with high rutin levels and an introgression region on chromosome 5 (T. Wells, unpublished). Due to the comparisons with neo-111, *S. habrochaites* NIL 3939 was included in this study for further characterisation on the suspicion that similar regulatory mechanisms may be responsible. Phenolic profiles of 3939 and neo-111 when grown under glasshouse conditions, however, indicated a differential regulatory control of phenolic biosynthesis (Figure 17). While high rutin levels were confirmed for both lines at 7 and 14 dpb, neo-111 showed levels of all other phenolics (except naringenin at 14 dpb) to remain at TA209 background levels. 3939, however, displayed far greater levels of rutin, apparently at the expense of

accumulations of other phenolics that were decreased at both 7 and 14 dpb. This indicated that regulatory control of both lines was different.

All three selected lines (neo-111, -123 and 3939) showed a reproducible high phenolic phenotype when cultivated under glasshouse conditions (Figure 17). This demonstrated that despite the previously observed effects of environment, these high rutin and high *p*-coumaric acid levels are stable in each of the assessed environments.

4 Optimisation of high throughput screening

4.1 Introduction

Simultaneous to the work conducted in the previous chapter, a method with greater throughput for the analysis of phenolic compounds using a UPLC-PDA platform was developed. This development of this method is described in this chapter, and thresholds of limitations were assessed using authenticated standard compounds. A second high throughput method for the separation of isoprenoid compounds was developed by P. Fraser (personal communication). Likewise this method was assessed in this chapter for limitation thresholds. An alternative extraction protocol for the extraction of isoprenoids was investigated.

These high throughput separation methods were later validated by applying to the screening of a population of colour mutant genotypes believed to possess variation in secondary metabolite profiles affecting colour. The phenolic and isoprenoid profiles of this colour mutant population were assessed by principal component analysis.

4.2 Development of high throughput UPLC method for separation of phenolics

4.2.1 Development and evaluation of method

A method for the separation of tomato fruit phenolics, specifically phenylpropanoids and flavonoids, by UPLC-PDA was adapted from the method for separation by HPLC-DAD that was used previously in section 3.2, and detailed in section 2.2.3.1. In order to prevent corrosion and to accommodate the greater sensitivity of the UPLC-PDA apparatus, mobile phase solvent A, comprising HPLC grade water containing hydrochloric acid, was substituted for MS grade water containing MS grade formic acid. The Waters Acquity UPLC-PDA system described in section 2.1.4 was used; however, initially the stationary phase comprised a column 5 cm in length but identical in all other parameters. For this column, a gradient of 12 min

was designed. During the gradient 17 standard compounds were eluted (Figure 18), 15 of which were separated. Adjusting the gradient for better resolution of rutin and isoferulic acid resulted in poorer resolution of naringenin and chalcone-naringenin further along the gradient.

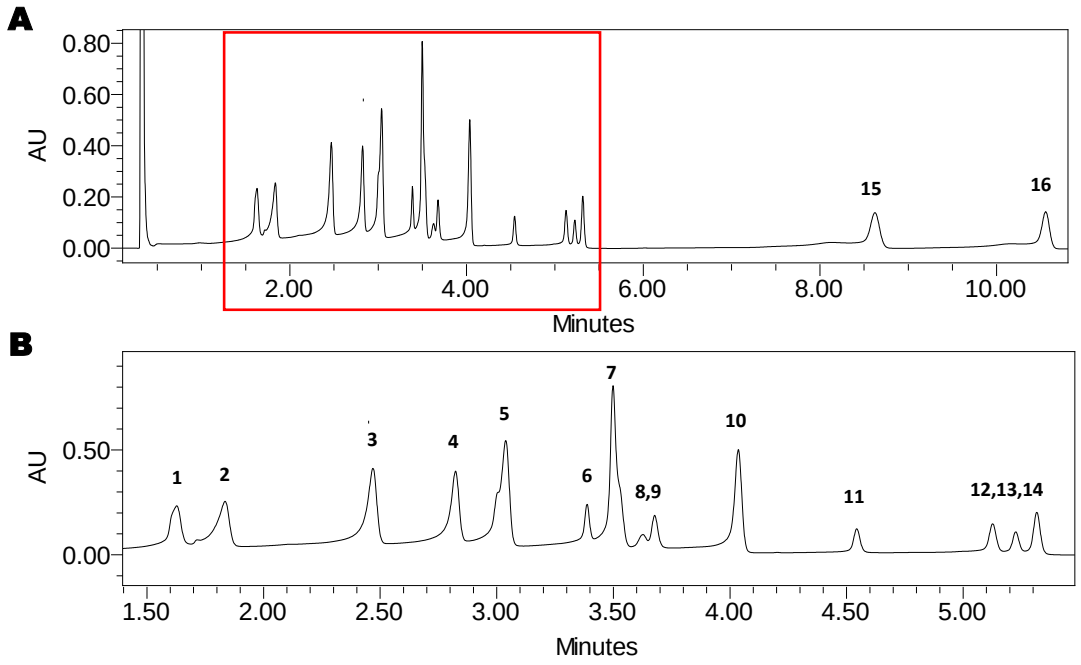

Figure 18 Method development for the separation of phenolic standards by UPLC-PDA

Standards separated using a 5 cm column, and used as development towards final method using 10 cm column. Peaks identified at 320 nm by PDA detection. (A) 16 peaks and (B) enlarged view of peaks in red box in (A). 1, chlorogenic acid; 2, caffeic acid; 3, *p*-coumaric acid; 4, ferulic acid; 5, coelution of rutin and isoferulic acid; 6, kaempferol-3-*O*-rutinoside; 7, *o*-coumaric acid; 8, naringin; 9, myricetin; 10, 3,4-dimethylcinnamic acid; 11, quercetin; 12, naringenin; 13, chalcone-naringenin; 14, kaempferol; 15, flavone; 16, 3-methoxyflavone.

The apparatus are able to accommodate two mobile phase solvent systems, used in parallel with one stationary phase. During the simultaneous optimisation of the second system for the separation of isoprenoids (section 4.3, P. Fraser), the stationary phase was substituted for a column of 10 cm in length, which is described in section 2.1.4. This resulted in greater backpressure and disrupted the optimised resolution seen in Figure 18.

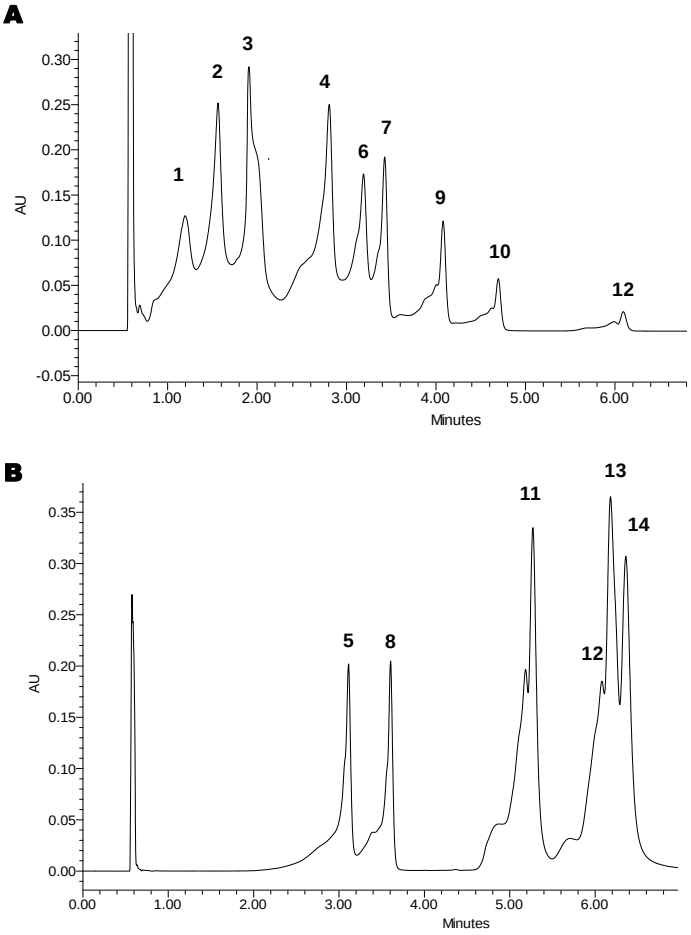

Figure 19 Separation of phenolic standards with high throughput UPLC-PDA method

Separation of 14 standards with method described in section 4.2.1 at (A) 320 nm and (B) 360 nm. 1, 3,4-dihydroxybenzoic acid; 2, chlorogenic acid; 3, caffeic acid; 4, *p*-coumaric acid; 5, rutin; 6, ferulic acid; 7, isoferulic acid; 8, kaempferol-3-*O*-rutinoside; 9, *o*-coumaric acid; 10, 3,4-dimethoxycinnamic acid; 11, quercetin; 12, naringenin; 13, chalcone-naringenin; 14, kaempferol.

In an attempt to better resolve flavonoid compounds and improve peak shape, the column temperature was firstly reduced to 30 °C and secondly to 25 °C; however, this decrease in column temperature resulted in poorer resolution of early eluting phenolic acids. A method was designed to include a column temperature gradient so that the column was at a higher temperature of 40 °C at the start of the solvent gradient for optimal phenolic acid resolution, but cooled to 30 °C towards the end of the gradient for optimal flavonoid resolution. While this resulted in some success, results were not reproducible since the speed at which the column cooled was affected by external temperature. On occasions therefore, the column was unable to cool to 30 °C by the end of the 12 min solvent gradient. Adjusting the method to exploit the Waters gradient

curve shapes facilitated in reducing the total gradient time to 9 min, as detailed by section 2.2.3.2. This allowed separation of rutin and isoferulic acid (Figure 19), which had not been achieved previously. Again, this compromised separation of naringenin and chalcone-naringenin; however, quantification at differing optimal wavelengths (280 and 360 nm, respectively) provided a solution.

4.2.2 Limitations of method

Calibration curves were conducted for 16 phenolic standards (Table 5) up to a maximum of 1.0 μg compound per injection. Limits of quantification, identification and detection were assessed, and are provided in Table 5 as compound amount per injection.

Table 5 Summary of chromatographic detection and limitations of phenolic standards for analysis by UPLC-PDA

Retention time is based on a chromatographic gradient for 9 min. Limitations provided are lower limits of quantification (LOQ), of identification (LOI), and of detection (LOD). Upper LOQ was recorded as > 1.0 μg in all cases, except 3,4-dimethoxycinnamic acid = 0.6 μg.

Compound standard	Wavelength of detection (nm)	Retention time [a] (min)	LOQ (μg)	LOI (ng)	LOD (ng)
3,4-Dihydroxybenzoic acid	280	1.19	0.03	0.5	0.1
Chlorogenic acid	320	1.53	0.15	5.0	2.0
Caffeic acid	320	1.80	0.04	5.0	0.5
p-Coumaric acid	320	2.78	0.04	1.0	0.2
Rutin	360	3.10	0.04	3.0	0.8
Ferulic acid	320	3.18	0.15	1.0	0.2
Isoferulic acid	320	3.43	0.15	1.0	0.2
m-Coumaric acid	280	3.48	0.15	4.0	0.5
Kaempferol-3-*O*-rutinoside	360	3.60	0.06	5.0	1.0
o-Coumaric acid	280	4.09	0.04	1.0	0.1
Salicylic acid [b]	320	4.16	0.08	4.0	1.0
3,4-Dimethoxycinnamic acid	320	4.70	0.01	0.2	<0.1
Quercetin	360	5.27	0.04	3.0	0.5
Naringenin	280	6.08	0.04	1.0	0.5
Chalcone-naringenin	360	6.17	0.04	4.0	0.5
Kaempferol	360	6.35	0.04	5.0	0.8

[a] approximate time provided. [b] used as internal standard

4.3 Development of high throughput UPLC method for separation of isoprenoids

4.3.1 Summary of method and assessment of limitations

The method was designed by P. Fraser (RHUL) and is described in section 2.2.3.3, comprising a gradient of 8 min. Within the scope of this study, a range of available authenticated standards were assessed for separation. Calibration curves were conducted for 15 isoprenoid and related compounds, and limitations were assessed. These are summarised in Table 6.

Table 6 Summary of chromatographic detection and limitations of isoprenoid standards for analysis by UPLC-PDA

Retention time is based on a chromatographic gradient for 8 min. Limitations provided are lower limits of quantification (LOQ), of identification (LOI), and of detection (LOD). Upper limit of quantification was recorded as > 1.0 μg in all cases, and is not shown.

Compound standard	Wavelength of detection (nm)	Retention time [a] (min)	LOQ (μg)	LOI (ng)	LOD (ng)
Neoxanthin	450	0.62	0.15	3.0	< 2.0
Violaxanthin	450	0.64	0.15	3.0	< 2.0
Antheraxanthin	450	0.82	0.15	5.0	< 2.0
Adonixanthin	450	0.95	0.20	10.0	2.0
Lutein	450	1.09	0.20	5.0	< 2.0
Zeaxanthin	450	1.10	0.30	10.0	2.0
Canthaxanthin	470	1.90	0.20	4.0	< 2.0
δ-Tocopherol	286	3.23	0.15	40.0	15.0
β-Cryptoxanthin	450	3.80	0.20	4.0	< 2.0
α-Tocopherol	286	3.93	0.15	40.0	15.0
Echinenone [b]	450	4.06	0.10	5.0	< 2.0
α-Tocopherol acetate [b]	286	4.39	0.15	20.0	10.0
δ-Carotene	450	5.22	0.15	5.0	< 2.0
ζ-Carotene	350	5.48	0.20	3.0	< 2.0
	400	5.48	0.20	3.0	5.0
β-Carotene	450	5.61	0.20	5.0	4.0

[a] approximate time provided. [b] chosen for their use as possible internal standards

4.3.2 Alternative extraction of isoprenoid compounds

Chloroform has long been used in the extraction of lipids from a range of tissues (Folch et al., 1957; Long et al., 2006); however, its substitution with MTBE has resulted in comparable extractions of lipids from *Escherichia coli*, *Caenorhabditis elegans* embryos, and mammalian tissue (Matyash et al., 2008). MTBE has been used here in the extraction of isoprenoid compounds on five tomato cultivars that include processing (TA209, M82, and San Marzano), fresh salad crop (Moneymaker), and high pigment ($hp2^j$) lines. Samples were analysed by UPLC-PDA to obtain isoprenoid profiles based on six compounds: phytoene, phytofluene, lycopene, β-carotene, lutein, and α-tocopherol (Figure 20).

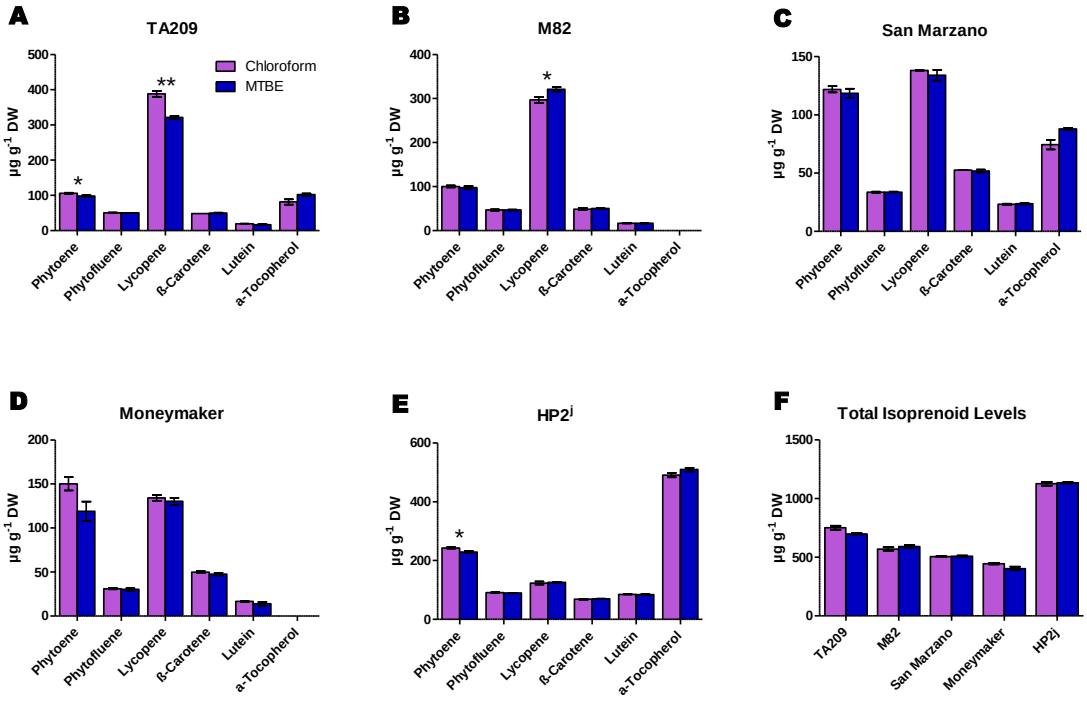

Figure 20 Isoprenoid profiles of five tomato cultivars using chloroform and MTBE for extraction.

Mean levels of isoprenoids detected by extraction with both chloroform and MTBE solvents. Error bars represent standard error of the mean (n=3). Total isoprenoid levels (F) are sum of six compounds shown within each cultivar (A to E). Levels of α-tocopherol for M82 and Moneymaker are below detection limits. Student's t-test was used to determine significant difference between two extraction methods at (*) $p \leq 0.05$ and (**) $p \leq 0.01$ limits.

For 24 out of 28 individual compound comparisons (Figure 20A to E) within a cultivar there were no significant differences in average amount of compound detected between the two extraction solvents. No significant differences were detected for four of the six isoprenoids (phytofluene, β-carotene, lutein, and α-tocopherol) for any cultivar. Further to this, two of the cultivars (San Marzano and Moneymaker) displayed no significant difference in isoprenoid profile (Figure 20C and D). Significantly higher amounts of phytoene ($p \leq 0.05$) were detected using chloroform for the extraction of material from TA209 and HP2^{j} cultivars (Figure 20A and E). Similarly, significantly more lycopene ($p \leq 0.01$) was detected in TA209 when using chloroform as the extraction solvent (Figure 20A). However, significantly more lycopene ($p \leq 0.05$) was detected in M82 using MTBE (Figure 20B). A comparison between the cultivars showed that there were no significant differences in total isoprenoid levels for any of the cultivars (Figure 20F).

One practical advantage of using MTBE over the use of chloroform is that MTBE, being less dense than the aqueous phase, forms the epiphase. The aqueous phase therefore separates MTBE extract from debris, which forms a pellet (Figure 21). Chloroform, by comparison, is more dense than the aqueous phase. Chloroform extract therefore forms the hypophase and debris accumulates at the interface of both phases. This difference therefore alleviates the need to disrupt cell debris during recovery of MTBE extract.

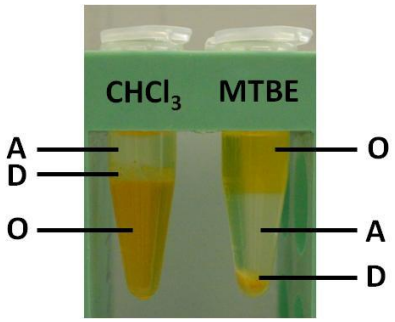

Figure 21 Extraction of isoprenoids from plant material using chloroform and MTBE.

Following centrifugation, chloroform containing isoprenoid extract comprises the hypophase and plant debris collects at the interface of the organic and aqueous phases. Where chloroform is substituted with MTBE, extract comprises the epiphase and plant debris forms a pellet below the aqueous hypophase. A is aqueous phase (methanol and Tris-HCl buffer in water), O is organic phase (chloroform or MTBE) containing isoprenoid extract, D is solid plant debris from which isoprenoids are extracted.

Conversely, one advantage of using chloroform is its greater volatility. It was noted through observation that the duration of time to take equivalent volumes of solvent to dryness by rotary evaporator was up to three times longer for MTBE than for chloroform (data not shown).

4.4 Validation of high throughput UPLC methods

The two previous methods, discussed in sections 4.2 and 4.3, which are designed for high throughput screening of, respectively, phenolic and isoprenoid profiles in large-scale populations, are assessed here for robustness. The CC colour mutant population was used since it was anticipated to contain a range of secondary metabolite pigment profiles. The population comprises 259 available CC genotypes and four elite lines (M82, Yarra, Nycos and 1107) used for comparison.

In CC genotypes, up to 13 phenolic compounds were detected (Table 7), nine of which were verified by authenticated standards. The remaining four were labelled as unknown compounds showing similarity to rutin or chlorogenic acid based on UV spectral shape. Phenolic profiles exhibiting high chlorogenic acid (CC# 6748, 4425 and 3306), high rutin (CC# 6068, 3466 and 41), and low level phenolics (CC# 294 and 1011) were detected (Table 7). Principle component analysis (PCA) of these data did not clearly cluster any genotypes, but instead separated most of the genotypes on the first principle component (PC1, representing 24.8 % of the variation, Figure 22A) in the direction of rutin or chalcone-naringenin (Figure 22B). Quality control samples were likewise distributed along PC1; however, elite cultivars 1107, Nycos and individual biological replicates of M82 displayed tighter clustering.

Profiles of isoprenoid and related hydrophobic compounds identified up to 15 compounds that were verified by authenticated standards (Table 8). Genotypes were identified that comprised high levels of the acyclic carotenes preceding lycopene (prolycopene, neurosporene and ζ-carotene; CC# 434, 1334 and 1367), high β-carotene (CC# 1337, 6631 and 6879) and the presence of chlorophylls (CC# 970, 6655 and 6748). All four elite lines and QC samples showed clustering when analysed by PCA (Figure 23A). One clearly defined cluster was formed comprising CC genotypes exhibiting changes to acyclic carotenes preceding lycopene (Figure 23B).

When datasets were combined, PCA continued to cluster genotypes based on profiles observed with isoprenoid and related compounds. M82 and QC clusters, however, were no longer clearly defined (Figure 24).

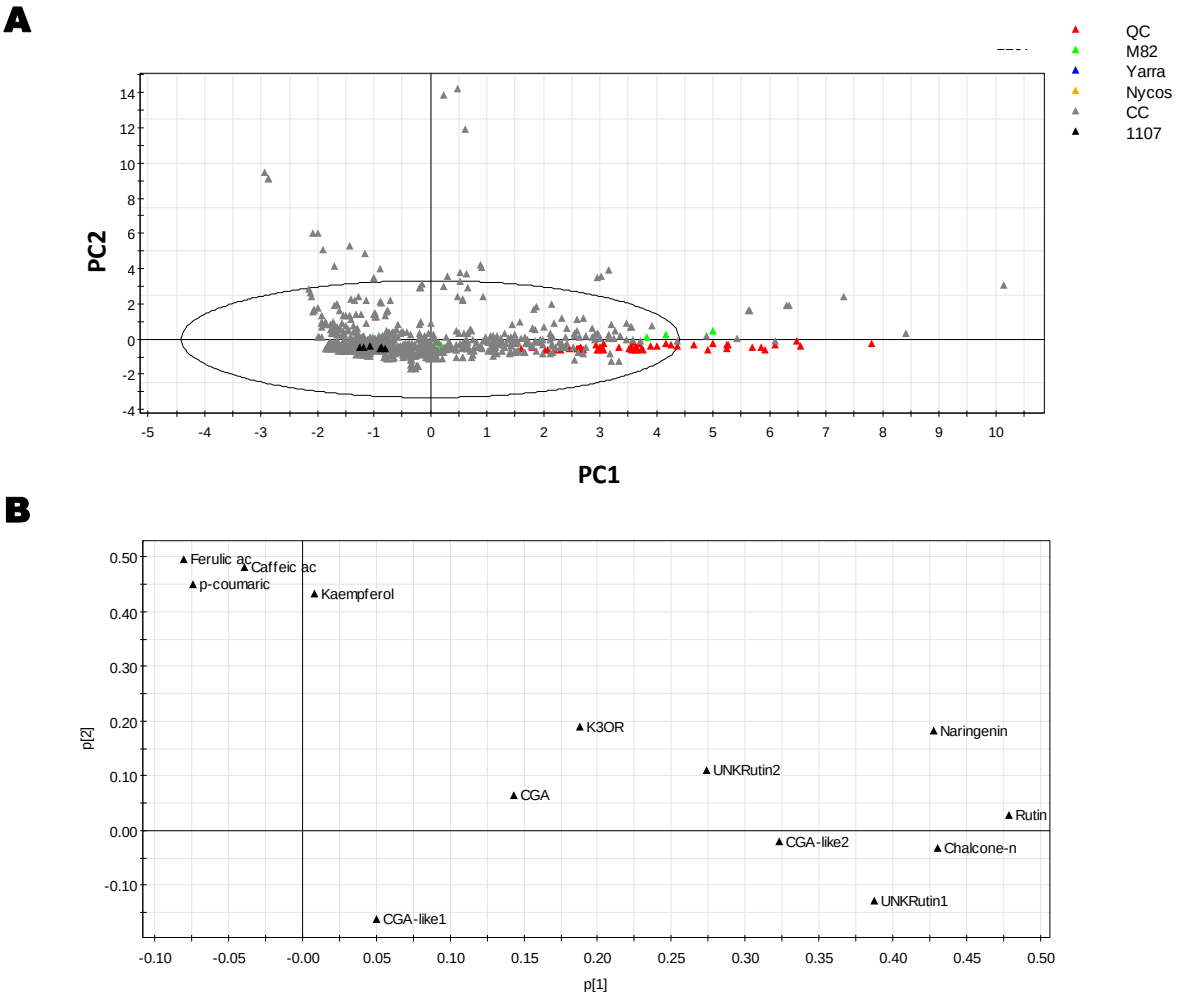

Figure 22 Principal component analysis of phenolic compounds in Core Collection Colour Mutant Population

(A) PCA score plot based on principal component (PC) 1 (24.8 %) and PC2 (14.0 %). Each accession is represented by three technical replicates. QC, quality control; CC, colour mutant accessions (grey). (B) Loadings plot for PCA. CGA, chlorogenic acid; chalcone-n, chalcone-naringenin; CGA-like1 and 2, unknown compounds similar to CGA based on spectral shape; UNK Rutin1 and 2, unknown compounds similar to rutin based on spectral shape; K3OR, kaempferol-3-*O*-rutinoside.

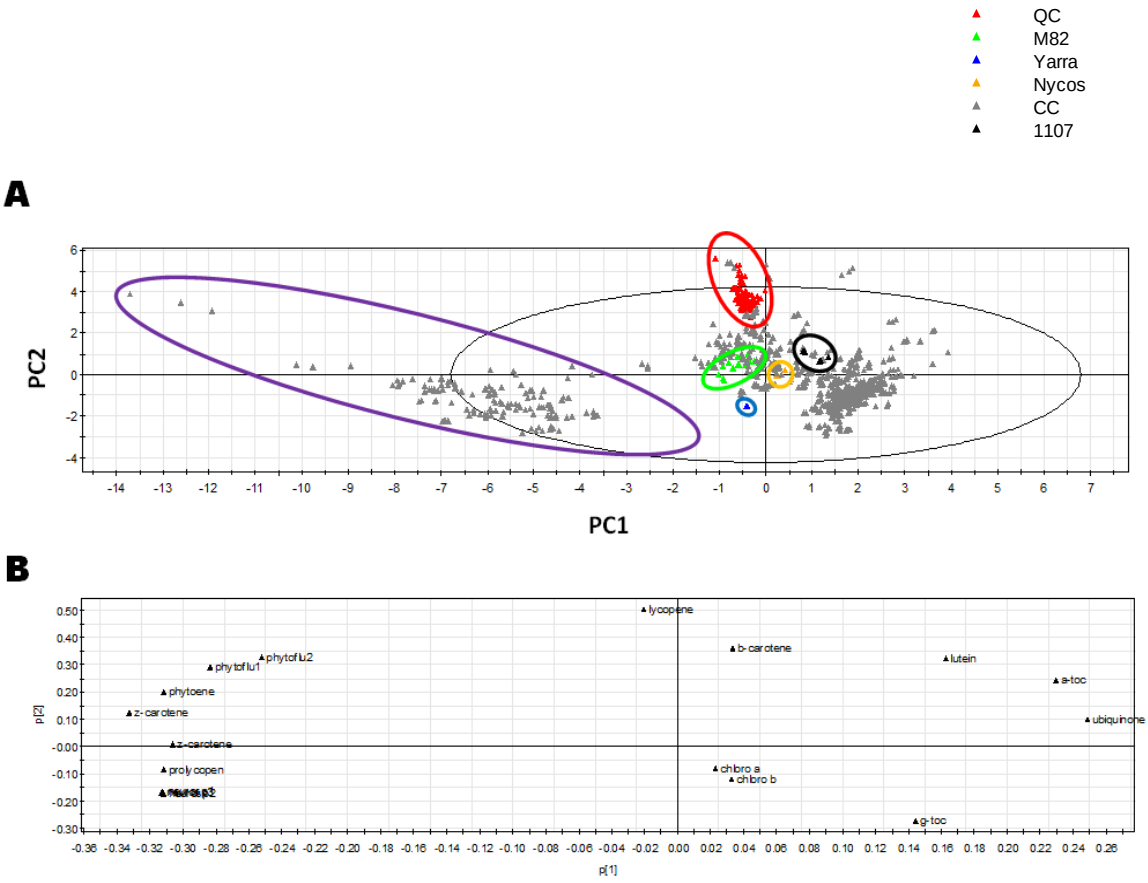

Figure 23 Principal component analysis of isoprenoid compounds in Core Collection Colour Mutant Population

(A) PCA score plot based on PC1 (45.2 %) and PC2 (17.5 %). Each accession represented by three technical replicates. QC, quality control (red); elite lines M82, Yarra, Nycos, and 1107 (green, blue, orange, and black, respectively); CC, colour mutant accessions (grey); identified cluster (purple). (B) Loadings plot for PCA. Phytoflu1 and 2, phytofluene isomers; z-carotene, ζ-carotene isomers; neurosp1, 2 and 3, neurosporene isomers; prolycopen, prolycopene; chloro a and b, chlorophyll a and b; g-toc, γ-tocopherol, a-toc, α-tocopherol.

Figure 24 (legend; figure overleaf)

(A) PCA score plot based on PC1 (26.8 %) and PC2 (15.0 %). Each accession represented by three technical replicates. QC, quality control (red); elite lines M82, Yarra, Nycos, and 1107 (green, blue, orange, and black, respectively); CC, colour mutant accessions (grey); identified cluster (purple). (B) Loadings plot for PCA. Phytoflu1 and 2, phytofluene isomers; z-carotene, ζ-carotene isomers; neurosp1, 2 and 3, neurosporene isomers; prolycopen, prolycopene; chloro a and b, chlorophyll a and b; g-toc, γ-tocopherol, a-toc, α-tocopherol; CGA, chlorogenic acid; chalcone-n, chalcone-naringenin; CGA-like1 and 2, unknown compounds similar to CGA based on spectral shape; UNK Rutin1 and 2, unknown compounds similar to rutin based on spectral shape; K3OR, kaempferol-3-*O*-rutinoside.

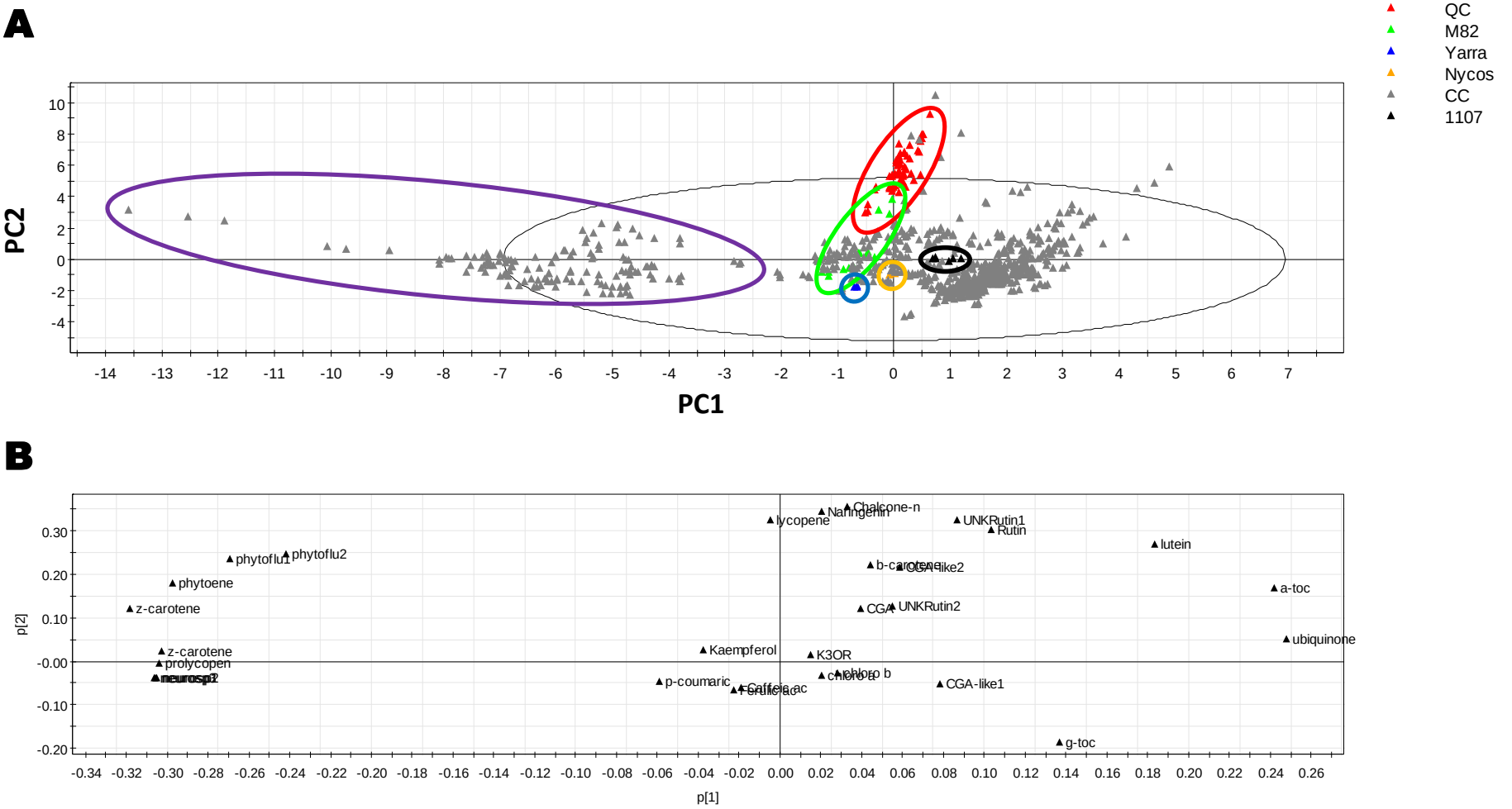

Figure 24 Principal component analysis of phenolic and isoprenoid levels in Core Collection Colour Mutant Population

Table 7 Phenolic content of available accessions from Core Collection Colour Mutant Population

Average values µg/g DW ± standard error of the mean (n=3) presented for 259 accessions, labelled by Core Collection accession number (CC#). N.D. indicates compound not detected by analytical platform. Trace indicates sample was detected, but at levels below limit of quantification. UNK represents unknown compounds and precedes an identity for which there is a similar UV spectral shape. CGA is chlorogenic acid. Nycos, Yarra, 1107, and M82 (1-5) represent elite lines cultivated alongside CC accessions. QC represents quality control replicates (n=50).

CC#	*p*-Coumaric acid	Ferulic acid	Caffeic acid	Chlorogenic acid	UNK CGA 1	UNK CGA 2	Naringenin	Chalcone-naringenin	Rutin	UNK rutin 1	UNK rutin 2	Kaempferol	Kaempferol-3-*O*-rutinoside
9	180.9 ± 11.4	618.9 ± 39.2	45.8 ± 5.1	702.8 ± 21.4	N.D.	131.5 ± 5.8	N.D.	N.D.	63.7 ± 2.3	22.1 ± 1.5	N.D.	N.D.	N.D.
10	N.D.	N.D.	N.D.	45.7 ± 0.9	N.D.	6.7 ± 1.7	30 ± 5.1	237.8 ± 13.2	504 ± 16.9	363.9 ± 12.3	N.D.	N.D.	N.D.
13	N.D.	N.D.	N.D.	79.8 ± 14.2	N.D.	11 ± 3.5	15.1 ± 1.8	N.D.	207.8 ± 3	25.7 ± 2.3	N.D.	N.D.	56 ± 0.8
27	N.D.	N.D.	N.D.	91.4 ± 3.4	19.1 ± 0.4	10.4 ± 0.6	N.D.	N.D.	166.3 ± 3.2	108.8 ± 6.2	N.D.	N.D.	N.D.
41	14.5 ± 1.9	N.D.	62.1 ± 4.9	131.7 ± 5.5	N.D.	63.3 ± 2.2	379.2 ± 53	522.7 ± 26.3	3544.8 ± 164	425.1 ± 19.2	63.9 ± 107	N.D.	191.1 ± 14.7
54	82.8 ± 6.8	62 ± 0.9	N.D.	40.8 ± 2.2	N.D.	15.5 ± 1	27.1 ± 3.6	N.D.	48.8 ± 11.5	12.1 ± 2.2	N.D.	N.D.	N.D.
83	N.D.	N.D.	N.D.	95 ± 1.9	26.8 ± 1.6	14.4 ± 1.1	N.D.	N.D.	104.7 ± 2.2	34.3 ± 4.4	N.D.	N.D.	N.D.
168	43.5 ± 2.7	N.D.	N.D.	78.1 ± 5.5	N.D.	29.9 ± 1.1	N.D.	N.D.	356.2 ± 7.2	169.4 ± 5.2	N.D.	N.D.	16.9 ± 4.1
171	N.D.	N.D.	N.D.	42.3 ± 2.5	N.D.	40.8 ± 2.6	462.6 ± 68.1	1225.8 ± 112.1	523 ± 20.2	24.1 ± 1.6	N.D.	N.D.	56.1 ± 2.5
181	58 ± 4.2	N.D.	N.D.	27.2 ± 2.8	N.D.	N.D.	18.1 ± 4.4	139.8 ± 13.7	717.4 ± 20.3	93 ± 3.5	N.D.	N.D.	41 ± 2.2
192	N.D.	N.D.	N.D.	42.5 ± 0.1	N.D.	N.D.	33.1 ± 8.5	153.3 ± 57.8	252.8 ± 9.6	110.2 ± 7.6	N.D.	N.D.	25.9 ± 3.1
195	24.9 ± 2.1	145.1 ± 12.2	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	209.7 ± 7.2	37.4 ± 3.2	N.D.	N.D.	21 ± 0.6
223	N.D.	N.D.	N.D.	64.6 ± 5	57.2 ± 3.9	35.9 ± 0.6	128.2 ± 19.6	461.2 ± 19.8	562.3 ± 22.9	85.1 ± 4.1	52.3 ± 62.8	N.D.	53.2 ± 4
258	N.D.	N.D.	N.D.	48.5 ± 5.6	24.7 ± 1.6	22.7 ± 0.7	N.D.	N.D.	254.6 ± 19	55.6 ± 3.4	N.D.	N.D.	45.1 ± 4.3
262	N.D.	N.D.	N.D.	641 ± 27.2	22 ± 2.2	59.9 ± 1.2	89.7 ± 15.4	1072.1 ± 74.9	802 ± 22.2	134.8 ± 5.9	N.D.	N.D.	112.3 ± 6.5
279	N.D.	N.D.	N.D.	25.6 ± 2.7	N.D.	N.D.	17.1 ± 3.6	7.3 ± 1.1	329.3 ± 10.5	109.2 ± 10	N.D.	N.D.	18.3 ± 0.7
280	N.D.	N.D.	80.1 ± 2.7	118.6 ± 3.3	34.9 ± 1.5	78 ± 2.5	N.D.	N.D.	514.2 ± 17	130.7 ± 1.9	N.D.	N.D.	N.D.
282	21.7 ± 1	19.7 ± 1.2	N.D.	37.4 ± 1.1	N.D.	7.6 ± 0.1	N.D.	N.D.	54.3 ± 1.6	36.4 ± 0.6	N.D.	N.D.	N.D.
294	N.D.	N.D.	N.D.	119.2 ± 2.9	N.D.	N.D.	N.D.	N.D.	47.6 ± 2.2	5 ± 1	N.D.	N.D.	N.D.
302	N.D.	N.D.	N.D.	52.3 ± 4.6	37.5 ± 2.4	13.5 ± 2.8	80.4 ± 21.9	481.5 ± 83.6	521.2 ± 72.1	77.4 ± 17.1	26.8 ± 45.2	N.D.	46.1 ± 4.1
325	N.D.	N.D.	N.D.	124.1 ± 26.2	81 ± 3.1	N.D.	50.1 ± 10.2	356.5 ± 78.9	637.4 ± 44.1	62.7 ± 5.6	N.D.	N.D.	N.D.
369	N.D.	N.D.	N.D.	17.9 ± 2.9	N.D.	N.D.	41.4 ± 8.8	161.2 ± 32	191.1 ± 31	51.1 ± 0.8	N.D.	N.D.	N.D.

Table 7 continued

375	N.D.	N.D.	39.8 ± 1.4	114.9 ± 6.5	N.D.	177.3 ± 10.5	61 ± 3.5	228.8 ± 16.1	1521.4 ± 84.3	131.5 ± 5.4	74.6 ± 61.5	N.D.	195.1 ± 9.5
383	N.D.	N.D.	N.D.	N.D.	N.D.	5.6 ± 0.2	N.D.	N.D.	102.3 ± 2.7	121.5 ± 2.1	N.D.	N.D.	N.D.
405	140.3 ± 1	124.6 ± 3.5	N.D.	39 ± 1.5	N.D.	16.4 ± 2.7	116.1 ± 7.4	341.1 ± 46.7	424.3 ± 20.5	62.5 ± 5.1	N.D.	N.D.	47 ± 0.7
412	N.D.	N.D.	N.D.	98.7 ± 4.3	N.D.	N.D.	114.8 ± 21.8	279.5 ± 44.3	1823.6 ± 27	226.2 ± 1.5	N.D.	N.D.	82.8 ± 2.5
434	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	45.5 ± 7.9	N.D.	101.2 ± 5.8	27.2 ± 1.3	N.D.	N.D.	N.D.
482	31.8 ± 6.8	414.7 ± 12.2	N.D.	148.8 ± 5.6	N.D.	352.1 ± 9.8	420.2 ± 4.4	304.7 ± 35.4	884.2 ± 10.4	29.7 ± 3	50.4 ± 41.8	N.D.	194.2 ± 8.8
488	13.4 ± 0.4	N.D.	N.D.	45.4 ± 5.2	N.D.	37.5 ± 5	100.3 ± 11.1	247.2 ± 43.2	2302.5 ± 69.3	323.9 ± 14.9	N.D.	N.D.	24.1 ± 2.2
489	44.1 ± 7.6	79.8 ± 2.1	28.7 ± 0.5	22.8 ± 1.1	N.D.	N.D.	N.D.	N.D.	104.2 ± 2.8	85.5 ± 4.4	N.D.	N.D.	9.2 ± 0.6
495	153.3 ± 18.8	N.D.	N.D.	59.2 ± 3.6	N.D.	N.D.	45.3 ± 4.8	N.D.	65.2 ± 2	11.6 ± 1.1	N.D.	N.D.	N.D.
503	N.D.	N.D.	N.D.	53.8 ± 2.8	22.6 ± 0.7	29.3 ± 1.1	N.D.	N.D.	124.5 ± 1.4	44.8 ± 2.8	N.D.	N.D.	N.D.
512	N.D.	N.D.	N.D.	62.3 ± 5.4	N.D.	20.1 ± 0.8	N.D.	N.D.	169.4 ± 4.5	18.4 ± 1.1	N.D.	N.D.	28.9 ± 2.7
523	N.D.	N.D.	N.D.	120.8 ± 8.5	N.D.	8.3 ± 1.5	50.9 ± 12.5	197.8 ± 17	222.1 ± 9.8	9.4 ± 0.8	N.D.	N.D.	N.D.
760	N.D.	133.4 ± 4.8	N.D.	146.6 ± 9.7	242.3 ± 15.6	17.7 ± 3.4	N.D.	27.2 ± 2.2	849.7 ± 18.4	127.9 ± 1.7	N.D.	N.D.	N.D.
797	N.D.	N.D.	N.D.	123.7 ± 3	98.3 ± 2.6	310.2 ± 10.8	44.3 ± 1.8	291.5 ± 59	1926.6 ± 44.9	206.3 ± 12.5	21.7 ± 23.9	N.D.	54.4 ± 3
809	N.D.	N.D.	N.D.	42.8 ± 1.8	13.7 ± 0.5	9 ± 0.2	26.6 ± 1.9	337.2 ± 41.8	554.9 ± 14.4	86.3 ± 3.6	N.D.	N.D.	89 ± 5.6
851	N.D.	N.D.	N.D.	20.7 ± 1.4	N.D.	N.D.	111.7 ± 11.3	458.2 ± 93.8	281 ± 1.2	32.4 ± 2.8	N.D.	N.D.	32.8 ± 1.6
861	N.D.	N.D.	N.D.	24.2 ± 3	N.D.	N.D.	119.6 ± 25.7	673.3 ± 147.4	359.1 ± 42.5	85.5 ± 7.3	N.D.	N.D.	14.6 ± 2.3
870	N.D.	N.D.	N.D.	49.3 ± 1.3	38.7 ± 1.6	24.8 ± 4.3	40.1 ± 2.5	892.5 ± 3.1	459.8 ± 12.3	71.6 ± 2.3	N.D.	N.D.	59.4 ± 0.8
874	N.D.	N.D.	N.D.	47 ± 0.3	40.6 ± 1.3	13.6 ± 1.3	56.7 ± 5.5	492.8 ± 137.6	380.2 ± 22.2	87.7 ± 2.7	N.D.	N.D.	43.4 ± 5.4
886	30.3 ± 1.6	50.9 ± 2.7	N.D.	42.2 ± 1.8	N.D.	N.D.	N.D.	N.D.	19 ± 3	14.8 ± 4.8	N.D.	N.D.	N.D.
896	35.9 ± 0.9	121.2 ± 6.5	19.7 ± 3.1	53.3 ± 6	N.D.	22.8 ± 7.7	N.D.	N.D.	69.7 ± 1.3	17.7 ± 1.2	N.D.	N.D.	9.9 ± 1.2
899	N.D.	N.D.	N.D.	630.1 ± 6.7	N.D.	77.1 ± 1.7	475.8 ± 61.9	2524.5 ± 198.8	683.4 ± 16.3	197.8 ± 9.5	N.D.	N.D.	39.5 ± 4.2
922	N.D.	N.D.	N.D.	63.6 ± 4.7	46.1 ± 0.6	27 ± 2.2	97.8 ± 16.5	828.5 ± 38.4	655.3 ± 17.9	225 ± 4.7	48.1 ± 64.7	N.D.	72.3 ± 0.2
923	N.D.	N.D.	N.D.	81.9 ± 3.7	N.D.	95 ± 6.2	N.D.	25.5 ± 4.4	659.4 ± 10.3	267.8 ± 2.5	N.D.	N.D.	49.9 ± 4.1
930	N.D.	N.D.	N.D.	95.8 ± 10.1	30.3 ± 2.9	58.5 ± 4.4	N.D.	N.D.	360.6 ± 11.3	391.1 ± 15	N.D.	N.D.	N.D.
933	10.9 ± 3	N.D.	N.D.	29.7 ± 3.2	N.D.	N.D.	N.D.	241.2 ± 27.6	256.7 ± 8.7	68.6 ± 3.2	N.D.	N.D.	16.7 ± 0.9
945	N.D.	N.D.	N.D.	852.2 ± 22.3	49.8 ± 0.9	40.8 ± 0.4	N.D.	9.8 ± 2.4	350.3 ± 4.9	91.2 ± 4.3	N.D.	N.D.	61.3 ± 4
948	N.D.	N.D.	N.D.	1903.1 ± 42.8	N.D.	35.2 ± 1.7	N.D.	N.D.	27.2 ± 3.1	14.2 ± 2.5	N.D.	N.D.	N.D.
949	347 ± 24.4	N.D.	56.8 ± 5.5	N.D.	N.D.	N.D.	113.9 ± 22.4	36.2 ± 5.5	402.7 ± 45.4	57.3 ± 15.4	N.D.	N.D.	59.1 ± 5.2
955	N.D.	N.D.	N.D.	57.8 ± 3.5	N.D.	64.2 ± 2.4	75.2 ± 27.9	357.6 ± 82.4	815.8 ± 19.7	441.8 ± 13.7	N.D.	N.D.	N.D.
959	N.D.	N.D.	N.D.	122.7 ± 6	218.2 ± 19.6	100.5 ± 9.3	N.D.	87.9 ± 14.6	345.7 ± 20.6	153.6 ± 10.6	N.D.	N.D.	N.D.
960	N.D.	N.D.	N.D.	121.4 ± 2.1	180.4 ± 4.1	91.3 ± 7.7	N.D.	103 ± 8.1	324.2 ± 3.6	182.3 ± 4.9	N.D.	N.D.	N.D.

Table 7 continued

963	N.D.	N.D.	N.D.	958.4 ± 31.9	N.D.	175.3 ± 3.3	116.7 ± 9.4	302.8 ± 23.1	1706.3 ± 36.8	121.6 ± 3.2	83.3 ± 107.6	N.D.	61.1 ± 1.4
970	N.D.	N.D.	N.D.	57.3 ± 5.3	N.D.	72.9 ± 8.5	68.7 ± 20.9	1033 ± 26.3	429.4 ± 31.2	213.1 ± 11.7	N.D.	N.D.	N.D.
971	31 ± 4.1	N.D.	N.D.	N.D.	N.D.	N.D.	238.5 ± 4.2	289.8 ± 49	368.3 ± 15.4	85.9 ± 7.5	19.8 ± 24.5	N.D.	17.7 ± 1.3
972	6.5 ± 0.6	N.D.	N.D.	67.8 ± 2.4	N.D.	N.D.	13.5 ± 2.1	83.2 ± 10.9	209.6 ± 2.9	39.2 ± 1	N.D.	N.D.	21.7 ± 0.3
975	N.D.	100.2 ± 4.9	107.6 ± 16	824.9 ± 51	18.2 ± 0.4	44.1 ± 5	N.D.	10.7 ± 2.4	235.2 ± 12.3	32.1 ± 3.6	N.D.	N.D.	40.4 ± 5.4
980	20.1 ± 2	26.5 ± 2.2	32.3 ± 2.8	28.2 ± 2.2	N.D.	N.D.	N.D.	N.D.	296.2 ± 4.4	264.1 ± 11.3	N.D.	N.D.	16.5 ± 3.4
1000	N.D.	N.D.	N.D.	49.9 ± 2.7	N.D.	N.D.	N.D.	N.D.	128.2 ± 3	178.6 ± 6.9	N.D.	N.D.	N.D.
1011	N.D.	N.D.	N.D.	46.3 ± 6.2	N.D.	N.D.	N.D.	N.D.	266 ± 20.5	91.7 ± 5	N.D.	N.D.	N.D.
1034	N.D.	N.D.	180 ± 0.9	96.5 ± 2	N.D.	68.3 ± 2	N.D.	N.D.	79.9 ± 4.1	23.7 ± 1.5	N.D.	N.D.	N.D.
1158	34.5 ± 2	102.7 ± 16.5	N.D.	30.8 ± 1.8	N.D.	N.D.	N.D.	N.D.	145.9 ± 4	47.3 ± 3.3	N.D.	N.D.	8.9 ± 0.4
1278	25.1 ± 3.9	N.D.	N.D.	46.1 ± 5.6	N.D.	N.D.	N.D.	60.1 ± 11.3	101.9 ± 4.5	63.4 ± 8	N.D.	N.D.	N.D.
1316	21.8 ± 1.1	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	126.3 ± 12.8	147.7 ± 3	16 ± 1.7	N.D.	N.D.	N.D.
1321	39.9 ± 2.9	N.D.	N.D.	42.9 ± 1.1	N.D.	15.7 ± 1.9	N.D.	N.D.	213 ± 7	109.3 ± 4.6	N.D.	N.D.	21.4 ± 1.7
1323	N.D.	N.D.	N.D.	54.3 ± 4.4	19.3 ± 1.2	14.3 ± 1.7	N.D.	49.6 ± 8.1	460.3 ± 23.4	90.8 ± 5	N.D.	N.D.	67.2 ± 2.6
1324	N.D.	N.D.	N.D.	22 ± 2.8	N.D.	5.6 ± 0.4	11.4 ± 1.4	N.D.	363.8 ± 10.7	140.4 ± 4.2	N.D.	N.D.	22.2 ± 2.3
1325	N.D.	N.D.	N.D.	67.8 ± 5.6	N.D.	22.4 ± 1.4	N.D.	N.D.	350.4 ± 17.9	157 ± 7.5	N.D.	N.D.	N.D.
1326	76.1 ± 4.1	N.D.	N.D.	42.1 ± 1	N.D.	13.1 ± 0.6	N.D.	N.D.	179.5 ± 3.6	83.2 ± 1.7	N.D.	N.D.	10.1 ± 1
1327	225.7 ± 3.2	109.8 ± 4	30.2 ± 1.9	71.8 ± 1.2	N.D.	79.4 ± 2	N.D.	N.D.	819.5 ± 8.4	313.5 ± 4.4	N.D.	N.D.	35.6 ± 1.7
1328	N.D.	N.D.	N.D.	62.2 ± 2.8	N.D.	20.2 ± 2.5	N.D.	N.D.	499.7 ± 4.6	188 ± 2.6	N.D.	N.D.	34.7 ± 3.2
1329	N.D.	N.D.	N.D.	436.6 ± 6.3	18.7 ± 0.9	38.4 ± 2.4	N.D.	N.D.	447 ± 25.1	120.3 ± 9.9	N.D.	N.D.	38.3 ± 2.2
1330	N.D.	N.D.	N.D.	20.7 ± 1.4	N.D.	6.7 ± 1.1	N.D.	N.D.	388.4 ± 3.5	123.1 ± 7.4	N.D.	N.D.	25.5 ± 1.9
1334	N.D.	N.D.	28.7 ± 1.1	87.1 ± 1.3	N.D.	67.4 ± 2.1	30.2 ± 1	314.9 ± 28.8	53.7 ± 2.1	100.6 ± 5.2	N.D.	N.D.	N.D.
1336	N.D.	N.D.	N.D.	67.4 ± 1.2	N.D.	34.1 ± 0.6	10.2 ± 0.7	271.5 ± 19.7	398.3 ± 20.8	80.2 ± 2.3	N.D.	N.D.	44.7 ± 2.9
1337	64.3 ± 0.2	N.D.	N.D.	69.1 ± 2.7	N.D.	9 ± 1.4	14.9 ± 1.8	49.4 ± 5.4	222.3 ± 7.4	36.9 ± 7.5	N.D.	N.D.	48.6 ± 2.7
1338	N.D.	N.D.	N.D.	45.2 ± 1.1	N.D.	N.D.	29.4 ± 7	112.1 ± 10.7	1387 ± 78.3	279.4 ± 11.2	64.8 ± 35.9	N.D.	19.7 ± 0.5
1339	N.D.	N.D.	N.D.	47.3 ± 0.9	N.D.	13.2 ± 1.7	N.D.	36.6 ± 7	311.9 ± 2.5	121.8 ± 10.9	N.D.	N.D.	36.9 ± 5.5
1340	72.7 ± 4	292.7 ± 5.1	47.3 ± 7.6	386.1 ± 20.4	N.D.	36.6 ± 0.8	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
1344	68.4 ± 1.3	251.5 ± 8.5	N.D.	56.1 ± 3.1	N.D.	13.1 ± 2.1	N.D.	N.D.	74.9 ± 4.9	25.3 ± 0.1	N.D.	N.D.	N.D.
1345	N.D.	N.D.	N.D.	131.5 ± 4.7	N.D.	52.2 ± 2.8	153.8 ± 41.2	664.2 ± 237.4	536.4 ± 35	198.5 ± 5.1	N.D.	N.D.	N.D.
1346	N.D.	N.D.	N.D.	37.7 ± 2.9	N.D.	N.D.	30.6 ± 1.4	147 ± 20.1	275.9 ± 16.4	49.1 ± 2.7	N.D.	N.D.	N.D.
1349	9.7 ± 1.8	N.D.	N.D.	40.2 ± 3	12.8 ± 1.4	7.4 ± 1.1	N.D.	N.D.	189.1 ± 9	79.4 ± 1.5	N.D.	N.D.	18.2 ± 1
1350	37.9 ± 1.1	N.D.	N.D.	57.4 ± 3.3	N.D.	20.3 ± 1.2	N.D.	N.D.	66 ± 0.7	56.6 ± 2.1	N.D.	N.D.	N.D.

Table 7 continued

1351	N.D.	N.D.	N.D.	43.3 ± 1.7	N.D.	15.5 ± 0.6	16.9 ± 2.1	95.8 ± 10.3	188.4 ± 7.6	17.5 ± 1.3	N.D.	N.D.	39.1 ± 2.6
1352	N.D.	N.D.	N.D.	60.9 ± 1.3	N.D.	N.D.	N.D.	N.D.	280.6 ± 11.7	61.4 ± 5.4	N.D.	N.D.	48.4 ± 4.3
1353	N.D.	N.D.	202.8 ± 5.6	691.3 ± 15.8	N.D.	21.9 ± 1.4	N.D.	N.D.	271.3 ± 7.6	93.8 ± 9.1	N.D.	N.D.	N.D.
1355	N.D.	N.D.	N.D.	790.5 ± 16.9	133.2 ± 2	106.2 ± 2.2	39.5 ± 1.9	762.3 ± 7	336 ± 28.2	86.6 ± 5.5	N.D.	N.D.	N.D.
1356	N.D.	N.D.	N.D.	96.3 ± 2.1	141.9 ± 4.1	74.5 ± 4.8	71.5 ± 22.2	702.3 ± 23.2	258.1 ± 9.2	79.9 ± 1.4	N.D.	N.D.	N.D.
1361	N.D.	N.D.	N.D.	33.1 ± 2.2	N.D.	N.D.	N.D.	N.D.	142.1 ± 11.2	23.5 ± 0.7	N.D.	N.D.	21.6 ± 3.9
1363	N.D.	N.D.	N.D.	76.4 ± 2.1	N.D.	38.1 ± 1.7	N.D.	69.8 ± 6.4	657.4 ± 7.9	264.5 ± 1.4	N.D.	N.D.	N.D.
1367	N.D.	N.D.	N.D.	45.8 ± 3.2	N.D.	N.D.	9.8 ± 2.4	117.4 ± 4.7	154.7 ± 3.2	24.3 ± 1.9	N.D.	N.D.	18.1 ± 1.5
1368	N.D.	N.D.	N.D.	78 ± 7	N.D.	7.2 ± 1.3	N.D.	22.4 ± 4.4	204.9 ± 6.8	67.4 ± 5.6	N.D.	N.D.	26.4 ± 3.5
1370	49.8 ± 2.1	N.D.	N.D.	43.6 ± 2	N.D.	N.D.	159.7 ± 2.7	1469.5 ± 442.8	980 ± 130.2	497.4 ± 38.2	37.1 ± 26	N.D.	32.8 ± 6.5
1371	10.3 ± 1.1	N.D.	N.D.	34.1 ± 1.5	N.D.	N.D.	N.D.	N.D.	167.8 ± 3.8	91.1 ± 3.1	N.D.	N.D.	N.D.
1374	N.D.	N.D.	N.D.	49.9 ± 0.8	N.D.	N.D.	N.D.	N.D.	137.3 ± 9.8	24.7 ± 4.2	N.D.	N.D.	N.D.
1377	59.9 ± 3.1	N.D.	N.D.	56.3 ± 0.7	N.D.	21.6 ± 3	60.7 ± 1.1	96.4 ± 3.5	293.9 ± 10.8	151.2 ± 12.2	N.D.	N.D.	N.D.
1378	N.D.	N.D.	28.2 ± 2	50.6 ± 1.1	N.D.	40 ± 1.4	48.5 ± 3.4	540.3 ± 67.9	225.8 ± 9.4	N.D.	N.D.	N.D.	43.1 ± 2.1
1380	N.D.	N.D.	N.D.	26.8 ± 2.5	N.D.	N.D.	73.7 ± 8.3	40.9 ± 4.6	505.3 ± 3.4	105 ± 4.7	N.D.	N.D.	27.7 ± 2.2
1631	N.D.	N.D.	N.D.	62.9 ± 1.6	N.D.	21.5 ± 0.5	109.3 ± 5	307.4 ± 7.7	155.1 ± 4.7	17.4 ± 6.1	N.D.	N.D.	N.D.
1637	N.D.	N.D.	N.D.	63.5 ± 1.9	N.D.	20 ± 0.5	196.1 ± 32.3	582.4 ± 88.5	632 ± 18.3	27.4 ± 2.1	138.7 ± 114.5	N.D.	179.8 ± 5.8
1719	N.D.	1156.2 ± 22.1	240.7 ± 1.8	853.6 ± 21	N.D.	19.5 ± 0.7	N.D.	N.D.	36.7 ± 0.5	6.6 ± 0.4	N.D.	N.D.	N.D.
1724	N.D.	N.D.	N.D.	163.9 ± 1.8	N.D.	25.7 ± 1	N.D.	N.D.	194.6 ± 6.5	28.7 ± 1.8	N.D.	N.D.	N.D.
1729	N.D.	N.D.	N.D.	397.6 ± 7.4	16.5 ± 1.7	31.4 ± 0.6	9 ± 5.3	314.7 ± 4.6	954.4 ± 12.9	268.3 ± 11.5	28.9 ± 27.2	N.D.	59.9 ± 2.5
1784	N.D.	N.D.	N.D.	2461.1 ± 109.2	N.D.	N.D.	N.D.	N.D.	128.3 ± 3.3	61.5 ± 3.7	N.D.	N.D.	N.D.
1836	N.D.	N.D.	37.1 ± 3.6	28.5 ± 0.7	N.D.	16.3 ± 1.3	231.8 ± 40.9	1468.9 ± 150.6	710.3 ± 30.6	13.3 ± 1.3	N.D.	N.D.	43.3 ± 4.5
1899	N.D.	N.D.	N.D.	1608.7 ± 33	N.D.	N.D.	N.D.	N.D.	33.1 ± 2	N.D.	N.D.	N.D.	N.D.
1913	N.D.	N.D.	N.D.	85.9 ± 7.9	N.D.	44.9 ± 2.2	61.1 ± 9	547.9 ± 121.8	1250.6 ± 38.3	380.8 ± 8.1	N.D.	N.D.	33.7 ± 3.5
1941	45.8 ± 0.6	N.D.	N.D.	61.4 ± 2.1	23.4 ± 1.3	27.8 ± 1.4	38.3 ± 8.4	223.3 ± 15.2	638.8 ± 7.3	27.8 ± 1.3	N.D.	N.D.	105.1 ± 2.1
1980	N.D.	N.D.	N.D.	58.2 ± 2.7	N.D.	21 ± 2	11.7 ± 3.2	N.D.	281.3 ± 3.7	76.6 ± 1.8	N.D.	N.D.	66.4 ± 2.2
1982	14.8 ± 0.4	N.D.	N.D.	44.4 ± 2.4	32.2 ± 2.1	N.D.	N.D.	68.5 ± 15.7	286.6 ± 19.8	62.7 ± 5.6	N.D.	N.D.	N.D.
1995	N.D.	N.D.	N.D.	93.6 ± 7.4	67.9 ± 3.5	63.9 ± 5.8	N.D.	55 ± 7.7	186.2 ± 13.6	46 ± 5.4	N.D.	N.D.	29.9 ± 5.4
2007	18.8 ± 1.9	N.D.	N.D.	43.9 ± 1.2	N.D.	9.6 ± 0.9	35.7 ± 3.7	199 ± 10.1	348.8 ± 22.1	61.5 ± 6	N.D.	N.D.	36.2 ± 3.5
2042	N.D.	N.D.	N.D.	122.2 ± 3.8	41 ± 0.7	92.4 ± 1.9	21.7 ± 4.1	122.5 ± 11.7	403.1 ± 9.6	149.4 ± 8.4	30.8 ± 24.1	N.D.	N.D.
2043	N.D.	N.D.	N.D.	26.7 ± 3.7	N.D.	N.D.	37.9 ± 10.2	192.7 ± 16.6	172.6 ± 4.3	14.7 ± 4.2	N.D.	N.D.	24 ± 2.9
2058	N.D.	N.D.	N.D.	670 ± 7.6	81.4 ± 1.3	166.4 ± 8.9	75.2 ± 4.7	250.6 ± 20.7	1591.4 ± 53.3	260.9 ± 6	N.D.	N.D.	118.9 ± 5.5

Table 7 continued

2065	N.D.	N.D.	N.D.	45.3 ± 2.2	36.8 ± 1	43.9 ± 1	76.7 ± 6.9	1046.8 ± 71.2	541.3 ± 31.7	114.5 ± 9.4	N.D.	N.D.	42.6 ± 5.9
2066	N.D.	N.D.	N.D.	61.1 ± 5.3	54.1 ± 7.9	30.5 ± 2	68.2 ± 13.5	698 ± 145.7	504.2 ± 57.7	106.5 ± 11.9	N.D.	N.D.	14.9 ± 0.7
2067	N.D.	N.D.	N.D.	49.5 ± 2.3	30.5 ± 1	21.3 ± 2	11.6 ± 4	96.4 ± 6.1	405.1 ± 19.8	83.2 ± 2	N.D.	N.D.	39.4 ± 3.2
2154	N.D.	N.D.	N.D.	7.6 ± 1.3	N.D.	N.D.	157.3 ± 14.8	797.5 ± 52.2	814.3 ± 49.4	177.6 ± 7.7	47.9 ± 43.5	N.D.	80.6 ± 4
2229	23.6 ± 0.9	N.D.	N.D.	34.7 ± 0.9	N.D.	31.9 ± 1.2	N.D.	N.D.	241.1 ± 6.3	123.1 ± 7.1	N.D.	N.D.	N.D.
2340	N.D.	N.D.	N.D.	64.4 ± 2.3	N.D.	56.5 ± 0.4	N.D.	70.5 ± 13.4	464.4 ± 23.7	48 ± 4.7	N.D.	N.D.	70.1 ± 0.4
2436	N.D.	N.D.	N.D.	141.4 ± 2.8	59.8 ± 1.6	204.1 ± 2.2	208 ± 27.9	533 ± 21.2	614.4 ± 9.3	213.5 ± 4.7	N.D.	N.D.	N.D.
2466	N.D.	N.D.	N.D.	589.9 ± 4.4	N.D.	109.8 ± 3.1	260.3 ± 20.4	1881.1 ± 103.5	385.4 ± 2.9	43.1 ± 0.9	N.D.	N.D.	57.7 ± 4.5
2553	N.D.	N.D.	N.D.	50.8 ± 2.5	75.2 ± 2.7	19.6 ± 1.7	17.9 ± 2.2	277.1 ± 21.7	725.8 ± 36.9	95.8 ± 6.9	N.D.	N.D.	14.5 ± 3
2569	N.D.	N.D.	N.D.	82.1 ± 4.4	52.6 ± 1.2	75.3 ± 4.5	162 ± 4.4	1474.3 ± 77.5	585.4 ± 3.2	70 ± 3	N.D.	N.D.	112.3 ± 6.5
2601	N.D.	N.D.	N.D.	34.5 ± 3.4	N.D.	17.9 ± 0.9	26.5 ± 1.5	240.9 ± 24.9	393.6 ± 9.6	32.6 ± 2.6	N.D.	N.D.	113.3 ± 1.8
2607	29.5 ± 3.2	19.3 ± 2.2	N.D.	65.9 ± 7.5	N.D.	14.6 ± 0.6	24.7 ± 4	58.8 ± 4.4	255.2 ± 25.9	23 ± 5.4	N.D.	N.D.	34.9 ± 2.5
2672	18.8 ± 0.7	N.D.	N.D.	76.2 ± 6.7	N.D.	79.4 ± 2.1	63.7 ± 2.6	371.2 ± 60.6	784.3 ± 23.8	76.6 ± 3.3	29.6 ± 26.1	N.D.	125.3 ± 2.9
2681	N.D.	260.3 ± 39	N.D.	33.9 ± 0.7	N.D.	18 ± 3.1	11.5 ± 3.4	109.2 ± 2	228.7 ± 5.7	64.7 ± 1	N.D.	N.D.	11.7 ± 1.1
2685	N.D.	N.D.	N.D.	37.5 ± 2.5	N.D.	N.D.	314.4 ± 11.5	2362.6 ± 43.3	1299 ± 26.2	87.8 ± 2.6	131.4 ± 136.1	N.D.	86 ± 3.5
2687	16.5 ± 1.2	N.D.	N.D.	60.8 ± 0.7	N.D.	28 ± 1.7	22.4 ± 1.8	248.9 ± 8.2	215.5 ± 9.7	9.6 ± 1.2	N.D.	N.D.	66.3 ± 0.9
2694	234.8 ± 6.2	263.8 ± 19.1	92.8 ± 14.2	52.6 ± 1.6	N.D.	30.4 ± 2.4	590.3 ± 30.6	52.1 ± 5.2	422.5 ± 3.6	50.7 ± 3.8	N.D.	36.2 ± 2.2	41.2 ± 1.8
2704	N.D.	N.D.	N.D.	641.5 ± 34.7	N.D.	162 ± 16.5	N.D.	N.D.	1012.4 ± 348.4	337.9 ± 89.7	N.D.	N.D.	34.4 ± 1.6
2705	N.D.	N.D.	N.D.	105.9 ± 4.1	N.D.	72.8 ± 2.1	312.9 ± 27.4	869.1 ± 149.1	842.1 ± 31.3	44.5 ± 4.9	40.6 ± 42.9	N.D.	134.8 ± 4.3
2707	41.7 ± 4.1	N.D.	N.D.	100.5 ± 1.6	122.1 ± 6	67.9 ± 1.8	47.5 ± 8.3	316.8 ± 17.5	169.7 ± 8.3	27.1 ± 2.9	N.D.	N.D.	N.D.
2710	N.D.	N.D.	N.D.	57.2 ± 4.4	44.3 ± 3.2	34.4 ± 1.6	80.1 ± 5.4	603.1 ± 85.9	463.2 ± 17.7	198.2 ± 0.3	23.3 ± 28.1	N.D.	49.9 ± 4.8
2715	N.D.	N.D.	N.D.	64.4 ± 2.4	N.D.	N.D.	N.D.	75.1 ± 1.2	368.9 ± 25.5	41.1 ± 4.5	N.D.	N.D.	32.6 ± 3.8
2721	N.D.	N.D.	N.D.	107.1 ± 1	82.8 ± 3.7	97.3 ± 3.5	58.1 ± 2.8	855.5 ± 54.1	489.3 ± 31.9	58.2 ± 6	N.D.	N.D.	105.7 ± 2.5
2723	32.7 ± 3.3	N.D.	N.D.	74.3 ± 2.8	N.D.	51.1 ± 1.8	N.D.	37.7 ± 9	1523.8 ± 64.4	195.6 ± 13.5	28.8 ± 31.6	N.D.	43.7 ± 4.4
2733	N.D.	N.D.	N.D.	64.9 ± 2.9	51.3 ± 1.8	27.2 ± 4	151.5 ± 21.3	996.6 ± 106.5	448.5 ± 18.5	142.4 ± 7	N.D.	N.D.	36.6 ± 3.4
2739	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	30.1 ± 4.9	271.6 ± 76.3	473.7 ± 68.2	24 ± 10.1	N.D.	N.D.	119.8 ± 18.4
2761	N.D.	N.D.	N.D.	90.1 ± 2.4	N.D.	20.5 ± 1.6	11.7 ± 3.3	N.D.	377.7 ± 5.8	103.3 ± 4.4	N.D.	N.D.	N.D.
2768	N.D.	N.D.	30.2 ± 1.3	629.1 ± 10	132.6 ± 1.4	292.5 ± 11.2	26.2 ± 4.3	122.9 ± 16.3	1298.5 ± 39.8	167.9 ± 5.8	N.D.	N.D.	N.D.
2790	N.D.	N.D.	N.D.	55.3 ± 3.8	N.D.	11.7 ± 1.5	N.D.	N.D.	26.5 ± 3.4	6.5 ± 0.2	N.D.	N.D.	N.D.
2872	N.D.	N.D.	N.D.	140.6 ± 1.5	78.2 ± 2.9	149.9 ± 7.6	54.9 ± 11.8	1016.4 ± 74.5	1436.5 ± 34.5	232.5 ± 4.5	50 ± 48.4	N.D.	88.1 ± 0.7
2903	N.D.	N.D.	N.D.	115 ± 4.7	28.2 ± 1	45.2 ± 1.7	133.8 ± 19.3	446.6 ± 35.2	1672.6 ± 67.1	219.1 ± 14.1	87.8 ± 71.2	N.D.	104.3 ± 5.9
2924	N.D.	N.D.	28.5 ± 1.7	95.2 ± 6.8	185.1 ± 8.9	78.3 ± 4.9	N.D.	38.4 ± 4.1	270.4 ± 9.9	111.3 ± 7.6	N.D.	N.D.	N.D.

Table 7 continued

2925	N.D.	N.D.	N.D.	73.9 ± 5.2	97.6 ± 8.3	53.3 ± 2	N.D.	9.4 ± 2	187.6 ± 13.9	36.1 ± 4.7	N.D.	N.D.	32.1 ± 6
2927	N.D.	N.D.	46.1 ± 1.4	515.7 ± 11.2	N.D.	34.4 ± 0.5	N.D.	88.8 ± 10.5	336.3 ± 7.2	141.7 ± 2.2	N.D.	N.D.	N.D.
2931	N.D.	N.D.	N.D.	20.3 ± 1.1	N.D.	11.1 ± 1.3	N.D.	N.D.	165.5 ± 81.3	52.4 ± 19.9	N.D.	N.D.	N.D.
2939	31.9 ± 2.9	N.D.	66.1 ± 4.2	40.5 ± 1.5	N.D.	21.3 ± 1.5	318.8 ± 25	2586.1 ± 194.4	572.3 ± 20.1	N.D.	24.2 ± 32.6	N.D.	132.1 ± 8.4
2968	N.D.	N.D.	N.D.	121.1 ± 6.7	84.4 ± 1.5	69.8 ± 4.3	364.8 ± 27.2	946.6 ± 332.3	854.4 ± 24.7	186 ± 11.6	46.8 ± 50.8	N.D.	N.D.
2979	N.D.	N.D.	N.D.	50.2 ± 2.8	N.D.	9.2 ± 0.4	N.D.	N.D.	226.1 ± 9.2	160.3 ± 6.1	N.D.	N.D.	N.D.
2989	N.D.	N.D.	N.D.	26.1 ± 0.6	N.D.	N.D.	N.D.	5.6 ± 1.3	117.7 ± 0.4	9.3 ± 0.7	N.D.	N.D.	15.7 ± 1.3
2995	N.D.	N.D.	N.D.	47.3 ± 3.6	153.5 ± 18.3	67.6 ± 2.9	203.8 ± 64.8	1455.6 ± 251.9	1081.5 ± 58.6	82.4 ± 8.2	N.D.	N.D.	294.8 ± 21
2997	N.D.	N.D.	N.D.	103.7 ± 4.7	65.8 ± 2.6	76.9 ± 0.9	97.9 ± 14.7	345.5 ± 13	418.1 ± 24.1	97.2 ± 3.3	N.D.	N.D.	N.D.
3052	19.1 ± 1.8	N.D.	N.D.	49.5 ± 2.3	N.D.	8.8 ± 1	6.3 ± 1.1	24.2 ± 2.9	297 ± 4	66.3 ± 4	N.D.	N.D.	31.1 ± 1.6
3229	97.1 ± 11.7	N.D.	N.D.	68.9 ± 3.2	N.D.	N.D.	135.3 ± 4.8	144.4 ± 16.7	155.9 ± 2.5	10.2 ± 0.6	N.D.	N.D.	N.D.
3294	N.D.	N.D.	N.D.	48.8 ± 1.2	52.2 ± 0.6	32.5 ± 0.8	11.9 ± 0.2	92.1 ± 5.2	361.8 ± 4.6	173.7 ± 0.7	N.D.	N.D.	N.D.
3306	N.D.	N.D.	N.D.	5189.3 ± 196.5	N.D.	N.D.	N.D.	203.1 ± 10.5	680.7 ± 25.9	150 ± 1.3	28.4 ± 44.7	N.D.	29.5 ± 0.2
3315	N.D.	N.D.	N.D.	74 ± 4	84.8 ± 5.1	25.5 ± 3.2	N.D.	72 ± 11.1	268.2 ± 22.3	159.2 ± 16.1	N.D.	N.D.	N.D.
3362	N.D.	N.D.	N.D.	27.1 ± 0.8	N.D.	N.D.	20.2 ± 5.9	156 ± 10.5	451.5 ± 17.4	135.3 ± 4.4	N.D.	N.D.	30.9 ± 2.1
3377	N.D.	N.D.	96.2 ± 41.1	44 ± 18.4	N.D.	N.D.	21.6 ± 10	179.6 ± 79.1	410.6 ± 176	60.4 ± 26.3	N.D.	N.D.	37.8 ± 16.4
3380	N.D.	N.D.	N.D.	111 ± 8.4	N.D.	14.5 ± 1.2	90.1 ± 7.4	358.8 ± 30.1	492.2 ± 16.3	117.1 ± 9.3	N.D.	N.D.	57.9 ± 2.8
3424	N.D.	N.D.	N.D.	118.8 ± 4.3	38.1 ± 2.4	15.2 ± 1.3	238 ± 45.8	1282.6 ± 218.3	991.9 ± 44.8	136.1 ± 9.8	77.5 ± 71	N.D.	84.6 ± 6
3443	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	73.3 ± 17.1	311.1 ± 30.3	82 ± 9.6	N.D.	N.D.	11 ± 1.1
3461	28 ± 1.2	N.D.	N.D.	109.6 ± 7.6	N.D.	N.D.	57.6 ± 1.4	139.1 ± 20.7	119 ± 3.6	12.5 ± 0	N.D.	N.D.	N.D.
3463	N.D.	N.D.	N.D.	26.5 ± 0.7	N.D.	19.9 ± 1	69.6 ± 14.9	448.8 ± 21.4	460.7 ± 18.4	91.7 ± 9.4	40.2 ± 34.4	N.D.	8.8 ± 0.9
3466	50.7 ± 5.3	N.D.	N.D.	104.9 ± 6.7	108.7 ± 6.4	377 ± 32.2	50.1 ± 11.2	431.9 ± 65.1	2560.2 ± 71.9	361.5 ± 16.8	72 ± 63.4	N.D.	119 ± 7.5
3470	N.D.	N.D.	N.D.	92.7 ± 7.1	N.D.	87 ± 4.1	52.9 ± 10	583.2 ± 83.4	412.7 ± 19.7	50.1 ± 0.6	N.D.	N.D.	72.1 ± 5.8
3485	N.D.	N.D.	N.D.	47.7 ± 5	N.D.	N.D.	39 ± 4.7	197.8 ± 20.5	265.4 ± 6.5	135.3 ± 4.1	N.D.	N.D.	14.6 ± 2.5
3495	N.D.	N.D.	N.D.	61.9 ± 1	25.3 ± 0.8	11.6 ± 0.3	39.4 ± 4.7	123.3 ± 48.9	398.2 ± 10.4	64.6 ± 1.3	N.D.	N.D.	23.4 ± 2.2
3509	86.8 ± 2.8	84.5 ± 4.2	33.1 ± 5.7	53.2 ± 1	N.D.	15.4 ± 0.7	N.D.	N.D.	182.2 ± 4.7	100.9 ± 5.2	N.D.	N.D.	25.8 ± 5.7
3510	N.D.	N.D.	N.D.	84.3 ± 6.8	31.2 ± 1.5	67.3 ± 38	N.D.	70.3 ± 7.7	379.2 ± 22.3	43.8 ± 1.9	N.D.	N.D.	68.8 ± 2.4
3511	N.D.	N.D.	N.D.	41.2 ± 0.3	N.D.	N.D.	138 ± 31.2	1015.4 ± 135.6	816.2 ± 23	300.2 ± 4.6	119.7 ± 123.5	N.D.	43.4 ± 2.5
3512	N.D.	N.D.	N.D.	64.7 ± 5.8	N.D.	30.3 ± 3.5	N.D.	N.D.	346.5 ± 19.3	127.2 ± 11.1	N.D.	N.D.	22 ± 1.8
3518	28.5 ± 0.9	N.D.	N.D.	N.D.	N.D.	72.9 ± 0.9	N.D.	N.D.	881.2 ± 39	132.5 ± 8.1	N.D.	N.D.	74.6 ± 5.3
3539	N.D.	33.5 ± 5.5	N.D.	99.7 ± 6.9	53.5 ± 6.2	149.9 ± 8.4	29.4 ± 1.5	300.6 ± 50.3	1183.1 ± 122.4	212.6 ± 7.4	47 ± 39.4	N.D.	52.4 ± 2.1
3545	N.D.	N.D.	50.5 ± 3.1	82.5 ± 6.9	N.D.	47 ± 3	20.2 ± 0.5	N.D.	215.8 ± 13.7	109.9 ± 6.7	N.D.	N.D.	N.D.

Table 7 continued

3546	N.D.	N.D.	N.D.	44.7 ± 1.7	N.D.	N.D.	182.8 ± 19	1775.8 ± 148.7	355.2 ± 7.2	54.8 ± 3	N.D.	N.D.	N.D.
3547	N.D.	N.D.	N.D.	68.2 ± 5.7	N.D.	50.7 ± 8.2	323.4 ± 54.1	394.2 ± 98.8	1926.4 ± 300.1	105.4 ± 9.2	N.D.	N.D.	113.7 ± 14.1
3552	N.D.	N.D.	N.D.	70.1 ± 1.5	25.7 ± 0.4	40.5 ± 1.9	105.5 ± 6.2	533 ± 58.5	311.9 ± 5.3	35.1 ± 2	N.D.	N.D.	40 ± 2.7
3602	N.D.	N.D.	N.D.	47.2 ± 5	41 ± 5.4	52.1 ± 3.7	30 ± 14.8	53 ± 3.1	151 ± 13.3	60.6 ± 3.3	N.D.	N.D.	15.1 ± 3.2
3603	N.D.	N.D.	N.D.	43.9 ± 1.4	N.D.	24.8 ± 1.9	55.8 ± 6.1	381.6 ± 24.1	252 ± 2.7	109.4 ± 3.1	N.D.	N.D.	N.D.
3618	N.D.	N.D.	N.D.	73 ± 1.2	31.4 ± 2.6	35.5 ± 2	22.6 ± 0.6	616.3 ± 44.5	772 ± 35.7	160 ± 5	N.D.	N.D.	48.8 ± 2.7
3630	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	56.2 ± 4.2	313.2 ± 53.7	330 ± 27.4	103 ± 9.6	N.D.	N.D.	N.D.
3658	N.D.	N.D.	N.D.	62.5 ± 6.3	N.D.	76.9 ± 6	65.3 ± 10.8	266.9 ± 5.1	825.8 ± 46.3	59 ± 3.2	N.D.	N.D.	70.3 ± 3.1
3672	N.D.	N.D.	N.D.	50.2 ± 0.6	N.D.	N.D.	125.1 ± 8.2	575.6 ± 163.9	305.6 ± 8.1	211 ± 10.2	N.D.	N.D.	N.D.
3688	N.D.	N.D.	N.D.	34.3 ± 3.4	N.D.	N.D.	N.D.	N.D.	211.8 ± 14.9	47.9 ± 3.3	N.D.	N.D.	16.9 ± 1.2
3690	N.D.	N.D.	N.D.	113 ± 18.4	N.D.	37.6 ± 8.5	N.D.	30.7 ± 7.9	469.4 ± 96.4	371.6 ± 74.9	N.D.	N.D.	N.D.
3691	N.D.	N.D.	N.D.	68.2 ± 1.7	16.4 ± 1	35.2 ± 1	N.D.	N.D.	281.6 ± 10.9	149.7 ± 11.6	N.D.	N.D.	18.8 ± 2.7
3712	N.D.	N.D.	N.D.	42 ± 1.5	N.D.	N.D.	222.1 ± 16.5	1144.2 ± 131.4	1006.2 ± 24.4	223.7 ± 7.2	65.2 ± 57.5	N.D.	N.D.
3730	N.D.	N.D.	N.D.	81.9 ± 9.3	60 ± 1.9	67.5 ± 6.1	N.D.	38.2 ± 4.6	154.5 ± 3.8	44.6 ± 4	N.D.	N.D.	N.D.
3973	23.8 ± 3.1	N.D.	46 ± 2.8	425.6 ± 24.3	N.D.	45.9 ± 2.5	91.9 ± 9.7	1114 ± 229.6	256.1 ± 13.6	N.D.	N.D.	N.D.	71.3 ± 3.3
3980	N.D.	N.D.	N.D.	45.9 ± 6.3	39.7 ± 0.7	N.D.	36.8 ± 3.8	63.4 ± 3	868.6 ± 31.1	263.5 ± 9.3	N.D.	N.D.	28.1 ± 3.4
4243	N.D.	N.D.	N.D.	55.2 ± 1.5	N.D.	38 ± 0.1	10.7 ± 0.8	165 ± 33.3	541.6 ± 40.7	262.5 ± 14.8	42.1 ± 45.3	N.D.	14.8 ± 0.3
4249	N.D.	N.D.	N.D.	91.4 ± 5.3	57.6 ± 1.9	106.1 ± 10.2	58.7 ± 2.6	363.6 ± 19.8	756 ± 32.9	222.4 ± 10.8	44.7 ± 33.9	N.D.	24.5 ± 2.2
4279	N.D.	N.D.	N.D.	1015.7 ± 37	82.3 ± 11.9	100.2 ± 10.8	N.D.	67.5 ± 0.8	407.8 ± 25.6	83.5 ± 7.7	N.D.	N.D.	40.5 ± 3.2
4425	N.D.	N.D.	180 ± 12.6	6087.3 ± 256.7	N.D.	N.D.	32.2 ± 10.1	252.9 ± 12.7	475 ± 24.6	84.1 ± 5.8	N.D.	N.D.	33.7 ± 3.7
4487	N.D.	N.D.	N.D.	39.5 ± 1.4	17.7 ± 0.9	33.1 ± 7.5	N.D.	56 ± 15.7	378.2 ± 27.5	61.7 ± 9.6	20.9 ± 19.1	N.D.	18.2 ± 2.7
4500	N.D.	N.D.	N.D.	85.8 ± 6.2	N.D.	32.4 ± 2.2	34 ± 8.7	403.9 ± 59.1	1459.9 ± 117.6	152.8 ± 3	39.9 ± 53.9	N.D.	107.4 ± 6.5
4507	10.6 ± 1.7	N.D.	N.D.	54.7 ± 4.6	N.D.	N.D.	N.D.	11.6 ± 2.6	310.8 ± 15.2	12.6 ± 1.3	N.D.	N.D.	67.1 ± 4
4514	N.D.	N.D.	N.D.	83 ± 1.9	33.9 ± 0.5	67.4 ± 1.4	17.2 ± 8.3	154.1 ± 17	915.3 ± 54.2	14 ± 2.8	125.9 ± 113.6	N.D.	27.3 ± 0.9
4528	N.D.	N.D.	N.D.	24.1 ± 4	N.D.	N.D.	8.9 ± 2.7	51.4 ± 14.7	105.2 ± 7.3	N.D.	N.D.	N.D.	25 ± 4.2
4530	32.5 ± 2.4	N.D.	N.D.	49 ± 1.4	N.D.	13.3 ± 0.6	12.7 ± 1.7	71.2 ± 11.8	514.9 ± 6.6	43.8 ± 1.4	N.D.	N.D.	121.2 ± 6.8
4536	N.D.	N.D.	N.D.	44.6 ± 0.9	N.D.	13.5 ± 0.5	N.D.	30.4 ± 9.3	407.9 ± 50.5	35.4 ± 2.2	N.D.	N.D.	124.8 ± 22.3
4545	N.D.	N.D.	N.D.	261 ± 25.7	N.D.	43.3 ± 2.1	35.5 ± 2.6	84.9 ± 8.6	660.8 ± 7.7	106.2 ± 2.2	N.D.	N.D.	52.2 ± 2.4
4546	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	20.1 ± 2.9	83.1 ± 12.2	486.6 ± 9.1	38.9 ± 2.7	27.8 ± 27.4	N.D.	45.7 ± 2.1
4547	161.6 ± 2.9	N.D.	N.D.	N.D.	N.D.	N.D.	155.4 ± 11	427.2 ± 34.6	730.7 ± 27.9	105 ± 8	61.6 ± 61.8	N.D.	130.5 ± 6.1
4548	N.D.	N.D.	N.D.	165.6 ± 8.5	72.7 ± 2.9	56 ± 1.4	57.9 ± 12.6	1008.6 ± 27	310.9 ± 8.7	84.6 ± 5.1	N.D.	N.D.	N.D.
4549	23.8 ± 1	N.D.	N.D.	102.8 ± 1.9	63.9 ± 1.9	274 ± 28.7	136.9 ± 41.9	1526.3 ± 242.2	1423.3 ± 65	189 ± 1.5	N.D.	N.D.	114.9 ± 5.1

Table 7 continued

4714	N.D.	N.D.	N.D.	49.5 ± 3.6	N.D.	N.D.	34.4 ± 3.9	130.9 ± 20	305.5 ± 15.2	206.3 ± 13	N.D.	N.D.	20.6 ± 2.8
5074	N.D.	N.D.	N.D.	82.7 ± 6	61.2 ± 5.9	64.1 ± 7.1	100.5 ± 26.9	751.1 ± 69	682.4 ± 41.2	12.6 ± 1.4	N.D.	N.D.	115.5 ± 9.3
5075	N.D.	N.D.	N.D.	138.3 ± 9.4	81.4 ± 3.5	42.9 ± 0.2	N.D.	73.2 ± 29.5	257 ± 19.6	34.2 ± 3.9	N.D.	N.D.	N.D.
5108	N.D.	N.D.	N.D.	17.8 ± 0.4	N.D.	N.D.	N.D.	N.D.	65.7 ± 9.3	28.2 ± 3.6	N.D.	N.D.	N.D.
5549	N.D.	N.D.	N.D.	59.6 ± 5.4	N.D.	23.4 ± 1.3	N.D.	N.D.	243.5 ± 11.6	44.4 ± 4.2	N.D.	N.D.	37.2 ± 2.3
5557	N.D.	N.D.	N.D.	42.6 ± 1.5	8.5 ± 1	13.6 ± 0.6	N.D.	N.D.	138.7 ± 9.5	48.3 ± 6.7	N.D.	N.D.	14.9 ± 4.5
5596	65.6 ± 8.9	N.D.	34.5 ± 5.6	72.9 ± 4.6	66.1 ± 7.8	45.7 ± 6.4	N.D.	N.D.	498.9 ± 51.5	153.6 ± 18.9	N.D.	N.D.	81.1 ± 7.7
5601	N.D.	91.9 ± 5.9	N.D.	55.5 ± 4.5	N.D.	26.7 ± 0.7	7.3 ± 1	68.7 ± 16.7	252 ± 4.5	20.9 ± 2.5	N.D.	N.D.	57.5 ± 1.5
5615	28.6 ± 6.5	N.D.	72 ± 3	755.6 ± 72	44.8 ± 4.2	20.3 ± 0.5	N.D.	N.D.	88.1 ± 6.8	19.2 ± 2.6	N.D.	N.D.	N.D.
5617	N.D.	259.8 ± 4	179.1 ± 8.4	80.8 ± 2.5	46.8 ± 2.8	192.4 ± 3.4	57.5 ± 5.5	912.1 ± 43.2	801.6 ± 21.5	42.9 ± 2.3	N.D.	N.D.	99.9 ± 3.1
5721	N.D.	N.D.	N.D.	68.2 ± 3.5	66.4 ± 2.2	179.8 ± 10.3	31.6 ± 3.4	222.1 ± 20.9	862.7 ± 15.9	133.4 ± 0.9	39.4 ± 40.3	N.D.	53.8 ± 3.5
5732	N.D.	N.D.	N.D.	49 ± 0.8	N.D.	31.4 ± 0.7	N.D.	N.D.	550.3 ± 12.2	250.6 ± 5.4	33.3 ± 46.6	N.D.	N.D.
6068	N.D.	N.D.	N.D.	519.2 ± 24.4	N.D.	127.8 ± 18.9	442.6 ± 78.3	3601.6 ± 982.1	2622.6 ± 439.4	304.8 ± 27	N.D.	N.D.	54.5 ± 13.7
6073	14.1 ± 0.9	N.D.	N.D.	46 ± 4.4	N.D.	24.5 ± 1.4	N.D.	120.6 ± 15.4	120.8 ± 7.2	19.9 ± 2.6	N.D.	N.D.	25.4 ± 2.4
6259	N.D.	N.D.	N.D.	50.7 ± 1.4	N.D.	41.2 ± 2.1	N.D.	N.D.	765.2 ± 30.5	479 ± 14.2	N.D.	N.D.	N.D.
6584	41.8 ± 1.1	N.D.	N.D.	58.2 ± 3	N.D.	19.2 ± 1.4	93 ± 2.3	1019.6 ± 153	375.5 ± 20.2	38.6 ± 7.3	N.D.	N.D.	51.2 ± 1.7
6588	N.D.	N.D.	N.D.	91.9 ± 3.4	N.D.	11.7 ± 0.6	39.2 ± 1.1	162.4 ± 13	164.1 ± 10.4	41.8 ± 6.2	N.D.	N.D.	N.D.
6589	N.D.	N.D.	N.D.	37.9 ± 3.8	8.8 ± 1.9	N.D.	116.5 ± 15.1	618.2 ± 31.5	1017.4 ± 55.6	230.3 ± 13.2	65 ± 104.7	N.D.	35.7 ± 1.7
6598	N.D.	N.D.	N.D.	865.6 ± 65.2	112.3 ± 5.9	697.5 ± 27.9	N.D.	N.D.	1039.2 ± 62.3	173.8 ± 15.4	N.D.	N.D.	N.D.
6628	18.1 ± 1.2	N.D.	N.D.	71.9 ± 1.6	10.6 ± 0.5	9.6 ± 1.1	N.D.	N.D.	110.9 ± 4.3	50.1 ± 4.8	N.D.	N.D.	8.3 ± 1.2
6631	N.D.	N.D.	N.D.	627.1 ± 13.2	N.D.	62.5 ± 3.5	59.7 ± 14.6	727.7 ± 73.7	2272 ± 126	49.5 ± 4.6	87.2 ± 100.6	N.D.	183.3 ± 13.1
6632	N.D.	N.D.	N.D.	65.3 ± 1.3	N.D.	11.4 ± 1.7	183.9 ± 12.9	1125.7 ± 157.2	289.7 ± 6.5	70.6 ± 6.8	N.D.	N.D.	N.D.
6633	N.D.	N.D.	N.D.	47.3 ± 10.1	N.D.	N.D.	25.7 ± 3.1	104.3 ± 21.4	245.6 ± 7.7	67 ± 4.7	N.D.	N.D.	N.D.
6647	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	327.2 ± 31.3	286.4 ± 71.7	1350.6 ± 79.8	291.2 ± 29.2	26.3 ± 19.2	N.D.	N.D.
6655	N.D.	199 ± 5.2	205.8 ± 10.6	37.2 ± 3.2	N.D.	N.D.	11.1 ± 1.5	9.2 ± 0.4	554.7 ± 13	88.7 ± 3.5	50.3 ± 39.7	N.D.	17.8 ± 2.4
6700	3.8 ± 1.4	N.D.	120.7 ± 11.6	58.9 ± 6.4	N.D.	20.2 ± 3.2	N.D.	19 ± 1.8	420 ± 33.6	134.7 ± 13.1	N.D.	N.D.	36.4 ± 3.8
6732	N.D.	N.D.	N.D.	36 ± 3.4	36.5 ± 2	N.D.	113.7 ± 12.7	760 ± 122.2	367.7 ± 11.3	29.9 ± 2.4	N.D.	N.D.	13.2 ± 1.4
6748	52.5 ± 10.4	N.D.	N.D.	2159 ± 150.6	N.D.	N.D.	N.D.	N.D.	1017.7 ± 93.6	221.4 ± 18.4	N.D.	N.D.	40 ± 5.9
6765	162.3 ± 4.4	N.D.	N.D.	123.6 ± 1.7	N.D.	5.7 ± 0.5	N.D.	N.D.	612.5 ± 8.9	172.4 ± 3.6	N.D.	N.D.	15.7 ± 1.5
6879	N.D.	N.D.	N.D.	41.6 ± 1.7	N.D.	16.2 ± 0.8	N.D.	130.8 ± 17.2	94.3 ± 0.1	16 ± 1.9	N.D.	N.D.	N.D.
6909	N.D.	142.7 ± 4.4	N.D.	54.5 ± 2.4	N.D.	16.3 ± 2.1	18.9 ± 4.6	17.7 ± 5.8	170.5 ± 5.2	38.7 ± 2.1	N.D.	N.D.	N.D.
6915	N.D.	N.D.	N.D.	29.3 ± 6.7	N.D.	N.D.	190.3 ± 32.7	1404.5 ± 181.3	317.3 ± 7.2	N.D.	N.D.	N.D.	48.2 ± 2.3

Table 7 continued

6928	N.D.	N.D.	N.D.	52.6 ± 1.1	N.D.	N.D.	78.7 ± 4.6	48.5 ± 2.4	717.3 ± 25.3	205 ± 3.6	N.D.	N.D.	48.2 ± 0.6
6940	N.D.	N.D.	22.7 ± 1	537.2 ± 10	125.3 ± 4.1	174.9 ± 4.2	46.8 ± 4.6	526.3 ± 60.8	858.8 ± 11.7	395.4 ± 11.4	N.D.	N.D.	72 ± 0.6
6954	83.7 ± 0.8	N.D.	152.7 ± 6.4	56.6 ± 1.4	N.D.	43.4 ± 6.3	52.2 ± 10.5	269.9 ± 60.3	695 ± 16.2	64.3 ± 7.1	43 ± 47.9	N.D.	278.1 ± 23.1
6981	N.D.	N.D.	N.D.	65.1 ± 3.4	N.D.	28.6 ± 1.5	77.9 ± 13.3	332.3 ± 20.4	273.9 ± 10.6	35.2 ± 4.1	N.D.	N.D.	N.D.
6982	N.D.	N.D.	N.D.	77.3 ± 2.4	86.7 ± 8.4	49.9 ± 2.4	53.4 ± 3.2	120.5 ± 28	423.6 ± 23.2	163.2 ± 9.2	N.D.	N.D.	N.D.
08/-08	27.5 ± 1	N.D.	N.D.	37.9 ± 3.3	N.D.	N.D.	N.D.	N.D.	268.5 ± 11.5	157.5 ± 18.7	N.D.	N.D.	N.D.
e1262	N.D.	N.D.	68 ± 6.6	1126.1 ± 65.2	N.D.	22 ± 2.8	N.D.	16.6 ± 3.7	269.8 ± 9.5	165.4 ± 5.8	N.D.	N.D.	N.D.
e1827	N.D.	N.D.	N.D.	51.1 ± 1.5	N.D.	37 ± 1.1	102 ± 14.1	557.1 ± 69.4	343.3 ± 3.3	39.9 ± 2.5	N.D.	N.D.	N.D.
e2180	N.D.	N.D.	35 ± 3.4	N.D.	N.D.	N.D.	684.5 ± 134.6	1716.8 ± 316	2505.8 ± 325.3	240.2 ± 25.6	308.6 ± 280.8	N.D.	102.9 ± 17.1
e2474	N.D.	N.D.	24.8 ± 0.5	41.5 ± 2.4	N.D.	17.3 ± 1.4	N.D.	90.6 ± 8.2	795.9 ± 24.7	348 ± 12.3	N.D.	N.D.	N.D.
e3406	N.D.	N.D.	N.D.	18 ± 1.4	N.D.	N.D.	108 ± 21.4	516.4 ± 81.5	736.7 ± 91.2	326.6 ± 43.3	N.D.	N.D.	15.4 ± 2.4
e3756	N.D.	N.D.	N.D.	35.9 ± 5	N.D.	16.6 ± 0.6	40.7 ± 14.4	322 ± 30.2	265.6 ± 7.4	80.9 ± 0.6	N.D.	N.D.	N.D.
e3756	N.D.	N.D.	8.3 ± 0.6	12 ± 1	N.D.	N.D.	211.3 ± 27.4	2229.6 ± 205.9	1201.2 ± 8.4	223.5 ± 3.1	115.2 ± 109.8	N.D.	11.2 ± 0.9
Nycos	25.4 ± 2.5	N.D.	23.1 ± 3.4	64.5 ± 1.1	6.6 ± 0.4	22.7 ± 0.8	17.5 ± 1.2	59.1 ± 17.7	185.9 ± 4.9	99.8 ± 4.2	N.D.	N.D.	N.D.
Yarra	N.D.	N.D.	N.D.	38.6 ± 2.3	N.D.	14.8 ± 1.7	59.9 ± 8.1	420.2 ± 51.4	314 ± 10.8	3.7 ± 0.9	21.1 ± 26.1	N.D.	27.7 ± 1.1
1107-1	N.D.	N.D.	N.D.	875.7 ± 84.5	N.D.	39 ± 3.6	11.7 ± 1.9	156.1 ± 7.9	225.4 ± 19.3	14.8 ± 2.4	N.D.	N.D.	N.D.
1107-2	13.8 ± 0.8	N.D.	N.D.	97.7 ± 7.8	42.9 ± 1	43.3 ± 0.8	61.2 ± 8.5	269.4 ± 20.8	227.1 ± 12.3	77 ± 2.8	N.D.	N.D.	N.D.
M82-1	N.D.	N.D.	N.D.	34.9 ± 1.6	N.D.	25 ± 2.3	43.5 ± 3.5	430.4 ± 33.1	459.6 ± 7.6	122.8 ± 6.3	N.D.	N.D.	N.D.
M82-2	N.D.	N.D.	19.9 ± 0.7	36.2 ± 0.4	N.D.	20.8 ± 0.6	232.2 ± 50.9	1971 ± 83.2	2450.6 ± 89.4	304.2 ± 10.7	91.6 ± 8.6	N.D.	21.9 ± 0.9
M82-3	N.D.	N.D.	N.D.	45.8 ± 1.6	N.D.	14.4 ± 0.9	57.6 ± 13.7	701.7 ± 39.5	572.4 ± 7.3	186 ± 2.9	N.D.	N.D.	11.8 ± 1
M82-4	N.D.	39 ± 3.5	N.D.	40.6 ± 1.7	N.D.	6.6 ± 0.6	N.D.	21.7 ± 2.1	105.9 ± 10.9	10.9 ± 1.7	N.D.	N.D.	6.2 ± 0.4
M82-5	N.D.	73.6 ± 1.6	22.7 ± 0.5	49.8 ± 2	N.D.	18.1 ± 2.1	20.9 ± 2.1	158.7 ± 21.8	288.9 ± 23.6	98.1 ± 8.4	N.D.	N.D.	N.D.
M82-6	N.D.	N.D.	N.D.	44.6 ± 2.1	N.D.	22.4 ± 0.8	96.9 ± 5	897.9 ± 108	508.2 ± 9.7	77.1 ± 0.8	26.8 ± 1.6	N.D.	12.4 ± 1
QC	N.D.	N.D.	N.D.	1076.6 ± 32.9	N.D.	167.1 ± 4.9	295.1 ± 17	2709.9 ± 107.3	1053.1 ± 38.5	361 ± 11.3	N.D.	N.D.	N.D.

Table 8 Isoprenoid content of available accessions from Core Collection Colour Mutant Population

Average values µg/g DW ± standard error of the mean (n=3) presented for 259 accessions, labelled by Core Collection accession number (CC#). N.D. indicates compound not detected by analytical platform. Trace indicates sample was detected, but at levels below limit of quantification. Nycos, Yarra, 1107, and M82 (1-5) represent elite lines cultivated alongside CC accessions. QC represents quality control replicates (n=75).

	Average amount (µg g^{-1} DW)														
		Phytofluene		cis-ζ-Carotene											
CC#	Phytoene	Isomer 1	Isomer 2	Isomer 1	Isomer 2	Neurosporene	Prolycopene	Lycopene	β-Carotene	Lutein	α-Tocopherol	γ-Tocopherol	Chlorophyll a	Chlorophyll b	Ubiquinone
9	165.2 ± 4.1	104.5 ± 3.1	62.8 ± 2.5	129 ± 5.4	140.2 ± 5.7	N.D.	N.D.	849.3 ± 32.4	80 ± 3.8	16.2 ± 0.5	N.D.	N.D.	N.D.	N.D.	80.1 ± 3.4
10	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	7.4 ± 0.4	100.6 ± 4	27.2 ± 1.1	348.8 ± 8.7	128.3 ± 2.5	N.D.	N.D.	162.1 ± 3.3
13	419.5 ± 0.9	148.4 ± 6	118.9 ± 3.5	207.6 ± 4.7	465.4 ± 0.8	177.1 ± 5.8	267 ± 4.9	N.D.	67.8 ± 1.9	4.5 ± 0.1	121.7 ± 3.7	N.D.	N.D.	N.D.	79.4 ± 1.7
27	45.6 ± 0.7	N.D.	61.1 ± 1.5	N.D.	N.D.	N.D.	N.D.	409.6 ± 8.6	91.7 ± 1.5	16.6 ± 0.7	173 ± 19	105.3 ± 3.2	N.D.	N.D.	92.3 ± 2.4
41	27.8 ± 0.6	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	27.6 ± 0.9	275.3 ± 8.5	84 ± 5	473.5 ± 15.2	110.1 ± 2.7	N.D.	N.D.	227.5 ± 5.5
54	387.4 ± 5	129 ± 2.8	84.6 ± 3.8	180.2 ± 7.9	345.9 ± 13.1	174.1 ± 8.2	278.5 ± 10.3	14.9 ± 0.3	108.5 ± 4.7	12.1 ± 0.6	125.9 ± 4.3	N.D.	N.D.	N.D.	78.4 ± 3.2
83	39.5 ± 0.2	N.D.	52.2 ± 0.3	N.D.	N.D.	N.D.	N.D.	445.5 ± 5.2	91.8 ± 1.9	12.9 ± 0.6	200.1 ± 1.6	107.2 ± 1.3	N.D.	N.D.	83.8 ± 0.5
168	110 ± 3.1	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	707 ± 8.1	170.7 ± 5.2	28.4 ± 1	178.1 ± 0.9	142.1 ± 3.2	N.D.	N.D.	110.3 ± 6.3
171	321.2 ± 6	141.1 ± 2.6	79.3 ± 1.5	183.3 ± 2.7	293.5 ± 6	166.8 ± 2.6	192 ± 4.3	6.9 ± 0.4	80.8 ± 1	6.9 ± 0.2	121.5 ± 4.1	87.4 ± 3.5	N.D.	N.D.	79.8 ± 2.8
181	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	60.9 ± 2	99 ± 3.1	22.6 ± 1	203 ± 4.9	71.2 ± 3.3	N.D.	N.D.	122.9 ± 3.8
192	23.8 ± 0.9	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	66.5 ± 1.2	281.9 ± 5.6	29.7 ± 0.7	229.5 ± 3	77.6 ± 3.2	N.D.	N.D.	147.9 ± 0.9
195	26.4 ± 0.2	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	414.7 ± 19.5	251.1 ± 40.1	25.7 ± 7.5	243.7 ± 47.8	88.5 ± 9.3	N.D.	N.D.	174.1 ± 27.3
223	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	119.7 ± 5.2	26.8 ± 0.6	258.7 ± 8.7	N.D.	N.D.	N.D.	146.4 ± 6.6
258	25.4 ± 0.6	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	24.9 ± 2.9	223.4 ± 15.9	125.7 ± 4.4	N.D.	127.7 ± 3	116.3 ± 5.7
262	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5.3 ± 0.2	83.5 ± 2.8	29.7 ± 1.2	264 ± 9.6	97.5 ± 2.7	N.D.	N.D.	126.1 ± 3.5
279	46.9 ± 1.4	55.7 ± 1.7	68.3 ± 1.7	135.9 ± 3.8	139.8 ± 3.7	N.D.	N.D.	711.5 ± 20.7	243.3 ± 4.8	50.7 ± 1.9	389.2 ± 9.4	N.D.	N.D.	143.4 ± 3.6	146.1 ± 4.4
280	66.4 ± 0.7	65.8 ± 0.7	77.5 ± 3.3	136.3 ± 5.4	144.7 ± 5.1	N.D.	N.D.	600 ± 13.7	178.6 ± 3.8	23.6 ± 0.5	342.6 ± 5	133.8 ± 3.7	N.D.	N.D.	108.4 ± 4.5
282	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	41.7 ± 1	77.3 ± 3.1	20.7 ± 0.3	192.8 ± 6.2	70 ± 4.6	N.D.	116.6 ± 5.6	118.1 ± 5.5
294	53.2 ± 1.1	56.9 ± 2	62.7 ± 1.9	130.6 ± 4.7	130.1 ± 4.6	N.D.	N.D.	436.6 ± 13.9	122.2 ± 2.1	3.3 ± 0.4	258 ± 2.5	103.3 ± 5.4	N.D.	N.D.	87.7 ± 2.5
302	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5.4 ± 0.1	116.2 ± 0.6	26.1 ± 0.8	215.4 ± 3.4	N.D.	N.D.	N.D.	126.4 ± 1.7
325	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	72.8 ± 2.7	7.9 ± 0.5	203.2 ± 5.3	110.9 ± 2.6	N.D.	N.D.	141.9 ± 4.8
369	102.7 ± 2.9	81.2 ± 2	83.8 ± 1.4	141 ± 2.8	145.9 ± 3.2	N.D.	N.D.	841.4 ± 22.3	115.5 ± 5.7	17.5 ± 0.6	152.2 ± 2.6	N.D.	N.D.	N.D.	105.4 ± 1.5
375	228.5 ± 4.7	119.9 ± 4.2	91.5 ± 2.3	161.2 ± 5.9	235.6 ± 7.4	179 ± 6.9	213.5 ± 6.6	29.7 ± 1.3	139.3 ± 4	18 ± 0.2	242.4 ± 7	134.9 ± 6	N.D.	N.D.	99.3 ± 2.7

Table 8 continued

383	27.6 ± 1.6	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	48.3 ± 2	149.6 ± 9.2	38.6 ± 2.2	251.6 ± 11.3	72.8 ± 3.8	N.D.	N.D.	132.7 ± 6.8
405	25.3 ± 1.6	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	11.3 ± 0.6	107.8 ± 3.9	32.7 ± 1.2	212.3 ± 11.4	89.2 ± 5.9	N.D.	N.D.	129.8 ± 5.4
412	85.9 ± 0.8	57.5 ± 1.4	77.4 ± 0.6	135.3 ± 3.5	133.3 ± 3.2	N.D.	N.D.	405.9 ± 8.7	64.3 ± 1.3	17.2 ± 0.5	241.7 ± 2.4	153.5 ± 7.2	N.D.	130 ± 3.1	100.9 ± 2.1
434	1394.4 ± 67.9	612.1 ± 28.2	148.2 ± 4.2	403.4 ± 15.6	995.1 ± 37.5	N.D.	930 ± 39.4	11.7 ± 0.3	92.1 ± 2.4	-2.3 ± 0.1	N.D.	N.D.	N.D.	N.D.	77 ± 6.3
482	29.3 ± 0.5	N.D.	51.4 ± 0.5	N.D.	N.D.	N.D.	N.D.	6.3 ± 0.2	415.8 ± 8.6	8.1 ± 0.2	172 ± 5.5	N.D.	N.D.	117.7 ± 1.1	83.6 ± 1
488	27.4 ± 0.8	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	90.4 ± 1.4	135.7 ± 0.8	77.5 ± 0	392.9 ± 4.3	N.D.	N.D.	N.D.	238.1 ± 2.2
489	34.2 ± 1.1	N.D.	55.7 ± 2	N.D.	N.D.	N.D.	N.D.	959.3 ± 44.5	130.5 ± 4.8	13.6 ± 0.4	146.1 ± 3.5	N.D.	N.D.	N.D.	124.2 ± 3.2
495	81.2 ± 0.6	68.2 ± 1.1	66 ± 1.1	130.4 ± 4.1	133.7 ± 3.9	N.D.	N.D.	1660 ± 29.7	181.3 ± 0.7	12.5 ± 0.9	223.1 ± 4.9	N.D.	N.D.	N.D.	127.8 ± 1.5
503	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	11.1 ± 0.3	205.8 ± 2.1	215.4 ± 2.1	N.D.	124.6 ± 2.8	128.9 ± 1.8
512	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	72.9 ± 2.5	9.6 ± 0.2	195.4 ± 2.8	N.D.	N.D.	N.D.	175.7 ± 7
523	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	7.6 ± 0.1	81.4 ± 2.8	18.1 ± 0.6	256.9 ± 4.6	85.8 ± 2.3	N.D.	N.D.	133.3 ± 2.1
760	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	100.9 ± 6.1	21.3 ± 1.6	346.1 ± 7.8	161.5 ± 1.5	N.D.	N.D.	170.1 ± 2.8
797	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5.2 ± 0.1	73.3 ± 1.8	39.7 ± 1.6	267.2 ± 10	N.D.	N.D.	N.D.	140.2 ± 4.3
809	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	63.4 ± 3.7	12.2 ± 0.5	86.7 ± 14.3	86.7 ± 3.4	N.D.	N.D.	87.4 ± 2.8
851	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	10.8 ± 0.3	90.5 ± 1.5	22.4 ± 0.8	383.8 ± 3.5	104.9 ± 0.7	N.D.	N.D.	205 ± 4.6
861	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	105.1 ± 2.8	46.1 ± 1.3	366.8 ± 8.5	85.5 ± 2.4	N.D.	N.D.	219 ± 5.2
870	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	76 ± 1.7	29.1 ± 1.1	326.1 ± 6.6	129.2 ± 3.4	N.D.	N.D.	173.8 ± 1.1
874	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	80 ± 1.4	32.3 ± 1.8	343.8 ± 4.6	81.5 ± 1.7	N.D.	N.D.	180.5 ± 0.5
886	122.6 ± 1.8	84.2 ± 2.5	60.9 ± 2.8	129.6 ± 6.1	133.1 ± 6.1	N.D.	N.D.	849 ± 18.7	93 ± 2.9	2.1 ± 0.3	104.1 ± 3.9	N.D.	N.D.	N.D.	82.3 ± 3.4
896	99.7 ± 1.4	N.D.	N.D.	170.1 ± 6.1	161 ± 8.6	N.D.	N.D.	706.1 ± 20.5	86.7 ± 7.9	5.9 ± 1.1	80.8 ± 3.6	80.2 ± 3.5	N.D.	N.D.	76.8 ± 4.2
899	55.9 ± 4.3	N.D.	66.3 ± 1.9	N.D.	N.D.	N.D.	N.D.	477.3 ± 16.9	159.5 ± 11.5	49.5 ± 7.1	223.3 ± 19.4	N.D.	N.D.	132.9 ± 2.8	105.4 ± 5.5
922	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	86.6 ± 1.8	21.1 ± 0.5	245 ± 8.3	145.3 ± 2.2	N.D.	N.D.	147.8 ± 2.8
923	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	6.1 ± 0.2	99.8 ± 2.1	23.1 ± 0.9	237.1 ± 4.2	95.5 ± 3.4	N.D.	116.7 ± 3.2	124.1 ± 1.8
930	86.2 ± 0.2	59.5 ± 1.1	69.1 ± 1.1	133.5 ± 2.6	138 ± 2.5	N.D.	N.D.	444.5 ± 7.6	433.2 ± 5.6	15.7 ± 0.6	247.8 ± 1.6	N.D.	N.D.	N.D.	112.7 ± 0.6
933	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	44.8 ± 0.5	80.9 ± 2.7	20.5 ± 0.1	246.3 ± 6.2	77.6 ± 5.2	N.D.	N.D.	194 ± 4.3
945	40.1 ± 0.8	N.D.	54.4 ± 1.1	125.1 ± 2.8	126.5 ± 2.9	N.D.	N.D.	27.8 ± 1.8	166 ± 3.3	17.6 ± 0.4	230.6 ± 2.3	111 ± 4.7	N.D.	N.D.	138.4 ± 2.5
948	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	74.3 ± 0.8	9.1 ± 0.2	171.4 ± 3.3	116.8 ± 1.5	N.D.	N.D.	153.6 ± 2.2
949	294.6 ± 29.4	105.2 ± 3.6	120.9 ± 15.2	211.6 ± 15.2	345.7 ± 24.7	211.6 ± 16.2	149.9 ± 6.1	13.3 ± 1.2	131.1 ± 10.8	12.3 ± 1.7	139.9 ± 5	110.1 ± 9.7	N.D.	116.8 ± 6	92 ± 5.5
955	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	129.2 ± 7.3	22.1 ± 2.8	316.6 ± 22.8	101.8 ± 3.4	N.D.	N.D.	126.4 ± 6
959	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	88.6 ± 6.2	24.7 ± 2.3	304 ± 24.6	69.6 ± 3.6	N.D.	N.D.	149.1 ± 12
960	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	86.5 ± 1	19.5 ± 0.3	264.1 ± 2.2	142.1 ± 1.1	N.D.	N.D.	143.4 ± 2.8
963	23.8 ± 0.4	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	25.6 ± 0.8	101.8 ± 1.1	32.1 ± 0.3	225.4 ± 2.9	126 ± 0.5	N.D.	N.D.	116.3 ± 3.7

Table 8 continued

970	25.3 ± 0.7	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	67.1 ± 1.7	15.5 ± 0.7	122.9 ± 2.1	96.3 ± 2	121.1 ± 3.5	123.3 ± 3.4	78.2 ± 1.1
971	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	11.4 ± 0.2	N.D.	24.4 ± 1.4	300.6 ± 18.5	175.6 ± 2.1	N.D.	N.D.	141.7 ± 4.7
972	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	18.9 ± 0.9	207.3 ± 10.3	181.9 ± 5.9	N.D.	N.D.	90 ± 1.9
975	43.3 ± 0.4	N.D.	61.5 ± 1.2	133.4 ± 3.6	133.2 ± 3.4	N.D.	N.D.	17.8 ± 0.2	114.2 ± 0.2	13.4 ± 0.3	213 ± 5.1	78.7 ± 1.7	N.D.	N.D.	121.1 ± 1.8
980	188.3 ± 1.9	118.4 ± 0	60.1 ± 0.6	143.4 ± 1.6	307.3 ± 1.2	162 ± 3.1	189.7 ± 1.7	10.1 ± 0.3	125.2 ± 2	21.8 ± 0.6	168.7 ± 1.7	89.7 ± 1.7	N.D.	N.D.	91.6 ± 1.3
1000	178.7 ± 11.9	112.4 ± 3.3	81.5 ± 4.2	142.3 ± 7.2	149.8 ± 7	N.D.	N.D.	1220.4 ± 108.7	155.1 ± 10	14.6 ± 2.6	170.1 ± 13.6	N.D.	N.D.	N.D.	107.1 ± 5.6
1011	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5.2 ± 0.2	114.1 ± 1.7	28.4 ± 2.1	298.8 ± 6.7	68.6 ± 1.8	N.D.	N.D.	187.8 ± 2.8
1034	52.1 ± 0.4	58.7 ± 1.4	61.3 ± 2	128.5 ± 6.3	129.6 ± 6.2	N.D.	N.D.	453.4 ± 10.3	128 ± 1.8	13.8 ± 0.9	183.3 ± 3.3	75.5 ± 4.3	N.D.	N.D.	95.6 ± 1.3
1158	25.5 ± 0.1	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	439.6 ± 4	188.7 ± 1.4	19.4 ± 0.3	130.5 ± 3	82.6 ± 3.9	N.D.	N.D.	113.2 ± 1.3
1278	141.1 ± 1.2	88.3 ± 6.7	93.1 ± 6.1	146 ± 2.7	151.7 ± 2.4	N.D.	N.D.	1012.8 ± 5	126 ± 1.5	11.6 ± 0.4	186.1 ± 0.3	N.D.	N.D.	N.D.	106.9 ± 2.1
1316	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	7 ± 0.3	92.6 ± 1.5	22 ± 0.9	291 ± 4.9	74.6 ± 2.4	N.D.	N.D.	161.5 ± 0.9
1321	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	68.5 ± 1.7	28 ± 0.5	227.3 ± 2.4	97.6 ± 1.4	N.D.	N.D.	121.3 ± 0.5
1323	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	61.2 ± 0.3	13.4 ± 1.8	150 ± 8.9	100.2 ± 6.3	N.D.	N.D.	104.6 ± 3.1
1324	102.5 ± 1.4	57.6 ± 1.1	64.3 ± 1.3	135.9 ± 3.7	165 ± 3.1	148.1 ± 4.3	150.2 ± 4	6.6 ± 0.5	67.2 ± 1.1	0.5 ± 0.1	N.D.	N.D.	N.D.	N.D.	N.D.
1325	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5.6 ± 0.1	62.5 ± 0.8	14.1 ± 0.7	246.3 ± 8.4	104.4 ± 1.4	N.D.	N.D.	134 ± 3.4
1326	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	6.1 ± 0.1	N.D.	15.1 ± 0.5	168.6 ± 3.3	107.9 ± 2.5	N.D.	N.D.	129.5 ± 2.5
1327	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	112 ± 1.6	44 ± 0.3	260.6 ± 1.7	161 ± 1.4	N.D.	N.D.	121.9 ± 2.4
1328	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	6.4 ± 0.2	N.D.	24.4 ± 0.8	284 ± 5.9	81.8 ± 3.5	N.D.	N.D.	164.3 ± 2.2
1329	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	60 ± 1.7	23 ± 0.2	305.2 ± 8.8	93.9 ± 1.1	N.D.	N.D.	143.5 ± 3.6
1330	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	62.9 ± 1.2	19 ± 0.4	237 ± 9	87.7 ± 1.7	N.D.	N.D.	156.6 ± 2.8
1334	452.5 ± 3.5	207.1 ± 2.2	110.7 ± 3.2	194.7 ± 3.3	541.5 ± 4.3	201.1 ± 3	239.2 ± 3	N.D.	85.9 ± 1.1	7.2 ± 0.2	181.7 ± 5.3	N.D.	N.D.	N.D.	96.2 ± 1.6
1336	28.4 ± 0.8	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	13.2 ± 0.6	434.3 ± 14.4	16.5 ± 0.5	302 ± 9.4	123.2 ± 6.1	N.D.	N.D.	138.9 ± 3.1
1337	41 ± 1.6	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	406.3 ± 19.3	490.1 ± 8.6	4.7 ± 0.4	82.5 ± 4.2	97.9 ± 4.1	N.D.	N.D.	98.4 ± 1.3
1338	24.8 ± 0.6	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	7.2 ± 0.2	159.4 ± 3.6	32.5 ± 1.8	383.2 ± 8.2	106.2 ± 1.4	N.D.	N.D.	222.6 ± 4.4
1339	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	10.7 ± 0.2	96.4 ± 2.1	25.5 ± 0.1	263 ± 1.7	114.6 ± 4.2	N.D.	N.D.	150.7 ± 2.1
1340	374.9 ± 1	169.3 ± 1.9	71.6 ± 2.6	167 ± 4.6	387.1 ± 7.8	161.5 ± 5.7	172.6 ± 4.2	N.D.	N.D.	1.3 ± 0.3	N.D.	N.D.	N.D.	N.D.	69.7 ± 1.7
1344	361 ± 9.1	139.3 ± 2.3	73.1 ± 1.2	174.1 ± 4.2	280.7 ± 4.2	162.7 ± 4	287.5 ± 4.6	11.5 ± 0.1	89.1 ± 1.7	5.9 ± 0.4	N.D.	79.4 ± 2	N.D.	N.D.	93.1 ± 1.9
1345	45.9 ± 0.8	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	77.3 ± 1.5	410.2 ± 5.3	31.3 ± 1.7	325.9 ± 6.5	78.4 ± 2.1	N.D.	N.D.	132.1 ± 3.8
1346	25.4 ± 0.7	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5.1 ± 0.1	110.9 ± 0.5	16 ± 0.5	269.9 ± 5.6	100.1 ± 2.5	N.D.	N.D.	231.2 ± 4.3
1349	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	10 ± 0.8	81.9 ± 1.8	33.5 ± 0.1	234.9 ± 6.5	90.4 ± 3.2	N.D.	N.D.	129.9 ± 3.3
1350	586.7 ± 17.3	233.4 ± 6.4	107.9 ± 3.4	178.1 ± 6.7	467.8 ± 14	173.2 ± 5.9	203.9 ± 6.4	N.D.	64.6 ± 2.3	1.5 ± 0.3	N.D.	N.D.	N.D.	N.D.	76.7 ± 2.6
1351	307.3 ± 1.8	139 ± 2.1	64.4 ± 1	161.5 ± 4.8	278.7 ± 11.4	156.2 ± 5	192.7 ± 7.7	N.D.	66.3 ± 2.1	1.1 ± 0	N.D.	N.D.	N.D.	N.D.	N.D.

Table 8 continued

1352	387 ± 2.4	198.8 ± 2.6	95.3 ± 1	227.7 ± 3	414 ± 3.8	203 ± 3.4	289.8 ± 2.2	22.2 ± 0.5	116.3 ± 1.8	5.8 ± 0.2	218.5 ± 4.7	N.D.	N.D.	N.D.	97.9 ± 1.8
1353	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	19.5 ± 1.7	164.3 ± 12.3	85.3 ± 0.5	N.D.	N.D.	110.6 ± 6.1
1355	212.9 ± 6.7	141.6 ± 3.6	67.4 ± 3.2	173 ± 3.8	269.6 ± 2.3	165.2 ± 6.1	188.7 ± 3.5	N.D.	130.6 ± 5.8	5.2 ± 0.8	174.7 ± 4.4	130.6 ± 6.7	N.D.	N.D.	80.1 ± 2.9
1356	264.5 ± 2.9	174.3 ± 1.8	80 ± 1.9	192.5 ± 2.8	341.6 ± 1.3	180.6 ± 2.6	233.1 ± 4.9	N.D.	222.2 ± 0.1	16.3 ± 0.3	217.8 ± 2.1	169.7 ± 1	N.D.	N.D.	94.4 ± 1.7
1361	353.2 ± 0.8	142.6 ± 3	82.2 ± 3.8	197.4 ± 7.5	300.7 ± 10	175.7 ± 9.5	193.3 ± 7.3	N.D.	85.4 ± 3.7	5.6 ± 0.1	135.4 ± 5.4	148.3 ± 5.4	N.D.	N.D.	78.3 ± 4.8
1363	356.7 ± 1.3	168.6 ± 6.6	92.4 ± 6	200.9 ± 2.2	462.5 ± 5.3	185.1 ± 3.2	236.1 ± 4	8.9 ± 0.2	124.9 ± 1.2	11.5 ± 0.3	204.3 ± 3.6	94.2 ± 1.6	N.D.	N.D.	101.8 ± 2.7
1367	642.7 ± 11.8	324.7 ± 17.4	110.7 ± 3.6	312.4 ± 11.5	964.4 ± 30.4	217.1 ± 11.3	361.2 ± 14.2	N.D.	117.4 ± 5	11.8 ± 0.6	156.3 ± 5.9	105.1 ± 5.5	N.D.	N.D.	91 ± 6.2
1368	476.1 ± 7.4	243.1 ± 3.8	95 ± 1.3	244 ± 2.1	575.7 ± 6.4	209.1 ± 1.1	292.5 ± 3.2	8.3 ± 0.3	100.6 ± 0.4	4.4 ± 0.4	154.5 ± 8.8	N.D.	N.D.	N.D.	89.5 ± 0.7
1370	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5 ± 0.1	105.6 ± 1.7	37.1 ± 0.9	339.9 ± 5.7	88.9 ± 1.6	N.D.	117.7 ± 3	174.6 ± 0.9
1371	27 ± 1.3	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	24.3 ± 1.3	152.4 ± 6.8	21.8 ± 0.8	209.6 ± 8.1	82.7 ± 2.7	N.D.	N.D.	122.2 ± 5.8
1374	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	63.8 ± 3.3	15.8 ± 0.3	294 ± 7.9	77.5 ± 3.4	N.D.	N.D.	174.6 ± 2.4
1377	361.4 ± 4.7	102.1 ± 2	104.9 ± 3	178.4 ± 5.3	269.2 ± 7.6	190.3 ± 5.8	284.5 ± 11.1	N.D.	128.4 ± 2	18.7 ± 0.7	205.8 ± 3.7	N.D.	N.D.	N.D.	111.2 ± 2.1
1378	296.9 ± 6.5	132.6 ± 2.3	64 ± 1.5	152.6 ± 3.6	253 ± 5.5	151.7 ± 3.6	168.5 ± 3.6	N.D.	61.8 ± 1.5	-0.8 ± 0.2	N.D.	N.D.	N.D.	N.D.	N.D.
1380	97.6 ± 4.8	80.6 ± 4.4	64.3 ± 3.7	132.8 ± 8.4	136 ± 8.6	N.D.	N.D.	1265.9 ± 51.7	165.6 ± 9.6	12.3 ± 0.6	165.1 ± 10.9	N.D.	N.D.	N.D.	103.7 ± 5.3
1631	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	9.4 ± 0.6	69.6 ± 3.6	6.7 ± 0.3	201.4 ± 4.6	71.6 ± 3.3	N.D.	N.D.	130.4 ± 4.6
1637	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	13.5 ± 0.4	190.8 ± 5.8	224.2 ± 6.4	N.D.	119.5 ± 1	107.3 ± 2
1719	24.9 ± 0.9	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5.4 ± 0.3	87.3 ± 0.9	5.6 ± 0.3	154.6 ± 9.9	77.2 ± 2	N.D.	N.D.	105.8 ± 0.3
1724	27.9 ± 1	N.D.	53.2 ± 2.2	N.D.	132.5 ± 5.7	N.D.	N.D.	78.1 ± 2.2	245.1 ± 1.3	10.2 ± 0.5	266.2 ± 5.4	82.2 ± 4.3	N.D.	N.D.	135.3 ± 0.4
1729	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	83.9 ± 1.9	16.1 ± 0.8	265.6 ± 12.3	104.9 ± 2.5	N.D.	N.D.	165.9 ± 1.9
1784	30.9 ± 0.2	N.D.	53.8 ± 0.5	N.D.	129.8 ± 0.8	N.D.	N.D.	17.2 ± 0.2	156 ± 2	16.1 ± 0.3	298.4 ± 3.5	115.7 ± 1	N.D.	N.D.	137.5 ± 1.1
1836	111.5 ± 0.6	95.3 ± 0.6	67 ± 0.3	129.6 ± 0.9	136.6 ± 1.1	N.D.	N.D.	736.3 ± 8.1	113.9 ± 0.5	15.3 ± 0.1	177.9 ± 10.1	N.D.	N.D.	N.D.	99.4 ± 1.9
1899	35.1 ± 0.3	N.D.	58.4 ± 0.3	133 ± 0.9	133.2 ± 1.3	N.D.	N.D.	414.8 ± 4.9	357.6 ± 4.8	21.3 ± 0.5	339.7 ± 2.1	103.8 ± 0.8	N.D.	N.D.	139.2 ± 0.3
1913	89 ± 0.6	71.1 ± 1.5	73.1 ± 1	134.4 ± 3.1	139.3 ± 3.1	N.D.	N.D.	546.9 ± 6.5	137.4 ± 1.6	25.7 ± 0.2	274.8 ± 9.9	86.7 ± 3.1	N.D.	N.D.	119.1 ± 1.7
1941	394.3 ± 5	163.5 ± 3.4	75.4 ± 3.2	197.9 ± 2.4	383.1 ± 3.5	175.5 ± 1	272.6 ± 1.4	9.1 ± 0.6	77.7 ± 1	4.3 ± 0.1	N.D.	141.7 ± 8.8	N.D.	N.D.	N.D.
1980	538.6 ± 19.1	226.7 ± 6.8	83.8 ± 3.8	220.1 ± 7.4	480.1 ± 15.4	178.1 ± 7.9	213.7 ± 7	7.7 ± 0.6	93.8 ± 3.6	7.9 ± 0.3	109.4 ± 3.4	N.D.	N.D.	N.D.	76.8 ± 2.3
1982	62.4 ± 0.7	69 ± 0.8	73.5 ± 1.1	135 ± 3	141.9 ± 2.6	N.D.	N.D.	427.7 ± 5.3	177.9 ± 2.2	23.5 ± 0.8	227.3 ± 4.8	202.1 ± 5.4	N.D.	N.D.	121 ± 5
1995	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	82.1 ± 0.6	13.7 ± 0.4	221.4 ± 1.8	157.4 ± 2.9	N.D.	N.D.	119.6 ± 1.2
2007	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	63.8 ± 0.3	15.8 ± 0.5	211.3 ± 4.7	123.7 ± 5.6	N.D.	N.D.	116.1 ± 1.3
2042	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	79.7 ± 2.7	16.7 ± 0.4	234.3 ± 4.2	129 ± 4.8	N.D.	N.D.	111.3 ± 2.5
2043	470 ± 7.6	179.4 ± 2.1	85.2 ± 4.4	180.6 ± 6.2	358 ± 8.7	172 ± 6.5	201.6 ± 5	N.D.	75.5 ± 2.5	4.5 ± 0.1	N.D.	N.D.	N.D.	N.D.	N.D.
2058	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	23.4 ± 0.8	242.7 ± 6.7	155.6 ± 1.5	N.D.	N.D.	143.4 ± 4.5
2065	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	96.2 ± 5.9	13.9 ± 0.5	286.2 ± 7.3	77.9 ± 5.9	N.D.	N.D.	178.4 ± 7.8

Table 8 continued

2066	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	110.7 ± 5.4	27.8 ± 2.4	272.9 ± 13.7	101.4 ± 8.5	N.D.	120.7 ± 4	182.5 ± 8.8
2067	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	89.3 ± 3.6	15.6 ± 0.8	211.7 ± 5.4	75.5 ± 3.5	N.D.	N.D.	140.6 ± 4
2154	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	152.4 ± 6.7	45.1 ± 4.5	483.9 ± 39.2	148.6 ± 13.6	N.D.	N.D.	293.9 ± 17.8
2229	125.7 ± 3.4	65.4 ± 1.2	67.2 ± 1.6	149.9 ± 3.6	166 ± 4.5	152.6 ± 3.2	144.9 ± 2.9	N.D.	88.7 ± 3.3	12.9 ± 0.7	170.5 ± 6.9	85.1 ± 0.6	N.D.	N.D.	88.6 ± 1.4
2340	165.1 ± 1.4	125.8 ± 5.6	92.4 ± 1.3	138.7 ± 6	156.2 ± 4.8	N.D.	N.D.	999.5 ± 4.3	149.9 ± 2.7	14.3 ± 1.3	188.8 ± 5.8	99.7 ± 4.5	N.D.	N.D.	108.5 ± 2.5
2436	90.7 ± 0.4	102.6 ± 1.2	71.1 ± 1.9	N.D.	131 ± 4.1	N.D.	N.D.	981.2 ± 25.2	677.3 ± 16.3	64.6 ± 2.5	448.6 ± 12.5	104.7 ± 1.2	N.D.	N.D.	166.5 ± 0.8
2466	76.9 ± 1	74.3 ± 0.6	60.8 ± 0.5	129.5 ± 1	137.7 ± 1.3	N.D.	N.D.	519.1 ± 2.5	136.7 ± 1.4	26 ± 0.2	232.5 ± 3.3	83.9 ± 1.3	N.D.	N.D.	111 ± 0.1
2553	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	71.8 ± 2.5	32.4 ± 0.6	344.5 ± 11.8	119.6 ± 4.6	N.D.	N.D.	199.5 ± 6.2
2569	56.3 ± 1.3	63.6 ± 1.3	56.6 ± 1	128.8 ± 2.4	133.6 ± 2.5	N.D.	N.D.	542.9 ± 12.2	174.9 ± 5.2	15.9 ± 0.6	139.1 ± 8	128.7 ± 1.9	N.D.	N.D.	107.7 ± 2.7
2601	292.2 ± 4.6	145.5 ± 2.2	70.6 ± 1.2	159.6 ± 4.1	319.1 ± 2.5	160.7 ± 3.9	198.2 ± 2.2	N.D.	64.6 ± 1	1.8 ± 0.1	114.1 ± 2.1	N.D.	N.D.	N.D.	72.5 ± 0.9
2607	102 ± 3.8	86.3 ± 3.1	69.9 ± 3.8	N.D.	144 ± 5.2	N.D.	N.D.	575.2 ± 23.4	114.1 ± 6.1	9.5 ± 0.4	153.5 ± 9.4	N.D.	N.D.	N.D.	98.3 ± 5.7
2672	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	18.5 ± 0.3	167 ± 5.4	192.9 ± 2	N.D.	123.8 ± 1.1	101 ± 0.6
2681	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	12.6 ± 0.5	130 ± 10.7	127.4 ± 2.4	N.D.	120.2 ± 2.2	85.2 ± 2.2
2685	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	14 ± 0.5	105.9 ± 0.6	30.7 ± 0.4	422.9 ± 5.9	84 ± 1.1	N.D.	N.D.	185.1 ± 1.4
2687	309.2 ± 4.1	113.2 ± 0.9	84.6 ± 2.3	168 ± 4.1	272.8 ± 2.1	158.2 ± 4.9	166 ± 4	N.D.	72.1 ± 1.5	3 ± 0.3	N.D.	N.D.	N.D.	N.D.	N.D.
2694	212.4 ± 5.8	74 ± 2.1	110.9 ± 5.3	191.9 ± 9	247 ± 9.7	190 ± 7.4	164.7 ± 6.1	N.D.	229.2 ± 9.4	44.1 ± 2.8	293.3 ± 23.2	N.D.	N.D.	134.4 ± 5.6	120.6 ± 6.9
2704	47.4 ± 1.9	56.5 ± 3.6	N.D.	N.D.	N.D.	N.D.	N.D.	450.6 ± 40.7	84.4 ± 4.6	32.7 ± 3.2	165.4 ± 1	N.D.	N.D.	102.3 ± 11.6	78 ± 4.4
2705	308.3 ± 5.2	90.8 ± 6.5	137.1 ± 4.2	173.3 ± 1.8	268.7 ± 1.5	166.4 ± 2.9	134.4 ± 2.3	N.D.	94.6 ± 1.5	7.3 ± 0.2	182.3 ± 54.4	83.6 ± 0.5	N.D.	N.D.	74.9 ± 1.1
2707	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	73.1 ± 2.2	11.8 ± 1	195.9 ± 5.8	159.4 ± 1.9	N.D.	N.D.	116.6 ± 2.9
2710	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	77.7 ± 1.9	23.2 ± 1	235.7 ± 5.6	113.2 ± 9.5	N.D.	N.D.	114.4 ± 1.6
2715	479 ± 19.5	142.9 ± 2.8	150.9 ± 11.6	238.8 ± 10.7	411.2 ± 15.9	201.5 ± 9.9	158.7 ± 4	N.D.	73.5 ± 2.3	5 ± 0.2	128.2 ± 6.7	N.D.	N.D.	N.D.	N.D.
2721	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	77.8 ± 0.6	20.4 ± 1.4	242.5 ± 16.9	144 ± 3.1	N.D.	N.D.	148.1 ± 3.3
2723	27.2 ± 0.4	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	15.6 ± 0.5	167.4 ± 3.7	61.4 ± 0.6	377.8 ± 11.7	87.8 ± 2.7	N.D.	N.D.	174.7 ± 3
2733	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	88.7 ± 2.1	15.5 ± 0.7	251.8 ± 5.7	130.1 ± 3	N.D.	N.D.	167.3 ± 2.9
2739	24.8 ± 0.7	75.6 ± 3.3	72.4 ± 3	129.2 ± 3.6	133.8 ± 3.7	N.D.	N.D.	5.8 ± 0.3	86.4 ± 1.9	22.7 ± 0.2	236.1 ± 5.4	85.2 ± 1.4	N.D.	N.D.	163.8 ± 2
2761	142.8 ± 0.5	89.2 ± 0.7	86.6 ± 1.5	138.9 ± 3	144.8 ± 3.5	N.D.	N.D.	1198.9 ± 8.3	299.8 ± 2.2	20.1 ± 0.3	298.4 ± 2.7	N.D.	N.D.	N.D.	136.7 ± 3.3
2768	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	130 ± 3	23.5 ± 0.9	236.9 ± 6.5	126.5 ± 3.9	N.D.	N.D.	135.7 ± 2.5
2790	116.8 ± 2.6	80.9 ± 0.4	75 ± 2.1	137.3 ± 5.6	147.2 ± 5.1	N.D.	N.D.	582.1 ± 7.7	132.2 ± 1	15.8 ± 0.4	198.7 ± 14.4	N.D.	N.D.	N.D.	101.5 ± 1.4
2872	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	88 ± 2.9	27.9 ± 0.1	229.4 ± 3.3	163.4 ± 6	N.D.	N.D.	128.9 ± 1.1
2903	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	126.4 ± 0.7	34.7 ± 0.2	326.1 ± 3	174.7 ± 9.6	N.D.	N.D.	174.6 ± 2.7
2924	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	81.8 ± 0.4	22 ± 1	274.9 ± 9.3	193.4 ± 4.1	N.D.	N.D.	131.8 ± 1.5
2925	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	75.1 ± 1.7	18 ± 0.9	282.9 ± 12.8	86.3 ± 0.9	N.D.	N.D.	108.1 ± 3.4

Table 8 continued

2927	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	23.5 ± 0.4	157.5 ± 4.6	138.5 ± 0.3	N.D.	140.3 ± 4.3	105.3 ± 2.2
2931	27.7 ± 0.4	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	9.8 ± 0.2	90 ± 2.5	21.5 ± 1.1	118.9 ± 5.7	79.1 ± 2	N.D.	N.D.	110.3 ± 0.8
2939	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	7.3 ± 0.5	95.2 ± 6.6	72.9 ± 2.2	N.D.	122.7 ± 2.6	85.3 ± 1.7
2968	55.3 ± 1.1	65 ± 1.2	69.4 ± 0.3	N.D.	137.2 ± 0.8	N.D.	N.D.	510 ± 8.8	164.9 ± 4.5	18.1 ± 0.6	212.2 ± 10	184.8 ± 11.6	N.D.	N.D.	120.4 ± 2.5
2979	61.3 ± 1.1	67.3 ± 1.2	62.1 ± 1.5	129.9 ± 4.4	132.2 ± 4.4	N.D.	N.D.	674.7 ± 2.4	208.8 ± 4.7	16.6 ± 0.5	201.9 ± 2.6	162.1 ± 0.6	N.D.	N.D.	138.3 ± 2.6
2989	99 ± 2.3	93.4 ± 2.6	72.7 ± 1.5	137.1 ± 4.4	146.3 ± 4.6	N.D.	N.D.	679.1 ± 13.6	113.8 ± 3	14.1 ± 0.8	151.4 ± 3.9	N.D.	N.D.	N.D.	91.7 ± 2.2
2995	24.8 ± 0.7	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	359 ± 26.3	15.5 ± 1.9	334.7 ± 16.3	104.6 ± 5.4	N.D.	N.D.	176.2 ± 6.7
2997	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	101.4 ± 1.4	20.3 ± 1.2	280.7 ± 13.3	133.4 ± 12.7	N.D.	N.D.	179.6 ± 5.8
3052	70.2 ± 1.1	72.4 ± 2	63.3 ± 2.5	144.4 ± 5.5	N.D.	N.D.	N.D.	1000.9 ± 12.4	100.6 ± 2.7	8.8 ± 0.3	156.1 ± 4.1	N.D.	N.D.	N.D.	100.4 ± 3.5
3229	99.3 ± 1	82 ± 1	73.2 ± 2.1	132.4 ± 4.5	136.4 ± 4.6	N.D.	N.D.	1010.4 ± 5.6	111.2 ± 1.3	9 ± 0.4	165.2 ± 4.4	N.D.	N.D.	N.D.	93.4 ± 1.9
3294	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	89.6 ± 1.8	28.5 ± 0.7	250.6 ± 4.2	180.7 ± 5.8	N.D.	N.D.	133.3 ± 1.6
3306	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	84.1 ± 1.4	27.5 ± 0.2	240.3 ± 1.4	97.7 ± 1.2	N.D.	N.D.	174.5 ± 0.2
3315	30.1 ± 0.6	N.D.	56.8 ± 1.2	N.D.	N.D.	N.D.	N.D.	39.5 ± 1	186.1 ± 1.4	17.8 ± 0.4	274.1 ± 2.3	151.4 ± 1.8	N.D.	N.D.	148.1 ± 3.4
3362	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	83.1 ± 3.1	17.4 ± 1.7	229.4 ± 14.5	N.D.	N.D.	N.D.	112.5 ± 6.8
3377	301.3 ± 3.9	157.5 ± 4	N.D.	167.7 ± 8.5	317.2 ± 13.7	160 ± 6.9	206.3 ± 8.7	N.D.	88.3 ± 2.9	9.5 ± 0.5	106.5 ± 4.8	N.D.	N.D.	N.D.	81.3 ± 2.4
3380	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	102.3 ± 3.5	30 ± 0.3	297.7 ± 5.6	99.7 ± 6.5	N.D.	N.D.	138.2 ± 4.9
3424	N.D.	121.7 ± 8.3	97.9 ± 3.9	133.6 ± 5.5	150.3 ± 6.6	N.D.	N.D.	N.D.	123.7 ± 1.3	27.7 ± 1.2	270.9 ± 4.1	152.6 ± 4.5	N.D.	N.D.	151.1 ± 1.6
3443	24.3 ± 0.2	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	30.6 ± 1.9	286.5 ± 8.1	31.4 ± 0.9	333.9 ± 8.1	N.D.	N.D.	N.D.	147.1 ± 2.6
3461	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5.6 ± 0.3	99.2 ± 3	15.1 ± 0.6	294.1 ± 6.6	95.8 ± 4.1	N.D.	N.D.	159.5 ± 4.6
3463	61 ± 1.4	60 ± 0.8	60.4 ± 0.8	125.6 ± 2.1	127.7 ± 1.9	N.D.	N.D.	51.3 ± 2.9	110.9 ± 10.2	9.2 ± 0.4	183.2 ± 4.5	97.3 ± 0.1	N.D.	N.D.	110.2 ± 2.1
3466	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	153.5 ± 9.3	32.1 ± 3.2	350 ± 24.9	N.D.	N.D.	N.D.	153.7 ± 6.5
3470	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	73.3 ± 0.5	13.2 ± 1.1	221.4 ± 10.7	144.1 ± 3.4	N.D.	N.D.	138.9 ± 6.9
3485	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	17.7 ± 0.4	74.3 ± 0.5	21.7 ± 0.5	203.7 ± 1.7	93.2 ± 1.5	N.D.	N.D.	124.1 ± 0.8
3495	94 ± 2.2	76.6 ± 1.6	75.1 ± 2.4	129.2 ± 4.6	139.2 ± 4.4	N.D.	N.D.	982.7 ± 26.5	147.7 ± 4.1	17.8 ± 1	201.2 ± 6.5	76.6 ± 1.8	N.D.	N.D.	99 ± 1.7
3509	71.1 ± 1.4	N.D.	67.7 ± 2.2	N.D.	N.D.	N.D.	N.D.	616.8 ± 13.9	154.1 ± 2.6	37.7 ± 0.7	94.1 ± 3.6	122.3 ± 2.9	N.D.	N.D.	92.1 ± 4.5
3510	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5 ± 0.2	104.8 ± 1.9	27.7 ± 0.8	321.5 ± 6.7	77.6 ± 2.4	N.D.	N.D.	183.1 ± 5
3511	28.9 ± 1	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	6.1 ± 0.2	312.9 ± 15.6	9.5 ± 0.8	256.7 ± 12.9	N.D.	N.D.	N.D.	130.2 ± 4.6
3512	132.8 ± 1.8	72.6 ± 1.5	75.8 ± 2.2	135.1 ± 4.4	141.9 ± 4.4	N.D.	N.D.	832.7 ± 12.3	128.8 ± 2.4	20 ± 0.6	235.6 ± 2.5	N.D.	N.D.	N.D.	102.5 ± 2.8
3518	114.5 ± 2.3	80.2 ± 0.9	75.5 ± 0.7	132.5 ± 3.6	137.6 ± 3.6	N.D.	N.D.	628.5 ± 5.6	148.6 ± 2.3	28.7 ± 1	215.6 ± 3.9	N.D.	N.D.	N.D.	105.9 ± 2.3
3539	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	129.4 ± 1.7	29.5 ± 0.3	235.4 ± 2.5	N.D.	N.D.	N.D.	132.6 ± 2.7
3545	87.3 ± 1.3	65.5 ± 0.9	81.5 ± 1.5	131 ± 1.9	137.5 ± 1.7	N.D.	N.D.	1072.4 ± 20.3	178.3 ± 1.8	20.5 ± 0.7	212.4 ± 2.3	97.5 ± 4.4	N.D.	N.D.	102 ± 3.3
3546	31.8 ± 0.2	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	40.8 ± 0.4	518.5 ± 3.8	8.8 ± 0	214.7 ± 1	N.D.	N.D.	N.D.	115.6 ± 1.1

Table 8 continued

3547	26 ± 0.4	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	7.7 ± 0.2	110.2 ± 3.8	27.2 ± 0.6	370.2 ± 16.2	N.D.	N.D.	N.D.	161.5 ± 8.3
3552	513.2 ± 5.6	177.3 ± 3.2	114.1 ± 2.1	198 ± 6.2	444.3 ± 3.6	171.4 ± 5.1	219.7 ± 4.1	N.D.	86.9 ± 1.8	4 ± 0.4	N.D.	N.D.	N.D.	N.D.	75.7 ± 3.2
3602	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	7.6 ± 0.3	88.8 ± 1.3	10.9 ± 0.6	190.7 ± 0.8	122.9 ± 1.2	N.D.	N.D.	116.3 ± 1.4
3603	119 ± 1.4	85.1 ± 6.9	82.5 ± 1.8	135.4 ± 4.5	145.5 ± 4.5	N.D.	N.D.	927.5 ± 1.8	175.8 ± 5.2	20 ± 1	196.4 ± 6.9	N.D.	N.D.	N.D.	131.9 ± 3.7
3618	128.6 ± 3.3	106 ± 2.1	71.2 ± 2	136.4 ± 4.7	162.5 ± 4.4	N.D.	N.D.	746.7 ± 15.3	196.2 ± 6.4	37.7 ± 1.4	289.8 ± 7.2	103.1 ± 6.1	N.D.	N.D.	119 ± 2
3630	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	103.8 ± 1.7	32.3 ± 0.5	350.9 ± 1.6	119.4 ± 7.3	N.D.	N.D.	203.1 ± 3.3
3658	27.8 ± 0.6	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	6 ± 0.2	397 ± 3.5	29.2 ± 0.7	333.3 ± 2.9	N.D.	N.D.	N.D.	165.6 ± 0.3
3672	155.7 ± 3.2	140.5 ± 4.1	71.5 ± 1.1	N.D.	149.5 ± 2.4	N.D.	N.D.	1235.9 ± 27.5	156.8 ± 4.6	19 ± 0.8	206 ± 13	N.D.	N.D.	N.D.	133.8 ± 3.3
3688	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	58 ± 2.2	7.7 ± 0.3	189.2 ± 11.7	68.1 ± 3.1	N.D.	N.D.	140.6 ± 5.9
3690	103.8 ± 0.8	63.7 ± 1.3	83.4 ± 0.7	136.8 ± 3.8	140 ± 3.4	N.D.	N.D.	813.9 ± 13.6	179.1 ± 3	16.8 ± 0.8	293.1 ± 8.5	181.2 ± 15.4	N.D.	N.D.	118 ± 1.3
3691	202.7 ± 5.7	120.6 ± 2	65.5 ± 1.4	158.2 ± 2.4	308.9 ± 3.9	174.9 ± 2.5	230.4 ± 2.3	12 ± 0.3	128.7 ± 1.1	18.7 ± 0.3	145.7 ± 7.2	124.1 ± 2.4	N.D.	N.D.	95.8 ± 0.5
3712	83.1 ± 1	72.9 ± 1.4	67.2 ± 1.3	125.4 ± 3.2	126.9 ± 3	N.D.	N.D.	89.8 ± 1.5	101.9 ± 1.9	26.4 ± 0.3	241.7 ± 2.4	N.D.	N.D.	N.D.	125 ± 0.2
3730	73.8 ± 0.6	62.1 ± 2.5	61.2 ± 2.8	128.7 ± 7.5	130.5 ± 7.5	N.D.	N.D.	30 ± 0.9	90.7 ± 2.8	14.7 ± 0.7	210.7 ± 2.2	146 ± 2.9	N.D.	N.D.	95.8 ± 2
3973	375.1 ± 7.1	136.6 ± 2.4	100.7 ± 1.7	171.2 ± 2.5	388.7 ± 2.7	176.1 ± 2.1	164.2 ± 4.5	N.D.	73.7 ± 1.2	5.9 ± 0.2	103.3 ± 1.3	N.D.	N.D.	N.D.	77.4 ± 1
3980	118.7 ± 1.8	95.2 ± 10.6	97.7 ± 6.4	137.5 ± 5.3	163.3 ± 4.9	N.D.	N.D.	1006.5 ± 17.4	259.1 ± 6.4	39.3 ± 2	303.2 ± 4.3	N.D.	N.D.	N.D.	136.7 ± 3.2
4243	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	118.6 ± 5.8	35.7 ± 2.1	357.6 ± 16.7	93.3 ± 3.3	N.D.	N.D.	170.7 ± 7
4249	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	103.1 ± 1.1	28.4 ± 1.1	265.9 ± 6	143.2 ± 11.4	N.D.	N.D.	139.5 ± 1
4279	25.9 ± 0.9	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5.6 ± 0.2	104.4 ± 3.7	21.8 ± 0.4	345.6 ± 0.9	75.2 ± 2.1	N.D.	N.D.	121 ± 2.5
4425	25.9 ± 0.7	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	6.6 ± 0.3	163 ± 5.2	19 ± 0.7	270.9 ± 8.2	157.9 ± 1.3	N.D.	N.D.	127.1 ± 3.2
4487	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	90.7 ± 3.6	27.9 ± 0.2	361.6 ± 9.5	107.8 ± 0.4	N.D.	N.D.	128.2 ± 3.6
4500	25.8 ± 0.4	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5.6 ± 0.2	113.9 ± 0.7	36.9 ± 0.3	415.5 ± 2.6	79.5 ± 1.6	N.D.	123.7 ± 1.9	161.9 ± 0.8
4507	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	76 ± 0.4	15 ± 0.9	260.6 ± 8.2	126 ± 4.6	N.D.	N.D.	147.4 ± 2.9
4514	28.4 ± 0.6	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	16.1 ± 0.6	185.1 ± 2.3	28.7 ± 0.4	256.9 ± 4.2	179.5 ± 6.8	N.D.	N.D.	137.1 ± 1.3
4528	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	84.7 ± 3.5	5.5 ± 0.2	214.5 ± 3.5	103.8 ± 2.5	N.D.	N.D.	194.6 ± 5.8
4530	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	8.5 ± 0.3	69.3 ± 1.8	23.1 ± 0.8	225.5 ± 4.3	124.9 ± 7.7	N.D.	N.D.	129.3 ± 3.8
4536	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	12 ± 0.4	73.6 ± 2.4	24.2 ± 0.4	266.1 ± 10.5	81.8 ± 3.8	N.D.	N.D.	166 ± 4.7
4545	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	27.9 ± 0.9	203.5 ± 9.2	132.4 ± 2	N.D.	N.D.	136.7 ± 3.7
4546	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	80.5 ± 1.3	21.3 ± 0.3	286.6 ± 3.7	110 ± 5.5	N.D.	N.D.	171.5 ± 0.3
4547	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5 ± 0.2	76 ± 4	16.7 ± 2.9	237.8 ± 31.8	122.3 ± 12.7	N.D.	N.D.	151.7 ± 13.9
4548	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	86.6 ± 2	30.6 ± 0.8	350.4 ± 6.3	148.9 ± 4.3	N.D.	N.D.	175.7 ± 4.6
4549	25.7 ± 1.2	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	90.1 ± 2.4	55.4 ± 5.7	377.6 ± 30	115.7 ± 2.1	N.D.	124.5 ± 5.6	151.6 ± 6.4
4714	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	13.8 ± 0.3	111.1 ± 1.6	29.2 ± 1	350.7 ± 7.8	74.5 ± 1.6	N.D.	N.D.	194.2 ± 2.7

Table 8 continued

5074	26 ± 0.6	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	82.4 ± 2.3	19.4 ± 0.2	233.6 ± 5.6	179.1 ± 2.4	N.D.	N.D.	154.3 ± 2.2
5075	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	90 ± 2.5	24.5 ± 0.2	317.4 ± 5.1	171.2 ± 2	N.D.	N.D.	241.4 ± 5.7
5108	24.8 ± 0.6	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	6.4 ± 0.2	80.6 ± 1.8	20.7 ± 0.3	290.4 ± 2.5	74.8 ± 0.9	N.D.	N.D.	127.4 ± 3.6
5549	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	72.8 ± 2.8	14.4 ± 0.7	212.9 ± 8.5	93.3 ± 1.8	N.D.	N.D.	114.1 ± 2.7
5557	100.6 ± 0.4	55.3 ± 1.4	60.2 ± 1.2	135.1 ± 3.2	140.1 ± 3.3	144.3 ± 3.9	144.7 ± 4	4.9 ± 0.2	63 ± 1.6	0.1 ± 0.2	N.D.	N.D.	N.D.	N.D.	N.D.
5596	274.7 ± 6.7	89.8 ± 2.5	114.3 ± 3.4	188.3 ± 2.2	255.8 ± 5.5	182.8 ± 1.8	147.5 ± 6.2	17.6 ± 0.1	95.8 ± 0.4	14.4 ± 0.5	134.5 ± 2.7	154 ± 8.1	N.D.	119.8 ± 3.2	100.8 ± 2.1
5601	142.2 ± 2.5	68.3 ± 2.9	60.4 ± 3	136.3 ± 7.4	145 ± 8.5	142 ± 7.6	141.4 ± 7.4	N.D.	N.D.	0.2 ± 0.2	N.D.	N.D.	N.D.	N.D.	N.D.
5615	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	93.4 ± 3	11.9 ± 0.2	241.6 ± 7.1	125.1 ± 4.4	N.D.	N.D.	146.4 ± 4.2
5617	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	82.4 ± 1.2	10 ± 0.3	131.1 ± 3	N.D.	N.D.	N.D.	271 ± 4.5
5721	60.6 ± 1	72.7 ± 1	58.5 ± 1.6	N.D.	136 ± 3.2	N.D.	N.D.	622.1 ± 8.4	239.8 ± 5.9	24.5 ± 0.3	317.5 ± 8.6	181.8 ± 5.3	N.D.	N.D.	123.9 ± 1.6
5732	133.1 ± 1.1	88.7 ± 2.7	91.1 ± 1.3	N.D.	N.D.	N.D.	N.D.	1299.2 ± 1.3	247.6 ± 2.9	28.5 ± 1.2	299 ± 1.7	N.D.	N.D.	N.D.	153.9 ± 0.3
6068	121.1 ± 3.3	72.4 ± 2.9	96 ± 2.2	138.6 ± 4.6	145.7 ± 4.7	N.D.	N.D.	623.5 ± 11.6	155.2 ± 5.8	53.9 ± 0.9	261.5 ± 8.2	N.D.	N.D.	N.D.	139 ± 4
6073	188.2 ± 6.1	74.1 ± 0.8	68.2 ± 0.6	147.9 ± 1.1	170.8 ± 1.3	143.6 ± 1.2	138.1 ± 1.1	N.D.	64 ± 0.5	5.5 ± 0.1	N.D.	N.D.	N.D.	N.D.	71.2 ± 0.6
6259	111.6 ± 5.2	72.2 ± 4.8	81.7 ± 2.4	132.1 ± 6.4	137.8 ± 6.3	N.D.	N.D.	621.5 ± 25.5	144.8 ± 3.6	26.5 ± 0.8	252.8 ± 9.5	109 ± 6.5	N.D.	N.D.	123.5 ± 6.4
6584	453 ± 7.1	171.5 ± 6.2	101.4 ± 3.6	189.4 ± 3.5	404.2 ± 3.1	167 ± 2.9	193.9 ± 2.1	N.D.	72.6 ± 2.3	1.3 ± 0.2	N.D.	N.D.	N.D.	N.D.	N.D.
6588	167.7 ± 2.8	109.9 ± 3	72.4 ± 1	157.7 ± 2.8	184.9 ± 2.9	146.2 ± 3.2	165.7 ± 3.4	528.2 ± 9.3	116.2 ± 1.2	15.5 ± 0.7	172.6 ± 3.3	N.D.	N.D.	N.D.	100.1 ± 2.5
6589	323.9 ± 6.9	136.3 ± 3.5	101.3 ± 2.6	190.5 ± 7.1	315.3 ± 7.5	176.9 ± 6.6	207.1 ± 6.4	366.9 ± 13.8	139.8 ± 5.5	26.5 ± 0.9	142.8 ± 5.5	N.D.	N.D.	N.D.	90.9 ± 2.9
6598	26.4 ± 0.3	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	6.9 ± 0.1	123.7 ± 4.7	27.9 ± 1.8	274.4 ± 13.9	169.9 ± 9.3	N.D.	N.D.	174 ± 4.5
6628	55.2 ± 2.5	55.7 ± 2.1	63.5 ± 2.1	131.4 ± 6.4	131.3 ± 6.5	N.D.	N.D.	723.7 ± 42.9	103.5 ± 5.1	12.3 ± 1.6	225.2 ± 12.5	113.6 ± 2.8	N.D.	N.D.	91.3 ± 2.3
6631	31.6 ± 0.1	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	6.8 ± 0.9	438.9 ± 16.8	23.7 ± 1.3	317.5 ± 10.7	80.1 ± 1.1	N.D.	N.D.	145.4 ± 1.2
6632	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	85.2 ± 1.9	17.6 ± 0.6	279.8 ± 5.2	77 ± 2.7	N.D.	N.D.	187.1 ± 3.5
6633	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	110.2 ± 2	29.1 ± 1.1	328.3 ± 4.2	87 ± 1	N.D.	N.D.	151.3 ± 0.8
6647	143.4 ± 2	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	1085.5 ± 9.9	224.7 ± 0.9	71.1 ± 0.5	414.3 ± 8	88.8 ± 2.5	N.D.	N.D.	199.1 ± 3.8
6655	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	11.5 ± 1	165.6 ± 7.6	N.D.	119.4 ± 2.3	125.4 ± 2.4	108.3 ± 0.9
6700	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	84.2 ± 2.6	12.3 ± 0.3	180.8 ± 0.8	93.2 ± 3	N.D.	N.D.	109.4 ± 2.8
6732	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	93.4 ± 1.8	10.7 ± 0.2	145.5 ± 2.5	182.6 ± 12.7	N.D.	N.D.	143.4 ± 1.2
6748	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5.1 ± 0.3	93.5 ± 3.3	46.5 ± 1.4	279.6 ± 10.6	98.8 ± 3.5	120.2 ± 5.2	126 ± 5.3	158.3 ± 3.9
6765	28.3 ± 0.3	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	10.9 ± 0.4	100.8 ± 0.8	38.6 ± 0.4	243.7 ± 1.3	181.9 ± 3.5	N.D.	N.D.	163.2 ± 1.1
6879	71.6 ± 1.8	92.1 ± 3.8	62.1 ± 3.9	N.D.	149.7 ± 10	N.D.	N.D.	537.4 ± 23	472 ± 18.3	5.9 ± 0.2	176.8 ± 9.2	83.7 ± 6.1	N.D.	N.D.	111.1 ± 5.9
6909	69.6 ± 0.3	62.3 ± 1.5	70.2 ± 0.9	127.5 ± 2.8	134.4 ± 3.1	N.D.	N.D.	658 ± 5	173.1 ± 4	25 ± 0.3	240.2 ± 0.5	87 ± 1.7	N.D.	N.D.	115.1 ± 0.3
6915	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	13.9 ± 0.2	66 ± 3	19.3 ± 0.3	290.7 ± 2.7	85.9 ± 3.3	N.D.	117 ± 4.5	133.5 ± 1.5
6928	193.2 ± 6.2	128.7 ± 2.5	89 ± 2.8	142.8 ± 2.4	148.5 ± 2.4	N.D.	N.D.	1824 ± 59.3	362.4 ± 10.3	48.4 ± 2.7	338.8 ± 9.7	N.D.	N.D.	N.D.	189.6 ± 4.6

Table 8 continued

6940	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	71.3 ± 1.2	30.9 ± 1.8	357.7 ± 9.6	153.4 ± 4.9	N.D.	N.D.	218.2 ± 4.8
6954	58.2 ± 0.3	61.5 ± 2.6	60.6 ± 1.9	143.4 ± 6	175.4 ± 5.3	159 ± 6.3	148.9 ± 7.1	7.6 ± 0.2	72.2 ± 3.1	4.6 ± 0.5	143.1 ± 3.7	N.D.	N.D.	N.D.	80.5 ± 3.8
6981	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	91 ± 1.2	15.3 ± 0.1	264.3 ± 1.2	71.9 ± 0.6	N.D.	N.D.	230.8 ± 0.9
6982	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	117.3 ± 3.4	39.2 ± 0.3	461.6 ± 8.9	152.1 ± 4.6	N.D.	N.D.	191.6 ± 3.9
08/-08	41.2 ± 0.3	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	21.2 ± 0.8	409.3 ± 4	19.8 ± 0.1	275.5 ± 3.4	N.D.	N.D.	N.D.	143.4 ± 1.4
e1262	28.4 ± 0.9	N.D.	55.5 ± 1.8	N.D.	N.D.	N.D.	N.D.	11 ± 0.3	169.7 ± 1.9	36.8 ± 0.8	280.3 ± 3.1	94.4 ± 1.1	N.D.	N.D.	157.4 ± 2.9
e1827	111.4 ± 1.7	92.5 ± 7.9	72.5 ± 3.7	133.8 ± 5.3	138.3 ± 5.3	N.D.	N.D.	732.6 ± 10	107.6 ± 4.7	7.3 ± 0.5	145.4 ± 0.4	N.D.	N.D.	N.D.	96.3 ± 3.1
e2180	323.8 ± 8.7	218.6 ± 23.9	128.7 ± 13.2	140.7 ± 4.3	187.9 ± 5.9	N.D.	N.D.	1686.9 ± 29.1	252.3 ± 6.9	51.6 ± 1.9	370.6 ± 14.5	N.D.	N.D.	N.D.	193.2 ± 3.4
e2474	61.8 ± 1.8	77.7 ± 2.8	63.1 ± 2.9	N.D.	140.5 ± 6.7	N.D.	N.D.	763.9 ± 23.9	156.9 ± 5.5	18.3 ± 0.8	250.8 ± 6.7	N.D.	N.D.	N.D.	175.4 ± 6.4
e3406	319.2 ± 4.9	161.5 ± 3.5	61.6 ± 1.2	179.8 ± 3.1	266 ± 6.2	163.8 ± 3.8	182.6 ± 3.6	7.8 ± 0.6	69.5 ± 2	4.2 ± 0.1	89.5 ± 2.5	N.D.	N.D.	N.D.	74 ± 1.4
e3756	105.8 ± 1.7	101.8 ± 1.3	66.6 ± 1.7	N.D.	N.D.	N.D.	N.D.	796.6 ± 5	127 ± 3	19.2 ± 0.6	222.7 ± 3.5	N.D.	N.D.	N.D.	117.6 ± 1.4
e3756	25 ± 0.2	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	5.7 ± 0.2	94.6 ± 1.1	12.8 ± 0.4	221.6 ± 2.2	N.D.	N.D.	N.D.	184.5 ± 2
Nycos	63.6 ± 2.4	70.1 ± 3.2	59.4 ± 0.5	N.D.	N.D.	N.D.	N.D.	94.7 ± 2.3	107.7 ± 3.2	18.6 ± 1	160.2 ± 7.3	N.D.	N.D.	N.D.	98.2 ± 9
Yarra	51.6 ± 0.7	69 ± 0.5	N.D.	N.D.	N.D.	N.D.	N.D.	41 ± 1.3	85.8 ± 1.4	5.9 ± 0.4	N.D.	N.D.	N.D.	N.D.	N.D.
1107-1	31.7 ± 0.9	N.D.	57.3 ± 2.2	135.6 ± 5.1	137.8 ± 5.4	N.D.	N.D.	81 ± 2.8	247.3 ± 4.9	50.5 ± 1.4	352.4 ± 4.2	108.4 ± 1.7	N.D.	129.8 ± 4.9	130.7 ± 3
1107-2	59.4 ± 0.9	71.3 ± 1.2	57.5 ± 1.7	N.D.	125.5 ± 3.8	N.D.	N.D.	917.8 ± 1.7	231.5 ± 1.4	18 ± 0.2	214.8 ± 1.2	133.4 ± 5	N.D.	N.D.	161.6 ± 3
M82-1	74.2 ± 0.9	75.4 ± 1.4	60.6 ± 2.6	127.7 ± 4.4	130.5 ± 5	N.D.	N.D.	381.5 ± 13.3	103 ± 0.8	22.3 ± 0.1	189.9 ± 3.1	N.D.	N.D.	N.D.	119.8 ± 5.8
M82-2	97.6 ± 3.7	88.9 ± 2.8	60.3 ± 2.7	131.9 ± 6	139.6 ± 6.4	N.D.	N.D.	490.6 ± 13.1	87.6 ± 3.2	14 ± 0.8	143.7 ± 2.4	N.D.	N.D.	N.D.	95.4 ± 6.9
M82-3	79.3 ± 2.1	79.8 ± 1.7	58.9 ± 1.1	126.3 ± 3.5	131.7 ± 3.8	N.D.	N.D.	388.2 ± 9.4	95.2 ± 2	19.5 ± 1	165.4 ± 4.9	N.D.	N.D.	N.D.	100.8 ± 1.8
M82-4	95.7 ± 2.1	84.1 ± 1	64.7 ± 1.2	134.6 ± 4.1	141.2 ± 3	N.D.	N.D.	538.4 ± 7.6	105.8 ± 0.6	19.3 ± 1.2	210.8 ± 11.2	N.D.	N.D.	N.D.	113.2 ± 3.3
M82-5	67.1 ± 2.4	70 ± 2.3	58.3 ± 3.4	135.5 ± 4.4	136.6 ± 5.2	N.D.	N.D.	404.1 ± 15.5	81.9 ± 2.2	3.6 ± 0.1	109.7 ± 6.8	N.D.	N.D.	N.D.	83 ± 1.4
QC	274.9 ± 3.5	170.8 ± 3.9	104 ± 4	N.D.	171.2 ± 1.1	N.D.	N.D.	1430 ± 20.2	202.6 ± 2.7	40 ± 0.5	301.2 ± 4.3	N.D.	N.D.	N.D.	122.3 ± 1

4.5 Discussion

4.5.1 High throughput screening of phenolic and colour profiles

HPLC has long been a successful analytical platform for the separation and subsequent quantification of plant phenolics (Galensa and Herrmann, 1980). In recent years, UPLC has offered the potential for greater high throughput separation of phenolics (Spacil et al., 2008). An established HPLC-DAD method for the separation of phenolic compounds from tomato (Melendez-Martinez et al., 2010) was used as the basis for developing a high throughput method using UPLC-PDA. The resulting method comprised a 9 min gradient that separated a mixture of 14 authenticated standards of phenylpropanoids and flavonoids, as shown by Figure 19. A study on a previously published phenolic separation method by UPLC (Spacil et al., 2008) highlighted the advantages of high throughput methods. In addition to the advantage of less time required for chromatographic separation, the study identified that smaller volumes of mobile phase were needed and smaller waste solvent volumes were therefore produced. This resulted in reduced financial costs and environmental impact (Spacil et al., 2008). These advantages are seen here too, and are substantial in the context of large-scale mapping population screening. Considering the *S. pennellii* advanced backcross population (BC_2, TA209 x LA1657) comprising 175 genotypes (Frary et al., 2004) as an example, if this were grown in triplicate for biological replicates and analysed in triplicate for technical replicates, the sample set (not accounting for parental lines or quality controls) would comprise 1575 samples. Screening this sample set for phenolics using the HPLC-DAD method used in section 3.2 would consume a volume of 63.0 l of mobile phase solvents and occupy 1050 h of chromatography. By comparison, the UPLC-PDA method discussed in section 4.2 would use only 5.7 l of solvent over a duration of 236.25 h chromatography.

Optimising resolution of both phenylpropanoids and flavonoids using a high throughput platform was challenging. The close proximity of many of the peak retention times shown in Figure 19 was a trade-off for the benefit of reducing chromatography time. Some peak shapes showed asymmetry, which was especially true of the flavonols and phenolic acids, as a result of rapid changes in mobile phase gradients. Early eluting peaks, especially, were broad, and this was likely an effect of methanol from the injection affecting the mobile phase composition (Spacil et al., 2008). However, the high throughput gradient was successful in the separation of caffeic and chlorogenic acids, previously shown to be difficult by UPLC due to similar

retention properties (Novakova et al., 2010). The method was also able to separate ferulic and isoferulic acids, which was not possible in previous attempts (Figure 18). While the limitations of this UPLC-PDA method were recognised, so are the benefits. The successful application of this method, together with the UPLC-PDA method for the separation of isoprenoid and related compounds was illustrated by section 4.4 where a population of colour mutants was screened.

Both high throughput methods (for phenolics and isoprenoid compounds) were successful in identifying secondary metabolite profile variation within the CC colour mutant population (Table 7 and Table 8). The identification of characteristic metabolite profiles thereby supported the validity of the high throughput method for screening of large-scale populations.

No clustering was identified by PCA based solely on phenolic profiles (Figure 22). In fact, separation of data points based on experimental error (shown by QC samples) showed a similar degree of variation to CC genotype data points separated by differences in phenolic profiles. This was not the case for PCA of isoprenoid and related compounds (Figure 23), where a distinct cluster was observed in which genotypes exhibited profiles similar to *tangerine* mutant (Isaacson et al., 2002). Furthermore, the cluster observed in Figure 23 was also visible when phenolic and isoprenoid datasets were combined (Figure 24). This indicated that the variation in isoprenoid and related compounds had a greater influence on the characterisation of CC genotypes than did the phenolic profiles. Although extreme changes in tomato phenolic profiles can alter perceived fruit colour (Ballester et al., 2010; Butelli et al., 2008b) many phenylpropanoids and flavonoid are either colourless or relative to isoprenoids exhibit minimal contribution to overall fruit colour due to limited distribution in fruit tissues (Ballester et al., 2010; Torres et al., 2005; Winkel-Shirley, 2001a). Since CC genotype accessions were selected for inclusion based on organoleptic (visual colour) properties rather than biochemical profiles, it is unsurprising that the greatest biochemical variation is observed for compounds such as carotenoids and chlorophylls that when altered contribute to fruit colour perception (Lewinsohn et al., 2005).

4.5.2 Alternative extraction method for isoprenoid compounds

The dangers and side effects associated with exposure to chloroform have been known since its decline in popularity as a general anaesthesia (Whitaker and Jones,

1965). Although high exposures in modern day use are largely restricted to industry and the laboratory, there is still a concern that prolonged or acute exposure can result in, among other side effects, cardiac toxicity, arrhythmia, renal cancer, pregnancy loss and birth defects (Nagano et al., 2006; Narotsky et al., 2011; Zhou et al., 2011).

A brief assessment as to whether MTBE could offer an alternative, less harmful, option to chloroform for the extraction of isoprenoids from plant tissue was investigated here, similar to that conducted previously by Matyash and colleagues (2008). These findings indicate that MTBE is a competent substitute for chloroform for the extraction of isoprenoids from plant material, and support previous findings with animal cells and bacteria (Matyash et al., 2008). For the majority of direct comparisons there was no significant difference between using MTBE and chloroform in isoprenoid quantification. Examples exist for both chloroform and MTBE where significantly greater amounts of phytoene or lycopene are detected, and these are the most highly abundant compounds. This might suggest that neither method offers a total extraction. It is possible that these values are near the limits of the extraction rather than differing as a result of the choice of solvent. As stated by Matyash and colleagues (2008), and supported by Figure 21, MTBE extraction in place of chloroform facilitates removal of the organic phase by pipette by reducing (a) the likelihood of transferring contaminating plant debris and (b) the residual loss of organic phase during transfer through the aqueous phase, known as 'dripping loss'. However, it was noticed here that MTBE requires greater time to take the extract to dryness by centrifugal evaporation. While both methods offer advantages, it is concluded here that extraction with chloroform is favoured due to the greater potential for high throughput analysis of large population sample sets. As a result, this extraction protocol will continue to be used for future analysis in this work.

5 Characterisation of QTL-containing lines

5.1 Introduction

This chapter described the characterisation of three lines from mapping populations that, as discussed in the previous chapter, possessed elevated levels of phenolic intermediates in ripe tomato fruit. These lines are *S. neorickii* BIL population (Fulton et al., 2000; Grandillo et al., 2011) genotypes neo-111 and -123, and *S. habrochaites* NIL population (Monforte and Tanksley, 2000) genotype 3939. Throughout this chapter these genotypes are hereby referred to solely by their accession numbers: neo-111, neo-123, and 3939.

Work described in the previous chapter successfully confirmed the high phenolic phenotype of neo-111 and -123 from replicated trials in the field and glasshouse environments. The high phenolic profile of 3939 previously identified by T. Wells (unpublished) was also confirmed in a glasshouse environment. In this chapter results are reported from genotypes that were cultivated once again in the glasshouse. Phenolic profiles were determined throughout fruit development and ripening. A better understanding of the perturbations to phenolic regulation of the plants was ascertained by investigation of phenolic composition in different plant tissue types. A metabolomic analysis expanded the investigation from a pathway-specific study to an attempt to understand the effects on global fruit metabolism, with a focus on isoprenoid compounds as well as non-targeted GC-MS analysis.

The impact of characterised metabolite perturbations (that are perceived as improvements with regards to desirable health-promoting traits) on broader fruit quality traits must first be determined before their potential application can be assessed. Commercially, there is no benefit from improving health-promoting traits if these are at a cost to organoleptic or agroeconomic quality traits that limit commercial success by, for example, lowering yield or hindering consumer preference. Equally, genotypes selected for improved health-promoting traits (or desirable metabolite profiles) may also possess improved organoleptic or agroeconomic traits.

With this in consideration, fruit physiological parameters, such as fruit size and colour, were assessed in this chapter in order to better determine organoleptic quality traits. Parameters including yield and fruit mass, as well as post-harvest properties were investigated in order to assess the agroeconomic quality traits of these genotypes. Finally, the antioxidant capacity of phenolic extracts was measured to assess whether or not these metabolite perturbations resulted in a measurable health-promoting characteristic. Together with the analyses of metabolite profiles, this chapter has identified some of the major effects that the wild relative QTL within the introgressed regions have on fruit development in the genotypes selected.

5.2 Detailed phenolic profile

5.2.1 Profile throughout fruit development

Genotypes were re-grown under glasshouse conditions. In order to assess changes in phenolic profiles through fruit development, fruit were harvested at four developmental stages, mature green (MG), breaker, turning, and ripe, and analysed by UPLC-PDA (Figure 25).

No *p*-coumaric acid was detected in TA209 recurrent parent (Figure 25A). However, levels of *p*-coumaric acid derivative were more abundant, and detected in all genotypes (Figure 25B). No change was observed in levels of *p*-coumaric acid and derivatives at MG stage when compared with TA209. Neo-123 exhibited the highest increases of all lines, which demonstrated significant differences at breaker, turning and ripe stages (between $p \leq 0.05$ and $p \leq 0.001$). Although 3939 and neo-111 showed significant increases in *p*-coumaric acid and its derivatives (between $p \leq 0.05$ and $p \leq 0.001$), respectively, these increases were not as prominent as in neo-123.

Caffeic acid was detected at low levels in ripe stage fruit, but not any other stage (Figure 25C). All lines showed significantly lower levels than TA209 (neo-123 $p \leq 0.05$; neo-111 and 3939 $p \leq 0.001$). Chlorogenic acid, however, was detected at all stages and in all lines. There were trend similarities between abundance of chlorogenic acid in each of the four lines throughout fruit development (Figure 25D). The greatest abundance was at breaker and turning stages resulting in relatively low levels at MG and ripe. Neo-111 displayed higher levels of chlorogenic acid than each of the lines at any stage, and these were significantly higher than TA209 at MG ($p \leq 0.05$), breaker ($p \leq 0.05$) and ripe ($p \leq 0.01$). Neo-123 and 3939 had lower than TA209 levels in early fruit ripening (significantly lower levels at breaker stage, $p \leq 0.05$); however, where 3939 replicated TA209 levels in early fruit development (MG) to late ripening (turning and ripe), neo-123 maintained levels below TA209 (significantly lower at ripe stage, $p \leq 0.01$).

There was large variation between biological replicates within each genotype for chalcone-naringenin and naringenin (Figure 25E and F), which was shown by SEM values, and was most prominent for neo-111 and -123. No significant differences were observed at any stage between any line and TA209. However, compared to TA209,

both *S. neorickii* BILs showed increases in naringenin-chalcone and naringenin accumulation. Neo-123 accumulated both compounds early, at breaker stage, whereas this occurred later in fruit ripening for neo-111. 3939 exhibited lower levels than TA209 at turning and ripe stages.

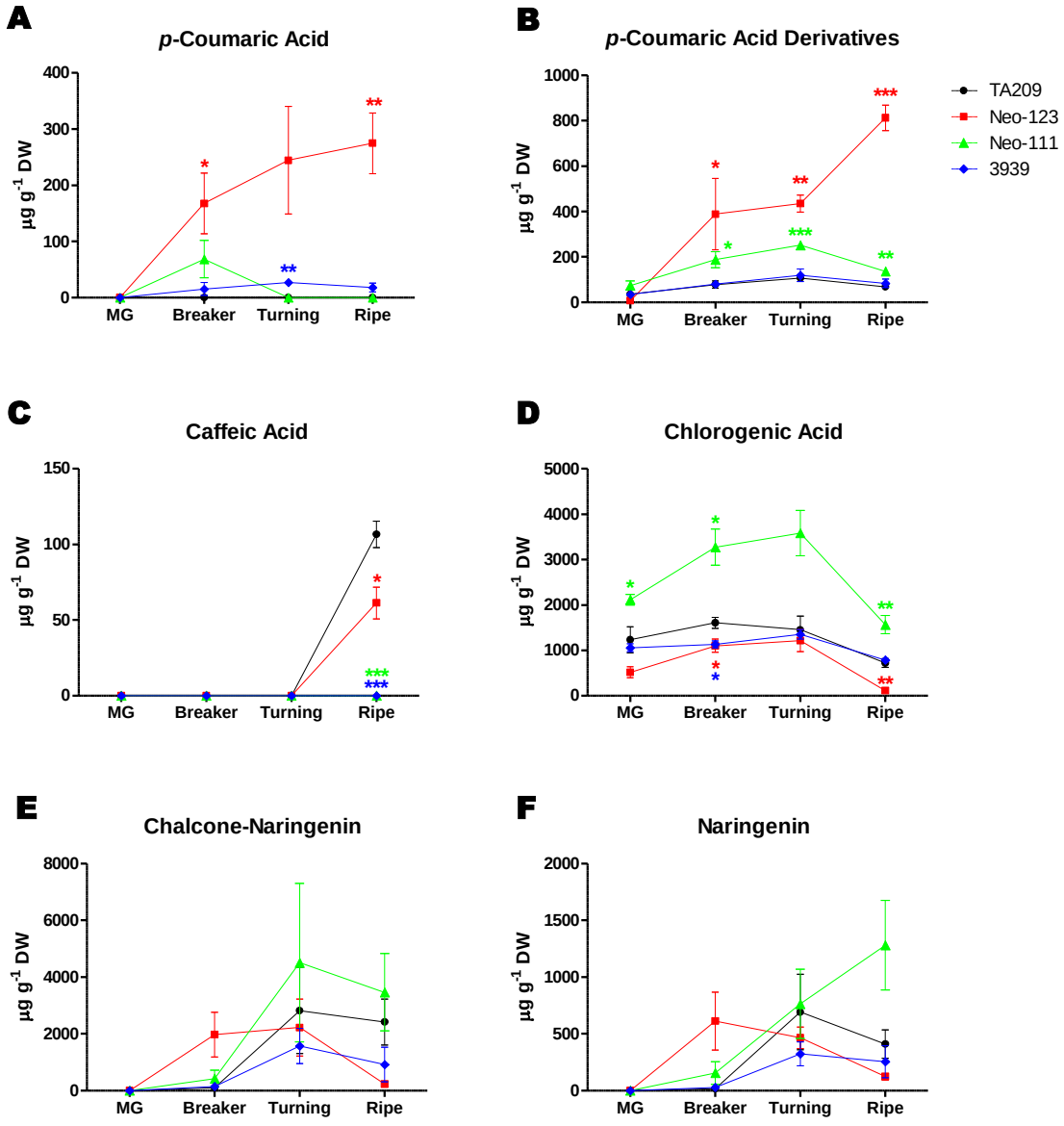

Figure 25 Profile of phenolic compounds at fruit development stages

Phenolic compounds detected by UPLC-PDA at four stages of development for fruit grown under glasshouse conditions. MG, mature green stage; UNK rutin-like 1 and 2 represent unknown compounds with UV spectral shapes similar to rutin. Error bars represent SEM for biological replicates, n=3 to 5. Values significantly different from TA209 parent are denoted by * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$). Zero values represent levels not detected by UPLC-PDA.

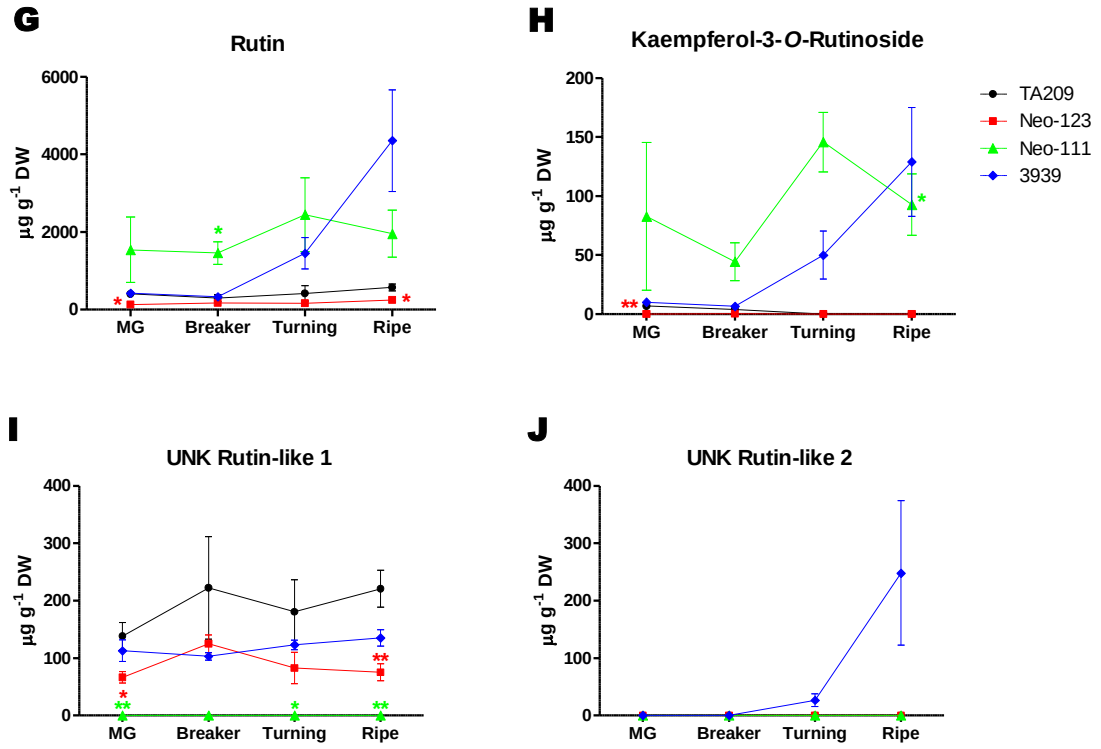

Figure 25 continued.

Detection of four compounds putatively named flavonol glycosides comprised rutin, kaempferol-3-*O*-rutinoside, and unknown compounds with UV spectral shape similar to rutin (UNK rutin-like 1 and 2) are shown by Figure 25G to J. Rutin was in greater abundance for all lines compared with each of the three other flavonol glycosides. Neo-123 exhibited levels lower than or equal to TA209 for all four compounds, and these decreases were significant (between $p \leq 0.05$ and $p \leq 0.01$) in some cases at MG (Figure 25G to I) and ripe (Figure 25G and I) stages. Neo-111 and 3939 accumulated flavonol glycosides throughout development differently. Neo-111 displayed elevated levels of rutin and kaempferol-3-*O*-rutinoside throughout fruit development, as shown by Figure 25G (significantly higher rutin at breaker stage; $p \leq 0.05$) and Figure 25H (significantly higher at ripe stage; $p \leq 0.05$). UNK rutin-like compounds were not detected in neo-111. By contrast, 3939 showed no significant difference from TA209 in flavonol glycoside accumulation in early fruit development (MG and breaker), but rapidly accumulated rutin, kaempferol-3-*O*-glycoside, and UNK rutin-like 2 later in fruit development (turning and ripe) to levels higher than TA209 (although not significant).

5.2.2 Profile in ripe fruit tissue types

Fruit at ripe stage were separated into three tissue types, as described in section 2.2.2. Analysis of these tissue types by UPLC-PDA showed that although phenylpropanoid compounds were found in all three tissue types (Figure 26A and B), flavonoid compounds were predominantly found in skin samples (Figure 26C to H). Levels of some flavonoids found in skin samples of neo-123 were equivalent to TA209 (chalcone-naringenin, naringenin and rutin, Figure 26C to E), but others in neo-123 were significantly reduced compared with TA209 (kaempferol-3-*O*-rutinoside, $p \leq 0.05$; UNK rutin-like 1, $p \leq 0.01$; Figure 26F and G). Significantly less chlorogenic acid was found in all tissue types of neo-123 ($p \leq 0.01$ to $p \leq 0.001$; Figure 26B); however, significantly more *p*-coumaric acid and derivatives were found in all tissue types ($p \leq 0.01$; Figure 26A).

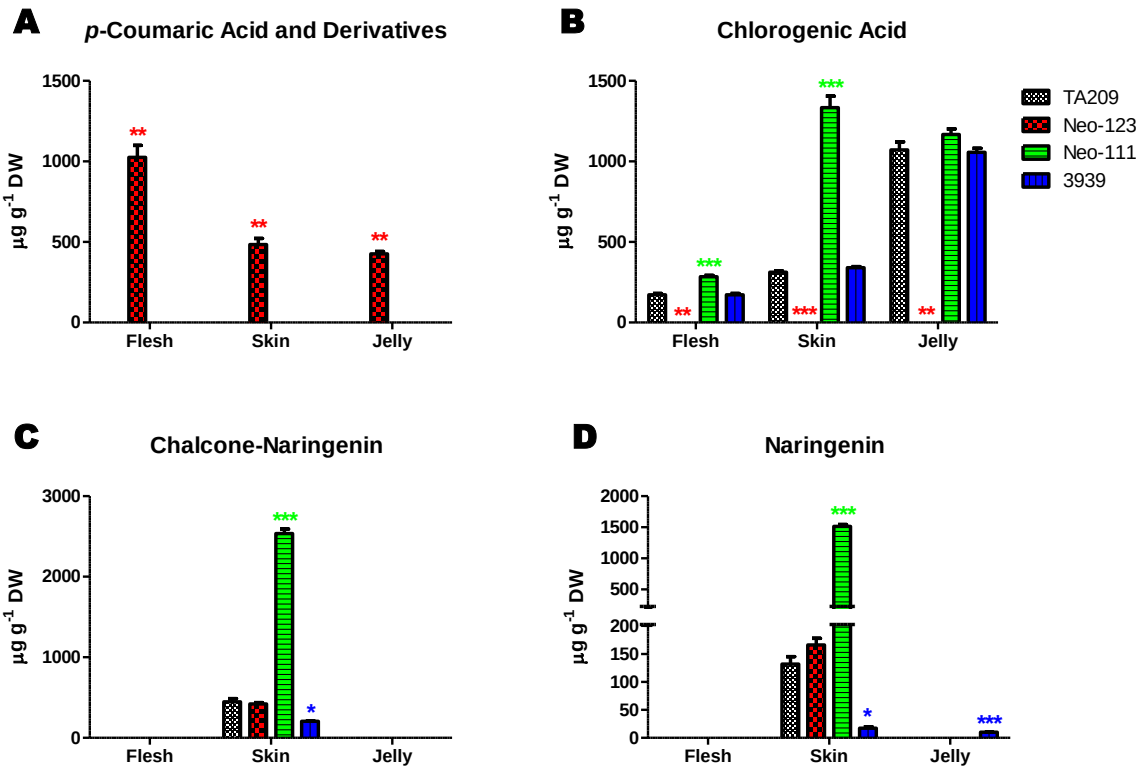

Figure 26 Profile of phenolic compounds in ripe fruit tissue types

Phenolic compounds detected by UPLC-PDA in fruit tissue types defined in section 2.2.2. UNK rutin-like 1 and 2 represent unknown compounds with UV spectral shapes similar to rutin. Error bars represent SEM for biological replicates, n=3. Values significantly different from TA209 parent are denoted by * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$). Zero values represent levels not detected by UPLC-PDA.

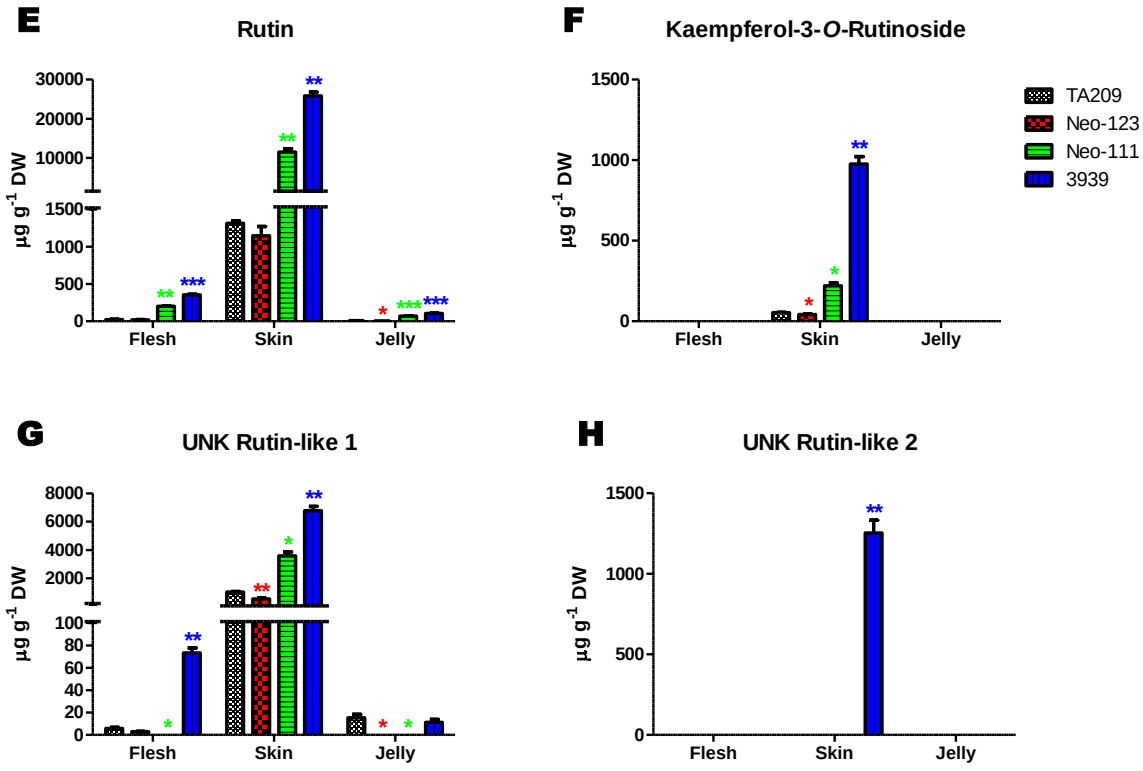

Figure 26 continued.

Levels of phenylpropanoids (*p*-coumaric acid and derivatives, and chlorogenic acid) were equivalent to TA209 in all tissue types of 3939 (Figure 26A and B). In skin samples of 3939, significantly lower levels of chalcone-naringenin and naringenin were found ($p \leq 0.05$; Figure 26C and D), but in contrast, significantly higher levels of naringenin were found in jelly of 3939 ($p \leq 0.001$; Figure 26D).

Neo-111 contained significantly more chlorogenic acid than TA209 in flesh and skin ($p \leq 0.001$), but not jelly samples (Figure 26B). Neo-111 also showed significantly more chalcone-naringenin and naringenin than TA209 in skin samples ($p \leq 0.001$), at levels higher than any of the four lines (Figure 26C and D).

Both high-rutin lines, neo-111 and 3939, showed significantly higher levels of flavonol glycoside and related compounds than TA209 ($p \leq 0.05$ to $p \leq 0.001$; Figure 26E to H), especially in skin samples. Both lines showed a significant increase in rutin levels in all tissue types (flesh $p \leq 0.01$ to $p \leq 0.001$; skin $p \leq 0.01$; jelly $p \leq 0.001$; Figure 26E).

5.2.3 Profile in leaf and flower tissues

Typical flower morphology is shown by Figure 27. Observations noted differences in florescence size for neo-123 (Figure 27B) and 3939 (Figure 27C), which were respectively smaller and larger than TA209 and neo-111 (Figure 27A and D).

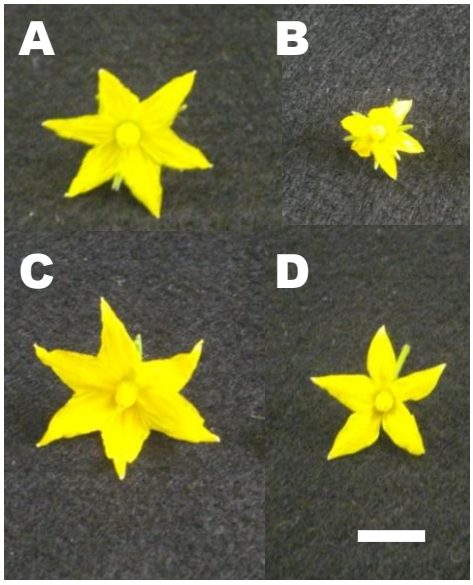

Figure 27 Typical flower morphology

Single florescence representing typical morphology from each line (A) TA209, (B) neo-123, (C) 3939, (D) neo-111. Scale bar represents 1 cm.

Despite there being no observed difference in colour, flowers were assessed for differences in phenolic profile, which largely comprised chlorogenic acid and flavonol glycosides (Figure 28). The compound of greatest abundance was unable to be identified by authenticated standard; however, it is labelled UNK naringenin-like because it is an unknown compound with a similar UV spectral shape to naringenin and was observed at a similar retention time by UPLC-PDA.

All lines showed a significant increase in chlorogenic acid level in flower tissue compared with TA209 ($p \leq 0.05$; Figure 28). Neo-123 and 3939 showed no other significant changes. Neo-111 was the only line to show significant changes in other compounds, which comprised a significant decrease in UNK naringenin-like ($p \leq 0.01$), and significant increases in rutin ($p \leq 0.01$), kaempferol-3-*O*-rutinside ($p \leq 0.05$), and UNK rutin-like 1 ($p \leq 0.05$).

Leaf material was similarly analysed by UPLC-PDA, and the two major compounds detected were chlorogenic acid and rutin (Figure 29). Although 3939 showed a notable increase in rutin compared with TA209, there were no significant differences between TA209 and any of the lines in phenolic content.

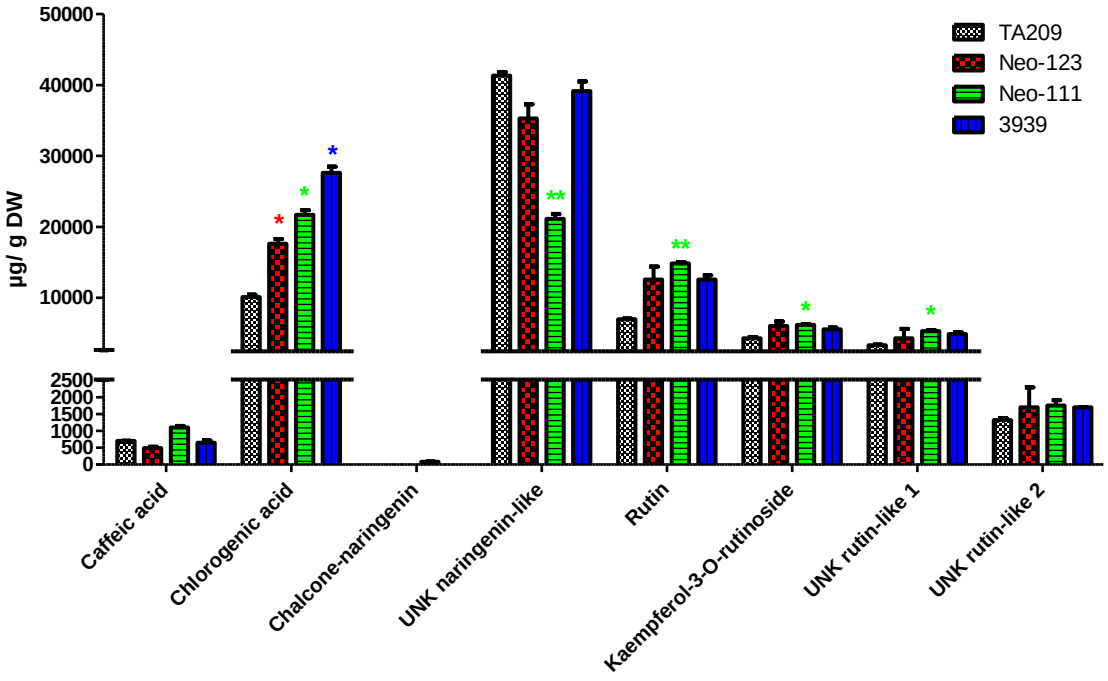

Figure 28 Phenolic content of flower tissue

Phenolic compounds detected by UPLC-PDA in flower tissue. UNK rutin-like 1 and 2, and UNK naringenin-like represent unknown compounds with UV spectral shapes similar to rutin and naringenin respectively. Error bars represent SEM, n=3 technical replicates where at least 20 flower were pooled per genotype. Values significantly different from TA209 are represented by * ($p \leq 0.05$), ** ($p \leq 0.01$), and *** ($p \leq 0.001$).

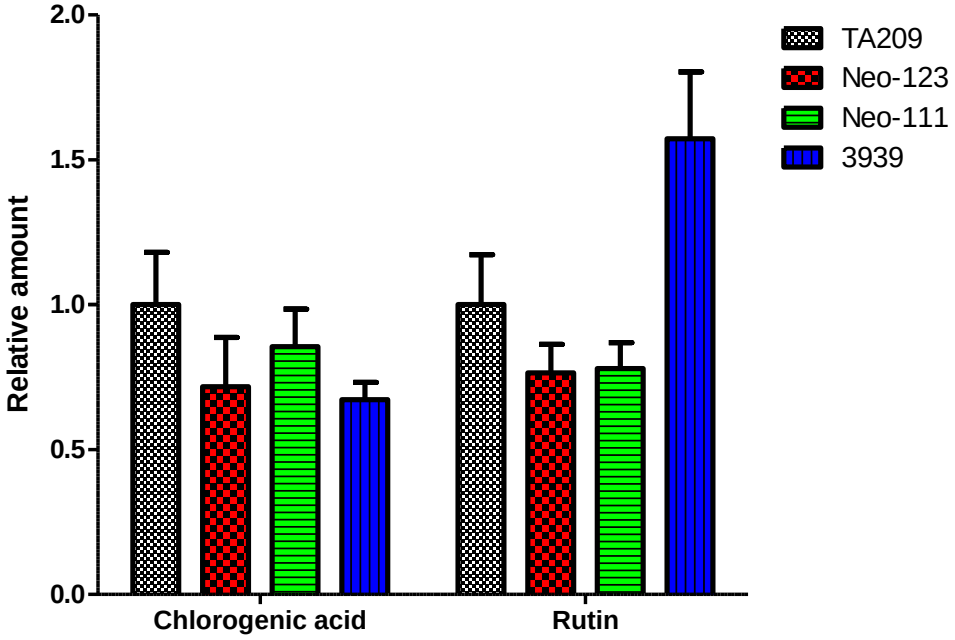

Figure 29 Relative levels of phenolic compounds in leaf material

Levels of two predominant compounds present in leaf material relative to TA209. No significant differences were observed between any line and TA209. Error bars represent SEM, n=5.

5.3 Metabolomic profile

5.3.1 Profile of isoprenoid and related compounds

A non-polar extraction method simultaneously extracted isoprenoids (namely carotenoids), chlorophylls, ubiquinone, and tocopherols. Identical biological material was used to that used previously for the developmental stages described in section 5.2.1. Extracts were analysed by UPLC-PDA. Isoprenoid levels throughout development are shown in Figure 30. Compound intermediates from early in the carotenogenesis pathway (phytoene to lycopene, Figure 30A to F) were not detected in any lines at MG stage and increased in abundance throughout fruit development and ripening. No significant difference was observed between 3939 and TA209 for any stage. Neo-111 showed the greatest difference from TA209. At ripe stage, significantly less phytoene ($p \leq 0.01$, Figure 30A), phytofluene isomers 1 ($p \leq 0.05$, Figure 30B) and 2 ($p \leq 0.001$, Figure 30C), and ζ-carotene isomer 2 ($p \leq 0.01$, Figure 30E) were seen in neo-111 compared with TA209. ζ-carotene isomer 1 (Figure 30D) was not detected in neo-111. Neo-123 displayed little change from TA209, and only one developmental stage was significantly ($p \leq 0.05$) different from TA209 (phytofluene isomer 1 at turning stage, Figure 30B).

Abundance of cyclic carotenes and xanthophylls throughout development are shown in Figure 30G to J. With the exception of δ-carotene, which was not detected in lines 3939 and TA209 (Figure 30G), the trends in abundance throughout development of all compounds were similar for all lines. In contrast to early carotenogenesis where neo-111 demonstrated the greatest changes compared with TA209, neo-123 showed the greatest increases for all four cyclic carotene and xanthophyll compounds. Neo-123 exhibited significant increases in δ-carotene levels at breaker ($p \leq 0.05$), turning ($p \leq 0.05$), and ripe ($p \leq 0.01$) stages (Figure 30G); β-carotene levels at breaker ($p \leq 0.01$), turning ($p \leq 0.001$), and ripe ($p \leq 0.001$) stages (Figure 30H); lutein levels at breaker, turning, and ripe ($p \leq 0.001$ in each case) stages (Figure 30I); and neoxanthin and violaxanthin at breaker ($p \leq 0.01$) and turning ($p \leq 0.05$) stages (Figure 30J). Neo-111 showed some increase compared with TA209. For each compound, neo-111 levels for at least one developmental stage were significantly increased (Figure 30G to J); however, increases did not match those of neo-123. 3939 again showed little change from TA209 levels except for a significant ($p \leq 0.05$) but relatively small decrease in β-carotene at turning stage (Figure 30H).

No γ-tocopherol was detected in TA209 or neo-111; however, in neo-123 and 3939, γ-tocopherol was at detectable levels at turning and ripe stages (Figure 31A). Although levels of α-tocopherol increased throughout fruit ripening and development, no significant differences were observed until ripe stage, where all lines showed significantly higher abundances than TA209 (neo-111 and 3939 $p \leq 0.05$; neo-123 $p \leq 0.001$; Figure 31B).

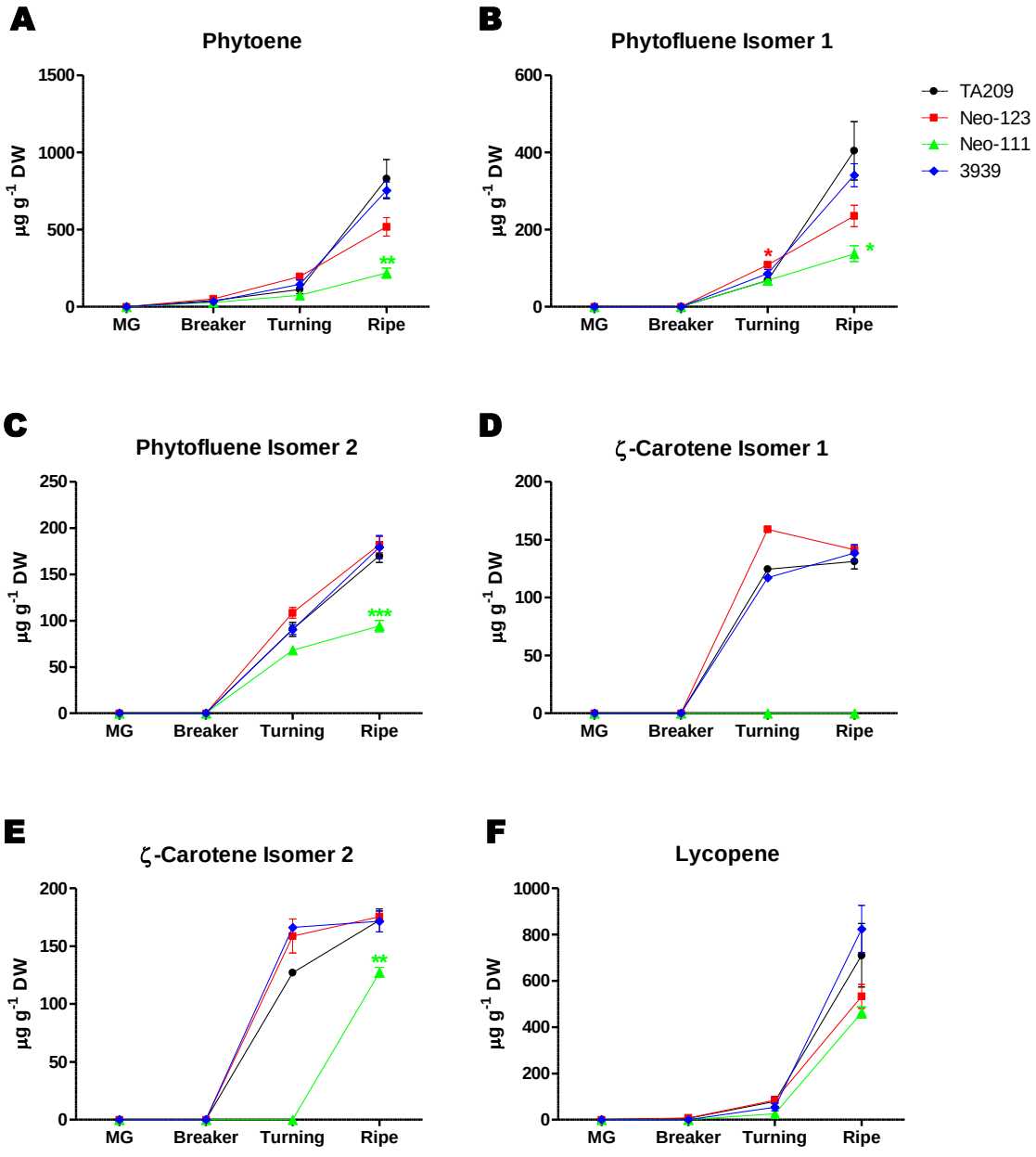

Figure 30 Profile of isoprenoid compounds throughout fruit development

Isoprenoid compounds detected by UPLC-PDA at four stages of development for fruit grown under glasshouse conditions. MG, mature green stage. Error bars represent SEM for biological replicates, n=3 to 5. Values significantly different from TA209 parent are denoted by * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$). Zero values represent levels not detected by UPLC-PDA.

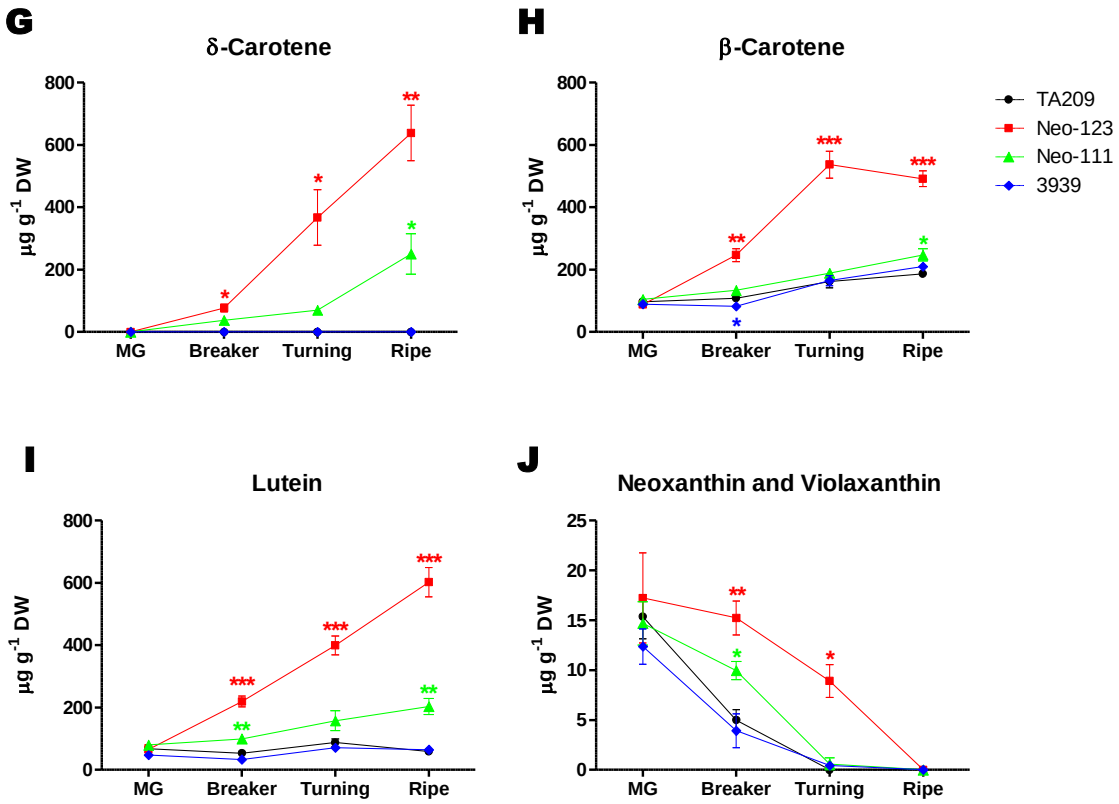

Figure 30 continued.

3939 levels of chlorophylls replicated those of TA209 throughout development and ripening (Figure 31C to F). The decreasing trend in chlorophylls levels throughout fruit development and ripening for neo-111 and -123 also followed that of TA209; however, a delay in reducing accumulation resulted in significantly more chlorophyll *a* (Figure 31C) for neo-111 at breaker stage ($p \leq 0.05$), and for chlorophyll *b* (Figure 31D) significantly higher levels at breaker stage for neo-111 and -123 ($p \leq 0.001$) and at turning stage for neo-123 ($p \leq 0.01$). UNK chlorophyll degradation product (Figure 31E) represents an unknown compound putatively named chlorophyll degradation product (personal communication P. Fraser) based on UV spectral shape. Significant increases were seen at ripe stage ($p \leq 0.01$) for neo-111 and at breaker ($p \leq 0.01$), turning ($p \leq 0.05$) and ripe ($p \leq 0.001$) stages for neo-123.

No significant changes in ubiquinone levels were observed between TA209 and neo-111 or 3939 (Figure 31F). Neo-123 showed significant differences throughout development and ripening, with a decrease at MG stage ($p \leq 0.05$), and increases at breaker ($p \leq 0.001$), turning ($p \leq 0.01$) and ripe ($p \leq 0.001$) stages.

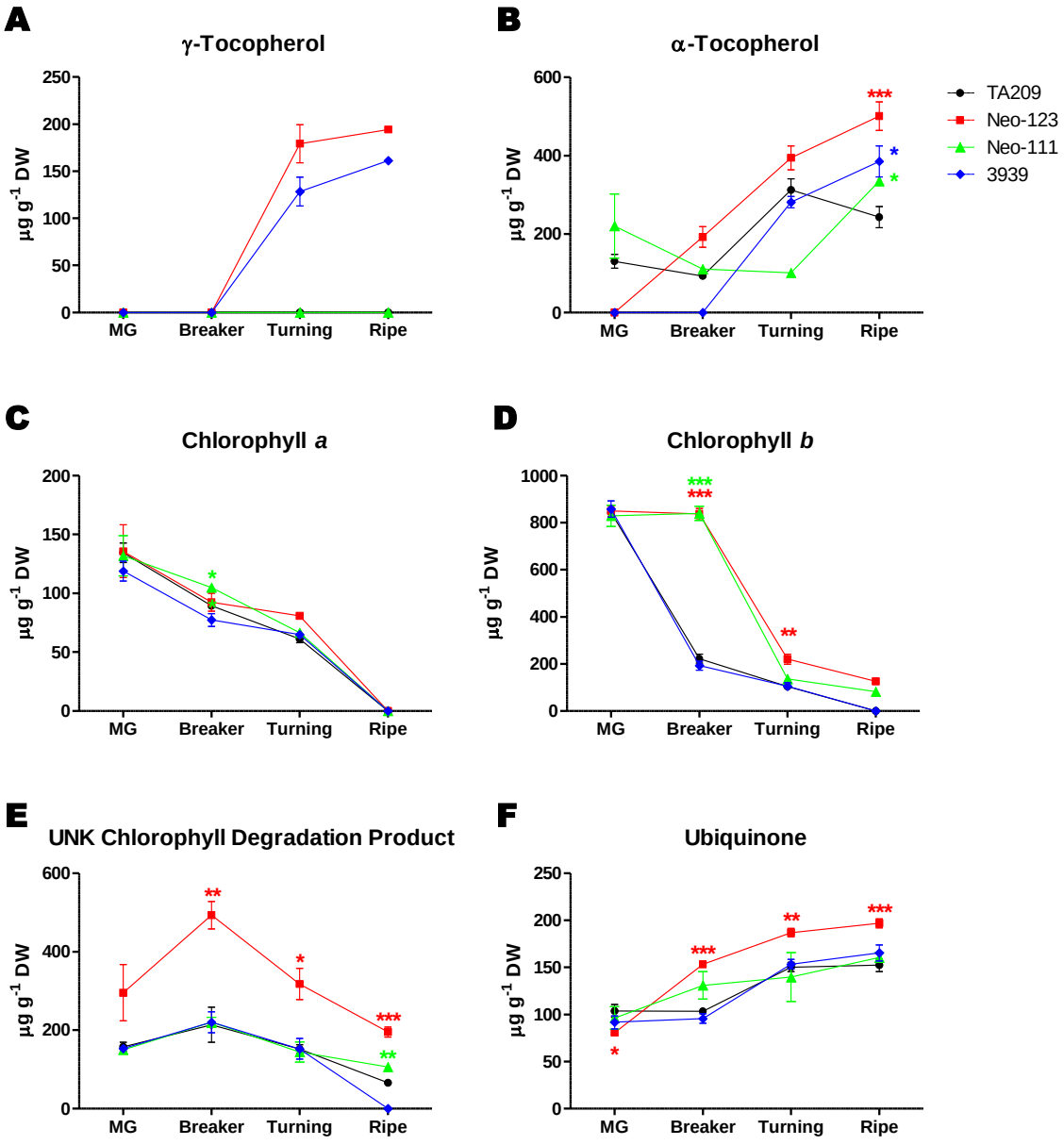

Figure 31 Profile of chlorophyll and other isoprenoid related compounds throughout fruit development

Compounds detected in non-polar extracts by UPLC-PDA at four stages of development for fruit grown under glasshouse conditions. MG, mature green stage; UNK chlorophyll degradation product represents unknown compound putatively identified based on UV spectral shape. Error bars represent SEM for biological replicates, n=3 to 5. Values significantly different from TA209 parent are denoted by * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$). Zero values represent levels not detected by UPLC-PDA.

5.3.2 Metabolomic analysis

Polar and non-polar extracts from each of the four development stages (MG, Br, T and R) for each of the four lines (TA209, neo-111, neo-123 and 3939) were analysed by GC-MS to represent a non-targeted metabolomic analysis. Quality control (QC) samples were included. Principal component analysis (PCA) separated the samples on the first and second principal component (PC1 and 2), and data are shown in Figure 32A. PC1 and 2 classified the samples into two major clusters that are predominantly separated by variation within PC1. The first cluster (encircled by dotted line) included all MG and Br samples, and aside from one outlier (one biological replicate of TA209 at Br stage), the cluster was tight. The second cluster (encircled by dashed line) included all T and R stage samples. This second cluster was sparser due to greater separation on both PC1 and PC2.

The loading plot (Figure 32B) indicated which variables were most influencing separation of these two major clusters. Some tricarboxylic acid cycle intermediates, sugars, organic acids, and fatty acids notably separated the clusters along PC1. 2-ketoglutaric acid, mannitol, oxalic acid, and behenic acid (docosanoic acid) were located on the loading plot in positive sector of PC1. These compounds were all found in higher abundance in R and T samples than in Br and MG samples. Conversely, aconitic acid, malic acid, butanoic acid, and malonic acid were all found in higher abundance in Br and MG samples than R and T, and were located on the loading plot in the negative sector of PC1 (the direction of the Br/MG cluster). Some unknown sugar variables (such as s13, s14 and disacch2) were also located in the negative sector of PC1. In these cases, compounds were only present in MG and Br samples, and not detected in R or T samples. Once again variables were shown to have the reverse effect; glucaric acid and aspartic acid were not detected in MG or Br samples, therefore contributed to the direction of the R/T cluster.

Each of the two major clusters was analysed in isolation by PCA to determine whether lines could be characterised based on GC-MS profiles alone. PCA results of MG and Br samples are shown in Figure 32C and the corresponding loading plot is shown in Figure 32D. PCA results of R and T samples are shown in Figure 32E and the corresponding loading plot is shown in Figure 32F.

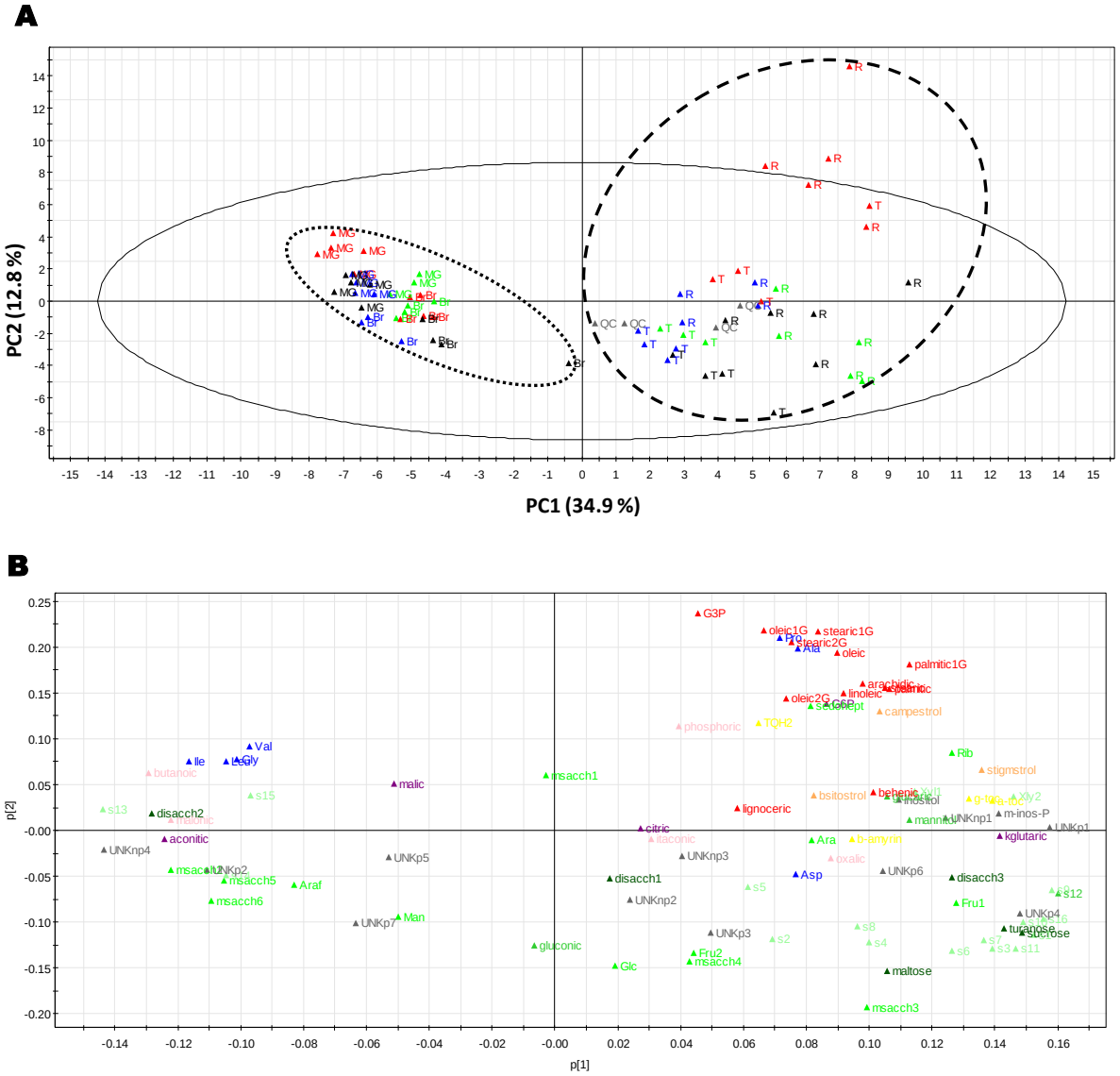

Figure 32 Discrimination of development stages for four lines, based on non-targeted metabolomic profile

(A) PCA for metabolites detected by GC-MS from polar and non-polar extractions of selected lines TA209 (black), neo-123 (red), neo-111 (green) and 3939 (blue) at the four development stages mature green (MG), breaker (Br), turning (T) and ripe (R). Quality control (QC) samples are shown in grey. Circle of dashes shows clustering of samples at T and R stages. Circle of dots shows clustering of samples at MG and Br stages. (B) Loading scatter plot for PCA variables. Sugars are shown in green: monosaccharides (bright), disaccharides (dark), sugar acid and alcohols (lime) and unknown and other sugars (pale). Other categories represented are amino acids (blue), organic and other acids (pink), sterols (orange), fatty acids and intermediates from triacylglyceride biosynthesis (red), intermediates from glycolysis and tricarboxylic acid cycle (violet), tocopherols and triterpenes (yellow), and other or unknown compounds (grey).

Legend continued overleaf.

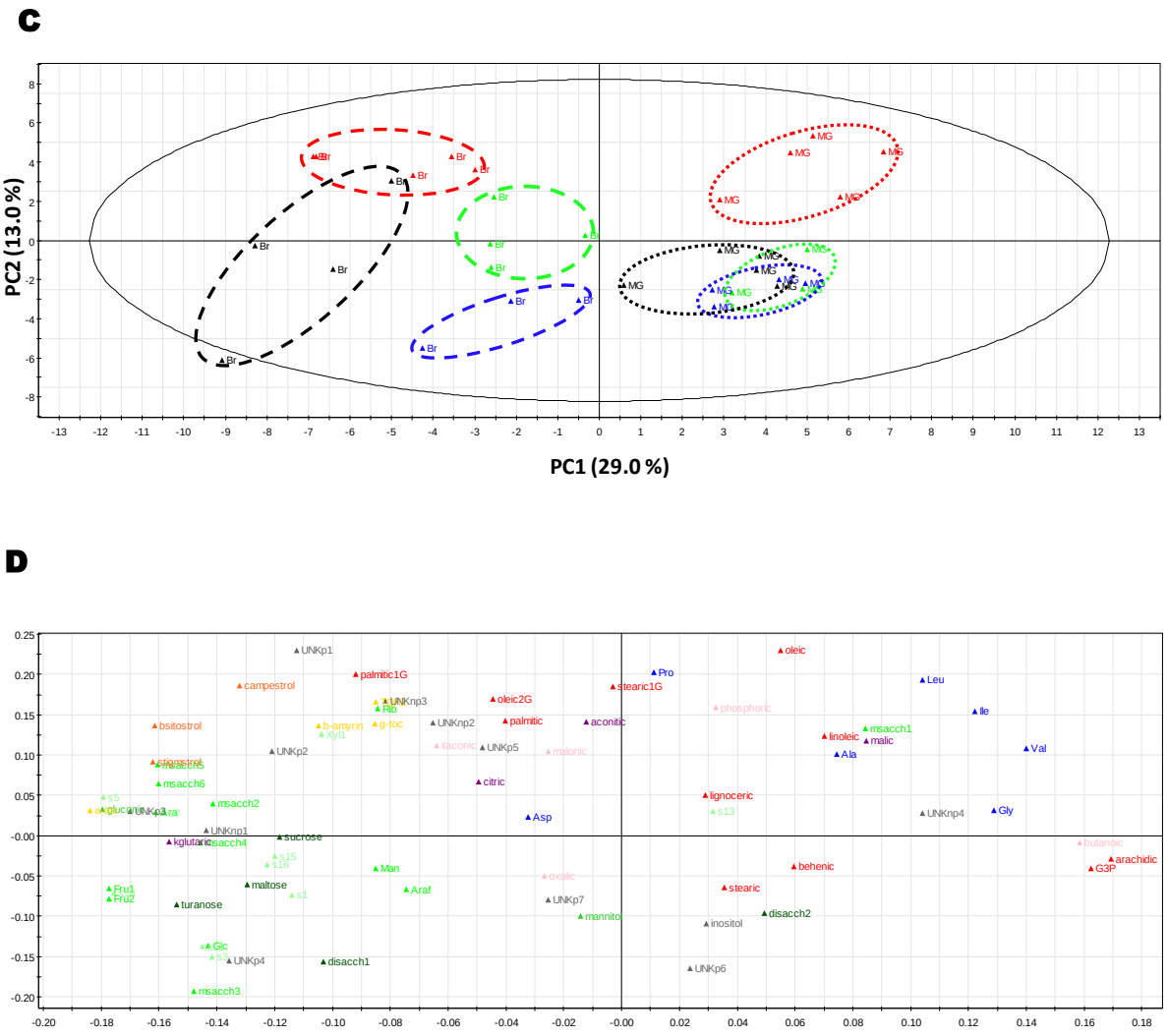

Figure 32 continued (legend continued from previous page)

Abbreviations are provided for unknown monosaccharides labelled 1 to 6 (msacch1 to 6), unknown disaccharides labelled 1 to 3 (disacch1 to 3), unknown compounds from polar extraction labelled 1 to 7 (UNKp1 to 7), unknown compounds from non-polar extraction labelled 1 to 4 (UNKnp1 to 4), unknown sugar labelled 1 to 16 (s1 to 16), 2-ketoglutaric acid (kglutaric), sedoheptulose (sedohept), glucose-6-phosphate (G6P), *myo*-inositol-phosphate (m-inos-P), glycerol-3-phosphate (G3P), γ-tocopherol (g-toc), α-tocopherol (a-toc), campesterol (campestrol), stigmasterol (stigmstrol), α-tocopherol hydroxyquinone (TQH2), β-sitosterol (bsitostrol), and β-amyrin (b-amyrin). The word 'acid' is omitted from labels (for example citric in place of citric acid). A fatty acid with the suffix 1G or 2G represents propyl or methyl ethyl fatty acid esters, respectively. The following widely accepted abbreviations for amino acids and sugars are used: alanine (Ala), aspartic acid (Asp), proline (Pro), valine (Val), leucine (Leu), isoleucine (Ile), glycine (Gly), mannose (Man), arabinose/arabinofuranose (Ara/Araf), xylose peaks 1 and 2 (Xyl1 and 2), ribose (Rib), fructose peaks 1 and 2 (Fru1 and 2), and glucose (Glc). (C) PCA, as described above, including only MG and Br samples with (D) accompanying loading plot, as described above. (E) PCA, as described above, including only T and R samples with (F) accompanying loading plot, as described above.

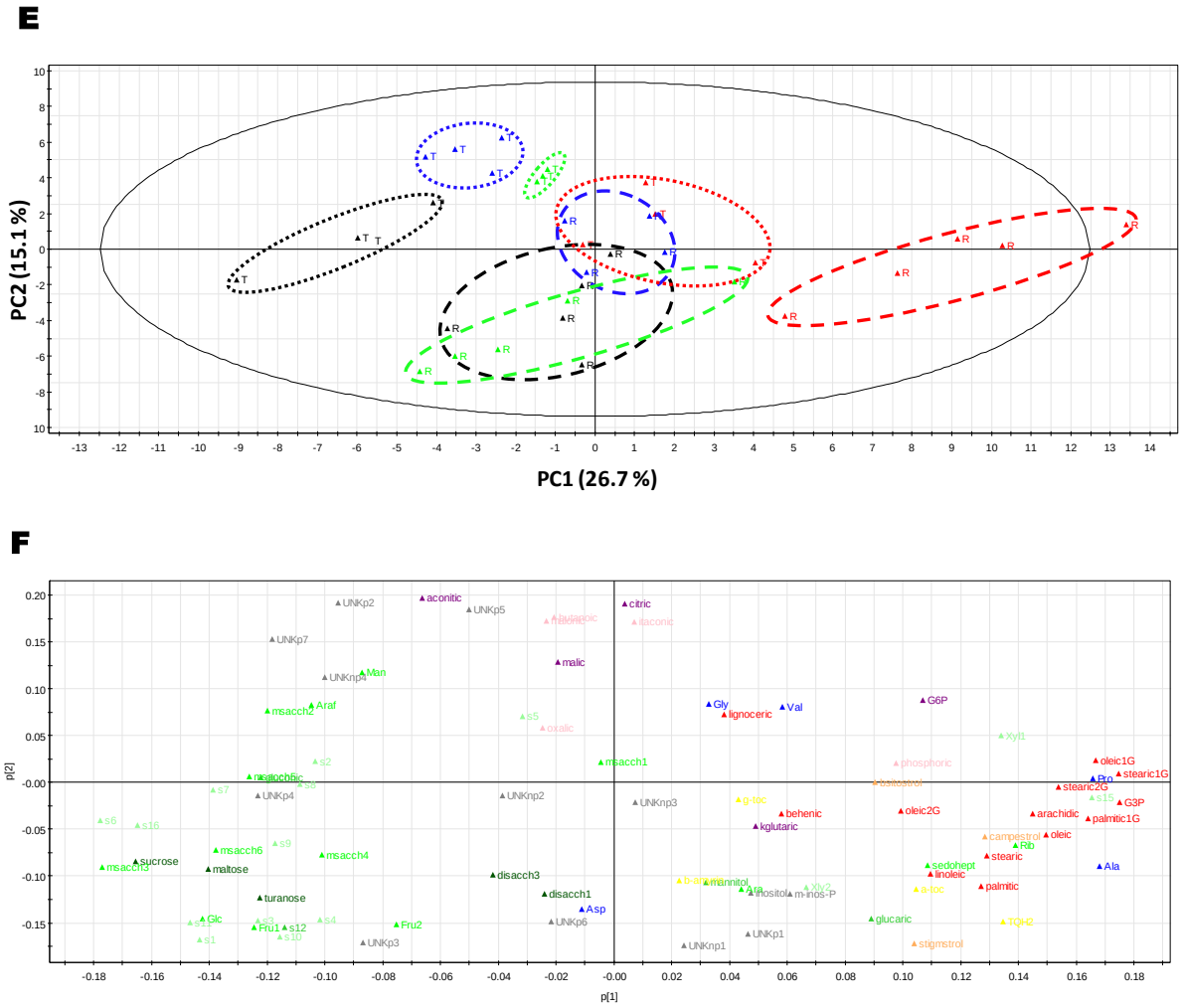

Figure 32 continued.

When MG and Br samples are analysed in isolation from T and R sample, a clearer intra-cluster relationship can be determined. MG clusters form in the positive sector of PC1 away from Br clusters in the negative sector of PC1 (Figure 32C). According to the loading plot, Br samples are influenced by many of the sugar variables; whereas, MG samples are influenced in the direction of many amino acids and fatty acids (Figure 32D). MG samples clearly show no separation of high rutin lines 3939 and neo-111 from TA209 according to GC-MS variables (Figure 32C) and all group in the negative cluster of PC2. Neo-123 samples, however, form a separate cluster in the positive sector of PC2. This is predominantly influenced by amino acid abundance in MG samples (Table 9), which have a significantly ($p \leq 0.05$) higher fold-change in neo-123 compared with TA209. These increases include leucine (2.3 fold), isoleucine (2.5 fold), valine (1.8 fold, but not significant), and proline (2.4 fold). The sugar, msacch1, and fatty acid, oleic acid (9-octadecenoic acid), are both present in neo-123 but not detected in TA209.

In contrast to MG samples, Br samples of neo-123 and TA209 did not form distinct clusters (Figure 32C), although this may in part be due to variation within TA209 and the outlying TA209 Br sample observed previously in Figure 32A. High rutin lines, 3939 and neo-111, both formed clusters distinct from one another and from TA209 (Figure 32C). The separation between TA209 and both high rutin lines was driven partly by sterol abundance (Figure 32D, Table 9) where significantly ($p \leq 0.05$) reduced fold-change was seen in high rutin lines compared with TA209. Campesterol was reduced to non-detectable levels in both high rutin lines. β-sitosterol was also not detected in 3939, and reduced in neo-111 (0.4 fold, not significant). Stigmasterol was differentially reduced in both lines: 0.3 fold in neo-111 ($p \leq 0.05$) and 0.6 fold in 3939 (no significance). Significant ($p \leq 0.05$) fold decreases in sugar compounds, such as maltose, s1, s5, s7 and msacch4, also contributed to the separation of high rutin lines from TA209 (Figure 32D, Table 9). Conversely, s13 was detected in high rutin lines, but not TA209.

There was no distinction between clusters of R and T (Figure 32E) as was seen previously between sample of MG and Br in Figure 32C. T samples of each genotype did however cluster apart from one another, as seen by the clusters encircled by dotted lines in Figure 32E. The direction of 3939, neo-111, -123 clusters away from TA209 appeared to be influenced by their relative (and significant, $p \leq 0.05$) decreases in sugars levels compared with TA209 (Table 9, and located in the negative sectors of both PC1 and 2, Figure 32F). Example variables include fructose isomers (Fru1 and 2), glucose (Glc), sucrose, and unknown sugars (msacch 2, 5, 6, s16), among others. The opposite (relative and significant increases in 3939, neo-111 and -123 compared with TA209) was observed for the amino acid glycine and for organic and fatty acids such as butanoic acid, itaconic acid, malonic acid, oxalic acid, and stearic1G (octadecanoic acid propyl ester; which most notably increased significantly by 11.3 fold in neo-111 and non-significantly by 24.9 fold in neo-123) (Figure 32F, Table 9). High rutin genotypes R samples did not cluster apart from TA209 (Figure 32E); neo-123 samples were the only R samples to differentially cluster away from TA209. Most noticeably from the loading plot (Figure 32F) this is in part contributed by relative levels of fatty acids and triacylglyceride biosynthesis intermediates. For example, stearic1G (octadecanoic acid propyl ester) increased significantly ($p \leq 0.05$) by 10.1 fold compared with 2.0 and 2.1 fold (no significance) for neo-111 and -3939, respectively (Table 9).

Table 9 Average levels of compounds detected by GC-MS, relative to TA209

Average (Av.) amounts ± SEM (n=3 to 5) detected by GC-MS from polar and non-polar extractions. MG, mature green stage. N.D. indicates compound was detected in neither genotype of interest nor TA209. 100* indicates compound was detected in genotype of interest but not detected in TA209. 0.0 ± 0.00 indicates compound was not detected in genotype of interest, but was detected in TA209. Compound abbreviations provided in legend of Figure 32. Values significantly ($p \leq 0.05$) different from TA209 are shown in bold, and highlighted red or green to indicate decrease or increase.

	RIPE			TURNING			BREAKER			MG		
	Neo-123	Neo-111	3939	Neo-123	Neo-111	3939	Neo-123	Neo-111	3939	Neo-123	Neo-111	3939
	Av. ± SEM	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**
Monosaccharides												
Araf	0.9 ± 0.05	0.9 ± 0.01	1.0 ± 0.04	**0.8 ± 0.03**	**0.6 ± 0.02**	1.0 ± 0.09	1.0 ± 0.13	0.8 ± 0.08	0.9 ± 0.03	**0.4 ± 0.08**	**0.5 ± 0.02**	0.7 ± 0.07
Ara	**0.4 ± 0.17**	**0.4 ± 0.05**	0.7 ± 0.28	100*	100*	N.D.	**0.4 ± 0.09**	**0.0 ± 0.00**	**0.0 ± 0.00**	N.D.	N.D.	N.D.
Fru1	0.9 ± 0.05	1.0 ± 0.07	1.0 ± 0.02	0.9 ± 0.03	**0.8 ± 0.01**	**0.9 ± 0.03**	1.0 ± 0.05	0.9 ± 0.04	1.0 ± 0.05	0.9 ± 0.04	0.9 ± 0.06	1.0 ± 0.01
Fru2	**0.8 ± 0.05**	1.3 ± 0.32	0.9 ± 0.04	0.9 ± 0.03	**0.8 ± 0.02**	**0.8 ± 0.04**	1.0 ± 0.07	0.9 ± 0.06	1.0 ± 0.11	0.8 ± 0.08	0.8 ± 0.10	0.9 ± 0.03
Glc	**0.7 ± 0.07**	0.9 ± 0.11	0.9 ± 0.06	**0.7 ± 0.05**	**0.7 ± 0.01**	0.8 ± 0.06	0.8 ± 0.08	0.9 ± 0.08	1.0 ± 0.07	**0.7 ± 0.08**	0.7 ± 0.12	1.0 ± 0.02
Man	0.6 ± 0.09	0.8 ± 0.21	0.7 ± 0.32	1.2 ± 0.17	0.8 ± 0.14	0.9 ± 0.26	0.9 ± 0.21	1.1 ± 0.20	1.3 ± 0.35	1.0 ± 0.26	0.8 ± 0.37	0.8 ± 0.20
Rib	**1.4 ± 0.00**	1.1 ± 0.14	1.3 ± 0.09	**2.0 ± 0.24**	**2.1 ± 0.15**	2.0 ± 0.46	1.5 ± 0.26	1.5 ± 0.50	0.8 ± 0.19	1.6 ± 0.31	0.9 ± 0.25	1.2 ± 0.12
sedohept	1.1 ± 0.38	0.1 ± 0.00	0.1 ± 0.07	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
msacch1	**1.7 ± 0.19**	**0.8 ± 0.00**	0.9 ± 0.10	0.3 ± 0.02	0.3 ± 0.00	0.3 ± 0.01	N.D.	100*	N.D.	100*	N.D.	N.D.
msacch2	**0.8 ± 0.07**	0.8 ± 0.08	1.1 ± 0.06	**0.8 ± 0.03**	**0.6 ± 0.01**	0.9 ± 0.04	1.0 ± 0.04	0.8 ± 0.06	1.0 ± 0.09	0.8 ± 0.06	**0.7 ± 0.07**	0.8 ± 0.04
msacch3	**0.6 ± 0.08**	1.1 ± 0.14	0.9 ± 0.06	0.7 ± 0.07	0.7 ± 0.04	0.7 ± 0.03	0.8 ± 0.10	0.8 ± 0.05	1.1 ± 0.09	**0.6 ± 0.11**	1.1 ± 0.17	1.0 ± 0.10
msacch4	**0.0 ± 0.00**	1.0 ± 0.19	**0.0 ± 0.00**	0.7 ± 0.15	1.4 ± 0.00	0.0 ± 0.00	0.7 ± 0.02	**0.0 ± 0.00**	0.6 ± 0.19	N.D.	100*	N.D.
msacch5	**0.5 ± 0.09**	0.7 ± 0.15	**0.6 ± 0.05**	0.7 ± 0.15	**0.6 ± 0.04**	0.7 ± 0.11	1.1 ± 0.08	0.9 ± 0.09	1.0 ± 0.05	0.8 ± 0.15	0.7 ± 0.14	0.7 ± 0.09
msacch6	**0.0 ± 0.00**	1.0 ± 0.09	**0.0 ± 0.00**	**0.0 ± 0.00**	**0.4 ± 0.00**	0.9 ± 0.18	0.9 ± 0.13	0.8 ± 0.05	0.9 ± 0.05	1.0 ± 0.15	0.8 ± 0.05	1.0 ± 0.24
Disaccharides												
maltose	0.5 ± 0.14	1.3 ± 0.39	0.6 ± 0.07	0.5 ± 0.06	0.5 ± 0.02	0.7 ± 0.07	**0.0 ± 0.00**	0.4 ± 0.04	**0.0 ± 0.00**	N.D.	N.D.	N.D.
sucrose	**0.5 ± 0.08**	0.8 ± 0.12	0.8 ± 0.05	**0.6 ± 0.05**	0.6 ± 0.03	0.6 ± 0.05	1.2 ± 0.12	0.0 ± 0.00	0.0 ± 0.00	N.D.	N.D.	N.D.
turanose	0.6 ± 0.11	1.0 ± 0.18	0.9 ± 0.11	0.8 ± 0.13	0.7 ± 0.03	0.7 ± 0.04	0.7 ± 0.16	0.6 ± 0.15	0.7 ± 0.22	0.8 ± 0.05	0.9 ± 0.00	0.8 ± 0.09
disacch1	**0.0 ± 0.00**	**0.0 ± 0.00**	**0.0 ± 0.00**	N.D.	N.D.	N.D.	**0.0 ± 0.00**	0.9 ± 0.10	1.4 ± 0.54	N.D.	N.D.	100*
disacch2	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	100*	100*	100*	0.8 ± 0.17	0.9 ± 0.17	1.2 ± 0.12
disacch3	0.6 ± 0.10	1.5 ± 0.33	0.6 ± 0.11	1.1 ± 0.07	0.6 ± 0.13	0.8 ± 0.47	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.

Table 9 continued

	RIPE			TURNING			BREAKER			MG		
	Neo-123	Neo-111	3939	Neo-123	Neo-111	3939	Neo-123	Neo-111	3939	Neo-123	Neo-111	3939
	Av. ± SEM	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**
Sugar acids and alcohols												
glucaric	0.7 ± 0.05	0.6 ± 0.13	**0.4 ± 0.12**	100*	100*	100*	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
gluconic	0.8 ± 0.09	0.9 ± 0.11	1.0 ± 0.04	1.0 ± 0.06	**0.8 ± 0.03**	0.9 ± 0.02	1.0 ± 0.06	0.8 ± 0.07	0.9 ± 0.07	1.0 ± 0.12	0.8 ± 0.12	0.9 ± 0.03
mannitol	1.1 ± 0.12	1.2 ± 0.27	1.3 ± 0.33	1.0 ± 0.28	1.1 ± 0.07	1.2 ± 0.12	0.9 ± 0.12	0.9 ± 0.23	1.1 ± 0.09	1.2 ± 0.32	1.1 ± 0.36	1.2 ± 0.17
s12	0.6 ± 0.14	1.1 ± 0.11	0.8 ± 0.03	0.9 ± 0.05	**0.6 ± 0.02**	**0.6 ± 0.04**	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Unknown and other sugars												
s1	**0.5 ± 0.12**	1.1 ± 0.05	0.8 ± 0.03	**0.5 ± 0.07**	**0.5 ± 0.04**	0.6 ± 0.06	**0.0 ± 0.00**	**0.0 ± 0.00**	**0.0 ± 0.00**	N.D.	N.D.	N.D.
s2	**0.0 ± 0.00**	1.1 ± 0.00	**0.0 ± 0.00**	**0.0 ± 0.00**	**0.0 ± 0.00**	2.0 ± 0.62	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
s3	**0.4 ± 0.08**	**1.3 ± 0.04**	1.1 ± 0.25	0.7 ± 0.07	0.7 ± 0.06	0.7 ± 0.08	0.4 ± 0.03	0.4 ± 0.07	0.7 ± 0.51	**0.0 ± 0.00**	**0.0 ± 0.00**	0.8 ± 0.16
s4	**0.0 ± 0.00**	**1.1 ± 0.00**	**0.0 ± 0.00**	**0.0 ± 0.00**	**0.0 ± 0.00**	1.0 ± 0.01	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
s5	**0.7 ± 0.05**	0.8 ± 0.13	1.0 ± 0.05	0.8 ± 0.11	4.8 ± 2.34	0.8 ± 0.06	0.7 ± 0.05	0.6 ± 0.01	**0.0 ± 0.00**	N.D.	N.D.	N.D.
s6	**0.0 ± 0.00**	1.2 ± 0.07	**0.7 ± 0.07**	0.7 ± 0.01	**0.6 ± 0.00**	**0.7 ± 0.01**	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
s7	**0.5 ± 0.10**	0.9 ± 0.18	**0.7 ± 0.03**	0.7 ± 0.12	0.9 ± 0.15	1.0 ± 0.33	0.6 ± 0.06	**0.0 ± 0.00**	**0.0 ± 0.00**	N.D.	N.D.	N.D.
s8	N.D.	100*	100*	0.7 ± 0.10	**0.0 ± 0.00**	0.7 ± 0.10	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
s9	0.8 ± 0.12	1.0 ± 0.16	1.1 ± 0.16	0.9 ± 0.13	1.0 ± 0.06	0.8 ± 0.08	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
s10	**0.4 ± 0.07**	**1.4 ± 0.04**	**0.7 ± 0.08**	0.7 ± 0.05	0.7 ± 0.05	0.6 ± 0.10	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
s11	**0.4 ± 0.08**	1.0 ± 0.23	0.8 ± 0.05	0.6 ± 0.11	0.7 ± 0.04	0.7 ± 0.07	0.4 ± 0.05	0.4 ± 0.03	0.5 ± 0.19	**0.0 ± 0.00**	**0.0 ± 0.00**	0.9 ± 0.07
s13	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	100*	100*	100*	0.8 ± 0.16	0.9 ± 0.22	1.0 ± 0.05
s14	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	0.7 ± 0.10	0.7 ± 0.16	0.7 ± 0.07	**0.0 ± 0.00**	**0.0 ± 0.00**	0.7 ± 0.16
s15	100*	N.D.	N.D.	N.D.	N.D.	N.D.	1.2 ± 0.27	1.0 ± 0.28	1.0 ± 0.10	0.8 ± 0.14	1.1 ± 0.12	1.1 ± 0.12
s16	**0.6 ± 0.05**	0.9 ± 0.14	**0.7 ± 0.03**	**0.7 ± 0.05**	**0.7 ± 0.04**	**0.7 ± 0.03**	0.4 ± 0.06	0.2 ± 0.01	**0.0 ± 0.00**	N.D.	N.D.	N.D.
Xyl1	1.2 ± 0.12	1.0 ± 0.17	1.0 ± 0.05	**1.5 ± 0.07**	**1.6 ± 0.09**	1.2 ± 0.06	1.1 ± 0.10	1.0 ± 0.21	0.7 ± 0.11	1.0 ± 0.13	0.7 ± 0.07	0.9 ± 0.08
Xly2	**1.3 ± 0.10**	**1.2 ± 0.07**	**0.0 ± 0.00**	**1.5 ± 0.02**	**0.0 ± 0.00**	1.2 ± 0.04	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
Amino acids												
Ala	1.8 ± 0.31	0.8 ± 0.18	0.8 ± 0.11	2.8 ± 0.61	1.0 ± 0.25	1.5 ± 0.16	0.4 ± 0.06	0.4 ± 0.12	0.5 ± 0.05	1.6 ± 0.17	0.5 ± 0.02	1.0 ± 0.19
Asp	1.4 ± 0.11	**10.4 ± 0.28**	0.8 ± 0.12	**0.0 ± 0.00**	17.8 ± 9.23	**2.6 ± 0.25**	N.D.	100*	N.D.	N.D.	N.D.	N.D.
Gly	1.2 ± 0.39	**0.6 ± 0.06**	2.8 ± 0.81	**1.7 ± 0.09**	**1.2 ± 0.00**	3.0 ± 0.90	1.0 ± 0.22	1.0 ± 0.24	1.6 ± 0.55	1.1 ± 0.09	**0.2 ± 0.04**	0.9 ± 0.19
Ile	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	100*	100*	100*	**2.5 ± 0.32**	**0.0 ± 0.00**	1.0 ± 0.12
Leu	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	100*	N.D.	N.D.	**2.3 ± 0.34**	**0.0 ± 0.00**	0.8 ± 0.12
Pro	**3.2 ± 0.53**	**0.1 ± 0.01**	1.7 ± 0.39	4.6 ± 1.40	**0.0 ± 0.00**	0.9 ± 0.19	0.6 ± 0.22	0.5 ± 0.03	**0.0 ± 0.00**	**2.4 ± 0.26**	**0.0 ± 0.00**	**0.0 ± 0.00**
Val	1.4 ± 0.13	0.7 ± 0.10	1.1 ± 0.06	1.1 ± 0.16	**0.4 ± 0.05**	0.8 ± 0.17	1.7 ± 0.39	1.0 ± 0.18	1.0 ± 0.13	1.8 ± 0.37	**0.2 ± 0.06**	**0.7 ± 0.06**

Table 9 continued

	RIPE			TURNING			BREAKER			MG		
	Neo-123	Neo-111	3939	Neo-123	Neo-111	3939	Neo-123	Neo-111	3939	Neo-123	Neo-111	3939
	Av. ± SEM	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**
Organic and other acids												
phosphoric	3.5 ± 1.43	1.6 ± 0.54	1.9 ± 0.62	2.0 ± 0.94	2.0 ± 0.48	1.9 ± 0.51	0.9 ± 0.24	0.7 ± 0.12	1.0 ± 0.06	2.9 ± 0.74	**0.0 ± 0.00**	1.0 ± 0.03
butanoic	100*	N.D.	100*	**3.1 ± 0.44**	**2.5 ± 0.00**	**10.8 ± 2.83**	1.1 ± 0.28	0.9 ± 0.29	1.6 ± 0.39	1.1 ± 0.13	0.5 ± 0.19	1.2 ± 0.10
itaconic	1.1 ± 0.17	0.9 ± 0.21	**1.6 ± 0.17**	100*	100*	100*	**0.4 ± 0.01**	**0.0 ± 0.00**	**0.0 ± 0.00**	100*	N.D.	N.D.
malonic	1.2 ± 0.38	0.4 ± 0.11	0.6 ± 0.07	**0.0 ± 0.00**	2.2 ± 0.45	**2.4 ± 0.26**	0.3 ± 0.08	0.6 ± 0.02	0.6 ± 0.14	**2.9 ± 0.52**	0.7 ± 0.16	1.5 ± 0.27
oxalic	0.9 ± 0.11	0.8 ± 0.10	0.5 ± 0.15	1.4 ± 0.60	1.4 ± 0.23	1.3 ± 0.24	0.7 ± 0.27	0.7 ± 0.26	0.5 ± 0.23	0.9 ± 0.33	0.3 ± 0.07	0.9 ± 0.61
Sterols												
bsitostrol	100*	100*	100*	1.6 ± 0.49	**0.0 ± 0.00**	1.4 ± 0.14	1.2 ± 0.19	0.4 ± 0.09	**0.0 ± 0.00**	N.D.	N.D.	N.D.
campestrol	1.5 ± 0.20	0.0 ± 0.00	1.4 ± 0.19	**2.8 ± 0.04**	0.7 ± 0.16	0.8 ± 0.10	1.1 ± 0.07	**0.0 ± 0.00**	**0.0 ± 0.00**	100*	N.D.	N.D.
stigmstrol	1.1 ± 0.19	0.7 ± 0.03	0.7 ± 0.15	1.6 ± 0.23	0.6 ± 0.02	0.8 ± 0.13	1.1 ± 0.15	**0.3 ± 0.10**	0.6 ± 0.01	0.6 ± 0.15	**0.4 ± 0.11**	0.8 ± 0.06
Fatty acid and triacylglyceride biosynthesis intermediates												
G3P	100*	100*	N.D.	100*	N.D.	N.D.	N.D.	N.D.	N.D.	1.1 ± 0.27	**2.3 ± 0.15**	1.1 ± 0.23
linoleic	1.7 ± 0.44	1.0 ± 0.29	0.6 ± 0.19	8.3 ± 3.72	1.9 ± 0.57	1.1 ± 0.25	1.1 ± 0.12	0.9 ± 0.08	0.7 ± 0.19	1.1 ± 0.09	1.3 ± 0.03	0.9 ± 0.08
oleic	100*	100*	100*	9.8 ± 4.81	**0.0 ± 0.00**	**0.0 ± 0.00**	100*	N.D.	N.D.	100*	N.D.	N.D.
oleic2G	1.0 ± 0.34	0.3 ± 0.11	0.8 ± 0.26	5.0 ± 1.42	1.2 ± 0.47	1.1 ± 0.20	1.7 ± 0.30	1.1 ± 0.15	0.6 ± 0.14	1.1 ± 0.14	1.4 ± 0.11	0.9 ± 0.08
oleic1G	100*	N.D.	N.D.	100*	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
behenic	1.2 ± 0.14	1.0 ± 0.30	**0.3 ± 0.01**	2.1 ± 0.38	8.7 ± 3.45	**0.0 ± 0.00**	N.D.	N.D.	N.D.	N.D.	100*	N.D.
arachidic	1.6 ± 0.44	0.4 ± 0.05	0.4 ± 0.10	**3.0 ± 0.41**	5.6 ± 1.50	1.8 ± 0.54	N.D.	N.D.	N.D.	0.8 ± 0.07	1.3 ± 0.41	**0.6 ± 0.05**
palmitic	1.8 ± 0.09	0.7 ± 0.15	0.5 ± 0.11	2.9 ± 0.84	1.3 ± 0.13	1.0 ± 0.16	1.3 ± 0.14	0.8 ± 0.07	0.8 ± 0.06	1.1 ± 0.04	1.2 ± 0.06	1.0 ± 0.04
palmitic1G	**1.7 ± 0.22**	0.5 ± 0.13	0.9 ± 0.23	3.6 ± 0.90	1.6 ± 0.21	1.2 ± 0.11	**1.9 ± 0.22**	1.3 ± 0.23	**0.6 ± 0.08**	1.2 ± 0.11	1.1 ± 0.22	0.7 ± 0.10
stearic	1.5 ± 0.25	0.9 ± 0.22	0.6 ± 0.04	1.8 ± 0.58	1.3 ± 0.03	0.9 ± 0.01	0.8 ± 0.05	0.9 ± 0.03	0.7 ± 0.01	1.1 ± 0.03	**1.3 ± 0.03**	1.2 ± 0.04
stearic2G	100*	N.D.	N.D.	100*	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
stearic1G	**10.1 ± 3.08**	2.0 ± 0.95	2.1 ± 0.70	24.9 ± 9.06	**11.3 ± 1.84**	**3.2 ± 0.61**	100*	100*	N.D.	**1.5 ± 0.10**	1.4 ± 0.64	**0.4 ± 0.05**
lignoceric	100*	100*	100*	1.3 ± 0.37	8.3 ± 1.86	0.8 ± 0.02	100*	100*	N.D.	0.8 ± 0.10	2.6 ± 1.58	1.1 ± 0.17

Table 9 continued

	RIPE			TURNING			BREAKER			MG		
	Neo-123	Neo-111	3939	Neo-123	Neo-111	3939	Neo-123	Neo-111	3939	Neo-123	Neo-111	3939
	Av. ± SEM	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**	**Av. ± SEM**
Glycolysis and tricarboxylic acid cycle intermediates												
G6P	100*	N.D.	N.D.	**3.0 ± 0.18**	**0.0 ± 0.00**	**1.6 ± 0.12**	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
kglutaric	0.9 ± 0.14	**0.7 ± 0.11**	1.0 ± 0.08	1.1 ± 0.10	0.7 ± 0.00	0.8 ± 0.09	1.0 ± 0.06	0.8 ± 0.02	1.2 ± 0.22	100*	100*	N.D.
aconitic	0.8 ± 0.19	0.2 ± 0.01	1.8 ± 0.48	0.8 ± 0.37	0.7 ± 0.31	1.7 ± 0.24	1.0 ± 0.15	1.0 ± 0.18	1.1 ± 0.10	**2.0 ± 0.31**	0.3 ± 0.21	1.1 ± 0.05
citric	1.1 ± 0.17	**0.7 ± 0.08**	1.1 ± 0.10	1.0 ± 0.18	0.9 ± 0.05	1.3 ± 0.09	0.9 ± 0.08	0.9 ± 0.10	1.0 ± 0.07	1.5 ± 0.18	0.8 ± 0.16	**1.2 ± 0.04**
malic	1.4 ± 0.35	0.8 ± 0.08	**0.6 ± 0.12**	**0.5 ± 0.13**	1.8 ± 0.43	1.0 ± 0.20	0.6 ± 0.16	0.7 ± 0.06	0.8 ± 0.11	3.6 ± 1.19	0.8 ± 0.10	1.2 ± 0.07
Tocopherols and triterpenes												
a-toc	**1.8 ± 0.23**	**1.4 ± 0.13**	0.9 ± 0.08	1.1 ± 0.20	0.7 ± 0.05	0.8 ± 0.09	1.1 ± 0.11	0.7 ± 0.15	0.8 ± 0.19	0.6 ± 0.18	0.5 ± 0.11	0.9 ± 0.17
g-toc	0.9 ± 0.11	**0.4 ± 0.05**	0.7 ± 0.11	**2.3 ± 0.22**	**0.3 ± 0.09**	1.1 ± 0.26	1.1 ± 0.38	0.3 ± 0.07	0.4 ± 0.13	**2.3 ± 0.33**	0.9 ± 0.04	**2.5 ± 0.48**
TQH2	**2.2 ± 0.27**	**1.3 ± 0.05**	1.4 ± 0.45	100*	N.D.	N.D.	100*	100*	100*	100*	N.D.	N.D.
b-amyrin	0.8 ± 0.17	0.8 ± 0.16	**0.2 ± 0.03**	0.8 ± 0.14	**1.4 ± 0.08**	1.4 ± 0.17	1.7 ± 0.25	0.9 ± 0.20	0.8 ± 0.29	0.9 ± 0.06	1.2 ± 0.60	0.8 ± 0.13
Other and unknown												
inositol	**1.5 ± 0.17**	**1.7 ± 0.14**	1.1 ± 0.29	1.1 ± 0.20	1.2 ± 0.06	1.1 ± 0.31	0.8 ± 0.16	0.9 ± 0.22	1.2 ± 0.04	1.0 ± 0.16	1.3 ± 0.29	1.1 ± 0.32
m-inos-P	2.0 ± 0.56	**2.2 ± 0.09**	1.4 ± 0.34	0.6 ± 0.09	**0.5 ± 0.04**	**0.5 ± 0.05**	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
UNKnp1	0.7 ± 0.13	0.6 ± 0.01	0.5 ± 0.06	0.9 ± 0.17	0.7 ± 0.03	0.7 ± 0.02	0.9 ± 0.04	0.8 ± 0.02	0.8 ± 0.07	1.0 ± 0.03	1.1 ± 0.01	0.9 ± 0.01
UNKnp2	N.D.	100*	N.D.	N.D.	100*	100*	**9.0 ± 1.94**	**9.2 ± 0.74**	**0.0 ± 0.00**	100*	100*	N.D.
UNKnp3	**4.1 ± 0.76**	**17.1 ± 5.64**	3.3 ± 1.92	**5.1 ± 0.79**	**13.9 ± 0.09**	5.7 ± 2.29	**16.4 ± 3.86**	13.8 ± 4.79	0.9 ± 0.12	**3.7 ± 0.96**	3.8 ± 1.18	**0.0 ± 0.00**
UNKnp4	N.D.	N.D.	N.D.	**0.0 ± 0.00**	**0.0 ± 0.00**	1.1 ± 0.36	1.3 ± 0.06	1.1 ± 0.05	1.0 ± 0.24	0.9 ± 0.10	1.3 ± 0.02	1.0 ± 0.11
UNKp1	1.0 ± 0.09	1.0 ± 0.09	0.7 ± 0.13	1.0 ± 0.04	1.0 ± 0.03	0.7 ± 0.14	1.2 ± 0.14	1.0 ± 0.04	**0.0 ± 0.00**	100*	N.D.	N.D.
UNKp2	0.7 ± 0.10	0.8 ± 0.08	1.1 ± 0.09	0.9 ± 0.15	1.0 ± 0.07	**1.3 ± 0.06**	0.8 ± 0.08	0.7 ± 0.05	0.9 ± 0.06	1.4 ± 0.13	0.7 ± 0.17	1.0 ± 0.08
UNKp3	0.8 ± 0.10	0.9 ± 0.10	**0.8 ± 0.03**	**0.8 ± 0.04**	0.8 ± 0.05	**0.8 ± 0.05**	0.9 ± 0.14	0.7 ± 0.03	0.9 ± 0.05	0.7 ± 0.10	0.5 ± 0.08	0.7 ± 0.03
UNKp4	**0.7 ± 0.11**	0.7 ± 0.13	1.0 ± 0.10	0.8 ± 0.08	0.8 ± 0.06	0.8 ± 0.06	0.5 ± 0.16	0.7 ± 0.14	1.1 ± 0.30	0.6 ± 0.09	1.2 ± 0.26	0.9 ± 0.20
UNKp5	N.D.	N.D.	N.D.	N.D.	100*	100*	100*	N.D.	N.D.	**0.0 ± 0.00**	**0.0 ± 0.00**	**0.0 ± 0.00**
UNKp6	0.9 ± 0.07	**1.5 ± 0.14**	1.0 ± 0.14	0.8 ± 0.06	0.8 ± 0.08	1.2 ± 0.28	0.5 ± 0.05	0.6 ± 0.16	0.9 ± 0.05	**0.7 ± 0.08**	**0.7 ± 0.02**	1.4 ± 0.23
UNKp7	N.D.	N.D.	N.D.	**0.0 ± 0.00**	0.7 ± 0.19	1.0 ± 0.10	0.2 ± 0.00	0.7 ± 0.13	0.9 ± 0.20	2.0 ± 0.43	**0.0 ± 0.00**	1.6 ± 0.26

Stearic2G (octadecanoic acid methyl ethyl ester) and oleic1G (9-octadecenoic acid propyl ester) were both detected in neo-123, but not in TA209 or either of the high rutin genotypes. There was a greater degree of significance in the changes to sugar levels of neo-123 compared with 3939 and neo-111 (Table 9) that when viewed with the loading plot (Figure 32F) also aids in explaining the cluster separation of neo-123. For example, ribose was significantly increased in neo-123 by 1.4 fold, but exhibited non-significant increases of 1.1 and 1.3 fold in neo-111 and 3939, respectively. Similarly, sugars such as Fru2, Glc and msacch3 are only shown to significantly decrease in neo-123.

5.4 Characterisation of fruit during ripening

Previous observations indicated that fruit ripening times may be affected in each of the lines when compared with TA209. As a result, a detailed investigation during fruit development and ripening monitored changes to fruit appearance, ripening times and phenolic profile.

5.4.1 Fruit ripening

Representative fruit for each of the four lines at four time points during fruit development and six time points during fruit ripening are shown in Figure 33. Little change was observed in fruit appearance during early fruit development, but some changes in size were seen towards 35 dpa and 0 dpb (breaker stage). A green-yellow breaker stage was observed in lines neo-111 and -123 (Figure 33C and D), whereas TA209 and 3939 exhibited a non-uniform red pigmentation at breaker stage (Figure 33A and B). While neo-111 and -123 fruit developed from orange/orange-green at early ripening and developed to red-orange throughout ripening, TA209 and 3939 changed from red-orange to deep red.

Fruit development and ripening time was quantified, and is shown in Figure 34. Both lines containing *S. neorickii* introgressed regions reached mature green stage in significantly fewer days than TA209 (neo-123 $p \leq 0.01$; neo-111 $p \leq 0.001$; Figure 34A). However, all three lines reached breaker stage significantly later than TA209 ($p \leq 0.001$, Figure 34B). A delay in ripening was observed therefore for all three lines, and this was greatest for neo-111 and neo-123.

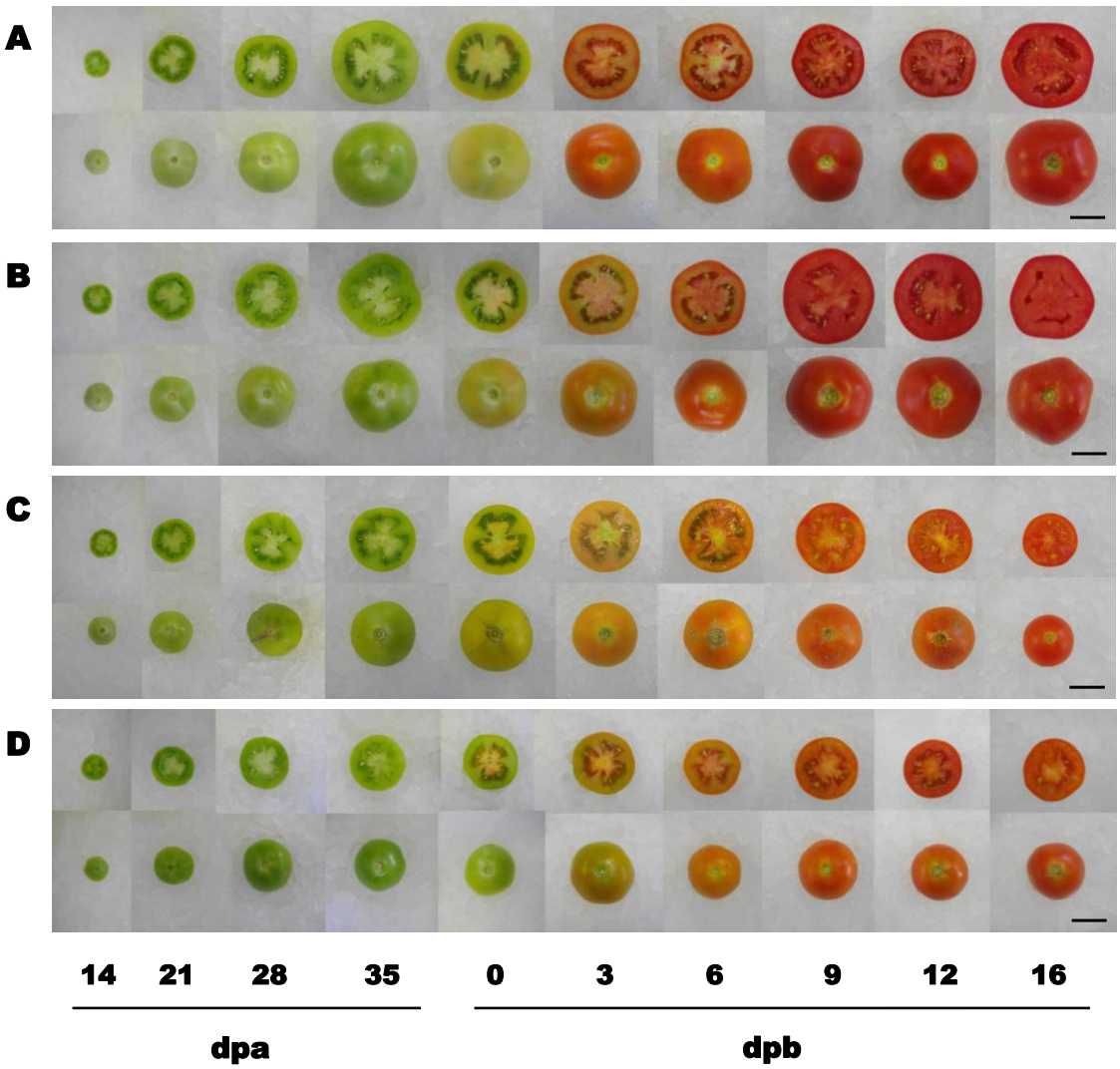

Figure 33 Fruit ripening series

Typical fruit morphology and pigmentation throughout fruit development and ripening at time points measured in days post anthesis (dpa) and breaker (dpb) for lines (A) TA209, (B) 3939, (C) neo-111, and (D) neo-123. Scale bars represent 2 cm.

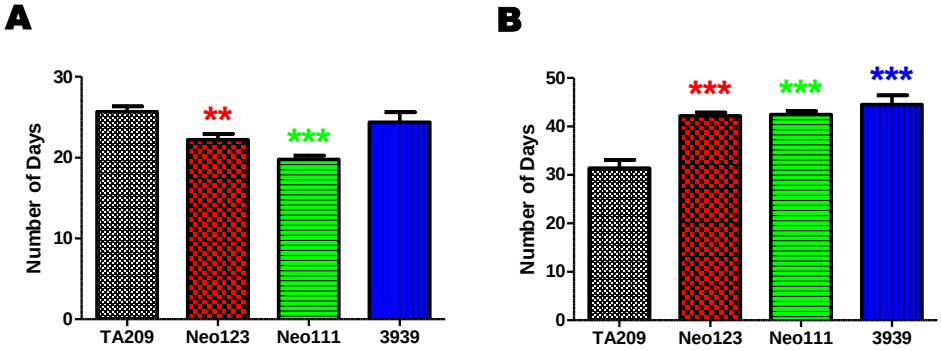

Figure 34 Time for fruit development and ripening

Number of days from anthesis to (A) mature green stage and (B) breaker stage. Error bars represent biological replicates: anthesis to mature green, n=14 to 26; anthesis to breaker, n=22 to 42. Values significantly different from TA209 are shown by * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$).

5.4.2 Analysis of phenolic profile throughout fruit ripening

Fruit at each time point were analysed by UPLC-PDA to assess phenolic profile at equivalent time points (Figure 35) rather than at equivalent fruit stages shown previously in Figure 25. Observed levels of *p*-coumaric acid were low, and were therefore combined with *p*-coumaric acid derivative in Figure 35A. Accumulation in neo-123 began at breaker stage (0 dpb) and increased to levels significantly higher than TA209 at 6 ($p \leq 0.01$), 12 ($p \leq 0.05$) and 16 ($p \leq 0.05$) dpb. Neo-111 and 3939 showed comparatively moderate accumulation of *p*-coumaric acid and derivatives, but at later time points (significant increased at 12 and 9 dpb, respectively; $p \leq 0.05$).

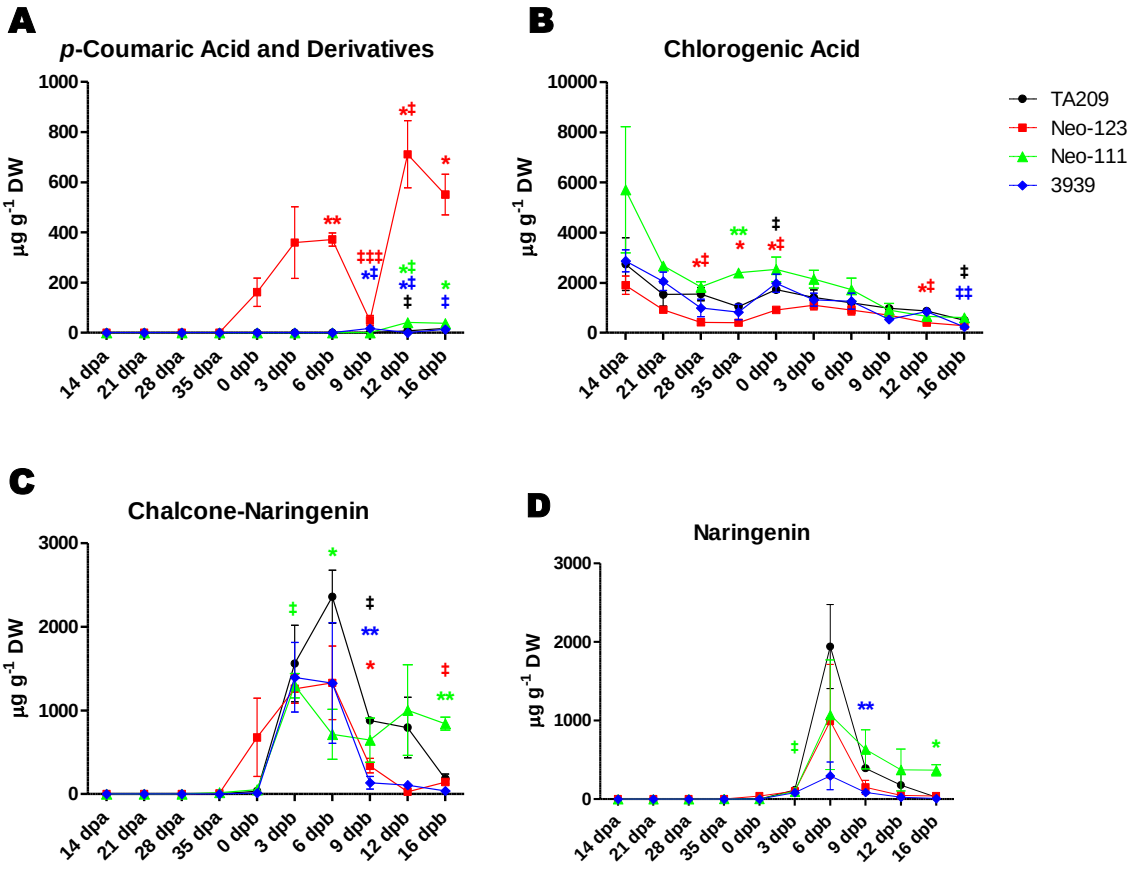

Figure 35 Phenolic profile throughout fruit development and ripening series

Phenolic compounds detected by UPLC-PDA throughout time course in days post anthesis (dpa) and breaker (dpb). MG, mature green stage; UNK rutin-like 1 and 2 represent unknown compounds with UV spectral shapes similar to rutin. Error bars represent SEM for biological replicates, n=3. At any one time point, values significantly different from TA209 parent are indicted by *, and values significantly different from the preceding time point of the same genotype are indicated by ‡. One, two and three replicated symbols denote significance thresholds $p \leq 0.05$, $p \leq 0.01$, and $p \leq 0.001$, respectively. Zero values represent levels not detected by UPLC-PDA.

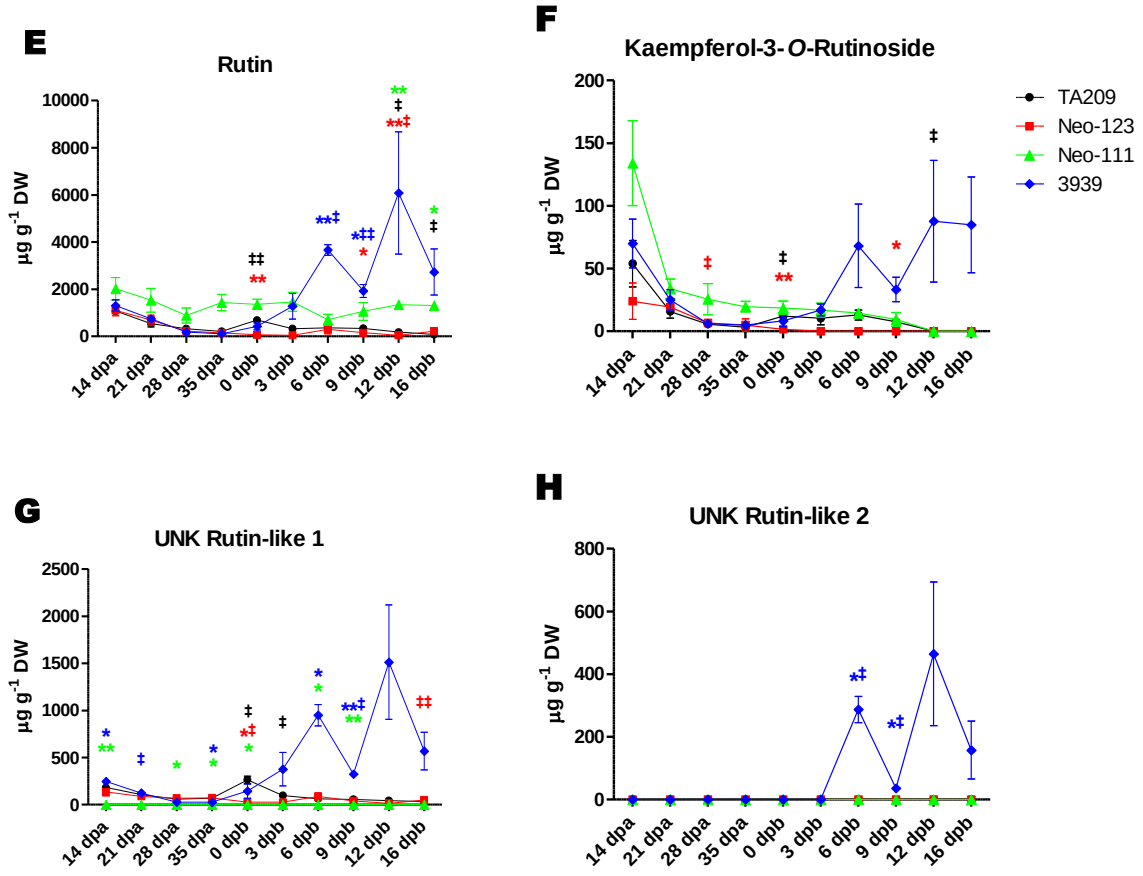

Figure 35 continued.

The trend in chlorogenic acid abundance was similar for all four lines (Figure 35B). High levels were found in early fruit development (14 dpa) relative to the remaining time points in the series for all lines. Abundance was reduced as fruit develops towards breaker stage (0 dpb), around which point levels were seen to increase slightly (but significantly so for TA209 and neo-123; $p \leq 0.05$). Abundance decreased once again towards ripe and overripe stages, and significant decreases between consecutive time points were observed with TA209 ($p \leq 0.05$), neo-123 ($p \leq 0.05$) and 3939 ($p \leq 0.01$). Throughout this trend, 3939 replicated TA209 and no significant differences were observed between these two lines. In contrast, during fruit development and ripening neo-111 exhibited higher levels than TA209 and earlier accumulation of chlorogenic acid, resulting in a significant increase compared with TA209 at 35 dpa ($p \leq 0.01$). Conversely, neo-123 displayed lower levels than TA209 at all time points, and showed significantly lower levels at 28 and 35 dpa, and 0 and 12 dpb ($p \leq 0.05$ in all cases).

Chalcone-naringenin and naringenin (Figure 35C and D) were not detected until approximately breaker stage (0 dpb for TA209 and 3939, and 35 dpa for neo-111 and -123). Peak abundance was seen during early ripening, at 3 to 6 dpb, after which time all lines showed a decrease in chalcone-naringenin and naringenin levels from 9 to 16 dpb. For TA209, this decrease reached levels almost below detection limits. 3939 and neo-123 decreased more rapidly than TA209, resulting in some significantly lower levels at 9 dpb (between $p \leq 0.05$ and $p \leq 0.01$). Although levels in neo-111 decreased through ripening, they remained significantly higher than TA209 at 16 dpb ($p \leq 0.01$ for chalcone-naringenin and $p \leq 0.05$ for naringenin).

Figure 35E to H represents rutin, kaempferol-3-*O*-rutinoside, and two unknown compounds with similar UV spectra to rutin. Likewise with Figure 25G to J, these were putatively and collectively named flavonol glycosides. Of the four compounds, rutin was observed in the greatest abundance. In TA209, rutin (Figure 35E), kaempferol-3-*O*-ruitnoside (Figure 35F), and UNK rutin-like 1 (Figure 35G) displayed a gradual decline in abundance throughout development and ripening except for a significant ($p<0.05$) increase spike at 0 dpb. For these three compounds, neo-123 maintained levels equivalent to or lower than TA209 throughout development and ripening but did not exhibit this spike at 0 dpb. UNK rutin-like 2 (Figure 35H) was not detected in either TA209 or neo-123. Flavonol glycoside accumulation occurred differently in each of the high-rutin lines neo-111 and 3939. Levels of rutin (Figure 35E) were consistently higher in neo-111 compared with TA209 throughout development, and significantly higher at late ripe stages 12 ($p \leq 0.01$) and 16 dpb ($p \leq 0.05$). Although neo-111 showed higher levels of kaempferol-3-*O*-rutinoside (Figure 35F) in early fruit development compared with TA209, there were no significant increases. Furthermore, UNK rutin-like 1 and 2 (Figure 35G and H) were not detected in neo-111. 3939, by comparison, largely replicated flavonol glycoside accumulation in TA209 in fruit development up to breaker stage (Figure 35E to H), with the exception of UNK rutin-like 1 at 14 and 35 dpa. Following breaker stage, in 3939 all four compounds increased throughout ripening to levels higher than TA209, and show significant increases for rutin, and UNK rutin-like 1 and 2 (between $p \leq 0.01$ and $p \leq 0.05$). Similar to the anti-spike seen for neo-123 *p*-coumaric acid and derivative levels at 9 dpb (Figure 35A), all four flavonol glycosides exhibited an anti-spike in abundance for 3939 at 9 dpb (Figure 35E to H).

5.5 Antioxidant activity of phenolic extract

The antioxidant capacity of phenolic extracts in methanol was assessed using the TEAC assay. Throughout fruit development (14 to 35 dpa) and at breaker stage (0 dpb) there were no significant differences in TEAC levels for any of the lines when compared with TA209 (Figure 36). All lines showed significantly higher TEAC levels at some point during fruit ripening compared with TA209. For neo-111 this was at 3 and 16 dpb ($p<0.05$ in each case); for neo-123 this was at 3 ($p\leq0.05$), 6 ($p\leq0.01$), and 12 dpb ($p\leq0.01$); and for 3939 this was at 12 dpb ($p\leq0.05$).

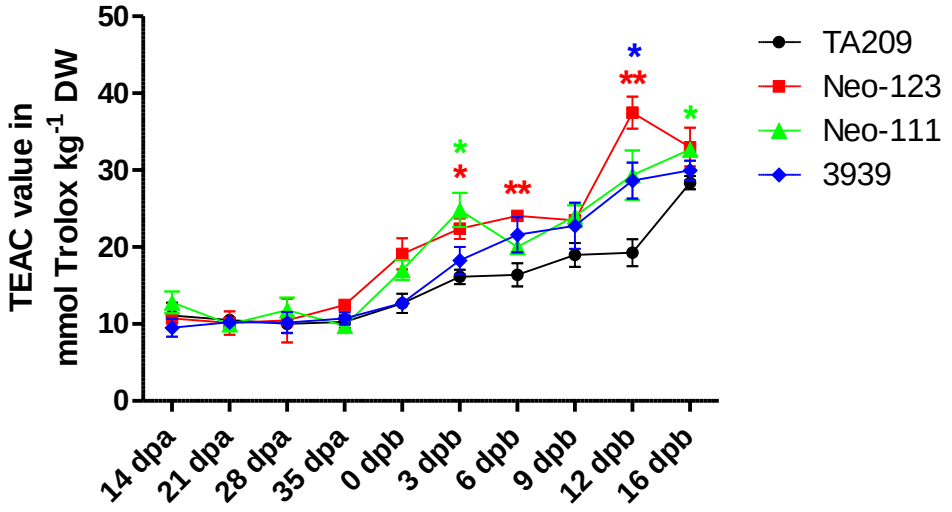

Figure 36 Antioxidant activity of polar extract throughout fruit development and ripening

Antioxidant activity is expressed as Trolox equivalent antioxidant capacity (TEAC). Error bars represent SEM, n=3. Values significantly different from TA209 are represented by * ($p\leq0.05$), ** ($p\leq0.01$), and *** ($p\leq0.001$).

Significant increases in antioxidant activities for high-rutin lines neo-111 ($p\leq0.001$) and 3939 ($p\leq0.01$) were limited to fruit skin samples (Figure 37). No significant change was observed in flesh and jelly samples for neo-111 and 3939. Neo-123 showed significant increases in antioxidant activity in both flesh and jelly samples ($p\leq0.001$) compared with TA209 (Figure 37), but no significant change in fruit skin samples. Analysis of antioxidant activity from polar extracts of flower tissue showed that no significant difference was observed between either neo-111 or -123 and TA209 (Figure 38). However, a significant increase was observed with 3939 ($p\leq0.05$).

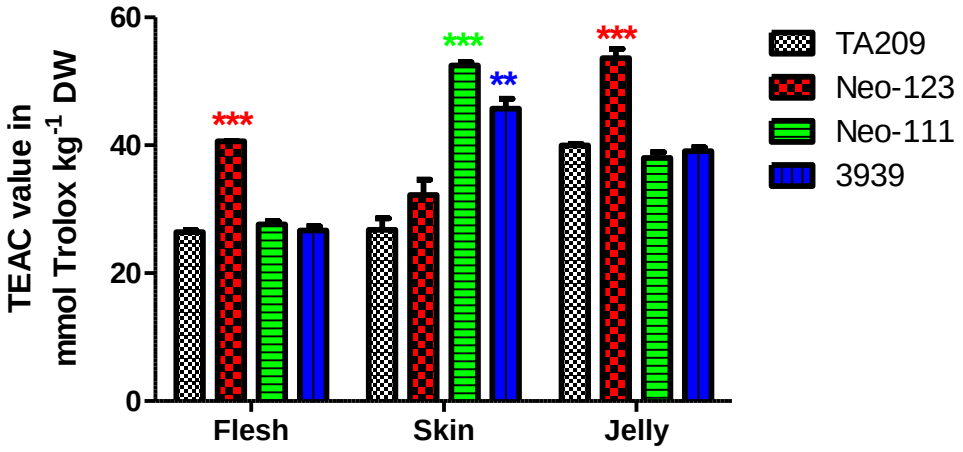

Figure 37 Antioxidant activity of polar extract from fruit tissue types

Antioxidant activity is expressed as Trolox equivalent antioxidant capacity (TEAC). Error bars represent SEM, n=3. Values significantly different from TA209 are represented by * ($p \leq 0.05$), ** ($p \leq 0.01$), and *** ($p \leq 0.001$).

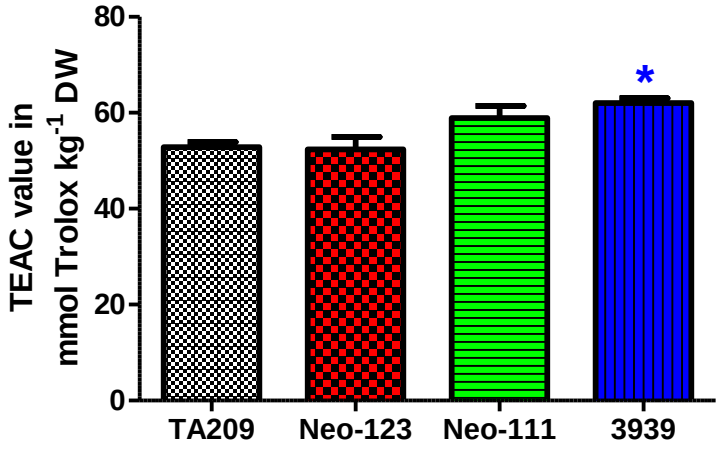

Figure 38 Antioxidant activity of polar extract from flower tissue

Antioxidant activity is expressed as Trolox equivalent antioxidant capacity (TEAC). Error bars represent SEM, n=3 technical replicates where at least 20 flower were pooled per genotype. Values significantly different from TA209 are represented by * ($p \leq 0.05$), ** ($p \leq 0.01$), and *** ($p \leq 0.001$).

5.6 Physiological fruit parameters

Typical fruit morphology at ripe stage is shown in Figure 39. Fruit from TA209 parent were ovate in shape and red in colour at ripe stage (Figure 39A). 3939 fruit were similar in shape and colour, but possessed a pointed blossom-end morphology in many cases (Figure 39D) that was not observed in any other line. Fruit from neo-123 were

similarly ovate in shape, but red to orange in colour (Figure 39B). Fruit from neo-111 were ovate to round in shape, and red to orange in colour (Figure 39C), and therefore the least similar to TA209 morphology. In addition, many fruit from neo-111 exhibited fruit cracking and micro-cracking. Fruit colour for all lines were uniform with no visible shoulders. Typically fruits contained two or three locules (as shown previously in Figure 33, section 5.4.1).

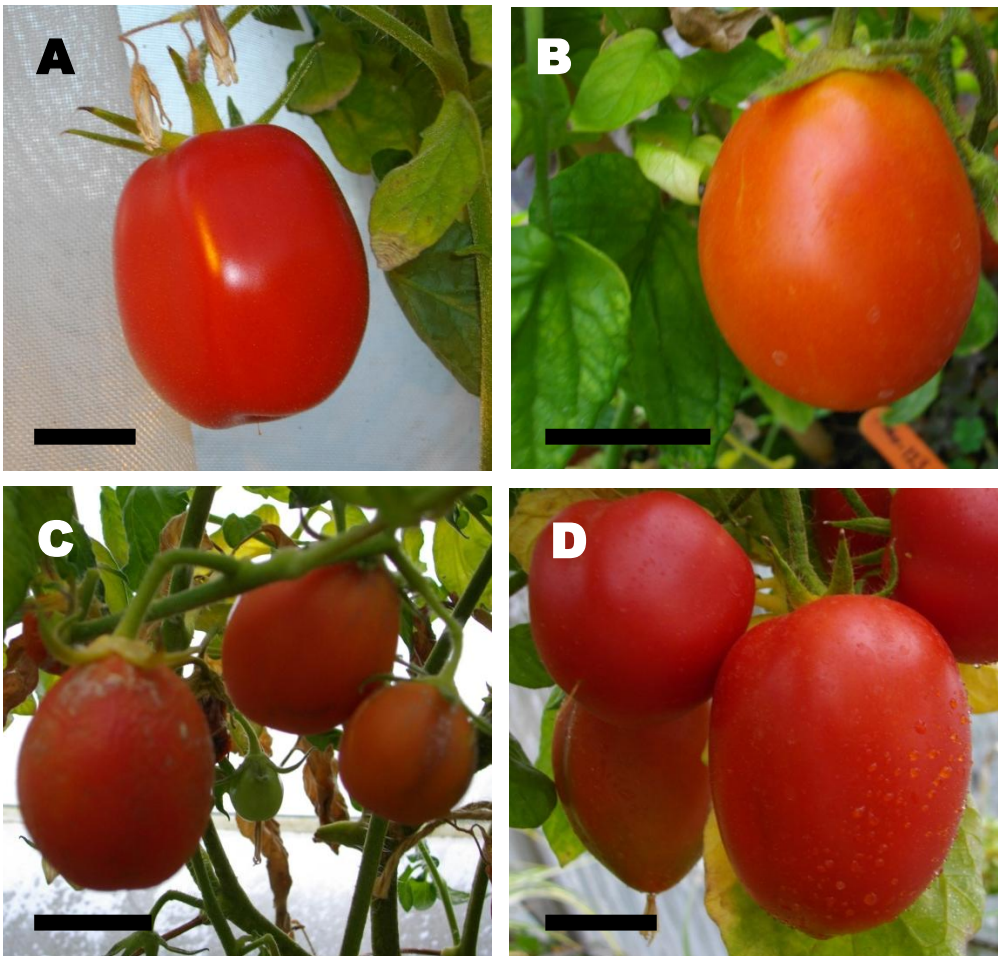

Figure 39 Typical fruit morphology

Typical intact fruit at ripe stage shown for lines (A) TA209, (B) neo-123, (C) neo-111, and (D) 3939. Scale bars represent 2 cm in each case.

The differences in fruit size between the lines was compared by quantification of fruit mass and diameter at time points throughout fruit development and ripening, and at ripe stage (Figure 40). For each of the lines fruit mass and diameter continued to increase throughout fruit development (Figure 40A and C) between 14 and 35 dpa. During this time the only line to show any difference from TA209 was neo-123, which

possessed fruit with significantly less mass ($p \leq 0.05$) at 28 dpa. At breaker stage (0 dpb) the only line to show any significant difference from TA209 was neo-111, where fruit had significantly greater mass ($p \leq 0.05$). While fruit mass and diameter continued to increase throughout fruit ripening (0 to 16 dpb) in lines TA209 and 3939, lines neo-111 and -123 showed either no change or a decrease. This resulted in significantly smaller fruit compared with TA209 for neo-111 at 9, 12 and 16 dpb ($p \leq 0.05$) and for neo-123 at 12 dpb ($p \leq 0.05$).

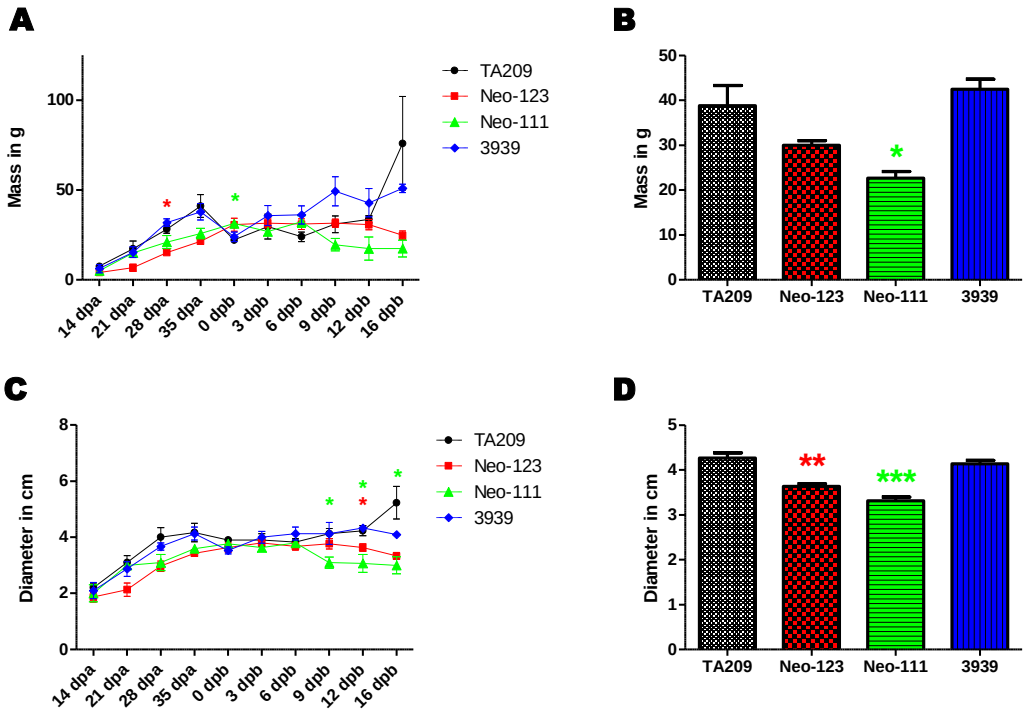

Figure 40 Comparison of fruit size

Fruit size determined by (A and B) fruit mass and (C and D) fruit diameter during fruit development and ripening (A and C) and at ripe and overripe stage (B and D). Error bars represent SEM based on biological replicates (A and C, n=3; B and D, n=14 to 20). Values significantly different from TA209 are shown by * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$).

Seed mass and number was assessed in ripe fruit for each line (Figure 41). No significant difference was detected in seed mass between any line and TA209 (Figure 41A). Additionally, there was no significant difference compared with TA209 in seed number per fruit for lines neo-123 and 3939 (Figure 41B); however, neo-111 fruit contained on average fewer than half the number of seeds of TA209 fruit, which was significantly different ($p \leq 0.05$).

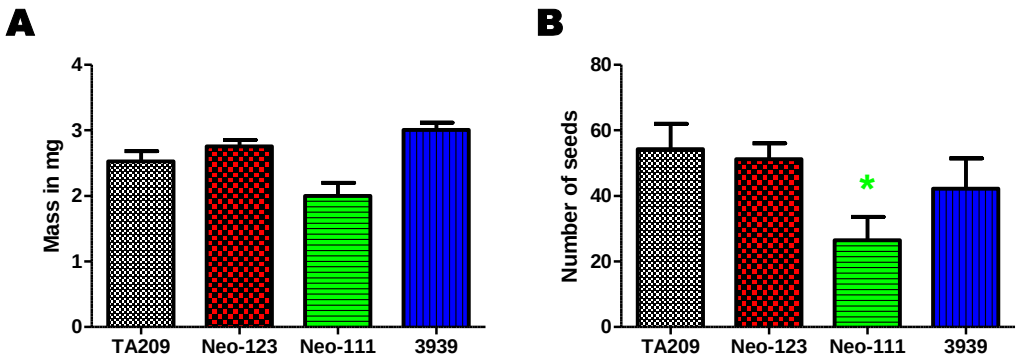

Figure 41 Comparison of seeds content in fruit

(A) Estimated mass per seed and (B) estimated number of seed per fruit for each line. Error bars represent SEM, where n=10 to 18. Values significantly different from TA209 are shown by * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$).

Throughout fruit development and ripening, fruit firmness and water content were monitored. At no point during development and ripening was fruit firmness significantly different from TA209 for any of the lines (Figure 42). Significant differences in water content were seen for all lines compared with TA209 (Figure 43). Except for at 14 dpa, neo-111 fruit contained less water at all time points, and significant differences were seen at 35 dpa, and at 0, 9 and 16 dpb ($p \leq 0.05$ in each case). Fruit from lines neo-123 and 3939 contained on average more water than TA209 throughout development and ripening except for 3939 at 16 dpb. Throughout this time scale, these increases were significant for both lines at 0, 3 and 6 dpb, and additionally for 3939 at 14 dpa ($p \leq 0.05$ in each case).

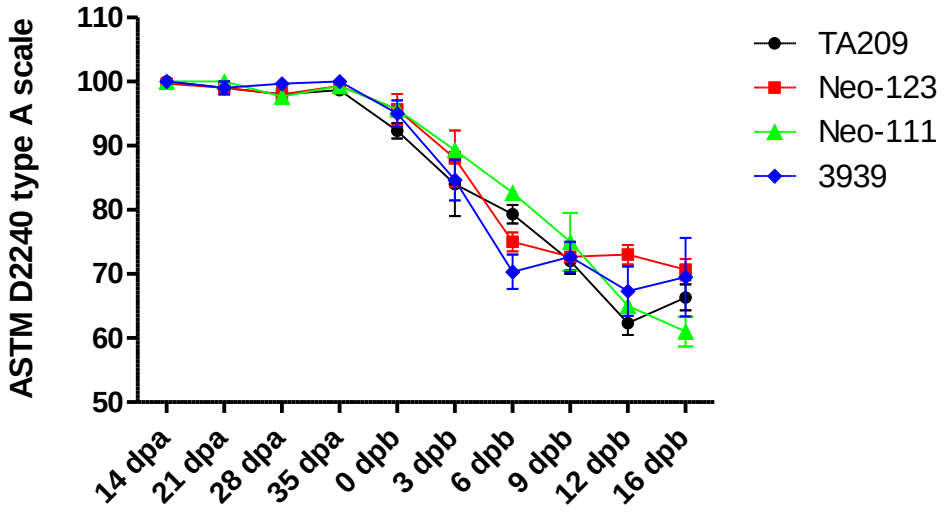

Figure 42 Fruit firmness throughout development and ripening

Fruit firmness units measured using the ASTM International D2240 (standard for hardness) type A scale ranging 0 to 100. Error bars represent SEM n=3. No values are significantly different ($p<0.05$) from TA209 at any stage.

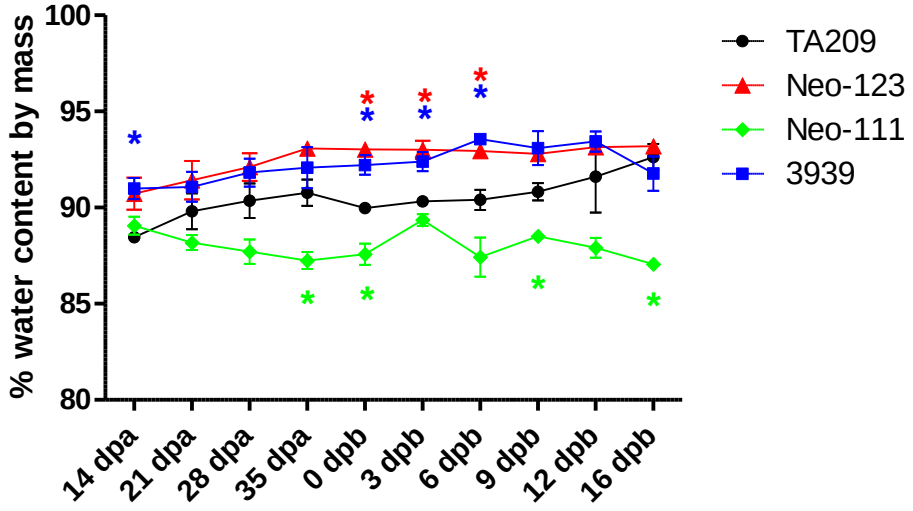

Figure 43 Comparison of water content throughout fruit development and ripening

Water content estimated by change in mass after freeze drying fruits and provided as percentage of initial mass. Error bars represent SEM, n=3. Values significantly different from TA209 are shown by * ($p\leq0.05$), ** ($p\leq0.01$), *** ($p\leq0.001$).

5.7 Post-harvest properties

In order to determine the effects of post-harvest conditions on each of the three lines and TA209, fruit were harvested at 6 dpb and stored for ten days (resulting in 6 dpb + 10 dph fruit). Comparisons were then made between 6 dpb + 10 dph fruit and two stages from the development and ripening series: 6 dpb and 16 dpb.

Fruit mass was monitored throughout post-harvest storage (Figure 44). Each line accumulatively lost mass throughout post-harvest storage at each two day interval. Identical conclusions were made when representing accumulative loss of mass in g (Figure 44A) and as a percentage of fruit mass at time of harvest (Figure 44B).

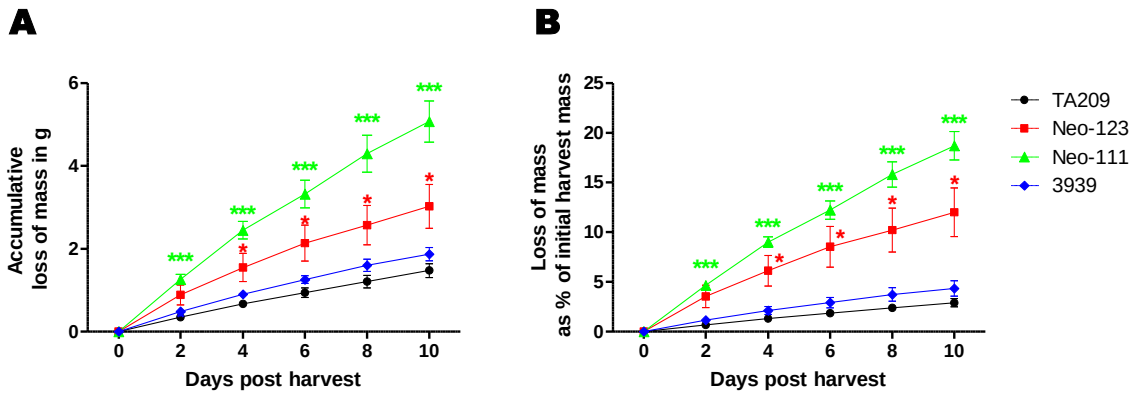

Figure 44 Loss of mass during post-harvest storage

Accumulative loss of mass expressed (A) in g and (B) as percentage of initial fruit mass for fruit stored post-harvest. Error bars represent SEM, n=6. Values significantly different are represented by * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$).

Fruit from 3939 showed no significant difference in loss of mass when compared with TA209. Although fruit from neo-123 showed no significant difference from TA209 at 2 dph, significantly greater mass was lost in neo-123 fruit compared with TA209 at all time points between 4 and 10 dph ($p \leq 0.05$). Neo-111 lost more mass than TA209 at all time points post-harvest, which was highly significant ($p \leq 0.001$).

Both *S. neorickii* BILs showed notable changes in morphology during post-harvest storage. Fruit from neo-111 displayed wrinkling of skin as early as 4 dph that resulted in the appearance of the epidermis shown in Figure 45, although internal fruit appearance seemed to be unaffected (Figure 45B).

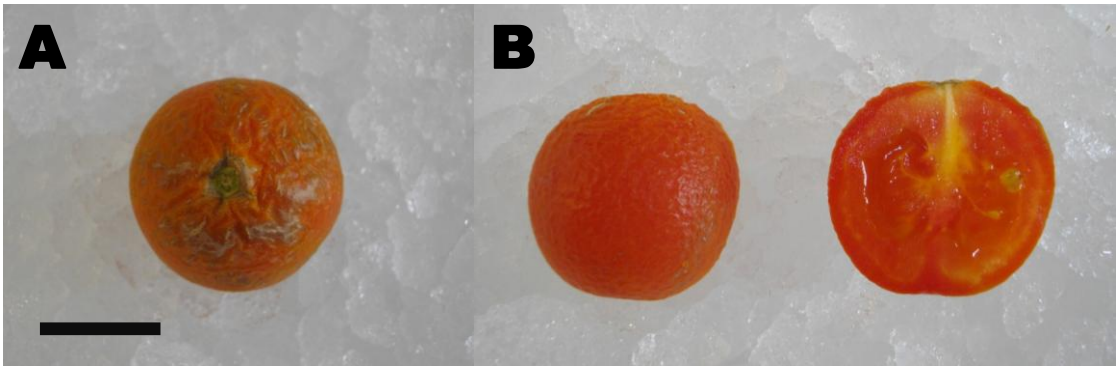

Figure 45 Neo-111 fruit after post-harvest storage

Wrinkled epidermis of neo-111 fruit following post-harvest storage (6 dpb + 10 dph) from (A) above and (B) side and cross section views. Scale bar represents 2 cm.

During post-harvest storage, fruit from neo-123 displayed an altered pigmentation, becoming more red in colour, as shown by Figure 46.

Figure 46 Neo-123 fruit before and after post-harvest storage

Typical change in pigmentation of neo-123 fruit before (6 dpb) and after (6 dpb + 10 dph) post harvest storage. Image represents two separate fruit. Scale bar represents 2 cm.

Despite the loss of mass observed in Figure 44 and the change in morphology to neo-111 (Figure 45), an assessment of water content expressed as a percentage of fruit mass found that none of the lines exhibited a significant decrease in water content during post-harvest storage (Figure 47).

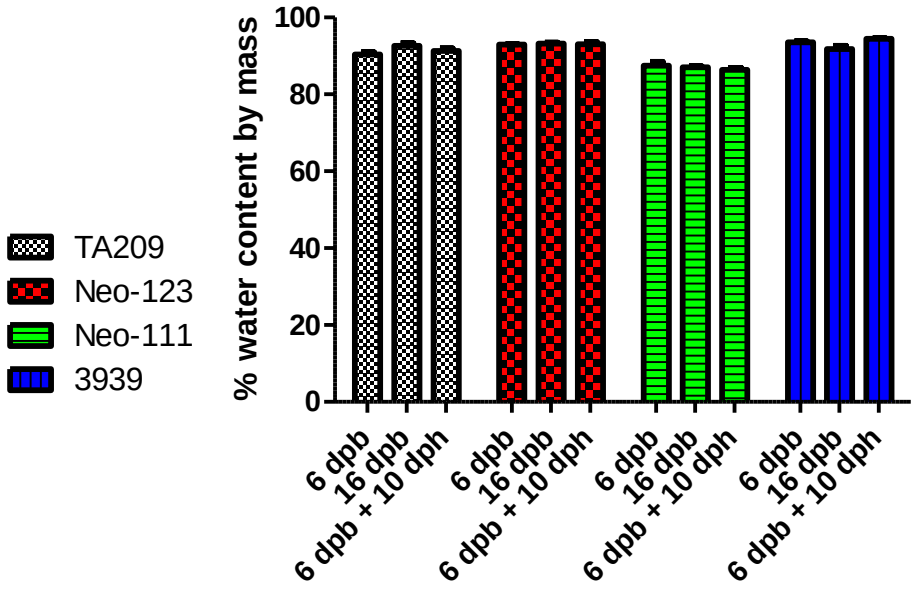

Figure 47 Comparison of water loss, measured post-harvest and on the vine

Estimated water content for fruit harvested at 6 and 16 days post breaker (dpb), and compared with fruit harvested at 6 dpb and stored for 10 days post harvest (dph). Error bars represent SEM: 6 and 16 dpb, n=3; 6 dpb + 10 dph, n=6. No significant decreases were observed between 6 dpb and 16 dpb or 6 dpb + 10 dph.

Fruit firmness was also assessed at these three stages (Figure 48). TA209 fruit were significantly less firm at 16 dpb when compared with 6 dpb ($p \leq 0.01$). An equivalent difference was seen between TA209 fruit at 6 dpb + 10 dph and 6 dpb (significance $p \leq 0.01$), but no significant difference was seen between fruits of identical time points stored off the plant (6 dpb + 10 dph) and allowed to remain on the vine of the plant (16 dpb). While fruit from neo-123 showed no significant difference between 6 and 16 dpb, fruit were significantly less firm post-harvest (6 dpb + 10 dph) when compared with both 6 and 16 dpb fruit ($p \leq 0.001$ and $p \leq 0.01$, respectively). On average, neo-111 showed the largest decrease in firmness between fruit at 6 and 16 dpb (significance $p \leq 0.01$). A further decrease in firmness was observed between 16 dpb and 6 dpb + 10 dph fruit (significance $p \leq 0.05$), which resulted in highly significant difference between fruit before and after post-harvest storage (comparing 6 dpb and 6 dpb + 10 dph, $p \leq 0.001$). No significant differences were observed for 3939 between any of the three stages.

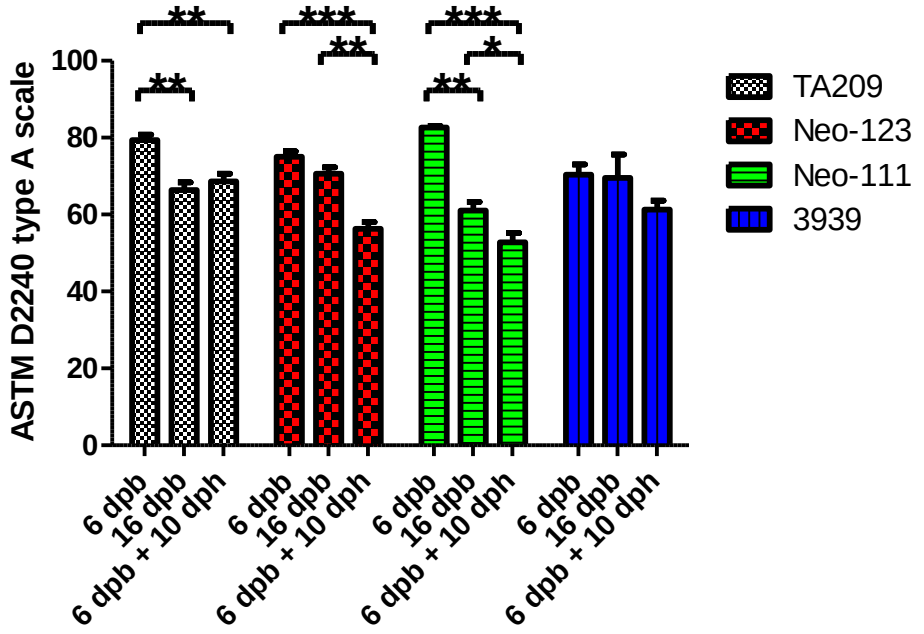

Figure 48 Comparison of fruit firmness, measured post-harvest and on the vine

Fruit firmness for fruit harvested at 6 and 16 days post breaker (dpb), and compared with fruit harvested at 6 dpb and stored for 10 days post harvest (dph). Units measured using the ASTM International D2240 (standard for hardness) type A scale ranging 0 to 100. Error bars represent SEM: 6 and 16 dpb, n=3; 6 dpb + 10 dph, n=6. Values significantly different are represented by * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$).

The effects of post-harvest storage conditions on levels of phenolic compounds are shown in Figure 49. Post-harvest storage (fruit at 6 dpb + 10 dph) shows little effect on phenylpropanoid levels (Figure 49A and B) when compared with fruit of equivalent age allowed to over-ripen on the vine (16 dpb). No significant differences were seen for *p*-coumaric acid and derivates or chlorogenic acid levels between 16 dpb and 6 dpb + 10 dph. This is supported by the significant decrease in chlorogenic acid levels observed between 6 and 16 dpb for TA209 ($p \leq 0.01$) also seen between 6 dpb and 6 dpb + 10 dph ($p \leq 0.001$; Figure 49B). However, significant increases in levels of *p*-coumaric acid and derivates observed for lines TA209, neo-111 and 3939 between 6 and 16 dpb ($p \leq 0.05$; Figure 49A), were not seen between 6 dpb and 6 dpb + 10 dph. Neo-123 was the only line to show a significant difference in levels of *p*-coumaric acid and derivates ($p \leq 0.001$) and chlorogenic acid ($p \leq 0.01$) between 6 dpb + 10 dph fruit and TA209 at 6 dpb + 10 dph.

Likewise, changes in chalcone-naringenin and naringenin that were observed between fruit at 6 and 16 dpb were also seen between fruit at 6 dpb and 6 dpb + 10 dph

(Figure 49C and D). In most cases these changes were non-significant decreases as fruit ripened/over-ripened (whether on the vine or post-harvest); however, for levels of chalcone-naringenin in TA209 fruit, both changes were significant decreases ($p \leq 0.05$). For naringenin levels, decreases were observed in post-harvest fruit (6 dpb + 10 dph) compared with 16 dpb fruit for all four lines (significant for TA209, neo-123 and -111; $p \leq 0.05$). None of the lines showed levels of chalcone-naringenin or naringenin at 6 dpb + 10 dph that were significantly different from TA209 at 6 dpb + 10 dph.

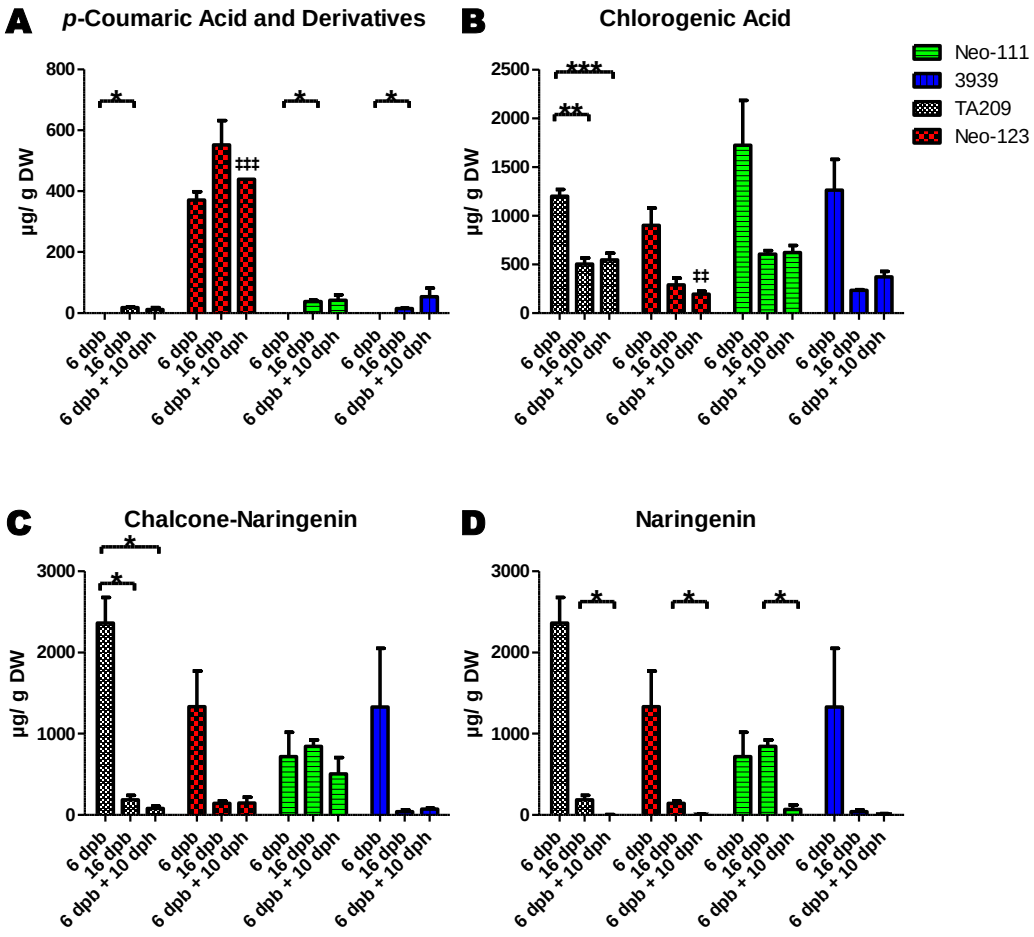

Figure 49 Comparison of phenolic profiles between post-harvest and on the vine fruit

Profile for phenolic compounds compare fruit at 6 days post breaker (dpb), at 16 dpb, and at 6 dpb then stored for 10 days post harvest (dph). Error bars represent SEM: 6 and 16 dpb, n=3; 6 dpb + 10 dph, n=6. Values significantly different are represented by one ($p \leq 0.05$), two ($p \leq 0.01$), and three ($p \leq 0.001$) symbols, where * show differences within one genotype, as indicated. Significant differences between 6 dpb + 10 dph for any line and 6 dpb + 10 dph for TA209 are indicated by ‡ symbol.

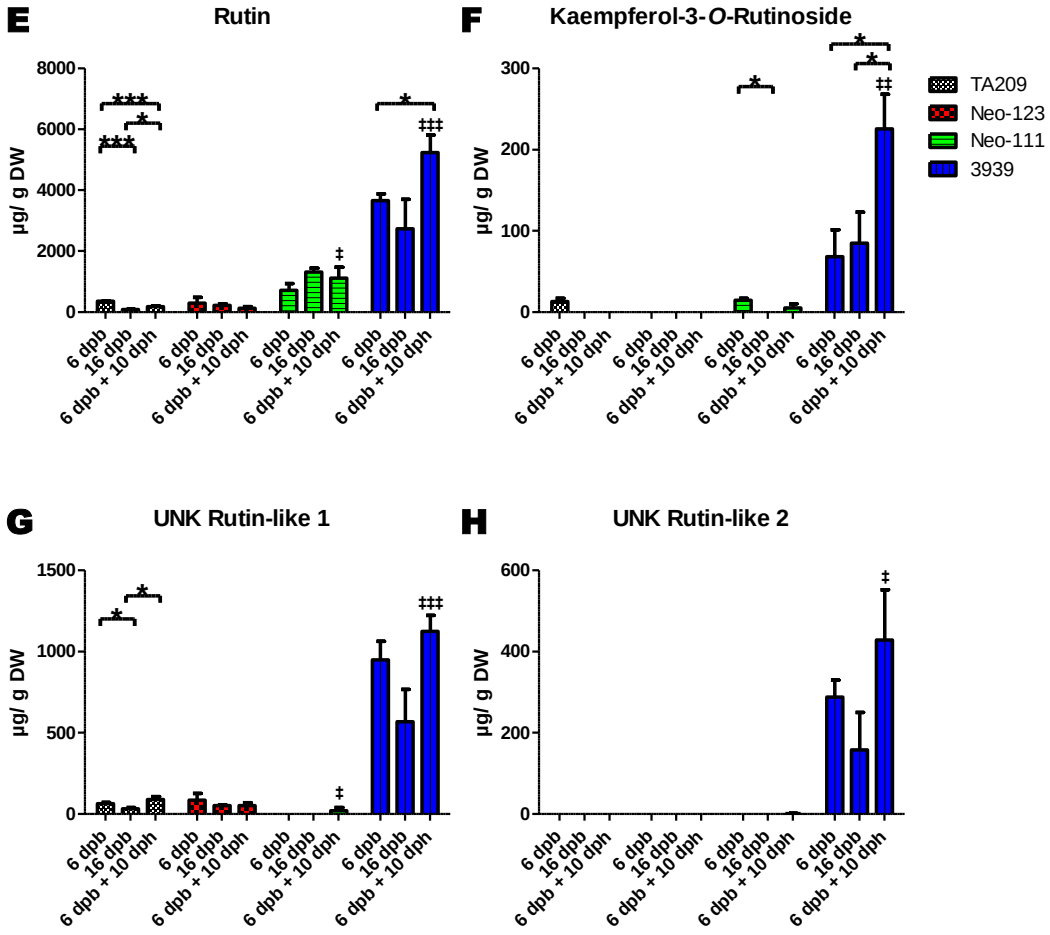

Figure 49 continued.

Levels of flavonol glycosides and related compounds were observed in greatest abundance in line 3939 at all three time points (Figure 49E to H). In many instances no significant differences were observed between post-harvest fruit (6 dpb + 10 dph) and either 6 or 16 dpb fruit. This was the case for rutin levels in neo-123 and -111 (Figure 49E), for kaempferol-3-*O*-rutinoside levels in all lines except 3939 (Figure 49F), for UNK rutin-like 1 levels in all lines except TA209 (Figure 49G), and for UNK rutin-like 2 levels in all lines (Figure 49H).

In contrast, some lines exhibited decreases or no change in levels of flavonol glycoside at 16 dpb compared with 6 dpb that were not reflected by equivalently low levels at 6 dpb + 10 dph. This was true firstly for rutin levels (Figure 49E) in line TA209 (significantly less at 16 dpb compared with 6 dpb, $p \leq 0.001$; significantly more

at 6 dpb + 10 dph compared with 16 dpb, $p \leq 0.05$) and in line 3939 (not significantly different except for more at 6 dpb + 10 dph compared with 6 dpb, $p \leq 0.05$); secondly for kaempferol-3-*O*-rutinoside levels (Figure 49F) in line 3939 (significantly more at 6 dpb + 10 dph compared with both 6 and 16 dpb, $p \leq 0.05$); thirdly for levels of UNK rutin-like 1 (Figure 49G) in line TA209 (significantly less at 16 dpb compared with 6 dpb, $p \leq 0.05$; significantly more at 6 dpb + 10 dph, compared with 16 dpb, $p \leq 0.05$) and in line 3939 (no significance); and lastly for levels of UNK rutin-like 2 (Figure 49H) in line 3939 (no significance).

The high *p*-coumaric acid and derivative levels of neo-123 and the high rutin levels of neo-111 and 3939 were maintained at significantly higher levels than TA209 following post-harvest storage (neo-123, $p \leq 0.001$, Figure 49A; neo-111, $p \leq 0.05$, Figure 49E; 3939, $p \leq 0.001$, Figure 49E).

Post-harvest storage additionally had no detrimental effects on antioxidant activity, as measured by TEAC assay. Both the significantly higher antioxidant activities seen in neo-123 at 6 dpb ($p \leq 0.01$) and in neo-111 at 16 dpb ($p \leq 0.05$), were observed at 6 dpb + 10 dph ($p \leq 0.001$ and $p \leq 0.05$, respectively; Figure 50). While three lines increased significantly in antioxidant activity between 6 and 16 dpb (TA209, $p \leq 0.01$; neo-111, $p \leq 0.001$; 3939, $p \leq 0.05$), all four lines showed a significant increase in antioxidant activity at 6 dpb + 10 dph compared with 6 dpb (TA209, $p \leq 0.01$; neo-123, $p \leq 0.001$; neo-111, $p \leq 0.001$; 3939, $p \leq 0.05$). Furthermore, no significant difference was shown between any lines at 16 dpb and at 6 dpb + 10 dph.

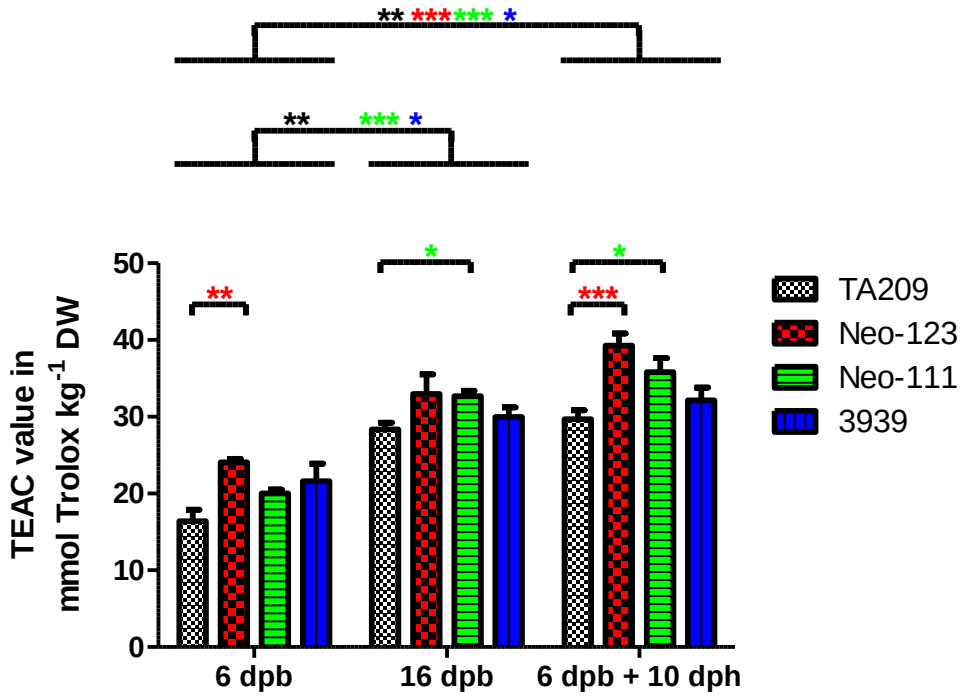

Figure 50 Effect of post-harvest storage on antioxidant activity

Comparison of antioxidant activity, expressed as Trolox equivalent antioxidant capacity (TEAC), for fruit at 6 and 16 days post breaker (dpb), and compared with fruit harvested at 6 dpb and stored for 10 days post harvest (dph). Error bars represent SEM: 6 and 16 dpb, n=3; 6 dpb + 10 dph, n=6. Significantly different values are represented by * ($p \leq 0.05$), ** ($p \leq 0.01$), and *** ($p \leq 0.001$), and are shown for differences between genotypes at any one stage and for differences for any one genotype between stages. No significant differences were observed for any genotype between 16 dpb and 6 dpb + 10 dph.

5.8 Discussion

The regulation of a complex trait such as the accumulation of intermediates within the phenolic biosynthetic pathway is complex (Quattrocchio et al., 2008). While previous studies have shown that alterations to relatively simple rate-limiting biosynthetic steps can improve accumulation of pathway intermediates (Muir et al., 2001), success has also been achieved by alterations to multiple regulatory units (Bovy et al., 2002; Butelli et al., 2008b; Luo et al., 2008; Schijlen et al., 2006) and therein lies the greater potential for wide-spread pathway manipulation. It has before been speculated that perturbations to phenolic profiles in mapping population accessions are the result of multiple QTL rather than single QTL or genes (Rousseaux et al., 2005). It is also well established that accessions within mapping populations may contain multiple QTL simultaneously affecting more than one trait (Paran et al., 1997). In this chapter some of these traits were characterised with the aim of better understanding the mechanism by which *S. neorickii* and *S. habrochaites* introgressed regions affect neo-111, -123 and 3939.

5.8.1 Characterisation of metabolism in high rutin lines neo-111 and 3939

5.8.1.1 Phenolic compounds

The accumulation of phenolic compounds throughout fruit development and ripening (Figure 25, Figure 35) clearly indicated that there is a differential regulatory control of phenolic biosynthesis in each of the high rutin lines. Phenylpropanoids and the early flavonoids naringenin and chalcone-naringenin showed few perturbations in 3939 throughout fruit development and ripening compared with TA209 (Figure 25A to F, Figure 35A to D). Those changes that were observed, although significant, were either comparatively moderate in abundance (such as *p*-coumaric acid at turning stage in Figure 25A and at 9 and 12 dpb in Figure 35A) or non-reproducible (such as significant decrease in chlorogenic acid at breaker stage in Figure 25D, compared with no change at 0 dpb in Figure 35B) and can therefore be explained by non-reproducible environmental effects (Rousseaux et al., 2005). Neo-111 by comparison displayed significant and reproducible increases in the accumulation of phenylpropanoids and flavonoids throughout fruit development and ripening (Figure 25B, D and F; Figure 35A to D), such as chlorogenic acid. This increased accumulation of chlorogenic acid throughout fruit development (Figure 25D, Figure 35B) is contrary to endogenous

tomato levels, which are expected to decline gradually throughout fruit development (Buta and Spaulding, 1997), as shown by TA209. The effect on flavonol glycoside accumulation was also shown to be different in each of the high rutin lines. Neo-111 exhibited elevated levels of rutin and sometimes K3OR throughout fruit development and ripening (Figure 25G and H, Figure 35E and F) and no detection of UNK rutin-like compounds; whereas, 3939 showed no deviation from TA209 levels during fruit development (MG, Br, and 14 dpa to 3 dpb) but after which time levels of rutin, K3OR and UNK rutin-like compounds increased beyond that of neo-111 levels (compare average rutin levels in ripe fruit of 7.6 fold increase in 3939 and 3.4 fold increase in neo-111; Figure 25G). Again, these profiles are contrary to expected trends of previously published phenolic accumulation (Buta and Spaulding, 1997).

The changes to phenolic profiles seen in whole ripe fruit exhibited a greater effect on phenolic profiles in isolated skin tissue than either flesh or jelly. Neo-111 whole ripe fruit displayed increases in chlorogenic acid (2.2 fold), naringenin (3.1 fold) and chalcone-naringenin (1.4 fold) levels (Figure 25). These changes were either non-significant or moderate in comparison with the far greater magnitudes detected in isolated skin tissue (chlorogenic acid 4.3 fold, naringenin 11.5 fold and chalcone-naringenin 5.7 fold; Figure 26). Likewise, increases in flavonol glycosides observed previously in ripe whole fruit of both neo-111 and 3939 (Figure 25, Figure 35) were detected at greater levels of fold change in isolated skin tissue, for example up to 19.7 fold rutin and 18.2 fold K3OR in 3939 (Figure 26). This is in accordance with previously published observations. Flavonoids, such as rutin, are known to accumulate predominantly in skin tissue (Navarro-Gonzalez et al., 2011; Torres et al., 2005); however, mass of skin tissue relative to flesh is low (5 % of whole fruit mass (Bovy et al., 2002)), and skin:fruit mass ratio depends on water content or fruit size (Levin, 2008). Therefore, by assessing compound abundance in dry weight mass (μg/ g DW), and additionally in isolated skin tissue samples, a more accurate measure of the effect on phenolic biosynthesis was ascertained in Figure 26.

Although flavonols usually accumulate in skin tissue (Torres et al., 2005), both neo-111 and 3939 also displayed flavonol glycoside increases in flesh and jelly tissue (Figure 26). Fold increases of rutin in flesh tissue (15.2 fold in 3939; 8.6 fold in neo-111) were comparable in magnitude to increases in skin tissue (19.6 fold in 3939; 8.8 fold in neo-111). An increase in flavonol glycoside content in fruit flesh is not guaranteed when flavonol glycoside content is increased in the skin. For example,

overexpression of a single biosynthetic step by Muir and colleagues (2001) increased flavonols in the skin by 78 fold with no increase in the flesh. Simultaneous increase in skin and flesh flavonol glycosides have been shown by manipulations of TFs (Bovy et al., 2002; Luo et al., 2008) and wild relative hybridisation (Willits et al., 2005). This supports the hypothesis that the QTL affecting flavonol glycoside accumulation in both neo-111 and 3939 is more likely a regulatory element than a biosynthetic step.

These data also indicate that the regulatory mechanism within neo-111 had more widespread influence on phenolic biosynthesis, the effects of which were seen throughout fruit development and ripening. The perturbation to regulation in 3939, however, was both targeted to flavonol biosynthesis specifically, and time specific for fruit ripening.

When the profile of phenolic compounds was extended to other plant tissue types, it was found that there were no significant differences in phenolics in leaf material (Figure 29). Plants are known to accumulate phenolics in tissues including the leaf when under biotic or abiotic stress (Lovdal et al., 2010; Mellway et al., 2009). The fact that no accumulation of phenolics was observed in leaf material supports the hypothesis that the changes to phenolic profiles seen here in fruit are due to genetic influence (introgressed regions from wild relatives), rather than as a result of plant stress. The effects on phenolic biosynthesis, however, were not restricted to fruit tissue, since both phenylpropanoids and flavonoid levels were disrupted in flower tissue (Figure 28).

5.8.1.2 Effects on the metabolome

The targeted nature of perturbations in 3939 was also illustrated by the isoprenoid and isoprenoid related compound profiles during fruit development shown in Figure 30 and Figure 31. Virtually all levels of isoprenoids, tocopherols and chlorophylls remained at TA209 levels throughout fruit development and ripening. Neo-111, however, displayed reduced levels of early biosynthetic carotenoids (Figure 30A, B, C and E), but elevated levels of cyclic carotenes and xanthophylls between Br and R stages (Figure 30G to J), as well as perturbations to chlorophylls (Figure 30C and D). This might imply that neo-111 contains one or several QTL with *S. neorickii* alleles that are capable of manipulating two independent pathways, the phenolic and the isoprenoid. Simultaneous manipulation of phenolic and isoprenoid pathways by one

single gene has been observed by overexpression of *CRY2* photoreceptor (Giliberto et al., 2005) and by down regulation of *DET1* (Enfissi et al., 2010) in tomato fruit. However, it is arguably more likely that these observed changed are due to the effects of multiple *S. neorickii* QTL alleles present in neo-111, especially when considering that it has been shown to possess multiple introgressed regions from *S. neorickii* according to RFLP and COSII marker data (EU Sol consortium, personal communication). Therefore, the observed changes in isoprenoid levels could be controlled by a *S. neorickii* allele at a different QTL to that which affects observed changes in phenolic levels. Simultaneous manipulation of independent pathways such as this has been reported previously in corn (Naqvi et al., 2009). In support of the hyposthesis that neo-111 may possess a *S. neorickii* QTL allele for isoprenoid accumulation is data from Enfissi (unpublished), where only 14 accessions within the *S. neorickii* BIL population were shown to possess reproducible isoprenoid accumulation, and therefore these accessions were assumed to contain *S. neorickii* QTL alleles responsible for these changes.

The clusters in Figure 32A, which group MG and Br samples away from T and R samples, reflect the rapid change in fruit metabolism between Br and T stages. Comparison with Figure 33 illustrates this rapidity (up to approximately three days), and possibly explains the outlying data point in Figure 32A (TA209 Br) as sampling error within this narrow harvest window of time.

Both neo-111 and 3939 showed no difference in metabolism according to PCA clustering of GC-MS non-targeted metabolites at MG and R stages (Figure 32C and E). Conversely, these same PCA showed that metabolism in both high rutin lines were differentially perturbed compared with TA209 at both Br and T stages. In both high rutin lines decreases in sugars were detected (Table 9), that include fructose, glucose, maltose and unknown sugars (s1, s5, s7, s16). A decrease in sterols was also detected in both neo-111 and 3939, at Br and T. These similarities in metabolomic fluctuations, unlike the previously observed differences in phenolic, isoprenoid and chlorophyll accumulations, indicated parallels in the regulation of neo-111 and 3939 for the first time, which are clearly demonstrated by Table 9 (for examples see s12, s7, oleic acid, amino acids at MG, organic acids at T). The comparison between neo-111 and 3939, however, may be the causation of these seemingly coordinated changes. The PCA clustering of 3939 (Figure 32C and E) coincides with the accumulation of flavonol glycosides (Figure 25, T stage; Figure 35, 3 to 6 dpb). Neo-111, however, did not

exhibit such an accumulation, since levels remain abundant throughout development and ripening. This does coincide with accumulation of carotenes and xanthophylls in neo-111, at Br to R stage (Figure 30G, I and J). The elevation of secondary metabolite synthesis in both 3939 and neo-111 (albeit for synthesis in independent pathways) may have had similar flux effects on central carbon metabolism, resulting in the similar parallel effects on metabolomic profiles in Figure 32 and Table 9.

5.8.2 Characterisation of metabolism in high p-coumaric acid line neo-123

5.8.2.1 Phenolic compounds

Neo-123 exhibited reproducible, high levels of *p*-coumaric acid and derivatives throughout fruit ripening that increased between breaker and ripe stages (Figure 25A and B, Figure 35A). During these same stages, levels of other phenylpropanoids and of flavonoids were sometimes shown to be significantly lower than TA209 (Figure 25C, D, G and I, Figure 35B, C, E and F). Flavonol glycoside levels were additionally reduced compared with TA209 at MG stage (Figure 25G to I), but this was not reproducible when analysed on a developmental and ripening series (Figure 35E to G). No other changes in phenolic compounds were reported at MG or early in fruit development. The most significant changes in *p*-coumaric acid and derivative, and in chlorogenic acid were seen in all tissue types (Figure 26), but less so in flower tissue than the high rutin lines (Figure 28). These data suggest that the effects on regulation of the phenolic pathway in neo-123 might be the result of a bottleneck in the flux beyond *p*-coumaric acid, such as an increase to the flux coefficients of 4CL or HCT, that prevents through-flow into either chlorogenic acid or flavonoid biosynthesis and therefore results in *p*-coumaric acid accumulation. Targeted overexpression or down regulation of key phenylpropanoid biosynthetic steps such as HQT or C4H in tomato has resulted in desired accumulation or deficiency of expected phenylpropanoid or flavonoid targets (Cle et al., 2008; Millar et al., 2007; Niggeweg et al., 2004). However, unexpected or non-correlative wider impacts within different branches of phenolic biosynthesis suggest plasticity of flux and more complex regulatory cross-talk throughout the phenolic network (Cle et al., 2008; Millar et al., 2007; Tanaka and Ohmiya, 2008) . It is likely, therefore, that the changes observed in neo-123, likewise to neo-111 and 3939, are the result of regulatory control, such as TF regulation.

5.8.2.2 Effects on the metabolome

On the one hand, early isoprenoid and chlorophyll *a* accumulation in neo-123 was unaffected throughout fruit development and ripening compared with levels in TA209 (Figure 30A to F, Figure 31C). On the other hand, carotene, xanthophylls, tocopherols, chlorophyll *b*, and ubiquinone were markedly increased throughout Br to R stages (Figure 30G to J, Figure 31A, B, D, F) to levels far exceeding changes observed in either of the high rutin lines. For reasons discussed previously in section 5.8.1.2 regarding changes in neo-111, this may similarly have been due to separate genetic factors within *S. neorickii* introgressed regions affecting these pathways independently. The carotenoid profiles in Figure 30 may be caused by similar mutations to those seen in *Delta* or *High Beta* carotenoid mutant tomato lines (Ronen et al., 1999; Ronen et al., 2000). Interestingly, neo-123 (as well as neo-111 mentioned previously) was one of only 14 accessions within the *S. neorickii* BIL population that was shown to possess reproducible isoprenoid accumulation, and therefore these accessions were hypothesised to contain *S. neorickii* QTL alleles responsible for manipulation of isoprenoids levels (Enfissi, unpublished).

The effect on fruit metabolome according to metabolites detected by GC-MS and represented by PCA appears to have had opposite results to those seen in either of the high rutin lines (Figure 32C and E). As previously shown, 3939 and neo-111 displayed no clustering by PCA at MG and R stages, but formed clusters distinct from TA209 at Br and T stages. Neo-123, however, was the only genotype to show distinct clusters at MG (Figure 32C) and R (Figure 32E) stages, and indicated no difference from TA209 at Br and T stages. The most distinct changes relative to TA209 at MG stage in neo-123 were to amino acids such as isoleucine, leucine and proline (Table 9), which showed increases of at least 2.3 fold, compared with the decreases seen in neo-111 and 3939. Metabolites contributing to separation of neo-123 from TA209 and high rutin lines at R stage in Figure 32E included a relative decrease in many sugar compounds such as the monosaccharides (for example, glucose, fructose and msacch2 to 6), most unknown sugars (s1 to 7), and some disaccharides (sucrose); and a relative increase in many fatty acid and triacylglyceride intermediates (for example oleic, 9-octadecenoic acid; stearic1G, octadecanoic acid propyl ester; and stearic2G, octadecanoic acid methyl ethyl ester) (Table 9). Variables from glycolysis and tricarboxylic acid cycle were unchanged in neo-123 at ripe stage, despite showing a decrease in intermediate compound abundance for both 3939 and neo-111. The cause

of this decrease seen previously in 3939 and neo-111 was hypothesised in section 5.8.1.2 to be due to the increased carbon source required for elevated levels of multiple phenolics and isoprenoids. It is hypothesised here, that an increased carbon influx is not a requirement of the phenolic metabolic profile observed in neo-123 since the accumulation is at the expense of accumulations in intermediates synthesised later in the pathways. This bottleneck theory is supported by profiles of neo-123 seen in Crop 1 and 2 previously (see section 3.2.3, Figure 16).

5.8.3 Characterisation of antioxidants

Polar extracts, containing phenolic compounds, were analysed for antioxidant activity using the TEAC assay (Re et al., 1999). No difference in antioxidant activity was observed in any genotype throughout fruit development; however, increased antioxidant activities were seen at various time points between 3 and 16 dpb (Figure 36) from whole fruit extracts, which corresponded with increased accumulation of phenolic compounds (Figure 35). This association is especially clear for flavonol glycosides in 3939 and *p*-coumaric acid and derivatives in neo-123, but less obvious for neo-111 since phenolic levels are high throughout fruit development and ripening (Figure 25, Figure 35).

In general, the greatest increases in antioxidant capacity were observed in tissue types exhibiting the largest increases in phenolic compound accumulation. Significantly higher antioxidant activity was seen in flesh and jelly tissue of neo-123 (Figure 37), where the highest change to phenolic profile was increases in *p*-coumaric acid and derivatives (Figure 26). The only flower extracts to show increased antioxidant activity was from 3939 (Figure 38), which exhibited the greatest fold increase of any compound in flower tissue (chlorogenic acid; Figure 28). In both high rutin lines, significantly higher antioxidant capacity was seen in skin tissue, but in neither jelly nor flesh (Figure 37). This corresponds to the location where the highest levels of flavonol glycosides were also observed (Figure 26). Neo-111, however, exhibited greater antioxidant capacity levels in skin than 3939 (Figure 37). This is surprising, considering 3939 exhibited higher levels of rutin, and far greater levels of flavonol glycosides in combination (inclusive of Figure 26E to H). It is possible, therefore, that elevated levels of other flavonoid compounds in skin tissue of neo-111 were contributing to overall antioxidant capacity, either additively or synergistically.

Chlorogenic acid, naringenin, and chalcone-naringenin are known antioxidants (Gonzalez and Nazareno, 2011; Spencer et al., 2005), therefore increased levels of these could be contributing to this effect. This hypothesis would also explain the high antioxidant capacity in neo-111 at 3 and 16 dpb (Figure 36) where no change in flavonol glycoside was observed (Figure 35).

Some reports in the literature argue that the presence of other hydrophilic compounds in phenolic extracts may contribute to total TEAC values (Willcox et al., 2003). For example, ascorbic acid is a known antioxidant and present in tomato at approximate levels of 4 mg/ 100 g FW (Willcox et al., 2003). Contradictory studies, however, have indicated that the contribution of ascorbic acid from tomato to total antioxidant activity can vary 1 to 51 % depending on season of cultivation (Rousseaux et al., 2005). Furthermore, tomatoes with the highest levels of phenolics have been shown to correlate with the highest antioxidant activities (Minoggio et al., 2003); rutin has been shown to be a more powerful free-radical scavenger than ascorbic acid, naringenin or chalcone-naringenin (Gonzalez and Nazareno, 2011); rutin levels have been shown to correlate greater to antioxidant capacity in tomato than either chlorogenic acid or chalcone-naringenin (Spencer et al., 2005); and both synergistic and antagonistic combinatory effects of phenolics on antioxidant capacity have been shown (Gonzalez and Nazareno, 2011).

It is likely therefore, that the highly abundant flavonol glycosides and *p*-coumaric acid contribute predominantly to antioxidant capacity from the phenolic extracts, but that other compounds contribute to total activity, especially in the case of neo-111. Any QTL identified for increased accumulation of these compounds therefore, will become a target for improvement of health-promoting traits.

5.8.4 Characterisation of post-harvest properties

Multiple parameters were assessed on post-harvest fruit (6 dpb + 10 dph) and compared with fruit at time of harvest (6 dpb) and equivalent aged fruit allowed to remain on the vine of the plant (16 dpb). Post-harvest fruit exhibited no detrimental effects to abundance of most phenolic compounds (Figure 49). Levels of *p*-coumaric acid and derivatives in neo-123 and levels of flavonol glycosides in neo-111 displayed no significant reduction during post-harvest storage. Further to this, 3939 was shown to

accumulate elevated levels of rutin and K3OR during post-harvest storage (Figure 49E and F). Phenolics have previously been shown to be increased in tomato fruit during post-harvest storage (Toor and Savage, 2006), believed to be a result of cell stress. This may have been caused by the storage temperatures. Antioxidant capacity for all three lines and TA209 was shown to increase both during post-harvest storage and when fruit were allowed to remain on the vine of the plant (Figure 50), which is consistent with maintaining high levels of *p*-coumaric acid and flavonol glycosides. Toor and Savage (2006) similarly showed that antioxidant activity increased during post-harvest storage of tomato, but that this was due in part to increasing ascorbic acid levels which may contribute to antioxidant activity (Willcox et al., 2003) as well as phenolic compounds.

Compared to these biochemical changes, fruit physiological parameters showed greater effects as a result of post-harvest storage. All fruit exhibited loss of mass at a steady rate post-harvest (Figure 44), and neo-111 showed the greatest loss. This rate in loss of mass is in accordance with previously published results (Jha and Matsuoka, 2005). Loss of mass was assumed to be water loss despite no detected significant loss of water detected in Figure 47. Both *S. neorickii* BILs and TA209 suffered greater effects to fruit firmness during post-harvest storage compared with fruit left on the vine of the plant (Figure 48), which is a well documented effect of post-harvest storage of tomato (de Castro et al., 2006; Jha and Matsuoka, 2005). Interestingly, 3939 better maintained firmness than TA209 during post-harvest storage. Fruit of 3939 appeared to exhibit no change in morphology during post-harvest storage; however, fruit of neo-111 showed low tolerance to post-harvest conditions in the form of wrinkling of epidermis (Figure 45). Although colour measurements were not recorded by quantitative means, it was observed that red pigmentation in neo-123 fruit increased during post-harvest storage. This is likely a result of increases in lycopene, which is known to accumulate post-harvest (Toor and Savage, 2006).

5.8.5 Characterisation of fruit physiology and ripening

Distinct differences in fruit physiology and ripening were observed in lines neo-111 and -123 compared with TA209. Both lines showed differences in colour throughout ripening (Figure 33), and ripe fruit were observed as more orange in colour than was TA209. This was likely due to the elevated carotenoid pigments detailed in Figure 30, which are known to contribute to tomato fruit colour (Lewinsohn et al.,

2005). Fruit from both lines were smaller in size (Figure 39, Figure 40C and D). Neo-123 fruit were also different in shape (Figure 39). Neo-111 displayed differences in fruit mass (Figure 40A and B) and seed number (Figure 41). Both lines showed no difference in seed mass (Figure 41), or fruit firmness throughout development and ripening (Figure 42) but showed differences in water content (Figure 43). Ripening development time to MG stage was more rapid in both lines; however, the ripening (measured in days until Br stage) was delayed in both lines as well as 3939.

3939 fruit, by comparison, showed no difference from TA209 in fruit size (Figure 40) or shape (Figure 39), seeds mass (Figure 41A) or number (Figure 41B), number of locules (Figure 33), fruit firmness (Figure 42), or colour throughout development and ripening (Figure 33). The only morphological difference between ripe fruit of 3939 and TA209 observed was the pointed blossom-end morphology shown in Figure 39D. This is unlikely to be related to any regulatory factors involved in phenolic biosynthesis; however, many genes in tomato have been identified that contribute to such morphology. One such gene, nipple tip (*n*), has been mapped to chromosome 5 (referenced by Barten et al., 1994). It is possible, therefore, that an allele such as *n* is located on the introgressed region of chromosome 5 from *S. habrochaites* wild relative and contributed to this difference in morphology seen in 3939. This example is illustrative of many of the morphological differences seen between both *S. neorickii* BILs and TA209. Introgressed wild relative chromosome regions have many times been shown to affect fruit morphology leading to the identification of QTL alleles for traits such as fruit weight (Prudent et al., 2009; Tanksley et al., 1982), size (Paterson et al., 1991), and shape (Ku et al., 1999); seed weight (Tanksley et al., 1982), parthenocarpy (Gorguet et al., 2008), and locule number (Barrero and Tanksley, 2004); and fruit ripening time (Lindhout et al., 1994). Due to the greater number of introgressed regions in both neo-111 and -123 compared with the single theoretical region in 3939, it is unsurprising that both neo-111 and -123 displayed greater morphological differences if they can be assumed to possess a greater number of unrelated wild QTL alleles from *S. neorickii*.

6 Transcriptomic analysis and candidate QTL identification

6.1 Introduction

In previous chapters, fruit from the *S. neorickii* BIL population (Fulton et al., 2000; Grandillo et al., 2011) was screened for perturbations to profiles of phenolic compounds. Neo-111 and -123 were hypothesised to contain *S. neorickii* alleles for QTL in chromosomes 5 and 10 that contribute to high rutin and *p*-coumaric acid accumulation, respectively. Comparisons were made between neo-111 and genotype 3939 from *S. habrochaites* NIL population (Monforte and Tanksley, 2000) due to its single introgressed region in chromosome 5 and high rutin accumulation. The phenolic profiles of these three genotypes were characterised throughout fruit development and within tissue types, and the effects on wider metabolism were identified. It was hypothesised that the observed changes were due to QTL for regulatory elements rather than biosynthetic genes due to the nature of the perturbations.

In this chapter data are presented from a newly developed platform for identifying changes in relative expression levels of known TFs (Rohrmann et al., 2011). This TF platform was chosen over otherwise available microarray platforms due to greater coverage and more recent inclusion of known TF regulatory elements. From these data, candidate TF promoter regions were sequenced for polymorphisms. Differences in sequence between parental genotypes are discussed for their potential role as regulatory elements for perturbations to phenolic profiles.

6.2 Analysis of relative expression levels of known transcription factors

Fruit material for lines TA209, neo-123, neo-111 and 3939 at turning stage were analysed from identical aliquots to those used previously in chaper 5. A transcription factor platform (Rohrmann et al., 2011) used qRT-PCR to assess levels of 1077 known transcription factor genes in neo-123, neo-111 and 3939 relative to TA209.

In total, 186 out of 1077 TFs showed significantly ($p \leq 0.05$) altered expression in at least one of the three lines when compared with TA209 (Figure 51). These TFs were classified across a range of TF families, including 24 TFs from MYB or MYB-related family, ten TFs from MADS-box family and two TFs from LIM family (Figure 51, Table 10). A greater number of TFs were affected in each of the *S. neorickii* BILs than in the *S. habrochaites* NIL. The number of TFs that showed significantly altered expression in neo-123 and -111 were 98 and 94, respectively, and 23 of these TFs showed altered expression in both lines. This is compared with only 29 TFs in total for line 3939 where expression was significantly altered. A larger proportion of those TFs that displayed altered expression in neo-123 were relative decreases compared with TA209. Conversely, the majority of TFs that showed altered expression in either of the high-rutin lines, neo-111 and 3939, exhibited relative increases compared with TA209. This being said, many relative decreases were observed for TFs affected in both neo-123 together with one of the two high-rutin lines (Figure 51).

Five TFs were perturbed in all three lines, and this included one MYB TF on chromosome 5 (Table 10A). Nine TFs were altered in both high-rutin lines, neo-111 and 3939 (Table 10A and B). In addition to the MYB TF previously mentioned on chromosome 5, these included two further TFs located on chromosome 5 (one MYB-related family and one orphan showing similarity to a LIM domain), and two MADS-box family TFs. 16 of the 186 perturbed TFs were located on chromosome 5 (highlighted in yellow, Table 10), and ten of these showed altered expression in one or both high-rutin lines, but not neo-123. The ten TFs on chromosome 5 with altered relative expression in neo-111 were all located on the same region of chromosome 5, approximately 61,370,000 bp to 64,250,000 bp.

Nine of the 186 perturbed TFs were located on chromosome 10 (highlighted in green, Table 10), and the majority were C2H2 or C3H zinc finger family TFs. Eight of these nine TFs displayed altered expression in neo-123 (seven of which were uniquely perturbed in neo-123, but in neither neo-111 nor 3939; Table 10G). Seven of the eight TFs on chromosome 10 that exhibited altered expression in neo-123 were located within a tight chromosomal region, approximately 59,300,000 bp to 63,300,000 bp.

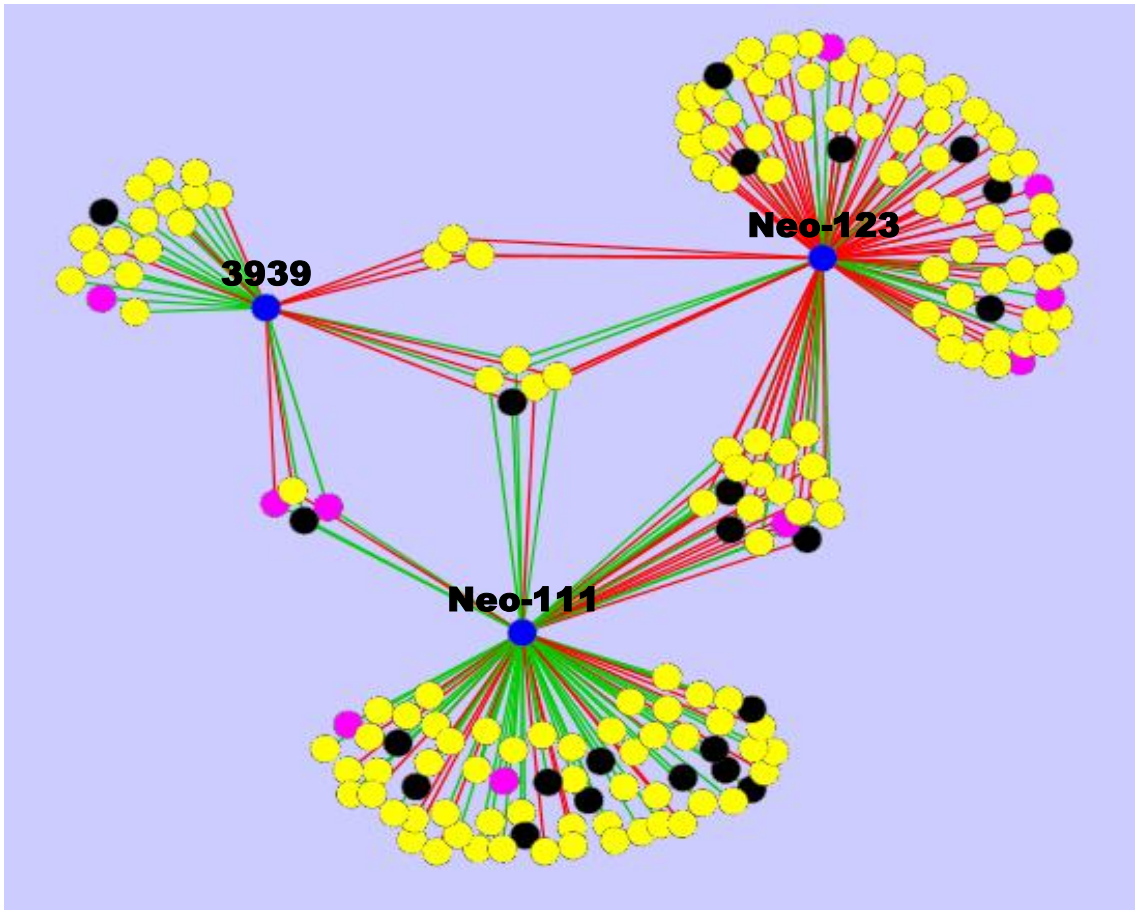

Figure 51 Frequency of differentially expressed transcription factors (TFs) at turning stage compared with TA209

Overview of TF genes (yellow, black and pink nodes) in lines neo-123, neo-111 and 3939 (blue nodes, as indicated) that exhibited significantly ($p \leq 0.05$) increased (green connector) or decreased (red connector) fold changes in turning stage fruit relative to TA209 levels (n=3). Black node specifies MYB and MYB-related family TFs; pink node specifies MADS-box family TFs; yellow node indicates all other TFs. Connector length not shown to scale.

Table 10 Transcription factor genes shown to be differentially expressed in selected lines at turning stage compared with TA209

Annotations for 186 TF genes with significantly ($p \leq 0.05$) altered expression at turning stage in at least one line relative to TA209. ([a]) represents TF annotations provided by J. Rohrmann (MPI-MP) for the TF platform (Rohrmann et al., 2011) and include gene of interest (GoI) identification, assigned TF family, and Plant tf DB accession number, found at ([b]) http://planttfdb.cbi.pku.edu.cn:9010/web/index.php?sp=le. Sequence from Plant tf DB was used in a BLAST search to generate ([c]) Sol Genomics Network (SGN, http://solgenomics.net/) unigene number, best BLAST match annotation, and ITAG2 location and accession number. From the unigene number ([d]) chromosome number was provided and position on tomato WSG chromosomes SL2.31 was detailed in base pairs (bp). Fold changes are represented as values > 1 in green indicating a fold increases and < -1 in red representing a fold decrease. Chromosomes 5 and 10 are highlighted yellow and green where applicable. Table 10A to G group TFs based on whether they were perturbed in one, two or three lines.

Table 10A TFs where relative expression was altered in all lines

		Fold Change			Plant tf DB	SGN unigene	ITAG2 acc. No.[c]	Chromosome Position[d]			
GoI[a]	TF-family[a]	Neo123	Neo111	3939	acc. No.[ab]	No.[c]		No.	Start (bp)	End (bp)	Best BLAST match annotation[c]
TA43657_4081	bZIP	-6.17	133.68	-6.01	PTLe00964.1	SGN-U564727	**Solyc04g081190**	4	62,783,956	62,785,022	*vsf-1*
TA50531_4081	C2C2-YABBY	-6.10	-4.91	-4.90	PTLe00236.1	SGN-U583545	**Solyc08g079100**	8	59,908,584	59,908,284	Protein YABBY 4 - Oryza sativa subsp. japonica (Rice)
TA49294_4081	MYB	-1.47	1.56	-1.18	PTLe00620.1	SGN-U564833	**Solyc05g052850**	5	62,183,794	62,182,759	AtMYB109, MYB109 | MYB109 (myb domain protein 109)
AI897649	NAC	3.33	22.48	16.02	PTLe00683.1	SGN-U586243	**Solyc02g088180**	2	44,916,125	44,916,816	no apical meristem (NAM) family protein
TA42160_4081	CCAAT	13.13	22.67	32.93	no hit found						

Table 10B TFs where relative expression was altered in both high-rutin lines neo-111 and 3939

		Fold Change			Plant tf DB	SGN unigene	ITAG2 acc. No.[c]	Chromosome Position[d]			
GoI[a]	TF-family[a]	Neo123	Neo111	3939	acc. No.[ab]	No.[c]		No.	Start (bp)	End (bp)	Best BLAST match annotation[c]
TA35979_4081	Orphans		-3.85	-4.11	PTLe00502.1	SGN-U575402	**Solyc05g052780**	5	62,111,300	62,111,760	similar to LIM domain-containing protein
TA39847_4081	MYB-related		3.58	1.54	no hit found	SGN-U582090	**Solyc05g054410**	5	63,485,076	63,484,526	telomere binding protein TBP1 [Nicotiana glutinosa]; MYB-like
TA52882_4081	MADS		4.52	-3.44	PTLe00561.1	SGN-U577229	**Solyc04g050380**	4	58,795,458	58,795,662	MADS16 [Solanum tuberosum], short vegetative phase
BF098196	MADS		6.53	9.49	PTLe00522.1	SGN-U572195	**Solyc12g056460**	12	47,763,257	47,763,503	MADS-box protein SOC1

Table 10C TFs where relative expression was altered in both *S. neorickii* BILs neo-111 and neo-123

GoI[a]	TF-family[a]	Fold Change Neo123	Fold Change Neo111	Fold Change 3939	Plant tf DB acc. No.[ab]	SGN unigene No.[c]	ITAG2 acc. No.[c]	Chromosome Position[d] No.	Chromosome Position[d] Start (bp)	Chromosome Position[d] End (bp)	Best BLAST match annotation[c]
TA43709_4081	MYB-related	**-78.96**	**-19.59**		PTLe00597.1	SGN-U578744	**Solyc12g008800**	12	2,143,983	2,144,362	Myb-like DNA-binding domain
BG123356	AUX/IAA	**-33.54**	**-43.30**		PTLe00153.1	SGN-U577682	**Solyc03g120500**	3	62,933,007	62,932,575	Auxin-responsive protein IAA27 - Arabidopsis thaliana
DB724399	MYB-related	**-31.65**	**-71.35**		PTLe00592.1	SGN-U583592	**Solyc02g064630**	2	30,301,738	30,302,086	MYBR6 [Malus x domestica]
DB701472	ABI3VP1	**-4.98**	**-5.37**		PTLe00012.1	SGN-U570617	**Solyc02g055370**	2	26,262,292	26,261,837	NP_194892.1 TF B3 family protein [Arabidopsis thaliana]
TA47976_4081	NAC	**-4.27**	**2.81**		PTLe00689.1	SGN-U586529			84,979,694	84,979,327	NAC domain-containing protein 18 - Arabidopsis thaliana
TA41646_4081	LIM	**-4.03**	**-33.05**		PTLe00505.1	SGN-U577047	**Solyc04g077780**	4	60,273,140	60,272,640	TF LIM - Nicotiana tabacum
EG553398	Orphans	**-3.68**	**-27.27**		PTLe00505.1	SGN-U577047	**Solyc04g077780**	4	60,273,140	60,272,640	transcription factor LIM
AI896489	G2-like	**-3.17**	**3.64**		PTLe00378.1	SGN-U603194	**Solyc06g061030**	6	35,460,800	35,460,963	Two-component response regulator-like - Arabidopsis
TA51320_4081	HMG	**-3.12**	**5.91**		PTLe00472.1	SGN-U587413	**Solyc08g082070**	8	62,139,069	62,139,398	high mobility group (HMG1/2) family protein
DB700269	C2H2	**-2.64**	**-2.06**		PTLe00279.1	SGN-U581348	**Solyc10g084180**	10	63,158,962	63,158,028	zinc finger (C2H2 type) family protein
TA47799_4081	MYB	**-2.38**	**-2.66**		PTLe00646.1	SGN-U584090	Solyc05g052610	5	61,964,071	61,964,513	L. esculentum coronatine-insensitive 1 (Coi1); MYB75
TA45987_4081	MADS	**-1.56**	**-1.70**		PTLe00539.1	SGN-U577167	**Solyc03g006830**	3	1,343,053	1,343,733	MADS-box transcription factor
DB716243	Jumonji	**1.86**	**-1.60**		no hit found						
TA50859_4081	SRS	**1.92**	**3.69**		PTLe00763.1	SGN-U586059	**Solyc11g064800**	11	47,080,717	47,080,330	hypothetical protein [Vitis vinifera]
TA44759_4081	Orphans	**1.97**	**3.48**		PTLe00188.1	SGN-U577301	**Solyc12g096500**	12	63,736,627	63,737,419	CONSTANS-like protein; zinc finger (B-box) family protein
TA49012_4081	PHD	**3.24**	**4.30**		PTLe00709.1	SGN-U577885	**Solyc03g097600**	3	63,986,072	63,986,645	PHD finger family protein
TA48227_4081	AP2-EREBP	**3.69**	**7.74**		PTLe00069.1	SGN-U577093	Solyc04g071770	4	56,339,213	56,339,745	Ethylene-responsive TF ABA REPRESSOR 1 - Arabidopsis
TA52198_4081	HB	**3.83**	**4.19**		no hit found						

Table 10D TFs where relative expression was altered in lines neo-123 and 3939

GoI[a]	TF-family[a]	Fold Change Neo123	Fold Change Neo111	Fold Change 3939	Plant tf DB acc. No.[ab]	SGN unigene No.[c]	ITAG2 acc. No.[c]	Chromosome Position[d] No.	Chromosome Position[d] Start (bp)	Chromosome Position[d] End (bp)	Best BLAST match annotation[c]
TA41497_4081	C2H2	**-5.88**		**-4.04**	PTLe00252.1	SGN-U577297	**Solyc06g062670**	6	35,943,648	35,942,979	Protein TRANSPARENT TESTA 1 - Arabidopsis thaliana
BP884065	ABI3VP1	**-3.19**		**-2.81**	PTLe00009.1	SGN-U572360	**Solyc01g108930**	1	87,824,994	87,825,350	transcriptional factor B3 family protein, contains Pfam
TA54849_4081	RWP-RK	**-2.91**		**-2.13**	PTLe00697.1	SGN-U565442	**Solyc04g082480**	4	63,714,645	63,715,222	Nin-like family, RWP-RK domain-containing protein

Table 10E TFs where relative expression was altered uniquely in line neo-111

GoI[a]	TF-family[a]	Fold Change Neo123	Fold Change Neo111	Fold Change 3939	Plant tf DB acc. No.[ab]	SGN unigene No.[c]	ITAG2 acc. No.[c]	Chromosome Position[d] No.	Chromosome Position[d] Start (bp)	Chromosome Position[d] End (bp)	Best BLAST match annotation[c]
TA44890_4081	bZIP		-696.17		PTLe00994.1	SGN-U565963	**Solyc04g078840**	4	61,096,187	61,095,461	bZIP, ripening-related
TA49955_4081	C3H		-10.60		no hit found						
TA47549_4081	BES1		-6.48		PTLe00186.1	SGN-U573386	**Solyc04g079980**	4	61,883,273	61,884,319	Arabidopsis thaliana BZR1 (BRASSINAZOLE-RESISTANT 1)
TA42990_4081	C2H2		-3.02		PTLe00271.1	SGN-U565583	**Solyc04g081370**	4	62,974,466	62,975,151	zinc finger (C2H2 type) family protein
TA56477_4081	ABI3VP1		-2.67		PTLe00004.1	SGN-U571435	**Solyc03g111500**	3	56,161,317	56,160,880	hypothetical protein
TA45247_4081	HSF		-2.67		PTLe00485.1	SGN-U573319	**Solyc03g026020**	3	7,812,546	7,813,427	Heat stress transcription factor B-2b - Arabidopsis thaliana
EG553758	LIM		-2.58		PTLe00502.1	SGN-U575402	**Solyc05g052780**	5	62,111,300	62,111,760	LIM domain protein WLIM2 [Nicotiana tabacum]
TA48958_4081	HB		-2.33		PTLe00442.1	SGN-U574959	**Solyc06g053220**	6	32,440,398	32,439,871	Homeobox-leucine zipper protein ATHB-12 - Arabidopsis
TA56528_4081	ABI3VP1		-2.21		PTLe00008.1	SGN-U574579	**Solyc03g082550**	3	56,058,661	56,058,176	nix bekannt /nothing known
TA40372_4081	ABI3VP1		-2.13		PTLe00025.1	SGN-U576732	Solyc02g090710	2	46,814,931	46,815,928	hypothetical protein
DB725360	C3H		-2.11		PTLe00297.1	SGN-U570264	**Solyc06g072720**	6	41,244,467	41,243,249	hypothetical protein
TA53559_4081	GRAS		-2.08		no hit found						
TA42903_4081	C2H2		-1.82		PTLe00267.1	Clone LE_HBa0088G08					
TA41694_4081	C2H2		-1.66		no hit found						
TA44358_4081	ABI3VP1		-1.60		PTLe00006.1	SGN-U570932	**Solyc04g064830**	4	55,121,884	55,122,291	NP_194897.1 TF B3 family protein [Arabidopsis thaliana]
TA38817_4081	AUX/IAA		-1.54		PTLe00141.1	SGN-U579749	**Solyc06g053840**	6	33,205,640	33,206,085	Auxin-responsive protein IAA4 - Arabidopsis thaliana
TA47844_4081	C2H2		-1.47		PTLe00245.1	SGN-U585104	**Solyc05g054650**	5	63,667,927	63,668,607	schwach /weak; zinc finger (C2H2 type) family protein
TA48873_4081	PHD		-1.40		PTLe00707.1	SGN-U566469	Solyc06g069360	6	39,497,780	39,498,219	PHD finger family protein, unnamed
TA38123_4081	CCAAT		1.24		PTLe00343.1	SGN-U565280	**Solyc01g079870**	1	71,516,933	71,516,245	Nuclear transcription factor Y subunit
TA42697_4081	MYB-related		1.28		PTLe00571.1	SGN-U585710	**Solyc01g079210**	1	70,815,241	70,815,741	myb family transcription factor
TA44843_4081	HSF		1.29		PTLe00491.1	SGN-U567833	**Solyc03g006000**	3	678,660	679,955	heat shock factor [Nicotiana tabacum]
DB695212	MADS		1.40		PTLe00550.1	SGN-U562777	**Solyc05g051830**	5	61,377,554	61,377,731	Weak homology with SGN Unigene: MADS-box family protein
TA43149_4081	MYB-related		1.46		PTLe00592.1	SGN-U583592	**Solyc02g064630**	2	30,301,738	30,302,086	TRB1 DNA-binding Arabidopsis; MYBR6 [Malus x domestica]
TA39163_4081	CCAAT		1.58		no hit found						
TA40671_4081	HSF		1.59		PTLe00494.1	SGN-U566892	**Solyc02g090820**	2	46,882,616	46,881,620	Heat shock factor protein
TA37697_4081	EIL		1.64		PTLe00360.1	SGN-U595854	Solyc06g073730	6	41,883,252	41,880,979	Protein ETHYLENE INSENSITIVE [Solanum lycopersicum]
TA45957_4081	NAC		1.66		PTLe00682.1	SGN-U583456	**Solyc11g008010**	11	2,211,630	2,210,889	nam-like protein 4
TA42758_4081	MYB-related		1.67		no hit found						putative SWI/SNF complex subunit SW13 [Arabidopsis
DB720178	NAC		1.72		PTLe00667.1	SGN-U584554	**Solyc05g041920**	5	64,255,354	64,255,003	nam-like protein 9
TA36312_4081	NAC		1.92		PTLe00654.1	SGN-U583015	**Solyc06g060230**	6	34,577,702	34,577,090	Nam-like protein 1 [Solanum lycopersicum]
BI924306	E2F-DP		2.03		PTLe00354.1	SGN-U571268	**Solyc02g087310**	2	44,333,799	44,334,014	DEL3 (DP-E2F-like 3); TF [Arabidopsis thaliana]
AW221946	bHLH		2.11		PTLe00935.1	SGN-U575250	**Solyc07g064040**	7	63,567,974	63,568,517	basic helix-loop-helix protein
TA36877_4081	HB		2.14		PTLe00460.1	SGN-U578015	**Solyc11g068950**	11	50,587,678	50,586,307	BEL1-related homeotic protein 11 [Solanum tuberosum]
BG129142	Orphans		2.16		PTLe00369.1	SGN-U585565	Solyc08g067190	8	58,309,962	58,309,406	Two-component response regulator-like APRR2 Arabidopsis
AB108840	EIL		2.19		PTLe00360.1	595854; SGN-U	**06g073730; Solyc06g0**	6			Protein ETHYLENE INSENSITIVE [Solanum lycopersicum]
TA44091_4081	HB		2.24		PTLe00464.1	SGN-U562744	**Solyc11g069470**	11	51,154,120	51,154,454	Homeobox-leu zipper protein REVOLUTA OS=Arabidopsis
TA39335_4081	CCAAT		2.28		PTLe00351.1	SGN-U576319	**Solyc01g096710**	1	79,487,523	79,486,732	unnamed; Dr1-associated corepressor OS=Homo sapiens
TA46164_4081	WRKY		2.40		no hit found						

Table 10E continued

GoI[a]	TF-family[a]	Fold Change Neo123	Fold Change Neo111	Fold Change 3939	Plant tf DB acc. No.[ab]	SGN unigene No.[c]	ITAG2 acc. No.[c]	Chromosome Position[d] No.	Chromosome Position[d] Start (bp)	Chromosome Position[d] End (bp)	Best BLAST match annotation[c]
TA50710_4081	E2F-DP		**2.61**		PTLe00353.1	SGN-U572319	**Solyc11g068800**	11	50,503,530	50,503,025	transcription factor [Nicotiana tabacum]
DB721891	NAC		**2.62**		PTLe00685.1	SGN-U583008	**Solyc04g009440**	4	2,857,707	2,858,404	NAC domain protein [Solanum lycopersicum]
TA53950_4081	bHLH		**2.98**		PTLe00909.1	SGN-U595579	**Solyc12g100140**	12	65,323,660	65,323,955	basic helix-loop-helix (bHLH) family protein
TA45646_4081	WRKY		**3.10**		no hit found						
TA37926_4081	HB		**3.20**		PTLe00446.1	SGN-U569793			43,681,608	43,680,888	Homeobox-leuc zipper protein HAT7 , putative - S. demissum
TA49515_4081	C2H2		**3.33**		PTLe00244.1	SGN-U585555	**Solyc04g080130**	4	61,978,353	61,977,138	zinc finger (C2H2 type) family protein
TA43058_4081	AUX/IAA		**3.45**		PTLe00158.1	SGN-U579410	**Solyc09g083280**	9	64,331,993	64,331,694	Nicotiana tabacum Nt-iaa2.3; Auxin-responsive
TA40720_4081	WRKY		**3.61**		PTLe00810.1	SGN-U565155	**Solyc02g093050**	2	48,578,475	48,579,250	Probable WRKY transcription factor 7 - Arabidopsis thaliana
TA45578_4081	HB		**3.79**		no hit found						
TA55227_4081	MYB		**5.45**		PTLe00609.1	SGN-U568692	**Solyc03g093890**	3	48,995,780	48,995,134	Myb4 - Oryza sativa; Myb-related protein
TA48563_4081	MYB		**5.60**		PTLe00626.1	SGN-U583773	**Solyc11g011050**	11	4,108,152	4,108,606	myb-related transcription factor [Solanum lycopersicum]
TA44917_4081	C2C2-Dof		**6.06**		PTLe00200.1	SGN-U563536	Solyc06g075370	6	43,177,840	43,176,714	Dof domain, zinc finger family protein [Solanum demissum]
TA46330_4081	WRKY		**7.36**		no hit found						
TA44918_4081	C2C2-Dof		**8.92**		PTLe00200.1	SGN-U563536	**Solyc06g075370**	6	43,177,840	43,176,714	Dof domain, zinc finger family protein [Solanum demissum]
TA55173_4081	WRKY		**8.98**		no hit found						
TA55123_4081	WRKY		**9.40**		no hit found						
TA48803_4081	MYB		**12.57**		PTLe00608.1	SGN-U571259	**Solyc09g008250**	9	1,719,216	1,718,891	MYB24 - Malus domestica (Apple) (Malus sylvestris)
TA52860_4081	HB		**14.66**		PTLe00465.1	SGN-U579283	**Solyc04g074700**	4	58,210,965	58,210,596	homeodomain S. lycopersicum; Homeobox-leuc zipper
TA55334_4081	WRKY		**14.66**		no hit found						
TA46084_4081	HMG		**21.86**		PTLe00475.1	SGN-U566908	Solyc08g082070	8	62,136,369	62,136,844	98b [Daucus carota]
TA49866_4081	MYB		**27.62**		PTLe00625.1	SGN-U573246	**Solyc04g014470**	4	4,709,641	4,708,639	MYB-related TF [Nicotiana tabacum]; Ar MYB43/20
AI490010	MYB-related		**33.21**		PTLe00605.1	SGN-U565825	**Solyc12g099140**	12	64,745,538	64,746,126	Cpm5 [Craterostigma plantagineum]; Myb-related protein
TA50786_4081	G2-like		**44.99**		PTLe00383.1	SGN-U580727	**Solyc07g045000**	7	55,383,654	55,383,269	Putative Myb TF At1g14600 - Arabidopsis thaliana
TA50816_4081	C2H2		**48.99**		PTLe00245.1	SGN-U585104	**Solyc05g054650**	5	63,667,927	63,668,607	zinc finger (C2H2 type) family protein
BP889238	MYB		**69.04**		PTLe00617.1	SGN-U602631	**Solyc08g079270**	8	60,046,530	60,046,878	Protein ODORANT1 - Petunia hybrida (Petunia); AtMYB85
TA49650_4081	NAC		**80.08**		PTLe00670.1	SGN-U566547	**Solyc08g006020**	8	784,693	784,899	weak match: no apical meristem (NAM) family protein
AW032656	MYB		**84.78**		PTLe00613.1	SGN-U565756	**Solyc05g053330**	5	62,595,457	62,596,082	MYB21 - in Arabidopsis responsible for flower development
TA39480_4081	MADS		**113.36**		PTLe00556.1	SGN-U568823	**Solyc04g081000**	4	62,655,426	62,655,746	floral homeotic protein DEFICIENS [Solanum lycopersicum]
AW651186	MYB		**150.43**		PTLe00614.1	SGN-U571619	**Solyc04g077260**	4	59,829,155	59,829,918	NM_125143.3 Arabidopsis thaliana MYB36

Table 10F TFs where relative expression was altered uniquely in line 3939

		Fold Change			Plant tf DB	SGN unigene	ITAG2 acc. No.[c]	Chromosome Position[d]			
Gol[a]	TF-family[a]	Neo123	Neo111	3939	acc. No.[ab]	No.[c]		No.	Start (bp)	End (bp)	Best BLAST match annotation[c]
TA42046_4081	C2C2-YABBY			-6.45	PTLe00238.1	SGN-U572646	**Solyc05g012050**	5	5,277,398	5,277,174	Protein CRABS CLAW - Arabidopsis thaliana
DB724814	C2H2			-5.84	PTLe00243.1	SGN-U581130	Solyc08g063040	8	49,787,313	49,787,578	Protein TRANSPARENT TESTA 1 - Arabidopsis thaliana
TA49590_4081	C2H2			-3.01	PTLe00738.1	SGN-U568128	**Solyc12g097070**	12	65,402,703	65,402,877	ASHH3 | SET domain-containing protein (ASHH3)
BP889685	HMG			1.79	PTLe00468.1	SGN-U580087	**Solyc02g055550**	2	27,343,160	27,342,793	high mobility group (HMG1/2), similar to HMG protein
AW030183	G2-like			2.23	PTLe00384.1	SGN-U582275	**Solyc06g008200**	6	2,083,265	2,083,945	GARP-G2-like
TA37643_4081	G2-like			2.28	PTLe00373.1	SGN-U583163	**Solyc01g108300**	1	87,423,526	87,423,927	GARP-G2-like
TA44796_4081	C2C2-Dof			2.39	PTLe00206.1	SGN-U570633	**Solyc05g007880**	5	2,321,050	2,320,538	Dof-type zinc finger domain-containing protein
TA46971_4081	MYB-related			2.53	PTLe00582.1	SGN-U577106	**Solyc03g098320**	3	54,112,848	54,111,842	Myb-like DNA-binding protein
AW932398	G2-like			2.91	PTLe00387.1	SGN-U586549	**Solyc10g085620**	10	64,062,467	64,062,150	MYR1 | MYR1 (MYB-RELATED PROTEIN 1)
DB714457	HSF			2.94	PTLe00493.1	SGN-U580800					Arabidopsis thaliana heat shock transcription factor
DV103875	HSF			2.97	PTLe00490.1	SGN-U578436	**Solyc09g065660**	9	59,475,187	59,475,794	AT-HSFA6B (Arabidopsis thaliana heat shock TF)
TA53654_4081	bHLH			3.29	PTLe00937.1	SGN-U567958	**Solyc04g006990**	4	707,913	708,438	basic helix-loop-helix (bHLH) family protein
TA55670_4081	MADS			3.43	PTLe00523.1	SGN-U583377	**Solyc02g089210**	2	45,668,222	45,668,480	transcription factor, MADS-box, FBP29
TA39409_4081	CCAAT			4.26	PTLe00328.1	SGN-U565880	**Solyc07g065500**	7	64,459,865	64,458,985	CCAAT-box binding TF subunit B (NF-YB) (HAP3) (AHAP3)
TA44534_4081	HSF			4.60	PTLe00489.1	SGN-U569512	**Solyc08g062960**	8	49,589,919	49,588,941	heat shock transcription factor family protein
DV103959	HSF			4.84	PTLe00490.1	SGN-U578436	**Solyc09g065660**	9	59,475,187	59,475,794	Arabidopsis thaliana heat shock transcription factor
DB706372	HSF			6.33	PTLe00489.1	SGN-U569512	**Solyc08g062960**	8	49,589,919	49,588,941	heat shock transcription factor family protein

Table 10G TFs where relative expression was altered uniquely in line neo-123

GoI[a]	TF-family[a]	Fold Change Neo123	Neo111	3939	Plant tf DB acc. No.[ab]	SGN unigene No.[c]	ITAG2 acc. No.[c]	Chromosome Position[d] No.	Start (bp)	End (bp)	Best BLAST match annotation[c]
TA37413_4081	HB	**-6218.4**			PTLe00463.1	SGN-U579121	**Solyc03g113270**	3	57,512,649	57,513,278	homeobox [Solanum lycopersicum]
TA50528_4081	C2C2-YABBY	**-6.12**			PTLe00237.1	SGN-U578286	Solyc11g071810	11	52,252,158	52,252,011	YABBY 2 - (Rice); FAS protein [S. lycopersicum]
TA52541_4081	C2H2	**-4.26**			PTLe00262.1	SGN-U582936	**Solyc06g065440**	6	37,237,999	37,239,067	Protein TRANSPARENT TESTA 1 - Arabidopsis thaliana
TA46868_4081	bHLH	**-4.21**			PTLe00895.1	SGN-U584664	**Solyc04g014360**	4	4,614,954	4,614,383	basic helix-loop-helix (bHLH) family protein
TA48385_4081	WRKY	**-4.14**			no hit found						
AW616045	C2C2-YABBY	**-3.71**			PTLe00240.1	SGN-U577176	**Solyc07g008180**	7	2,920,895	2,921,108	NM_179750.2 Arabidopsis thaliana YAB5 (YABBY5)
TA55076_4081	AUX/IAA	**-3.63**			PTLe00161.1	SGN-U577682	**Solyc03g120500**	3	62,933,007	62,932,575	Auxin-responsive protein IAA16 - Arabidopsis thaliana
DB701475	WRKY	**-3.63**			PTLe00813.1	SGN-U565158	**Solyc01g079360**	1	70,961,525	70,961,156	WRKY transcription factor IIe-1 [Solanum lycopersicum]
TA44922_4081	C2H2	**-3.50**			PTLe00279.1	SGN-U581348	**Solyc10g084180**	10	63,158,962	63,158,028	zinc finger (C2H2 type) family protein
TA53587_4081	ABI3VP1	**-3.41**			PTLe00023.1	SGN-U584903	Solyc01g106230	1	85,952,116	85,952,568	hypothetical
TA53555_4081	TUB	**-3.41**			PTLe00791.1	SGN-U562976	**Solyc02g085130**	2	42,690,907	42,691,746	tubby-like protein [Cicer arietinum]
TA44923_4081	C2H2	**-3.39**			PTLe00279.1	SGN-U581348	**Solyc10g084180**	10	63,158,962	63,158,028	zinc finger (C2H2 type) family protein
TA50343_4081	MYB	**-3.18**			PTLe00636.1	SGN-U583772	**Solyc06g065100**	6	36,998,577	36,999,089	FaMYB1; Myb-related
BE458482	AUX/IAA	**-3.17**			PTLe00154.1	SGN-U577682	**Solyc03g120500**	3	62,933,007	62,932,575	Auxin-responsive protein IAA30 - Oryza sativa
DB678724	WRKY	**-3.17**			PTLe00813.1	SGN-U565158	**Solyc01g079360**	1	70,961,525	70,961,156	WRKY transcription factor IIe-1 [Solanum lycopersicum]
TA42784_4081	C2C2-YABBY	**-3.17**			PTLe00239.1	SGN-U577176	**Solyc07g008180**	7	2,920,895	2,921,108	NM_179750.2 Arabidopsis thaliana YAB5 (YABBY5)
TA40070_4081	MADS	**-3.03**			PTLe00543.1	SGN-U577952	**Solyc03g114840**	3	58,768,314	58,767,986	MADS-box protein 1 [Solanum lycopersicum]
TA52425_4081	Trihelix	**-2.97**			PTLe00805.1	SGN-U572185	Solyc11g005380	11	304,702	306,379	GTL1 (Arabidopsis thaliana); GT-2 (Arabidopsis thaliana)
TA42487_4081	C3H	**-2.85**			PTLe00313.1	SGN-U586184	**Solyc02g021760**	2	14,320,157	14,319,007	None
TA43235_4081	bZIP	**-2.78**			PTLe00965.1	SGN-U569399	**Solyc08g006110**	8	849,846	849,543	Glycine max bZIP TFr bZIP61; basic leu zipper O2 homolog 2
TA44542_4081	C3H	**-2.76**			PTLe00293.1	SGN-U569683	**Solyc12g017410**	12	6,661,285	6,660,850	Zinc finger CCCH domain-containing protein
TA50552_4081	MYB	**-2.62**			PTLe00627.1	SGN-U586047	**Solyc02g088190**	2	44,931,809	44,932,641	MYBAS2 (Nicotiana tabacum); MYB39 Arabidopsis thaliana
TA38804_4081	HB	**-2.59**			PTLe00462.1	SGN-U568678	**Solyc02g091930**	2	47,732,359	47,731,686	Homeobox-leucine zipper protein
TA46957_4081	bHLH	**-2.56**			PTLe00903.1	SGN-U570946	**Solyc07g043580**	7	54,828,538	54,829,717	SRL2, PIF4 | PIF4 (PHYTOCHROME INTERACTING FACTOR
TA47146_4081	HB	**-2.44**			PTLe00462.1	SGN-U568678	**Solyc02g091930**	2	47,732,359	47,731,686	Homeobox-leucine zipper protein HAT22 OS=Arabidopsis
TA39762_4081	TUB	**-2.43**			PTLe00784.1	SGN-U562978	Solyc03g033980	3	56,305,590	56,306,247	Tubby-like F-box protein 2 OS=Arabidopsis thaliana
DB685091	C2C2-Dof	**-2.36**			PTLe00209.1	SGN-U574638	**Solyc06g076030**	6	43,593,107	43,592,267	not good; Dof zinc finger protein [Solanum tuberosum]
TA50691_4081	NAC	**-2.35**			PTLe00659.1	SGN-U568168	**Solyc05g009840**	5	4,060,707	4,060,331	ANAC008 | ANAC008 (Arabidopsis NAC domain protein 8)
TA40865_4081	NAC	**-2.15**			PTLe00653.1	SGN-U585288	**Solyc12g056790**	12	48,229,255	48,230,443	nam-like protein 6 [Petunia x hybrida]
TA45060_4081	C3H	**-2.14**			no hit found						
TA44950_4081	ULT	**-2.14**			PTLe00807.1	SGN-U586255	**Solyc07g054450**	7	60,096,168	60,096,482	Protein ULTRAPETALA 1 - Arabidopsis thaliana
TA44951_4081	ULT	**-2.11**			PTLe00807.1	SGN-U586255	**Solyc07g054450**	7	60,096,168	60,096,482	Protein ULTRAPETALA 1 - Arabidopsis thaliana
TA45117_4081	AP2-EREBP	**-2.11**			PTLe00033.1	SGN-U580203	**Solyc02g093150**	2	48,658,361	48,658,969	Floral homeotic protein APETALA 2 - Arabidopsis thaliana
TA43586_4081	C3H	**-2.07**			PTLe00306.1	SGN-U582363	Solyc10g080260	10	60,915,830	60,918,389	zinc finger (CCCH-type) family protein
TA39499_4081	NAC	**-2.05**			PTLe00663.1	SGN-U568609	**Solyc07g063420**	7	63,130,403	63,131,096	NAC domain protein [Solanum lycopersicum]
TA47162_4081	MADS	**-2.05**			PTLe00517.1	SGN-U576965	**Solyc11g032100**	11	21,523,505	21,523,839	Agamous-like MADS-box protein AGL12 OS=Arabidopsis
TA44543_4081	C3H	**-2.03**			PTLe00293.1	SGN-U569683	**Solyc12g017410**	12	6,661,285	6,660,850	Zinc finger CCCH domain-containing protein ZFN1 Arabidopsis
TA35771_4081	CSD	**-2.03**			PTLe00796.1	SGN-U585676	**Solyc01g096470**	1	79,318,921	79,319,653	transcription factor [Arabidopsis thaliana]

Table 10G continued

GoI[a]	TF-family[a]	Fold Change Neo123	Fold Change Neo111	Fold Change 3939	Plant tf DB acc. No.[ab]	SGN unigene No.[c]	ITAG2 acc. No.[c]	Chromosome Position[d] No.	Chromosome Position[d] Start (bp)	Chromosome Position[d] End (bp)	Best BLAST match annotation[c]
TA44254_4081	HSF	-1.94			PTLe00495.1	SGN-U568618	**Solyc07g040680**	7	46,702,910	46,703,900	heat stress transcription factor HSFA9 [Helianthus annuus]
AW035599	C2H2	-1.89			no hit found						
TA42335_4081	C3H	-1.82			PTLe00322.1	SGN-U582330	**Solyc09g074640**	9	61,953,252	61,952,718	Zinc finger CCCH domain-containing protein
TA51093_4081	bHLH	-1.80			PTLe00929.1	SGN-U584819	Solyc05g050560	5	59,857,562	59,858,491	basic helix-loop-helix (bHLH) family protein
TA47630_4081	C2C2-GATA	-1.79			PTLe00224.1	SGN-U582438	**Solyc01g106030**	1	85,775,557	85,775,984	GATA transcription factor 27 OS=Arabidopsis thaliana
TA52404_4081	CPP	-1.78			PTLe00352.1	SGN-U570452	**Solyc07g020710**	7	13,283,585	13,284,023	tesmin/TSO1-like CXC domain-containing protein
AW621859	C2H2	-1.78			PTLe00267.1	SGN-U572521	**Solyc10g077110**	10	59,317,149	59,317,436	transcription factor IIIA [Solanum lycopersicum]
TA46556_4081	C2H2	-1.78			no hit found						
TA40860_4081	GRAS	-1.73			PTLe00414.1	SGN-U574351	**Solyc07g052960**	7	58,691,740	58,693,296	DELLA protein GAIP - Cucurbita maxima; scarecrow TF
TA46404_4081	Orphans	-1.71			no hit found						
TA42169_4081	Trihelix	-1.67			PTLe00795.1	SGN-U586349	**Solyc02g076810**	2	36,499,618	36,500,935	NP_191422.2 transcription factor [Arabidopsis thaliana]
TA49291_4081	MYB	-1.64			PTLe00602.1	SGN-U584837	Solyc09g014250	9	5,741,135	5,741,709	MYB transcription factor MYB51 [Glycine max]
TA46743_4081	bHLH	-1.50			PTLe00911.1	SGN-U565924	**Solyc06g065040**	6	36,946,789	36,947,234	NP_174776.1 (bHLH) family protein [Arabidopsis thaliana]
TA36724_4081	HSF	-1.46			PTLe00493.1	SGN-U580800			33,333,459	33,332,668	Heat stress TF A-6b OS=Arabidopsis thaliana
BP886085	C3H	-1.46			PTLe00294.1	SGN-U571812	**Solyc10g078750**	10	59,801,318	59,801,633	KH domain-containing protein / zinc finger (CCCH type) family
					see above	SGN-U586223	**Solyc01g014850**	1	15,587,394	15,586,532	KH domain-containing protein / zinc finger (CCCH type) family
TA46956_4081	C3H	-1.28			PTLe00299.1	SGN-U571675	**Solyc01g110490**	1	88,946,348	88,946,702	dihydrouridine synthase family protein
BM535903	AP2-EREBP	-1.22			PTLe00111.1	SGN-U571523	**Solyc09g059510**	9	50,197,235	50,196,981	Ethylene-responsive At4g13040 - Arabidopsis thaliana
TA43946_4081	MYB-related	1.44			no hit found						At1g09710/F21M12_10 [Arabidopsis]
TA43691_4081	FHA	1.44			PTLe00367.1	SGN-U569699	Solyc10g006390	10	1,009,671	1,008,632	Smad nuclear-interacting protein 1 OS=Mus musculus
TA45707_4081	E2F-DP	2.11			PTLe00354.1	SGN-U571268	**Solyc02g087310**	2	44,333,799	44,334,014	E2L2, E2FF, DEL3 | DEL3 (DP-E2F-like 3); transcription factor
DB720071	C3H	2.19			PTLe00296.1	SGN-U585861	**Solyc12g009390**	12	2,677,279	2,676,950	zinc finger (CCCH-type) family protein [Arabidopsis thaliana]
TA41510_4081	Orphans	2.43			PTLe00198.1	SGN-U565146	**Solyc07g006630**	7	1,496,685	1,495,923	CONSTANS-like protein [S. lycopersicum]; zinc finger (B-box)
TA56675_4081	Orphans	2.51			no hit found						
BI932393	GeBP	2.66			PTLe00426.1	SGN-U571553	**Solyc07g052900**	7	58,639,822	58,640,176	storekeeper protein [Solanum tuberosum]
					see above	SGN-U571555	**Solyc07g052830**	7	58,612,938	58,614,117	storekeeper protein [Solanum tuberosum]
DV935828	NAC	2.83			PTLe00654.1	SGN-U583015	**Solyc06g060230**	6	34,577,702	34,577,090	Nam-like protein 1 [Solanum lycopersicum]
BG126128	HB	3.03			PTLe00457.1	SGN-U600479	**Solyc03g098200**	3	53,993,320	53,993,130	HDG11 (HOMEODOMAIN GLABROUS11) Arabidopsis
TA39350_4081	Trihelix	3.15			PTLe00805.1	SGN-U572185	**Solyc11g005380**	11	304,702	306,379	GTL1; GT-2 (Arabidopsis); trihelix, putative
TA48591_4081	MYB-related	3.25			PTLe00612.1	SGN-U563576	**Solyc05g055240**	5	64,118,776	64,119,244	I-box binding factor [Solanum lycopersicum]; myb family
TA41657_4081	C3H	3.29			PTLe00291.1	SGN-U581830	**Solyc03g111580**	3	56,265,060	56,264,510	zinc finger (CCCH-type) family protein
BT012856	MYB	3.81			PTLe00612.1	SGN-U563576	**Solyc05g055240**	5	64,118,776	64,119,244	Mybl; I-box binding factor [Solanum lycopersicum]
TA52727_4081	Orphans	3.91			no hit found						NP_194461.1 B-Box-type zinc finger
TA43773_4081	MYB-related	6.59			PTLe00589.1	SGN-U573609	**Solyc10g084370**	10	63,267,559	63,267,383	MYB transcription factor [Camellia sinensis]
DB711260	MADS	23.02			PTLe00558.1	SGN-U596175	**Solyc03g019710**	3	6,653,924	6,654,120	TDR8 [S. lycopersicum]; Agamous-like MADS-box Arabidopsis
DB715635	MADS	29.98			PTLe00558.1	SGN-U596175	**Solyc03g019710**	3	6,653,924	6,654,120	TDR8 [S. lycopersicum]; Agamous-like MADS-box Arabidopsis

A subset of five TFs was selected as potential QTL identities based on expression levels, chromosome locations, and TF family. These are summarised in Table 11 and labelled TF1 to TF5 for further reference. TF1 and TF2 were selected because fold change in each TF relative to TA209 was correlated in directionality in both high rutin lines (Table 10B). Additionally, the locations of TF1 and TF2 were shown to be in close proximity to one another on chromosome 5. TF3, TF4 and TF5 were each selected for their relatively high and unique fold change in neo-111, neo-123 and 3939, respectively, as well as chromosome location. TF3 was selected because it showed the greatest relative increase in expression of all TFs located on chromosome 5 and perturbed in neo-111 but neither neo-123 nor 3939 (Table 10E). TF4 was selected because it showed the greatest relative increase in expression of all TFs located on chromosome 10 and perturbed in neo-123 but neither neo-111 nor 3939 (Table 10G). And TF5 was selected because it showed the greatest relative decrease in expression of all TFs located on chromosome 5 and perturbed in 3939 but neither neo-111 nor neo-123 (Table 10F).

Table 11 Candidate QTL, labelled TF1 to 5

Gene of interest, family, and reference are taken from Table 10. Chr. = chromosome location.

	Chr.	Gene of interest	TF Family	Reference
TF1	5	TA39847_4081	MYB-related	Table 10B
TF2	5	TA35979_4081	Orphan, LIM domain	Table 10B
TF3	5	AW032656	MYB	Table 10E
TF4	10	TA43773_4081	MYB-related	Table 10G
TF5	5	TA42046_4081	C2C2-YABBY, CRABS-CLAW	Table 10F

6.3 Analysis of candidate gene promoter regions for sequence polymorphisms

Primers were designed with the aim of amplifying, cloning, and sequencing upstream genomic regions, assumed to contain the promoters, of each of the TF genes for potential sequence polymorphisms between TA209 (*S. lycopersicum*), *S. neorickii*, and *S. habrochaites*.

Optimisation of PCR for TF3 promoter region resulted in PCR products using TA209 and *S. habrochaites* genomic templates; however, no product was shown using template for *S. neorickii* (Table 12). Since *S. neorickii* was the genotype of interest for TF3 (corresponding to the change in expression in neo-111) no further analysis was conducted on the TF3 promoter region. Likewise, PCR optimisation resulted in product when TA209 and *S. neorickii* genomic templates were used to amplify the promoter region of TF5, but not when using *S. habrochaites* template. Again, since 3939 showed a change in TF5 expression, no further analysis was conducted on the promoter region.

Table 12 Summary of candidate TF promoter regions sequenced for polymorphisms

Red and green arrows indicate significant decreases or increases relative to TA209, reported previously in Table 10. Black horizontal bar represents no significant difference from TA209. Successful amplification represented is by 'tick' mark when PCR product was shown using genomic DNA from TA209, *S. neorickii* (Neo) and *S. habrochaites* (Hab). 'X' shows where no product was found as a result of PCR.

Promoter	Chr	TF expression			PCR amplification of wild relative			
		3939	Neo 111	Neo 123	TA209	Neo	Hab	
TF1 (MYB-related)	5	↑	↑	–	✓	✓	X	Sequenced
TF2 (LIM)	5	↓	↓	–	✓	✓	✓	Sequenced
TF3 (MYB)	5	–	↑	–	✓	X	✓	
TF4 (MYB-related)	10	–	–	↑	✓	✓	✓	Sequenced
TF5 (Crabs-claw)	5	↓	–	–	✓	✓	X	

Amplification of TF1 promoter region was successful using TA209 and *S. neorickii* genomic template, but not *S. habrochaites*. Despite both wild relative genomic sequences being of interest as a result of changes in expression in 3939 and neo-111, further analysis on TA209 and *S. neorickii* alone was conducted. Sequence analysis, however, resulted in no consensus sequence, which was likely the result of non-specific primer binding. Therefore, no further analysis was conducted on the TF1 promoter region.

The promoter regions of both TF2 and TF4 were successfully amplified by PCR, and products were cloned, and sequenced. The results for TF2 promoter region are shown in Figure 52. Sequence alignment by a pairwise alignment algorithm (Corpet,

1988) indicated that consensus sequences existed between both wild relatives and TA209 (indicated in red, Figure 52) for large segments of the genomic sequences upstream from TF2. Polymorphisms existed between each of the wild relative sequences and that of TA209. Notable differences were the insertion mutations such as *S. habrochaites* genome sequence nucleotides labelled number 486 to 496 and *S. neorickii* genome sequence number 131 to 135. There existed, however, many point mutations (such as nucleotide labelled number 62), some deletions (such as number 785), and some insertions (such as number 465 to 467) that were identical in both wild relative genomes and difference from the TA209 sequence (Figure 52).

Figure 52 Sense strand of TF2 promoter sequence in TA209, *S. neorickii*, and *S. habrochaites*

Sequences of cloned genomic regions upstream of TF2 in TA209, *S. neorickii*, and *S. habrochaites* genotypes. Consensus sequences as a result of pairwise alignment (Corpet, 1988) are shown in red. Sequence polymorphisms in wild relatives compared with TA209 are shown in blue or black. Green and blue boxes indicate conserved binding motif identified by promoter motif analysis and flanking regions of polymorphisms, respectively (Zhou, personal communication). Red arrow indicates approximate coding region and direction of TF2 exon.

It was hypothesised that these regions of sequence polymorphisms in the assumed promoter region of TF2 may have affected expression of TF2 by disrupting or enhancing binding of a TF protein. Sequences were sent to N. Zhou (Syngenta) for promoter motif analysis. Results indicated that a sequence of 12 nucleotides and located upstream of TF2 (indicated by green box, Figure 52) was a target binding motif for LFY TF protein, and homologous for binding motifs in *AtCRC* and *LaCRC* (Zhou, personal communication). This target binding motif was identical in all three genotypes. Further to this, the sequence was flanked by point polymorphisms and two larger regions of polymorphism (Zhou, personal communication) upstream (approximately 230 bp) and downstream (approximately 130 bp), which are indicated by blue boxes (Figure 52).

Sequence results for the TF4 promoter region are shown in Figure 53. Sequence alignment, again shown by pairwise alignment (Corpet, 1988), indicated that large sections of genomic sequence close to the coding region of TF4 were identical in TA209, *S. neorickii* and *S. habrochaites*. Contrary to the TF2 promoter region, however, differences in sequences were more common further upstream from TF4, where many sequence polymorphisms were unique to either *S. neorickii* or *S. habrochaites* (Figure 53). Three target binding motifs were identified following promoter motif analysis by N. Zhou (Syngenta), and these are indicated in Figure 53 by green boxes. Although upstream of TF4, all three binding motifs were located on the antisense strand relative to the coding sequence of TF4, and no TF binding motif was identified in the sense strand (Zhou, personal communication; Figure 53). One of these three was located within the coding region of TF4 (60 to 69 bp, Figure 53), but exhibited no sequence polymorphisms between TA209 and either wild relative. The second was located at 93 to 110 bp (Figure 53), and showed an insertion in *S. neorickii* sequence, but no polymorphism in *S. habrochaites*. The third binding motif was located at 646 to 653 bp (Figure 53), and exhibited two different polymorphisms in *S. neorickii* and *S. habrochaites* genome sequences.

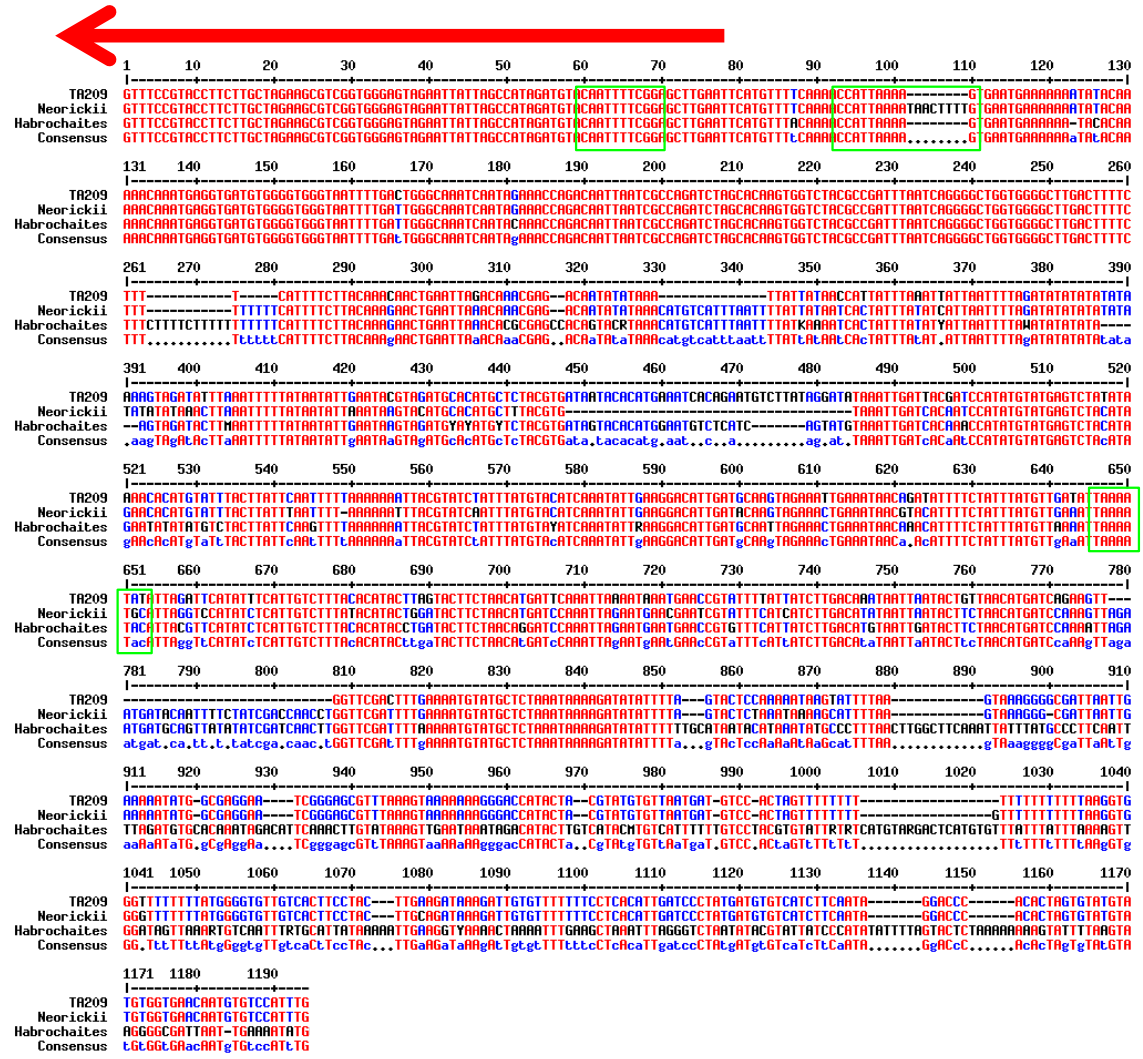

Figure 53 Antisense strand of TF4 promoter sequence in TA209, *S. neorickii*, and *S. habrochaites*

Sequences of cloned genomic regions upstream of TF4 in TA209, *S. neorickii*, and *S. habrochaites* genotypes. Consensus sequences as a result of hierarchical clustering (Corpet, 1988) are shown in red. Sequence polymorphisms in wild relatives compared with TA209 are shown in blue or black. Green boxes indicate binding motifs identified by promoter motif analysis (Zhou, personal communication). Red arrow indicates approximate coding region and direction of TF4 on corresponding sense strand.

Sequence alignments for TF2 promoter and TF4 promoter regions (Figure 52 and Figure 53, respectively) were verified by ClustalX 2.1 (Larkin et al., 2007). In both cases, sequences were aligned and comparably highlighted regions of polymorphisms were identified (Figure 54).

A

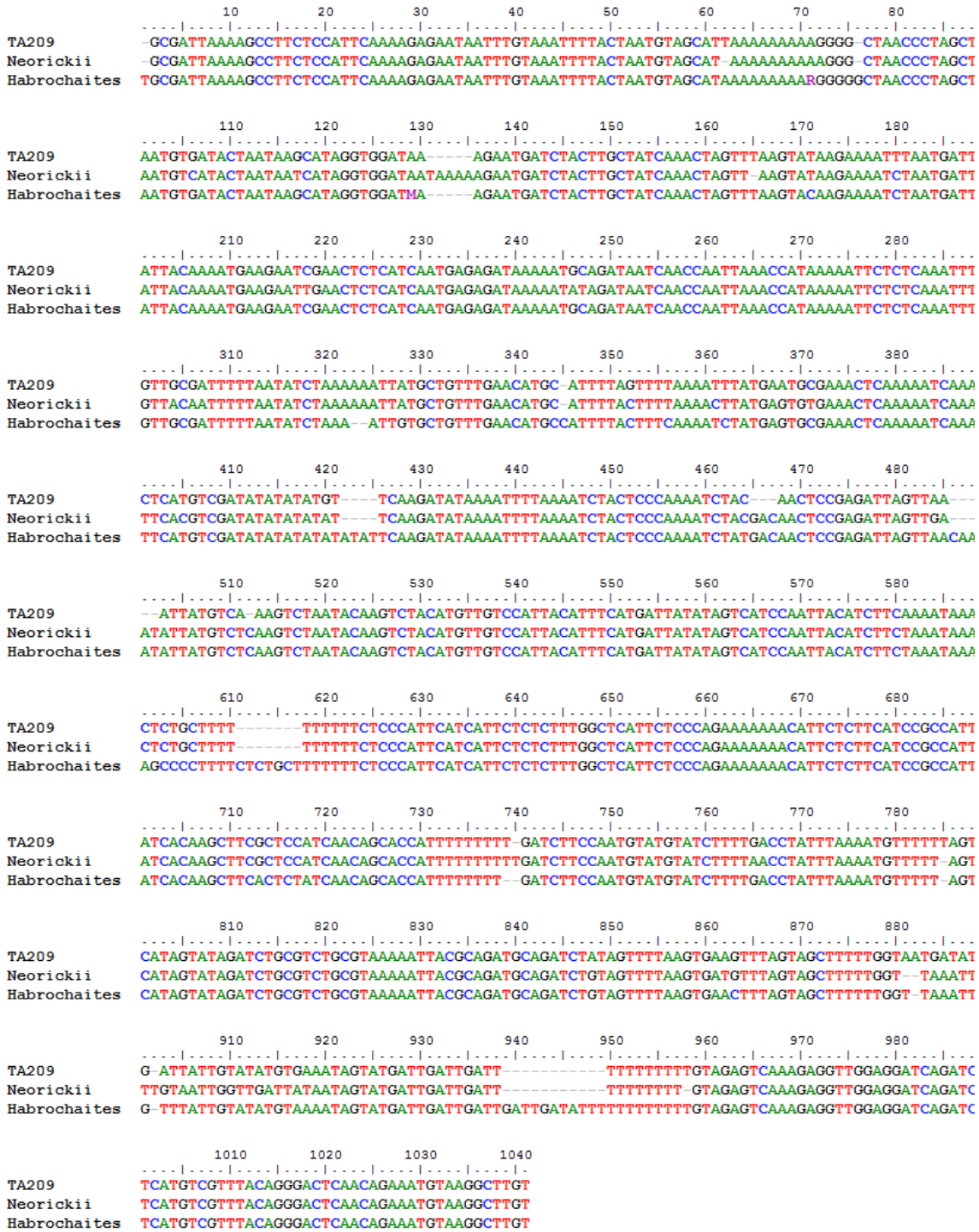

Figure 54 Alignments of (A) TF2 promoter and (B) TF4 promoter sequences from genotypes TA209, *S. neorickii* and *S. habrochaites* using ClustalX 2.1

Sequence alignments generated in support of alignments shown in Figure 52 and Figure 53.

B

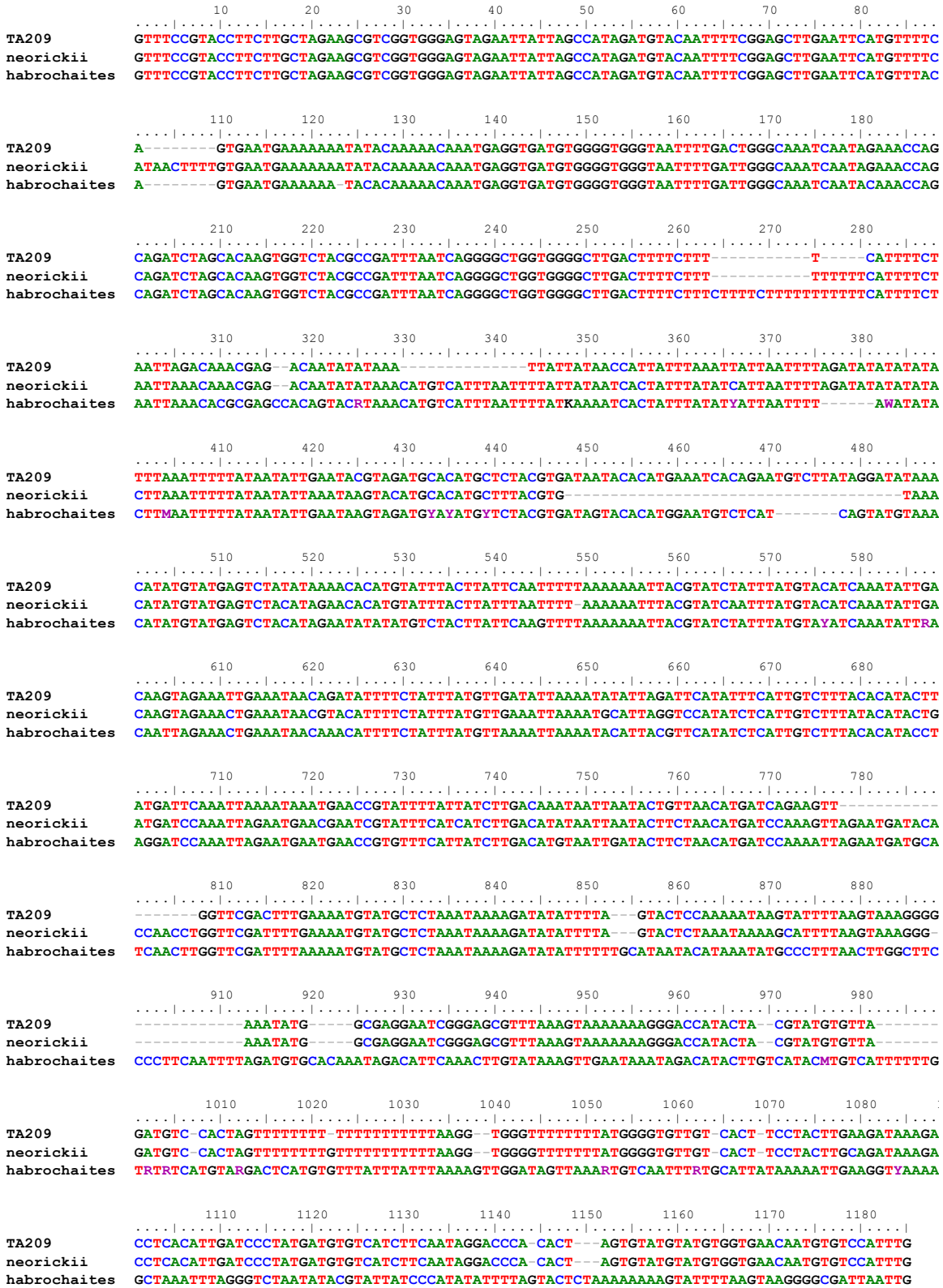

Figure 54 continued.

6.4 Discussion

A far greater number of TFs was shown to be perturbed in expression in either *S. neorickii* BILs than the *S. habrochaites* NIL, 3939 (Figure 51). This occurred with the greater number of morphological changes observed in neo-111 and -123 compared with 3939, such as fruit size, and the effects on post-harvest properties and isoprenoid accumulations discussed in section 5.8. Although speculative, it was hypothesised that the greater number of introgressed regions from wild relative genomes in neo-111 and -123 compared with 3939 could have resulted in a greater number of wild relative QTL alleles affecting a wider range of independent phenotypes and biochemical pathways. This hypothesis is supported in section 6.2 therefore, by the greater number of perturbed TF expression levels.

A wealth of information is available in the 186 TFs that have shown changes in expression relative to TA209 in lines neo-111, -123 and 3939 (Figure 51). Of particular interest are those TFs that are perturbed in two or more lines (Table 10A to D) since these may be involved in the disruption of commonly altered phenotypes compared with TA209, such as levels of phenolic compounds. What is not known, is whether these relative changes in TF expressions are the direct result of regulatory elements from wild relative introgressed regions (that is to say, sequence polymorphisms directly affecting expression) and are therefore causing phenotypic difference in the elite line (TA209) background, or whether they are endogenous responses to phenotypic changes themselves the result of introgressed genomic sequences (and therefore no difference in sequence would be present). Herein lies a limitation of this wealth of data, and a detailed study of most of the TFs could arguably be a valuable use of resources. For example, three of the perturbed TFs have shown similarity to *TT1* in Arabidopsis. These are TA41497_4081 on chromosome 6 (reduced in expression in neo-123 and 3939; Table 10), TA52541_4081 on chromosome 6 (reduced expression in neo-123; Table 10G), and DB724814 on chromosome 8 (reduced in expression in 3939; Table 10F). *TT1* is known to affect regulation of phenolics and result in reduced proanthocyanidin accumulation in seed testa (Chopra et al., 2008; Shirley, 1996). In order to manage these data, further analysis was therefore focused on TFs located within chromosomes 5 and 10, identified in chapter 3, and limited to a subset of TFs selected by a ‘trial and error’ approach.

Of all the TFs showing significantly altered expression in one or both high rutin lines, neo-111 and 3939, 12 were located on chromosome 5 (Table 10). Three of these 12 showed altered expression in both high rutin lines (these three also existed out of a possible nine TFs with perturbed expression in both high rutin lines, Table 10A and B combined). The first of these three, TA49294_4081 (Table 10A), was a MYB family TF. Although MYB family TFs have been linked with regulation of phenolic metabolism in tomato (Bovy et al., 2002; Luo et al., 2008; Mathews et al., 2003), directionality of change in expression was different in each high rutin line. TF TA49294_4081 was observed to increase in neo-111 by 1.56 fold; however, expression in 3939 was opposite to neo-111 (decrease of 1.18 fold) and was also altered in neo-123 (decrease of 1.47 fold). Furthermore, these fold changes, although significant, were all below two fold relative change in expression from TA209. The second and third TF in this group were better candidates for QTL identity. TA35979_4081 in Table 10B (TF2, Table 11) was an Orphan TF; however, a BLAST search showed it to contain a domain similar to that found in LIM family. LIM family TFs have been shown to be involved in phenylpropanoid metabolism in tobacco and eucalyptus, especially with regards to regulation of or correlation with early phenylpropanoid and lignin biosytnetic genes (Kawaoka and Ebinuma, 2001; Kawaoka et al., 2000; Negishi et al., 2011). TF2 decreased in expression for both neo-111 and 3939 (3.85 and 4.11 fold, respectively). Likewise, TA39847_4081 in Table 10B (TF1, Table 11) showed a similar change in expression in both neo-111 and 3939 (increased expression of 3.58 and 1.54 fold, respectively), and was shown to be a MYB-related family TF. TF1 and TF2 were therefore suggested as possible QTL candidates.

As detailed in section 6.2, TF3, TF4 and TF5 were selected on the basis of their chromosome location and degree of change in expression in each of the candidate lines. MYB TF family are known phenolic regulators (Bovy et al., 2002; Luo et al., 2008; Mathews et al., 2003) and, although not involved in phenolic biosynthesis, YABBY TF family has been shown previously to be involved in the regulation of tomato fruit development (Cong et al., 2008).

The promoter regions were targeted for the identification of polymorphisms in sequence that may affect TF expression. Mutations within the promoter region of isoprenoid biosynthetic genes have been shown to affect isoprenoid accumulation in select genotypes of tomato mapping populations (E. Enfissi, P. Fraser, D. Heldt, unpublished). Further to this, sequence specificities and polymorphisms in promoter

regions of TFs such as MYB have been shown to segregate with phenolic accumulation and regulate biosynthesis (Sidorenko et al., 2000; Takos et al., 2006). Those sequences that were not amplified by PCR (Table 12) are likely due to mutations on the site of primer binding since only those genotypes where expression was perturbed were unable to be amplified (for example TF3 promoter in *S. neorickii* where expression was shown to be altered in neo-111, and TF5 promoter in *S. habrochaites* where expression was shown to be altered in 3939). These suggested mutations may also be responsible for the perturbed expression levels observed by affecting transcription of the TF genes. This could be verified by shifting the target amplicon site allowing primers to bind to an alternative sequence. Unfortunately time constraints did not permit this to be verified by these means.

TF4 promoter region exhibited polymorphisms at binding motif regions believed to bind regulatory elements (Zhou, personal communication). These polymorphisms were different in *S. neorickii* and *S. habrochaites* (Figure 53). Since only neo-123 (that is to say *S. neorickii* introgressed genome region) exhibited the change in *p*-coumaric acid levels, if TF4 promoter region is involved in regulation of this trait then only the sequence polymorphism of *S. neorickii* and not *S. habrochaites* would underlie this QTL allele.

Sequence analysis of TF2 promoter region indicated that some polymorphisms were identical in both *S. neorickii* and *S. habrochaites* genomes. It was at first surprising that no differences were identified in such large regions of these sequences considering these wild relatives are classified in different ancestral groups (*S. neorickii* exists in the Arcanum group of *Lycopersicon* section, whereas *S. habrochaites* exists in the Eriopersicon group) (Grandillo et al., 2011). However, previous links between *S. neorickii* and *S. habrochaites* have been identified. A QTL allele for the resistance to powdery mildew caused by *Oidium lycopersici* was identified in *S. neorickii* accession G1.1601, and found to map to the same region of chromosome 6 where an existing resistance gene was previously identified in *S. habrochaites* (Bai et al., 2003). Three putative QTL alleles were identified in *S. neorickii* for resistance to tomato grey mould caused by *Botrytis cinerea* (Finkers et al., 2008), and each of them was assigned to a putative homologous locus in *S. habrochaites* (Finkers et al., 2007a; Finkers et al., 2007b). Additionally, Crabs-claw promoter regions have been reported as being highly conserved between plant species and involved in TF networks in nectary development (Lee et al., 2005). It is feasible, therefore, that the polymorphisms present in TF2

promoter region are contributory to wider changes via an unknown TF regulatory network that affects traits common to both neo-111 and 3939. The TF2 promoter region therefore, remains the strongest candidate from this study for possible influence on phenolic profile regulation in *S. neorickii* BIL population.

7 Conclusion

7.1 Summary

Ripe tomato fruit from a field cultivated trial of the *S. neorickii* BIL population (Fulton et al., 2000; Grandillo et al., 2011), comprising 142 accessions, were screened by HPLC-DAD for phenylpropanoid and flavonoid compounds. Data were compared with a previously screened field trial (E. Enfissi, unpublished). The trends in metabolite levels across the population showed poor reproducibility between the two field cultivated crops. This could have been caused by a change in environmental stresses affecting phenolic production differently in subsequent years (Lovdal et al., 2010; Paterson et al., 1991; Rousseaux et al., 2005). Alternatively, this could have been due to segregation distortion or the presence of heterozygous inserts from *S. neorickii* (as shown by RFLP and COSII markers available within the EU Sol consortium) segregating over subsequent generations resulting in loss of introgressed segments and therefore the wild species phenotype (Paterson et al., 1991; Xu et al., 1997).

Of the assessed compounds, rutin and *p*-coumaric acid showed better reproducibility between the two crops grown in the field than other metabolites. This was especially evident for BILs possessing the higher relative increases (≥2 fold rutin and ≥5 fold *p*-coumaric acid). Trends in chlorogenic acid and chalcone-naringenin showed poor consistency between subsequent field grown crops; however, some individual BILs were reproducible. It is possible that levels of chalcone-naringenin were close to limits of quantification and detection in some lines, which explains inconsistent results.

Crude, non-quantitative observations were used to identify associations between high relative levels of rutin or *p*-coumaric acid and RFLP marker data showing chromosome 'clusters' of *S. neorickii* introgressed regions. These associations were identified in chromosomes 5 and 10 for BILs possessing high relative rutin and *p*-coumaric acid levels, respectively, in both field cultivated trials. Two genotypes, neo-111 and -123, were selected as candidate lines to represent introgressed *S. neorickii* genome regions on these chromosomes, respectively, due to the presence of homozygous inserts according to RFLP markers. When later compared with COSII

marker set, neo-111 additionally possessed a chromosome 5 homozygous introgressed region; however, neo-123 according to COSII markers possessed no chromosome 10 introgressed region. It was hypothesised therefore, that there might have existed an introgressed region in chromosome 10 of neo-123 that fell between two COSII markers, which was not detected by the COSII marker set.

A previous study (T. Wells, unpublished) showed *S. habrochaites* NIL (Monforte and Tanksley, 2000) genotype 3939 with an introgressed region in chromosome 5 to display a similar high rutin phenotype to neo-111. Line 3939 was therefore included in this study for comparison. 3939, neo-111 and -123 genotypes were replicated in the UK cultivated under glasshouse conditions. Despite the change in environment and alterations to levels of some phenolic compounds, high rutin and *p*-coumaric acid accumulations seen previously were replicated. This reproducibility reflected the stability of the wild QTL allele, which is not guaranteed when using mapping populations (Rousseaux et al., 2005).

To address the need for future high throughput analytical methodologies for screening of large-scale populations, such as mapping populations, a UPLC-PDA method was developed for the separation and quantification of tomato fruit phenylpropanoids and flavonoids. Results were comparable to published methodologies (Novakova et al., 2010; Spacil et al., 2008). Together with a UPLC-PDA high throughput method for the separation and quantification of tomato isoprenoid and related compounds developed by P. Fraser (unpublished), this method was assessed for limitations, and validated by screening a colour mutant population. The combined isoprenoid and phenolic profiles demonstrated some clustering when represented by PCA and individual genotypes with characteristic metabolite profiles were identified. MTBE was assessed to be a suitable alternative to chloroform for the extraction of isoprenoids from tomato, and results were consistent with those previously published (Matyash et al., 2008).

Lines neo-111, -123 and 3939 were characterised for their phenolic profile throughout fruit development and ripening. Neo-123 accumulated high levels of *p*-coumaric acid and derivatives in all fruit tissue types after fruit had developed. Levels increased throughout ripening at the detrimental cost to the accumulation of other phenolic compounds. An increase in the accumulation of carotenoids and isoprenoid

related compounds occurred at similar times to *p*-coumaric acid in fruit ripening. The widespread metabolome, however, was shown to be disrupted at MG and R stages.

3939 showed similar patterns to neo-123 in accumulating phenolic compounds through ripening; however, flavonol glycosides rather than *p*-coumaric acid were increased in all tissue types. Aside from flavonol glucosides, 3939 exhibited virtually no difference from TA209 in profiles of other phenolics, isoprenoid or related compounds; however, PCA revealed perturbations to metabolism during Br and T stages. Increases in phenolics were observed during post-harvest storage of 3939.

Neo-111 was the only line to show disruption to metabolism of phenolics at all stages of fruit development and ripening. A range of phenolics in addition to flavonol glycosides were perturbed, and this was seen in all tissue types. Following fruit development, levels of compounds in early carotenogenesis were decreased relative to TA209, while later pathway intermediates were increased. Neo-111 exhibited similar effects to wider metabolism as 3939 according to PCA.

A possible association was observed between high phenolic content and increased antioxidant capacity in polar extracts. This was especially true for 3939 and neo-123. Antioxidant activities in tissue types of neo-111 however, indicated that in this line there existed a combinatory effect of flavonol glycosides and other phenylpropanoid or flavonoid compounds on antioxidant capacity.

All genotypes displayed delayed ripening times compared with TA209, but this was more extreme in neo-111 and -123. 3939 largely resembled fruit morphology indistinguishable to TA209. Neo-111 and -123 showed many physiological distinctions to TA209. Neo-111 and -123 additionally suffered greater detriment to morphology during post-harvest storage.

This illustrated that although these genotypes possess wild relative QTL alleles desirable for commercially important and health-promoting traits, fully exploiting these alleles would require breeding programmes that transfer desirable alleles into commercial lines without genetic drag of the undesirable traits.

A TF platform (Rohrmann et al., 2011) identified 186 TFs that displayed elevated or decreased expression at turning stage in at least one genotype relative to TA209. While these data offered great potential for detailed investigation of unknown regulatory networks, limited resources resulted in further analysis on a subset of five

TFs. For three of these, no PCR product could be amplified, resulting in two TFs that were sequenced and subsequently analysed by promoter motif analysis (N. Zhou, Syngenta). TF4 was identified as a possible regulatory element in the manipulation of phenolic profiles in neo-123. TF2 was highlighted as the more likely candidate for regulation of flavonol glycosides in 3939 and neo-111.

7.2 Impact summary

Data presented in this work support the hypothesis that alleles from *S. neorickii* and *S. habrochaites* wild relatives exist that affect the accumulation of intermediate compounds within the phenolic biosynthetic pathway. In this regard, it is suggested that these alleles localise to QTL for health related fruit quality traits, and can therefore potentially be utilised for the improvement of fruit quality traits in cultivated germplasm.

On the condition that the speculative association between enhanced accumulation of specific phenolic intermediates and wild relative introgressed regions were verified, these QTL may have a beneficial impact on future breeding programmes. These wild relative alleles for health related QTL could be incorporated into elite cultivars in isolation and without inheriting detrimental parental phenotypes, such as the organoleptic trait of small fruit size seen in this work.

By increasing the current scientific knowledge of the mechanisms by which the accumulation of specific metabolite intermediates occurs, this work and others like it have the potential to benefit current domesticated tomato cultivars beyond the current restrictions imposed by the limited genepool.

7.3 Future work

Phenolic profile data from a screen of *S. neorickii* BIL population in Crops 1 and 2 were used together with RFLP marker data to identify possible associations between increases in rutin and *p*-coumaric acid levels and introgressed chromosome regions from the wild relative *S. neorickii*. From these, candidate BILs neo-111 and -123 were selected to represent introgressions within these regions of chromosome 5 and 10. Greater confidence could be achieved in these associations by statistical validation, such

as QTL mapping, as has been shown recently (Chapman et al., 2012), with either RFLP or COSII markers. QTL mapping would statistically determine the threshold of association between the introgressed regions and accumulation of rutin, *p*-coumaric acid, or any of the other quantified phenolics. It is possible that multiple QTL may exist for the accumulation of any one compound.

The available markers could be used to genotype those biological replicates grown in Crops 1 and 2 from neo-111 and neo-123 to confirm whether the RFLP and COSII theoretical introgressed regions were in fact present. Were resources available, markers could also be sought within the region of chromosome 10 of neo-123 where no introgressed region was identified by COSII markers to ascertain whether neo-123 is still a candidate BIL for introgressions in chromosome 10. Fruit was made available from a third crop (Crop 3) and a backcross crop (section 2.2.2, Table 4) for the previously nominated BILs. Phenolic profiling of these select lines would ascertain whether the phenotypes were replicated once again in a field trial and whether they were dominant when backcrossed to the TA209 parent.

Direct comparisons could be made between the UPLC-PDA methods validated here and their corresponding HPLC methods. Accessions from the CC colour mutant population could be used for this purpose.

The wealth of data produced from the metabolomic and TF platform analyses could be better exploited alongside transcriptomic data. For this purpose microarray analyses were conducted at Syngenta using identical plant material to that in this study (see sections 2.2.7.2 and 2.2.7.5). Results could be used to better infer coordination between seemingly independent pathways as has been shown previously (Enfissi et al., 2010). The identification of sequence polymorphisms in this study represents only a small proportion of the potential regulatory networks involved in fully characterising these genotypes were further resources available. In the same regard, work on TF2 and TF4 could be furthered by sequence analysis of the gene coding regions, and by determining which genotype is present in lines neo-111, -123 and 3939. Expression analysis by qRT-PCR could be used to determine expression levels of TF2 and TF4 throughout fruit development stages MG, Br and R in addition to those shown here at T stage. This could be used to correlate with available metabolite data. Were the accessions backcrossed, a study could assess if co-segregation between sequence polymorphism and phenotype occurs.

From these data, the project could be further reaching if the selected QTL alleles were introduced into an elite background in isolation (in other words, without the entire introgressed regions of neo-111, -123 and 3939). This could be achieved by GM or breeding strategies, and if successful offers the opportunity for commercial exploitation and benefit of optimising health-promoting traits in elite tomato cultivars.

8 References

Aharoni,A., De Vos,C.H.R., Wein,M., Sun,Z.K., Greco,R., Kroon,A., Mol,J.N.M., and O'Connell,A.P. (2001). The strawberry FaMYB1 transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. Plant Journal *28*, 319-332.

Apel,W. and Bock,R. (2009). Enhancement of Carotenoid Biosynthesis in Transplastomic Tomatoes by Induced Lycopene-to-Provitamin A Conversion. Plant Physiology *151*, 59-66.

Ashrafi,H., Kinkade,M.P., Merk,H.L., and Foolad,M.R. (2012). Identification of novel quantitative trait loci for increased lycopene content and other fruit quality traits in a tomato recombinant inbred line population. Molecular Breeding *30*, 549-567.

Asins,M.J. (2002). Present and future of quantitative trait locus analysis in plant breeding. Plant Breeding *121*, 281-291.

Atkinson,N.J., Dew,T.P., Orfila,C., and Urwin,P.E. (2011). Influence of Combined Biotic and Abiotic Stress on Nutritional Quality Parameters in Tomato (Solanum lycopersicum). Journal of Agricultural and Food Chemistry *59*, 9673-9682.

Bai,Y.L., Huang,C.C., van der Hulst,R., Meijer-Dekens,F., Bonnema,G., and Lindhout,P. (2003). QTLs for tomato powdery mildew resistance (Oidium lycopersici) in Lycopersicon parviflorum G1.1601 co-localize with two qualitative powdery mildew resistance genes. Molecular Plant-Microbe Interactions *16*, 169-176.

Ballester,A.R., Molthoff,J., de Vos,R., Hekkert,B.T.L., Orzaez,D., Fernandez-Moreno,J.P., Tripodi,P., Grandillo,S., Martin,C., Heldens,J., Ykema,M., Granell,A., and Bovy,A. (2010). Biochemical and Molecular Analysis of Pink Tomatoes: Deregulated Expression of the Gene Encoding Transcription Factor S1MYB12 Leads to Pink Tomato Fruit Color. Plant Physiology *152*, 71-84.

Barrero,L.S. and Tanksley,S.D. (2004). Evaluating the genetic basis of multiple-locule fruit in a broad cross section of tomato cultivars. Theoretical and Applied Genetics *109*, 669-679.

Barten,J.H.M., Scott,J.W., and Gardner,R.G. (1994). Characterization of Blossom-End Morphology Genes in Tomato and Their Usefulness in Breeding for Smooth Blossom-End Scars. Journal of the American Society for Horticultural Science *119*, 798-803.

Bernacchi,D., Beck-Bunn,T., Eshed,Y., Inai,S., Lopez,J., Petiard,V., Sayama,H., Uhlig,J., Zamir,D., and Tanksley,S. (1998a). Advanced backcross QTL analysis of tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from Lycopersicon hirsutum and L-pimpinellifolium. Theoretical and Applied Genetics *97*, 170-180.

Bernacchi,D., Beck-Bunn,T., Eshed,Y., Lopez,J., Petiard,V., Uhlig,J., Zamir,D., and Tanksley,S. (1998b). Advanced backcross QTL analysis in tomato. I. Identification of

QTLs for traits of agronomic importance from Lycopersicon hirsutum. Theoretical and Applied Genetics *97*, 381-397.

Bieza,K. and Lois,R. (2001). An Arabidopsis mutant tolerant to lethal ultraviolet-B levels shows constitutively elevated accumulation of flavonoids and other phenolics. Plant Physiology *126*, 1105-1115.

Bombarely,A., Menda,N., Tecle,I.Y., Buels,R.M., Strickler,S., Fischer-York,T., Pujar,A., Leto,J., Gosselin,J., and Mueller,L.A. (2011). The Sol Genomics Network (solgenomics.net): growing tomatoes using Perl. Nucleic Acids Research *39*, D1149-D1155.

Bonguebartelsman,M. and Phillips,D.A. (1995). Nitrogen Stress Regulates Gene-Expression of Enzymes in the Flavonoid Biosynthetic-Pathway of Tomato. Plant Physiology and Biochemistry *33*, 539-546.

Bovy,A., de Vos,R., Kemper,M., Schijlen,E., Pertejo,M.A., Muir,S., Collins,G., Robinson,S., Verhoeyen,M., Hughes,S., Santos-Buelga,C., and van Tunen,A. (2002). High-flavonol tomatoes resulting from the heterologous expression of the maize transcription factor genes LC and C1. Plant Cell *14*, 2509-2526.

Brewer,M.S. (2011). Natural Antioxidants: Sources, Compounds, Mechanisms of Action, and Potential Applications. Comprehensive Reviews in Food Science and Food Safety *10*, 221-247.

Bugos,R.C., Chiang,V.L., Zhang,X.H., Campbell,E.R., Podila,G.K., and Campbell,W.H. (1995). Rna Isolation from Plant-Tissues Recalcitrant to Extraction in Guanidine. Biotechniques *19*, 734-737.

Burbulis,I.E. and Winkel-Shirley,B. (1999). Interactions among enzymes of the Arabidopsis flavonoid biosynthetic pathway. Proceedings of the National Academy of Sciences of the United States of America *96*, 12929-12934.

Buta,J.G. and Spaulding,D.W. (1997). Endogenous levels of phenolics in tomato fruit during growth and maturation. Journal of Plant Growth Regulation *16*, 43-46.

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

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

Caldana,C., Scheible,W.R., Mueller-Roeber,B., and Ruzicic,S. (2007). A quantitative RT-PCR platform for high-throughput expression profiling of 2500 rice transcription factors. Plant Methods *3*.

Carli,P., Barone,A., Fogliano,V., Frusciante,L., and Ercolano,M.R. (2011). Dissection of genetic and environmental factors involved in tomato organoleptic quality. Bmc Plant Biology *11*.

Carrera,J., del Carmen,A.F., Fernandez-Munoz,R., Rambla,J.L., Pons,C., Jaramillo,A., Elena,S.F., and Granell,A. (2012). Fine-Tuning Tomato Agronomic Properties by Computational Genome Redesign. Plos Computational Biology *8*.

Causse,M., Friguet,C., Coiret,C., Lepicier,M., Navez,B., Lee,M., Holthuysen,N., Sinesio,F., Moneta,E., and Grandillo,S. (2010). Consumer Preferences for Fresh Tomato at the European Scale: A Common Segmentation on Taste and Firmness. Journal of Food Science *75*, S531-S541.

Causse,M., Saliba-Colombani,V., Lesschaeve,I., and Buret,M. (2001). Genetic analysis of organoleptic quality in fresh market tomato. 2. Mapping QTLs for sensory attributes. Theoretical and Applied Genetics *102*, 273-283.

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

Chen,G.Q. and Foolad,M.R. (1999). A molecular linkage map of tomato based on a cross between Lycopersicon esculentum and L-pimpinellifolium and its comparison with other molecular maps of tomato. Genome *42*, 94-103.

Chopra,S., Hoshino,A., Boddu,J., and Iida,S. (2008). Flavonoid pigments as tools in molecular genetics. In The Science of Flavonoids, E.Grotewold, ed. (NY: Springer Science), pp. 147-173.

Cle,C., Hill,L.M., Niggeweg,R., Martin,C.R., Guisez,Y., Prinsen,E., and Jansen,M.A.K. (2008). Modulation of chlorogenic acid biosynthesis in Solanum lycopersicum; consequences for phenolic accumulation and UV-tolerance. Phytochemistry *69*, 2149-2156.

Conceicao,L.F.R., Ferreres,F., Tavares,R.M., and Dias,A.C.P. (2006). Induction of phenolic compounds in Hypericum perforatum L. cells by Colletotrichum gloeosporioides elicitation. Phytochemistry *67*, 149-155.

Cong,B., Barrero,L.S., and Tanksley,S.D. (2008). Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nature Genetics *40*, 800-804.

Corpet,F. (1988). Multiple Sequence Alignment with Hierarchical-Clustering. Nucleic Acids Research *16*, 10881-10890.

Davies,K.M. (2007). Genetic modification of plant metabolism for human health benefits. Mutat. Res. *622*, 122-137.

Davuluri,G.R., van Tuinen,A., Fraser,P.D., Manfredonia,A., Newman,R., Burgess,D., Brummell,D.A., King,S.R., Palys,J., Uhlig,J., Bramley,P.M., Pennings,H.M.J., and Bowler,C. (2005). Fruit-specific RNAi-mediated suppression of DET1 enhances carotenoid and flavonoid content in tomatoes. Nature Biotechnology *23*, 890-895.

de Castro,J.P.A., Nick,C., Milagres,C.D., Mattedi,A.P., Marim,B.G., and da Silva,D.J.H. (2010). Genetic diversity among tomato's subsamples for pre-breeding. Crop Breeding and Applied Biotechnology *10*, 74-82.

de Castro,L.R., Cortez,L.A.E., and Vigneault,C. (2006). Effect of sorting, refrigeration and packaging on tomato shelf life. Journal of Food Agriculture and Environment *4*, 70-74.

Dwyer,J. (1995). Overview - Dietary Approaches for Reducing Cardiovascular-Disease Risks. Journal of Nutrition *125*, S656-S665.

Enfissi,E.M.A., Barneche,F., Ahmed,I., Lichtle,C., Gerrish,C., McQuinn,R.P., Giovannoni,J.J., Lopez-Juez,E., Bowler,C., Bramley,P.M., and Fraser,P.D. (2010). Integrative Transcript and Metabolite Analysis of Nutritionally Enhanced DE-ETIOLATED1 Downregulated Tomato Fruit. Plant Cell *22*, 1190-1215.

Eshed,Y., Abu-Abied,M., Saranga,Y., and Zamir,D. (1992). Lycopersicon esculentum lines containing small overlapping introgressions from L. pennellii. Theoretical and Applied Genetics *83*, 1027-1034.

Eshed,Y., Gera,G., and Zamir,D. (1996). A genome-wide search for wild-species alleles that increase horticultural yield of processing tomatoes. Theoretical and Applied Genetics *93*, 877-886.

Eshed,Y. and Zamir,D. (1994). A Genomic Library of Lycopersicon Pennellii in Lycopersicon-Esculentum - A Tool for Fine Mapping of Genes. Euphytica *79*, 175-179.

Eshed,Y. and Zamir,D. (1995). An Introgression Line Population of Lycopersicon Pennellii in the Cultivated Tomato Enables the Identification and Fine Mapping of Yield-Associated Qtl. Genetics *141*, 1147-1162.

Finkers,R., Bai,Y.L., van den Berg,P., van Berloo,R., Meijer-Dekens,F., ten Have,A., van Kan,J., Lindhout,P., and van Heusden,A.W. (2008). Quantitative resistance to Botrytis cinerea from Solanum neorickii. Euphytica *159*, 83-92.

Finkers,R., Finkers,R., van Heusden,A.W., Meijer-Dekens,F., van Kan,J.A.L., Maris,P., and Lindhout,P. (2007a). The construction of a Solanum habrochaites LYC4 introgression line population and the identification of QTLs for resistance to Botrytis cinerea. Theoretical and Applied Genetics *114*, 1071-1080.

Finkers,R., van den Berg,P., van Berloo,R., ten Have,A., van Heusden,A.W., van Kan,J.A.L., and Lindhout,P. (2007b). Three QTLs for Botrytis cinerea resistance in tomato. Theoretical and Applied Genetics *114*, 585-593.

Folch,J., Lees,M., and Sloane Stanley,G.H. (1957). A simple method for the isolation and purification of total lipides from animal tissues. J. Biol. Chem. *226*, 497-509.

Frary,A., Fulton,T.M., Zamir,D., and Tanksley,S.D. (2004). Advanced backcross QTL analysis of a Lycopersicon esculentum x L. pennellii cross and identification of possible orthologs in the Solanaceae. Theoretical and Applied Genetics *108*, 485-496.

Fraser,P.D. and Bramley,P.M. (2004). The biosynthesis and nutritional uses of carotenoids. Progress in Lipid Research *43*, 228-265.

Fridman,E., Carrari,F., Liu,Y.S., Fernie,A.R., and Zamir,D. (2004). Zooming in on a quantitative trait for tomato yield using interspecific introgressions. Science *305*, 1786-1789.

Fridman,E., Liu,Y.S., Carmel-Goren,L., Gur,A., Shoresh,M., Pleban,T., Eshed,Y., and Zamir,D. (2002). Two tightly linked QTLs modify tomato sugar content via different physiological pathways. Molecular Genetics and Genomics *266*, 821-826.

Fulton,T.M., BeckBunn,T., Emmatty,D., Eshed,Y., Lopez,J., Petiard,V., Uhlig,J., Zamir,D., and Tanksley,S.D. (1997). QTL analysis of an advanced backcross of Lycopersicon peruvianum to the cultivated tomato and comparisons with QTLs found in other wild species. Theoretical and Applied Genetics *95*, 881-894.

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

Fulton,T.M., Grandillo,S., Beck-Bunn,T., Fridman,E., Frampton,A., Lopez,J., Petiard,V., Uhlig,J., Zamir,D., and Tanksley,S.D. (2000). Advanced backcross QTL analysis of a Lycopersicon esculentum x Lycopersicon parviflorum cross. Theoretical and Applied Genetics *100*, 1025-1042.

Galensa,R. and Herrmann,K. (1980). Analysis of Flavonoids by High-Performance Liquid-Chromatography. Journal of Chromatography *189*, 217-224.

Garcia-Martinez,S., Andreani,L., Garcia-Gusano,M., Geuna,F., and Ruiz,J.J. (2006). Evaluation of amplified fragment length polymorphism and simple sequence repeats for tomato germplasm fingerprinting: utility for grouping closely related traditional cultivars. Genome *49*, 648-656.

Gautier,H., Diakou-Verdin,V., Benard,C., Reich,M., Buret,M., Bourgaud,F., Poessel,J.L., Caris-Veyrat,C., and Genard,M. (2008). How does tomato quality (sugar, acid, and nutritional quality) vary with ripening stage, temperature, and irradiance? Journal of Agricultural and Food Chemistry *56*, 1241-1250.

Giliberto,L., Perrotta,G., Pallara,P., Weller,J.L., Fraser,P.D., Bramley,P.M., Fiore,A., Tavazza,M., and Giuliano,G. (2005). Manipulation of the blue light photoreceptor cryptochrome 2 in tomato affects vegetative development, flowering time, and fruit antioxidant content. Plant Physiology *137*, 199-208.

Gonzalez,E.A. and Nazareno,M.A. (2011). Antiradical action of flavonoid-ascorbate mixtures. Lwt-Food Science and Technology *44*, 558-564.

Gonzali,S., Mazzucato,A., and Perata,P. (2009). Purple as a tomato: towards high anthocyanin tomatoes. Trends in Plant Science *14*, 237-241.

Goodfellow, M. Identification and biochemical characterisation of metabolite quantitative trait loci associated with consumer health traits in *Solanum pennellii* introgression lines. 2008. Royal Holloway, University of London.
Ref Type: Thesis/Dissertation

Gorguet,B., Eggink,P.M., Ocana,J., Tiwari,A., Schipper,D., Finkers,R., Visser,R.G.F., and van Heusden,A.W. (2008). Mapping and characterization of novel parthenocarpy QTLs in tomato. Theoretical and Applied Genetics *116*, 755-767.

Grandillo,S., Chetelat,R., Knapp,S., Spooner,D., Peralta,I., Cammareri,M., Perez,O., Termolino,P., Tripodi,P., Chiusano,M.L., Ercolano,M.R., Frusciante,L., Monti,L., and Pignone,D. (2011). *Solanum* sect. *Lycopersicon*. In Wild Crop Relatives: Genomic and Breeding Resources. Vegetables, C.Kole, ed. Springer), pp. 129-215.

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

Halket,J.M., Waterman,D., Przyborowska,A.M., Patel,R.K.P., Fraser,P.D., and Bramley,P.M. (2005). Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. Journal of Experimental Botany *56*, 219-243.

Hertog,M.G.L., Feskens,E.J.M., Hollman,P.C.H., Katan,M.B., and Kromhout,D. (1993). Dietary Antioxidant Flavonoids and Risk of Coronary Heart-Disease - the Zutphen Elderly Study. Lancet *342*, 1007-1011.

Isaacson,T., Ronen,G., Zamir,D., and Hirschberg,J. (2002). Cloning of tangerine from tomato reveals a carotenoid isomerase essential for the production of beta-carotene and xanthophylls in plants. Plant Cell *14*, 333-342.

Jha,S.N. and Matsuoka,T. (2005). Determination of post-harvest storage life of tomato fruits. JOURNAL OF FOOD SCIENCE AND TECHNOLOGY *42*, 526-529.

Jones,S. (2001). Hybridism. In Almost Like a Whale: The Origin of Species Updated, (London: Black Swan), pp. 223-250.

Joseph,J.A., Shukitt-Hale,B., Denisova,N.A., Bielinski,D., Martin,A., Mcewen,J.J., and Bickford,P.C. (1999). Reversals of age-related declines in neuronal signal transduction, cognitive, and motor behavioral deficits with blueberry, spinach, or strawberry dietary supplementation. Journal of Neuroscience *19*, 8114-8121.

Kabelka,E., Franchino,B., and Francis,D.M. (2002). Two loci from Lycopersicon hirsutum LA407 confer resistance to strains of Clavibacter michiganensis subsp michiganensis. Phytopathology *92*, 504-510.

Kawaoka,A. and Ebinuma,H. (2001). Transcriptional control of lignin biosynthesis by tobacco LIM protein. Phytochemistry *57*, 1149-1157.

Kawaoka,A., Kaothien,P., Yoshida,K., Endo,S., Yamada,K., and Ebinuma,H. (2000). Functional analysis of tobacco LIM protein Ntlim1 involved in lignin biosynthesis. Plant Journal *22*, 289-301.

Kok,E.J. and Kuiper,H.A. (2003). Comparative safety assessment for biotech crops. Trends in Biotechnology *21*, 439-444.

Kootstra,A. (1994). Protection from Uv-B-Induced Dna-Damage by Flavonoids. Plant Molecular Biology *26*, 771-774.

Koukol,J. and Conn,E.E. (1961). The metabolism of aromatic compounds in higher plants. IV. Purification and properties of the phenylalanine deaminase of Hordeum vulgare. J. Biol. Chem. *236*, 2692-2698.

Kreuzaler,F. and Hahlbrock,K. (1975). Enzymic synthesis of an aromatic ring from acetate units. Partial purification and some properties of flavanone synthase from cell-suspension cultures of Petroselinum hortense. Eur. J. Biochem. *56*, 205-213.

Kreuzaler,F., Ragg,H., Fautz,E., Kuhn,D.N., and Hahlbrock,K. (1983). UV-induction of chalcone synthase mRNA in cell suspension cultures of Petroselinum hortense. Proc. Natl. Acad. Sci. U. S. A *80*, 2591-2593.

Ku,H.M., Doganlar,S., Chen,K.Y., and Tanksley,S.D. (1999). The genetic basis of pear-shaped tomato fruit. Theoretical and Applied Genetics *99*, 844-850.

Labate,J.A., Grandillo,S., Fulton,T., Munos,S., Caicedo,A.L., Peralta,I., Ji,Y., Chetelat,R., Scott,J.W., Gonzalo,M.J., Francis,D., Yang,W., van de Knaap,E., Baldo,A.M., Smith-White,B., Mueller,L.A., Prince,J.P., Blanchard,N.E., Storey,D.B., Stevens,M.R., Robbins,M.D., Wang,J.-F., Liedl,B.E., O'Connell,M.A., Stommel,J.R., Aoki,K., Iijima,Y., Slade,A.J., Hurst,S.R., Loeffler,D., Steine,M.N., Vafeados,D., McGuire,C., Freeman,C., Amen,A., Goodstal,J., Facciotti,D., Van Eck,J., and Causse,M. (2007). Tomato. In Genome Mapping and Molecular Breeding in Plants: Vegetables, C.Kole, ed. Springer), pp. 1-125.

Larkin,M.A., Blackshields,G., Brown,N.P., Chenna,R., McGettigan,P.A., McWilliam,H., Valentin,F., Wallace,I.M., Wilm,A., Lopez,R., Thompson,J.D., Gibson,T.J., and Higgins,D.G. (2007). Clustal W and clustal X version 2.0. Bioinformatics *23*, 2947-2948.

Lee,J.Y., Baum,S.F., Alvarez,J., Patel,A., Chitwood,D.H., and Bowman,J.L. (2005). Activation of CRABS CLAW in the nectaries and carpels of Arabidopsis. Plant Cell *17*, 25-36.

Levin,R. (2008). The Quality is in the Breeding. In The Red Bodyguard: The Amazing Health Promoting Properties of the Tomato, (Cambridge: Icon Books), pp. 125-132.

Lewinsohn,E., Sitrit,Y., Bar,E., Azulay,Y., Ibdah,M., Meir,A., Yosef,E., Zamir,D., and Tadmor,Y. (2005). Not just colors - carotenoid degradation as a link between pigmentation and aroma in tomato and watermelon fruit. Trends in Food Science & Technology *16*, 407-415.

Li,J.Y., Oulee,T.M., Raba,R., Amundson,R.G., and Last,R.L. (1993). Arabidopsis Flavonoid Mutants Are Hypersensitive to Uv-B Irradiation. Plant Cell *5*, 171-179.

Lin,J.-K. and Weng,M.-S. (2008). Flavonoids as nutraceuticals. In The Science of Flavonoids, E.Grotewold, ed. (NY: Springer Science), pp. 213-238.

Lindhout,P., Vanheusden,S., Pet,G., Vanooijen,J.W., Sandbrink,H., Verkerk,R., Vrielink,R., and Zabel,P. (1994). Perspectives of Molecular Marker Assisted Breeding for Earliness in Tomato. Euphytica *79*, 279-286.

Lippman,Z.B., Semel,Y., and Zamir,D. (2007). An integrated view of quantitative trait variation using tomato interspecific introgression lines. Current Opinion in Genetics & Development *17*, 545-552.

Lois,R. and Buchanan,B.B. (1994). Severe Sensitivity to Ultraviolet-Radiation in An Arabidopsis Mutant Deficient in Flavonoid Accumulation .2. Mechanisms of Uv-Resistance in Arabidopsis. Planta *194*, 504-509.

Long,M., Millar,D.J., Kimura,Y., Donovan,G., Rees,J., Fraser,P.D., Bramley,P.M., and Bolwell,G.P. (2006). Metabolite profiling of carotenoid and phenolic pathways in mutant and transgenic lines of tomato: Identification of a high antioxidant fruit line. Phytochemistry *67*, 1750-1757.

Lovdal,T., Olsen,K.M., Slimestad,R., Verheul,M., and Lillo,C. (2010). Synergetic effects of nitrogen depletion, temperature, and light on the content of phenolic compounds and gene expression in leaves of tomato. Phytochemistry *71*, 605-613.

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

MacDonald,M.J. and D'Cunha,G.B. (2007). A modern view of phenylalanine ammonia lyase. Biochemistry and Cell Biology-Biochimie et Biologie Cellulaire *85*, 759.

Martin,C. and PazAres,J. (1997). MYB transcription factors in plants. Trends in Genetics *13*, 67-73.

Martinez-Valverde,I., Periago,M.J., Provan,G., and Chesson,A. (2002). Phenolic compounds, lycopene and antioxidant activity in commercial varieties of tomato (Lycopersicum esculentum). Journal of the Science of Food and Agriculture *82*, 323-330.

Mathews,H., Clendennen,S.K., Caldwell,C.G., Liu,X.L., Connors,K., Matheis,N., Schuster,D.K., Menasco,D.J., Wagoner,W., Lightner,J., and Wagner,D.R. (2003). Activation tagging in tomato identifies a transcriptional regulator of anthocyanin biosynthesis, modification, and transport. Plant Cell *15*, 1689-1703.

Matyash,V., Liebisch,G., Kurzchalia,T.V., Shevchenko,A., and Schwudke,D. (2008). Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. Journal of Lipid Research *49*, 1137-1146.

Mayer,J.E., Pfeiffer,W.H., and Beyer,P. (2008). Biofortified crops to alleviate micronutrient malnutrition. Current Opinion in Plant Biology *11*, 166-170.

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

Melendez-Martinez,A.J., Fraser,P.D., and Bramley,P.M. (2010). Accumulation of health promoting phytochemicals in wild relatives of tomato and their contribution to in vitro antioxidant activity. Phytochemistry *71*, 1104-1114.

Mellway,R.D., Tran,L.T., Prouse,M.B., Campbell,M.M., and Constabel,C.P. (2009). The Wound-, Pathogen-, and Ultraviolet B-Responsive MYB134 Gene Encodes an R2R3 MYB Transcription Factor That Regulates Proanthocyanidin Synthesis in Poplar. Plant Physiology *150*, 924-941.

Millar,D.J., Long,M., Donovan,G., Fraser,P.D., Boudet,A.M., Danoun,S., Bramley,P.M., and Bolwell,G.P. (2007). Introduction of sense constructs of cinnamate 4-hydroxylase (CYP73A24) in transgenic tomato plants shows opposite effects on flux into stem lignin and fruit flavonoids. Phytochemistry *68*, 1497-1509.

Miller,J.C. and Tanksley,S.D. (1990). Rflp Analysis of Phylogenetic-Relationships and Genetic-Variation in the Genus Lycopersicon. Theoretical and Applied Genetics *80*, 437-448.

Mink,P.J., Scrafford,C.G., Barraj,L.M., Harnack,L., Hong,C.P., Nettleton,J.A., and Jacobs,D.R. (2007). Flavonoid intake and cardiovascular disease mortality: a prospective study in postmenopausal women. American Journal of Clinical Nutrition *85*, 895-909.

Minoggio,M., Bramati,L., Simonetti,P., Gardana,C., Iemoli,L., Santangelo,E., Mauri,P.L., Spigno,P., Soressi,G.P., and Pietta,P.G. (2003). Polyphenol pattern and antioxidant activity of different tomato lines and cultivars. Annals of Nutrition and Metabolism *47*, 64-69.

Minutolo,M., Amalfitano,C., Evidente,A., Frusciante,L., and Errico,A. (2012). Polyphenol distribution in plant organs of tomato introgression lines. Nat. Prod. Res.

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

Mol,J., Grotewold,E., and Koes,R. (1998). How genes paint flowers and seeds. Trends in Plant Science *3*, 212-217.

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

Muir,S.R., Collins,G.J., Robinson,S., Hughes,S., Bovy,A., De Vos,C.H.R., van Tunen,A.J., and Verhoeyen,M.E. (2001). Overexpression of petunia chalcone isomerase in tomato results in fruit containing increased levels of flavonols. Nature Biotechnology *19*, 470-474.

Nagano,K., Kano,H., Arito,H., Yamamoto,S., and Matsushima,T. (2006). Enhancement of renal carcinogenicity by combined inhalation and oral exposures to chloroform in male rats. Journal of Toxicology and Environmental Health-Part A-Current Issues *69*, 1827-1842.

Naqvi,S., Zhu,C.F., Farre,G., Ramessar,K., Bassie,L., Breitenbach,J., Conesa,D.P., Ros,G., Sandmann,G., Capell,T., and Christou,P. (2009). Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct

metabolic pathways. Proceedings of the National Academy of Sciences of the United States of America *106*, 7762-7767.

Narotsky,M.G., Best,D.S., McDonald,A., Godin,E.A., Hunter,E.S., III, and Simmons,J.E. (2011). Pregnancy loss and eye malformations in offspring of F344 rats following gestational exposure to mixtures of regulated trihalomethanes and haloacetic acids. Reprod. Toxicol. *31*, 59-65.

Navarro-Gonzalez,I., Garcia-Valverde,V., Garcia-Alonso,J., and Periago,M.J. (2011). Chemical profile, functional and antioxidant properties of tomato peel fiber. Food Research International *44*, 1528-1535.

Negishi,N., Nanto,K., Hayashi,K., Onogi,S., and Kawaoka,A. (2011). Transcript abundances of LIM transcription factor, 4CL, CAld5H and CesAs affect wood properties in Eucalyptus globulus. Silvae Genetica *60*, 288-296.

Ness,A.R. and Powles,J.W. (1997). Fruit and vegetables, and cardiovascular disease: A review. International Journal of Epidemiology *26*, 1-13.

Niggeweg,R., Michael,A.J., and Martin,C. (2004). Engineering plants with increased levels of the antioxidant chlorogenic acid. Nature Biotechnology *22*, 746-754.

Nijveldt,R.J., van Nood,E., van Hoorn,D.E.C., Boelens,P.G., van Norren,K., and van Leeuwen,P.A.M. (2001). Flavonoids: a review of probable mechanisms of action and potential applications. American Journal of Clinical Nutrition *74*, 418-425.

Novakova,L., Spacil,Z., Seifrtova,M., Opletal,L., and Solich,P. (2010). Rapid qualitative and quantitative ultra high performance liquid chromatography method for simultaneous analysis of twenty nine common phenolic compounds of various structures. Talanta *80*, 1970-1979.

Pairoba,C.F. and Walbot,V. (2003). Post-transcriptional regulation of expression of the Bronze2 gene of Zea mays L. Plant Molecular Biology *53*, 75-86.

Paran,I., Goldman,I., and Zamir,D. (1997). QTL analysis of morphological traits in a tomato recombinant inbred line population. Genome *40*, 242-248.

Paterson,A.H., Damon,S., Hewitt,J.D., Zamir,D., Rabinowitch,H.D., Lincoln,S.E., Lander,E.S., and Tanksley,S.D. (1991). Mendelian Factors Underlying Quantitative Traits in Tomato - Comparison Across Species, Generations, and Environments. Genetics *127*, 181-197.

Peer,W.A. and Murphy,A.S. (2008). Flavonoids as signal molecules. In The Science of Flavonoids, E.Grotewold, ed. (NY: Springer Science), pp. 239-268.

Peralta,I., Spooner,D., and Knapp,S. (2008). Taxonomy of Wild Tomatoes and their Relatives (Solanum sect. Lycopersicoides, sect. Juglandifolia, sect. Lycopersicon; Solanaceae). Systematic Botany Monographs *84*, 1-186.

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

Quattrocchio,F., Baudry,A., Lepiniec,L., and Grotewold,E. (2008). The regulation of flavonoid biosynthesis. In The Science of Flavonoids, E.Grotewold, ed. (NY: Springer Science), pp. 97-122.

Quattrocchio,F., Wing,J., van der Woude,K., Souer,E., de Vetten,N., Mol,J., and Koes,R. (1999). Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color. Plant Cell *11*, 1433-1444.

Rausher,M.D. (2008). The evolution of flavonoids and their genes. In The Science of Flavonoids, E.Grotewold, ed. (NY: Springer Science), pp. 175-212.

Re,R., Pellegrini,N., Proteggente,A., Pannala,A., Yang,M., and Rice-Evans,C. (1999). Antioxidant activity applying an improved ABTS radical cation decolorization assay. Free Radical Biology and Medicine *26*, 1231-1237.

Rick,C.M., Kesicki,E., Fobes,J.F., and Holle,M. (1976). Genetic and Biosystematic Studies on 2 New Sibling Species of Lycopersicon from Inter-Andean Peru. Theoretical and Applied Genetics *47*, 55-68.

Rohrmann,J., Tohge,T., Alba,R., Osorio,S., Caldana,C., McQuinn,R., Arvidsson,S., van der Merwe,M.J., Riano-Pachon,D.M., Mueller-Roeber,B., Fei,Z.J., Nesi,A.N., Giovannoni,J.J., and Fernie,A.R. (2011). Combined transcription factor profiling, microarray analysis and metabolite profiling reveals the transcriptional control of metabolic shifts occurring during tomato fruit development. Plant Journal *68*, 999-1013.

Ronen,G., Carmel-Goren,L., Zamir,D., and Hirschberg,J. (2000). An alternative pathway to beta-carotene formation in plant chromoplasts discovered by map-based cloning of Beta and old-gold color mutations in tomato. Proceedings of the National Academy of Sciences of the United States of America *97*, 11102-11107.

Ronen,G., Cohen,M., Zamir,D., and Hirschberg,J. (1999). Regulation of carotenoid biosynthesis during tomato fruit development: Expression of the gene for lycopene epsilon-cyclase is down-regulated during ripening and is elevated in the mutant Delta. Plant Journal *17*, 341-351.

Rosler,J., Krekel,F., Amrhein,N., and Schmid,J. (1997). Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. Plant Physiology *113*, 175-179.

Rousseaux,M.C., Jones,C.M., Adams,D., Chetelat,R., Bennett,A., and Powell,A. (2005). QTL analysis of fruit antioxidants in tomato using Lycopersicon pennellii introgression lines. Theoretical and Applied Genetics *111*, 1396-1408.

Ruijter,J.M., Ramakers,C., Hoogaars,W.M.H., Karlen,Y., Bakker,O., van den Hoff,M.J.B., and Moorman,A.F.M. (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Research *37*.

Rushton,P.J., Somssich,I.E., Ringler,P., and Shen,Q.X.J. (2010). WRKY transcription factors. Trends in Plant Science *15*, 247-258.

Ryan,K.G., Swinny,E.E., Winefield,C., and Markham,K.R. (2001). Flavonoids and UV photoprotection in Arabidopsis mutants. Zeitschrift fur Naturforschung C-A Journal of Biosciences *56*, 745-754.

Saliba-Colombani,V., Causse,M., Langlois,D., Philouze,J., and Buret,M. (2001). Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. Theoretical and Applied Genetics *102*, 259-272.

Sanchez-Rodriguez,E., Ruiz,J.M., Ferreres,F., and Moreno,D.A. (2012). Phenolic profiles of cherry tomatoes as influenced by hydric stress and rootstock technique. Food Chemistry *134*, 775-782.

Sato,S., Tabata,S., Hirakawa,H., Asamizu,E., Shirasawa,K., Isobe,S., Kaneko,T., Nakamura,Y., Shibata,D., Aoki,K., Egholm,M., Knight,J., Bogden,R., Li,C.B., Shuang,Y., Xu,X., Pan,S.K., Cheng,S.F., Liu,X., Ren,Y.Y., Wang,J., Albiero,A., Dal Pero,F., Todesco,S., Van Eck,J., Buels,R.M., Bombarely,A., Gosselin,J.R., Huang,M.Y., Leto,J.A., Menda,N., Strickler,S., Mao,L.Y., Gao,S., Tecle,I.Y., York,T., Zheng,Y., Vrebalov,J.T., Lee,J., Zhong,S.L., Mueller,L.A., Stiekema,W.J., Ribeca,P., Alioto,T., Yang,W.C., Huang,S.W., Du,Y.C., Zhang,Z.H., Gao,J.C., Guo,Y.M., Wang,X.X., Li,Y., He,J., Li,C.Y., Cheng,Z.K., Zuo,J.R., Ren,J.F., Zhao,J.H., Yan,L.H., Jiang,H.L., Wang,B., Li,H.S., Li,Z.J., Fu,F.Y., Chen,B.T., Han,B., Feng,Q., Fan,D.L., Wang,Y., Ling,H.Q., Xue,Y.B.A., Ware,D., McCombie,W.R., Lippman,Z.B., Chia,J.M., Jiang,K., Pasternak,S., Gelley,L., Kramer,M., Anderson,L.K., Chang,S.B., Royer,S.M., Shearer,L.A., Stack,S.M., Rose,J.K.C., Xu,Y.M., Eannetta,N., Matas,A.J., McQuinn,R., Tanksley,S.D., Camara,F., Guigo,R., Rombauts,S., Fawcett,J., Van de Peer,Y., Zamir,D., Liang,C.B., Spannagl,M., Gundlach,H., Bruggmann,R., Mayer,K., Jia,Z.Q., Zhang,J.H., Ye,Z.B.A., Bishop,G.J., Butcher,S., Lopez-Cobollo,R., Buchan,D., Filippis,I., Abbott,J., Dixit,R., Singh,M., Singh,A., Pal,J.K., Pandit,A., Singh,P.K., Mahato,A.K., Dogra,V., Gaikwad,K., Sharma,T.R., Mohapatra,T., Singh,N.K., Causse,M., Rothan,C., Schiex,T., Noirot,C., Bellec,A., Klopp,C., Delalande,C., Berges,H., Mariette,J., Frasse,P., Vautrin,S., Zouine,M., Latche,A., Rousseau,C., Regad,F., Pech,J.C., Philippot,M., Bouzayen,M., Pericard,P., Osorio,S., del Carmen,A.F., Monforte,A., Granell,A., Fernandez-Munoz,R., Conte,M., Lichtenstein,G., Carrari,F., De Bellis,G., Fuligni,F., Peano,C., Grandillo,S., Termolino,P., Pietrella,M., Fantini,E., Falcone,G., Fiore,A., Giuliano,G., Lopez,L., Facella,P., Perrotta,G., Daddiego,L., Bryan,G., Orozco,M., Pastor,X., Torrents,D., van Schriek,K.N.V.M., Feron,R.M.C., van Oeveren,J., de Heer,P., daPonte,L., Jacobs-Oomen,S., Cariaso,M., Prins,M., van Eijk,M.J.T., Janssen,A., van Haaren,M.J.J., Jo,S.H., Kim,J., Kwon,S.Y., Kim,S., Koo,D.H., Lee,S., Hur,C.G., Clouser,C., Rico,A., Hallab,A., Gebhardt,C., Klee,K., Jocker,A., Warfsmann,J., Gobel,U., Kawamura,S., Yano,K., Sherman,J.D., Fukuoka,H., Negoro,S., Bhutty,S., Chowdhury,P., Chattopadhyay,D., Datema,E., Smit,S., Schijlen,E.W.M., van de Belt,J., van Haarst,J.C., Peters,S.A., van Staveren,M.J., Henkens,M.H.C., Mooyman,P.J.W., Hesselink,T., Van Ham,R.C.H.J., Jiang,G.Y., Droege,M., Choi,D., Kang,B.C., Kim,B.D., Park,M., Kim,S., Yeom,S.I., Lee,Y.H., Choi,Y.D., Li,G.C., Gao,J.W., Liu,Y.S., Huang,S.X., Fernandez-Pedrosa,V., Collado,C., Zuniga,S., Wang,G.P., Cade,R., Dietrich,R.A., Rogers,J., Knapp,S., Fei,Z.J., White,R.A., Thannhauser,T.W., Giovannoni,J.J., Botella,M.A., Gilbert,L., Gonzalez,R., Goicoechea,J.L., Yu,Y., Kudrna,D., Collura,K., Wissotski,M., Wing,R., Schoof,H., Meyers,B.C., Gurazada,A.B., Green,P.J., Mathur,S., Vyas,S., Solanke,A.U., and Kumar,R. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. Nature *485*, 635-641.

Schijlen,E., De Vos,C.H.R., Jonker,H., van den Broeck,H., Molthoff,J., van Tunen,A., Martens,S., and Bovy,A. (2006). Pathway engineering for healthy phytochemicals

leading to the production of novel flavonoids in tomato fruit. Plant Biotechnology Journal *4*, 433-444.

Seeram,N.P., Adams,L.S., Hardy,M.L., and Heber,D. (2004). Total cranberry extract versus its phytochemical constituents: Antiproliferative and synergistic effects against human tumor cell lines. Journal of Agricultural and Food Chemistry *52*, 2512-2517.

Sevenier,R., van der Meer,I.M., Bino,R., and Koops,A.J. (2002). Increased production of nutriments by genetically engineered crops. Journal of the American College of Nutrition *21*, 199S-204S.

Shih,C.H., Chen,Y.L., Wang,M.F., Chu,I.K., and Lo,C. (2008). Accumulation of isoflavone genistin in transgenic tomato plants overexpressing a soybean isoflavone synthase gene. Journal of Agricultural and Food Chemistry *56*, 5655-5661.

Shirley,B.W. (1996). Flavonoid biosynthesis: 'New' functions for an 'old' pathway. Trends in Plant Science *1*, 377-382.

Sidorenko,L.V., Li,X.G., Cocciolone,S.M., Chopra,S., Tagliani,L., Bowen,B., Daniels,M., and Peterson,T. (2000). Complex structure of a maize Myb gene promoter: functional analysis in transgenic plants. Plant Journal *22*, 471-482.

Slimestad,R., Fossen,T., and Verheul,M.J. (2008). The flavonoids of tomatoes. Journal of Agricultural and Food Chemistry *56*, 2436-2441.

Spacil,Z., Novakova,L., and Solich,P. (2008). Analysis of phenolic compounds by high performance liquid chromatography and ultra performance liquid chromatography. Talanta *76*, 189-199.

Spelt,C., Quattrocchio,F., Mol,J.N.M., and Koes,R. (2000). anthocyanin1 of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. Plant Cell *12*, 1619-1631.

Spencer,J.P.E., Kuhnle,G.G.C., Hajirezaei,M., Mock,H.P., Sonnewald,U., and Rice-Evans,C. (2005). The genotypic variation of the antioxidant potential of different tomato varieties. Free Radical Research *39*, 1005-1016.

Stracke,R., Werber,M., and Weisshaar,B. (2001). The R2R3-MYB gene family in Arabidopsis thaliana. Current Opinion in Plant Biology *4*, 447-456.

Takos,A.M., Jaffe,F.W., Jacob,S.R., Bogs,J., Robinson,S.P., and Walker,A.R. (2006). Light-induced expression of a MYB gene regulates anthocyanin biosynthesis in red apples. Plant Physiology *142*, 1216-1232.

Tanaka,Y. and Ohmiya,A. (2008). Seeing is believing: engineering anthocyanin and carotenoid biosynthetic pathways. Current Opinion in Biotechnology *19*, 190-197.

Tanksley,S.D. and McCouch,S.R. (1997). Seed banks and molecular maps: Unlocking genetic potential from the wild. Science *277*, 1063-1066.

Tanksley,S.D., Medinafilho,H., and Rick,C.M. (1982). Use of Naturally-Occurring Enzyme Variation to Detect and Map Genes-Controlling Quantitative Traits in An Interspecific Backcross of Tomato. Heredity *49*, 11-25.

Tohge,T., Nishiyama,Y., Hirai,M.Y., Yano,M., Nakajima,J., Awazuhara,M., Inoue,E., Takahashi,H., Goodenowe,D.B., Kitayama,M., Noji,M., Yamazaki,M., and Saito,K. (2005). Functional genomics by integrated analysis of metabolome and transcriptome of Arabidopsis plants over-expressing an MYB transcription factor. Plant Journal *42*, 218-235.

Toledo-Ortiz,G., Huq,E., and Quail,P.H. (2003). The Arabidopsis basic/helix-loop-helix transcription factor family. Plant Cell *15*, 1749-1770.

Toor,R.K. and Savage,G.P. (2006). Changes in major antioxidant components of tomatoes during post-harvest storage. Food Chemistry *99*, 724-727.

Torres,C.A., Davies,N.M., Yanez,J.A., and Andrews,P.K. (2005). Disposition of selected flavonoids in fruit tissues of various tomato (Lycopersicon esculentum Mill.) genotypes. Journal of Agricultural and Food Chemistry *53*, 9536-9543.

Tucker,G. (2003). Nutritional enhancement of plants. Current Opinion in Biotechnology *14*, 221-225.

Vauzour,D., Rodriguez-Mateos,A., Corona,G., Oruna-Concha,M.J., and Spencer,J.P.E. (2010). Polyphenols and Human Health: Prevention of Disease and Mechanisms of Action. Nutrients *2*, 1106-1131.

Verhoeyen,M.E., Bovy,A., Collins,G., Muir,S., Robinson,S., De Vos,C.H.R., and Colliver,S. (2002). Increasing antioxidant levels in tomatoes through modification of the flavonoid biosynthetic pathway. Journal of Experimental Botany *53*, 2099-2106.

Vogt,T. (2010). Phenylpropanoid biosynthesis. Molecular Plant *3*, 2-20.

Weisshaar,B. and Jenkins,G.I. (1998). Phenylpropanoid biosynthesis and its regulation. Current Opinion in Plant Biology *1*, 251-257.

Whitaker,A.M. and Jones,C.S. (1965). Report OF 1500 CHLOROFORM ANESTHETICS ADMINISTERED WITH A PRECISION VAPORIZER. Anesth. Analg. *44*, 60-65.

Willcox,J.K., Catignani,G.L., and Lazarus,S. (2003). Tomatoes and cardiovascular health. Critical Reviews in Food Science and Nutrition *43*, 1-18.

Willits,M.G., Kramer,C.M., Prata,R.T.N., De Luca,V., Potter,B.G., Steffens,J.C., and Graser,G. (2005). Utilization of the genetic resources of wild species to create a nontransgenic high flavonoid tomato. Journal of Agricultural and Food Chemistry *53*, 1231-1236.

Winkel,B.S.J. (2008). The biosynthesis of flavonoids. In The Science of Flavonoids, E.Grotewold, ed. (NY: Springer Science), pp. 71-96.

Winkel-Shirley,B. (2001a). Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiology *126*, 485-493.

Winkel-Shirley,B. (2001b). It takes a garden. How work on diverse plant species has contributed to an understanding of flavonoid metabolism. Plant Physiology *127*, 1399-1404.

Xu,Y., Zhu,L., Xiao,J., Huang,N., and McCouch,S.R. (1997). Chromosomal regions associated with segregation distortion of molecular markers in F-2, backcross, doubled haploid, and recombinant inbred populations in rice (Oryza sativa L). Molecular & General Genetics *253*, 535-545.

Yu,O. and Jez,J.M. (2008). Nature's assembly line: biosynthesis of simple phenylpropanoids and polyketides. Plant Journal *54*, 750-762.

Zamir,D. (2001). Improving plant breeding with exotic genetic libraries. Nature Reviews Genetics *2*, 983-989.

Zhao,Z.H. and Moghadasian,M.H. (2008). Chemistry, natural sources, dietary intake and pharmacokinetic properties of ferulic acid: A review. Food Chemistry *109*, 691-702.

Zhou,Y., Wu,H.J., Zhang,Y.H., Sun,H.Y., Wong,T.M., and Li,G.R. (2011). Ionic mechanisms underlying cardiac toxicity of the organochloride solvent trichloromethane. Toxicology *290*, 295-304.