Molecular analysis of red colouration in 'Bon Rouge' pear *(Pyrus communis L.)*

by

Marlene Geraldine du Preez

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy, Department of Biotechnology, University of the Western Cape, South Africa.

August 2018

Supervisor: Prof. A. Christoffels

Co-supervisor: Dr. D. J. G. Rees

Acknowledgements

I would like to express my sincerest appreciation to my supervisors, Dr. D. J. G. Rees and Prof. A. Christoffels, for their support during the course of my PhD studies. In particular I would like to thank Dr. D. J. G. Rees for financial, infrastructural and research mentoring support, and Prof. A. Christoffels and his research group for bioinformatics support during the latter part of my studies. Without their support, this PhD study would not have been possible.

I acknowledge the financial support from the NRF, SA and the UWC Research fund, without which this research could not have been conducted.

I would also like to extend my thanks and appreciation to the members of the Biotechnology Department at UWC for their continued support throughout the research for this PhD study.

To my parents for their unwavering support throughout my academic career, and for instilling in me and my siblings, the value of a good education as a basis for a successful life. I would like to acknowledge the contribution of my late mother and father who encouraged me without fail, and for providing the financial means to afford their seven children a University education. To my siblings for their continued support and encouragement, I extend my heartfelt thanks.

To my husband and daughter for always believing in me, and their endless encouragement and emotional support during the course of my PhD study, I would like to say, thank you.

Dedication

This PhD thesis is dedicated to my late mother and father, my siblings, my husband and my daughter.

Table of contents

SUPPLIED IN ELECTRONIC FORMAT

List of figures

Page numbers

List of tables

Page number

List of appendixes

List of acronyms

AAA ATPase	ATPases associated with various cellular activities
ABA	Abscisic acid
ABP	Auxin binding proteins
ACC	1-aminocyclopropane carboxylic acid
AGO	ARGONAUTE
ACCO	1-aminocyclopropane carboxylic acid oxidase
AMT	Anthocyanin methyltransferases
AN1	Anthocyanin 1
AN11	Anthocyanin 11
AN2	Anthocyanin 2
AN4	Anthocyanin 4
AN9	Anthocyanin 9
ANL2	Anthocyaninless 2
ANR	Anthocyanidin reductase
ANS	Anthocyanidin synthase
ANT	Adenine nucleotide transporter
AOS	Active Oxygen Species
AP2	APETALA2
AP2/ERF-like	Apetala 2/Ethylene Response Factor1-like
APase	Acid phosphatase
APS	Adenosine 5-phosphosulfate

ARC	Agricultural Research Council
AREB	ABA response element binding protein
ARF	alpha-L-Arabinofuranosidase
ART	Anthocyanin rhamnosyl transferase
BLAST	Basic Local Alignment Search Tool
BSC	Bundle sheath cells
Bz	Bronze
bZIP	basic domain/leucine zipper
C4H	Cinnamate 4-hydroxylase
C4L	Cinnamate 4-ligase
CAM	Crassulacean acid metabolism
CBF	C-Repeat Binding Protein
CBF	Cold Binding Factor
CBL	Calcineurin B-like
CCDs	Carotenoid cleavage dioxygenases
cDNA	complementary DNA
CHI	Chalcone isomerase
CHLs	Chloroplastic lipocalins
CHS	Chalcone synthase
COP1	Constitutively photomorphogenic1
COP9	Constitutively photomorphogenic 9
CRE	Cytokinin receptor
CREB	Cytokinin receptor binding

CRT3	Calreticulin3
CsC	Castanea sativa Cystatin
CSN	COP9 signalosome
CSN5	CSN subunit 5
ctDNA	Calf thymus DNA
CTR1	Chymotrypsin-like protease
dCAPS	derived Cleaved Amplified Polymorphic Sequence
DCL1	Dicer-like 1
DCT	Plastidic dicarboxylic acid transporter
DDM1	Decreased DNA methylation1
DFR	Dihydroflavonol 4-reductase
DiT1	2-oxoglutarate/malate translocator
DNA	Deoxyribonucleic acid
DREB	Dehydration-responsive element binding
EBF	F Box proteins
EGFR	Epidermal growth-factor receptor
EGO	Enhancer of glp-1
EH	Epoxide hydrolase
EIL	Ethylene-Insensitive3-like
EIN2	Ethylene-Insensitive2
EIN3	Ethylene-Insensitive3
ELIP	Early Light Inducible Proteins
ERF1	Ethylene Response Factor1-like

EST	Expressed sequence tag
ET	Ethylene
F3',5'H	Flavonoid 3',5'-hydroxylase
F3'H	Flavonoid 3'-hydroxylase
F3H	Flavonoid 3-hydroxylase
FAD	Flavin adenine dinucleotide
$FADH_2$	Flavin adenine dinucleotide reduced
FHT	Flavanone 3-hydroxylase
FLS	Flavonol synthase
GDT	Dicarboxylic acid transporters
GEF	Guanine nucleotide exchange factors
GL	GLABRA
GOGAT	Glutamine/2-oxoglutarate aminotransferase
GST	Glutathione S-transferase
HEN	HUA ENHANCER1
HD-GL2	Homeodomain Glabra2
HLH	Helix-loop-helix
HPLC	High Performance Liquid Chromatography
HY5	Elongated hypocotyl5
HYL1	HYPONASTIC LEAVES1
ICE1	Inducer of CBF Expression1
ICXS	Increased chalcone synthase expression
INFRUITEC	Institute for Fruit Technology

JA	Jasmonic acid
KEGG	Kyoto Encyclopedia of Genes and Genomes
LAR	Leucoanthocyanidin reductase
LCAT	Lecithin Cholesterol Acyltransferase
LC MS	Liquid Chromatography Mass Spectrometry
LDOX	Leucoanthocyanidin dioxygenase
Lil3	Light inducible–like protein 3
LLR	Leucine-rich repeat
LQY1	Low Quantum Yield Of Photosystem I
LZF1	Light-Regulated Zinc Finger1
MAPK	Mitogen-activated protein kinase
MeJA	Methyl jasmonic acid
METS	Mitochondrial energy transfer signature
miRNA	microRNA
mRNA	messenger RNA
mRNAseq	messenger RNA sequencing
MS-MS	Mass Spectrometry- Mass Spectrometry (tandem MS)
mtCP	Mitochondrial carrier protein
MUM4	Mucilage Modified 4
NAD	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide (reduced)
NADP	Nicotinamide adenine dinucleotide phosphate
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)

NCBI	National Centre for Biotechnology Information
NMR	Nuclear Magnetic Resonance
OGs	Polygalacturonides
OHPs	One-helix proteins
OMT	Oxoglutarate-malate translocator
PAs	Proanthocyanidin pigments
PAL	Phenylalanine ammonia-lyase
PAPS	Adenosine 3'-phosphate 5'-phosphosulphate
PAZ	Piwi-Argonaute-Zwille
PGI	Polygalacturonase 1
PGIP	Polygalacturonase inhibitor protein
PGs	Endopolygalacturonases
PiC	Phosphate transporter
PS II	Photosystem II
PTGS	Post-Transcriptional Gene Silencing
Pvpp1	Phaseolus vulgaris plant protein phosphatase 1
Q	Ubiquinone
QDE	Quelling defective
QH_2	Ubiquinol
QRS	Glutaminyl tRNA synthetase
R2R3 MYB	Repeat 2, Repeat 3 MYB
RAD23	Radiation sensitive23
RCC1	Regulator of Chromatin Condensation 1

RDE	RNAi-deficient
RdRP	RNA-dependent RNA polymerases
RFLP	Restriction Fragment Length Polymorphism
RHM	Rhamnose synthase
RI	Expresssion ratio
RISC	RNA-Induced Silencing Complex
RNA	Ribonucleic Acid
RNAi	RNA interference
RNase	Ribonuclease
RNS1	Ribonucleases 1
RNS2	Ribonucleases 2
ROP1	Rho GTPases of plant
ROS	Reactive Oxygen Species
RP	Regulatory particle
RPKM	Reads Per Kilobase of exon model per Million mapped reads
RPT6	Regulatory Particle 6A
RT	Rhamnosyl transferase
SA	Salicylic acid
SAPs	Stress-associated proteins
SAR	Systemic acquired resistance
SDE	Silencing defective
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	SDS Polyacrylamide Gel Electrophoresis

SEPs	Stress enhanced proteins
siRNA	Short interfering RNA
SOD	Superoxide dismutase
SOS	Salt-Overly-Sensitive
SQD2	Sulfolipid synthase
SQDG	Sulfoquinovosyl diacylglycerol
SQDX	Bacterial sulfolipid synthase
SQR	Succinate-coenzyme Q reductase
SRP	Signal Recognition Particle
STH3	Salt Tolerance Homolog 3
TAD1	Temperature associated defensin1
TF	Transcription Factor
TGS	Transcriptional Gene Silencing
TIL	Temperature-Induced Lipocalins
TT	Transparent Testa
TTG	Transparent Testa Glabrous
TZF	CCCH Zinc Finger
Ub	Ubiquitin
UDP	Uridine diphosphate
UFGT	UDP-Glucose Flavonoid: 3-O-Glucosyltransferases
UP	Uncoupling Protein
UPS	Ubiquitin/26S Proteasome System
USP15	Deubiquitinase

UV-B	Ultra violet-B
UVR8	UV resistance locus
V-ATPase	Vacuolar-type H^+-transporting adenosine triphosphatase
V-PPase	Vacuolar H^+pyrophosphatase
VCP	Valosin-containing protein
VDEs	Violaxanthin de-epoxidases
VHAc"	Vacuolar H^+-ATPase c"
WD	Tryptophan-Aspartic acid
WRKY	Tryptophan-Arginine-Lysine-Tyrosine
YGM-3	3-(6,6^1-cafeylferulylsophoroside)-5-glucoside of cyanidin
YGM-6	3-(6,6^1-cafeylferulylsophoroside)-5-glucoside of peonidin
ZEPs	Zeaxanthin epoxidases
ZFP-like	Zinc finger protein-like

ABSTRACT

The 'Bon Rouge' pear cultivar was developed from a bud mutation on a 'Bon Chretien' pear tree. The latter is characterised by green fruit skin and leaves, while 'Bon Rouge' is characterised by red leaves and red fruit skin as a result of the production of anthocyanin and other pigments. Branch forming buds on 'Bon Rouge' trees often revert to the parent phenotype producing green leaves and fruit skin. The occurrence of both phenotypes on the same tree presents a unique model to study gene expression associated with anthocyanin production in a similar genetic background under the same set of environmental condition.

To elucidate the difference in the underlying molecular mechanism for anthocyanin production in 'Bon Rouge' and its reverted phenotype, we performed a comparative gene expression analysis using differential display (DD), and whole transcriptome analysis by mRNA sequencing (mRNAseq). The aim of this strategy was to identify a controlling element responsible for the difference in anthocyanin production between the two phenotypes of 'Bon Rouge'. Additionally, we characterised the pigment profiles of the two phenotypes by High Performance Liquid Chromatography (HPLC), Liquid Chromatography Mass Spectrometry (LC-MS) and Nuclear Magnetic Resonance NMR.

Differential expression analysis between the two phenotypes identified a number of genes associated with the stress response in plants, and one that could be linked to anthocyanin production. A number of the structural genes for anthocyanin production were upregulated in the red compared to the green phenotype. Furthermore, pigment characterisation confirmed the presence of the red anthocyanin pigment in 'Bon Rouge' and its absence in the green phenotype.

Key words: 'Bon Rouge', pear, mRNAseq, differential display, anthocyanin, High Performance Liquid Chromatography (HPLC), Liquid Chromatography Mass Spectrometry (LC-MS), Nuclear Magnetic Resonance NMR.

RESEARCH STATEMENT

The 'Bon Rouge' pear cultivar originated in South Africa and is a mutation of the Bon Chretien cultivar. 'Bon Rouge' is characterised by brownish-red leaves at bud break and its red fruit skin colour. However, branches of 'Bon Rouge' pear trees can revert to produce green coloured fruit that is characteristic of its Bon Chretien parent. The frequency of the localised bud mutations that result in the green coloured fruit of reverted 'Bon Rouge' pear tree branches is too high to be spontaneous and therefore we hypothesise that a transposable element is involved in the reversion process. The stabilisation of fruit colour in this important cultivar will have significant advantages for the 'Bon Rouge' and possibly other cultivars which rely heavily on this qualitative trait for desirability on local and export markets.

Climacteric fruit are subject to the daily and seasonal ranges of temperature, light conditions and water supply. This has implications for a variety of characteristics, including red fruit skin colour that is a value added property in many fruit cultivars including those under commercial production in the Western Cape region. The production of colour pigment, namely anthocyanin, can also be induced under other conditions of stress such as pathogen attack and photo-oxidative stress implicating anthocyanin production in a variety of signalling pathways which intersect at various points. Consequently, the regulation of anthocyanin production is complex. Furthermore, unravelling the interaction of

genes responsible for the regulatory and biosynthetic pathways in anthocyanin production is compounded by the fact that production of these pigments is developmentally regulated in a tissue specific manner with some variation between divergent plant species. In unravelling the complexity of anthocyanin production, the identification of genes that are responsible for this desired fruit skin colour which can be manipulated to allow for the development of a more stable fruit skin colour phenotype under storage, and variable temperature and water supply conditions, should prove extremely valuable for the fruit production industry. Additionally, improvement in the selection of clonal material for propagation from stock material should decrease the propagation of grafted materials that do not exhibit the desired phenotypic trait.

Red fruit skin colour is a value added quality in many fruit crops. Pear is an important commercial fruit crop for export and local markets, and red pears and apples have higher commercial value than their green counterparts. Fruit skin colour is produced by a blend of carotenoids, chlorophyll and anthocyanin (Lancaster, 1992) with the latter contributing to red skin colour. Pre-selection of fruit skin colour at the seedling stage would be highly advantageous in the deciduous fruit tree industry (Cheng *et al*., 1996) where development to full production can take up to five years. Flavonoids such as anthocyanin are plant-specific compounds (Harborne, 1976) that accumulate in almost all tissue from mosses to flowering plants (Koes *et al*., 1994) and are classified as secondary metabolites (natural products) with primary ecological functions (Croteau *et al*.,

2000). As common colorants in flowers, they attract pollinators and have been implicated in male sterility through lack of accumulation of flavonol, an intermediate in the anthocyanin biosynthetic pathway, in pollen (Napoli *et al*., 1999). Recent nutritional and epidemiological studies have described various beneficial health effects of anthocyanin consumption for protection against certain cancers, cardiovascular disease and ageing (Sehitoglu *et al*., 2014, Pérez-Hernández *et al*., 2016). Anthocyanin production is also elicited in response to nutrient deficiencies such as phosphate starvation, and pathogen attack (Dixon and Paiva, 1995; Zakhleniuk *et al*., 2001). The genes encoding the biosynthetic enzymes of the anthocyanin pathway and their respective regulators have been well characterised, and show high similarity between related species (Sainz *et al*., 1997a). The structural genes of the anthocyanin biosynthetic pathway in pear has been characterised (Thilo *et al*., 2007) but a limited number of the regulatory genes have thus far been identified (Pierantoni *et al*., 2010).

To elucidate the difference in the underlying molecular mechanism for anthocyanin production in 'Bon Rouge' and its reverted phenotype, we are performing a comparative gene expression analysis using differential display, and whole transcriptome analysis by RNA sequencing (mRNAseq). The aim of this strategy is to identify a controlling element responsible for the difference in anthocyanin production between the two phenotypes of 'Bon Rouge'.

I declare that '**Molecular Studies in 'Bon Rouge' pears**' is my own work, that it has not been submitted before for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged as complete references.

Marlene Geraldine du Preez August 2018

CHAPTER 1

A REVIEW OF THE LITERATURE

1.1 Introduction

Phenylpropanoids are versatile compounds that serve unique and general functions both in plants and animals. They are products of a secondary metabolic pathway that uses products from the primary shikimic acid pathway in plants. Phenylpropanoids constitute a large number of compounds such as flavonols, phytoalexins, and flavonoids such as anthocyanidins and anthocyanins that have varied functions in both plants and animals. In plants, one of the major functions of anthocyanin is to impart colour to flowers, vegetables, leaves and fruit (Kong *et al*., 2003). Furthermore, they play an important role in the attraction of animals for pollination and seed dispersal and as such, play a significant role in the co-evolution of plant-animal interactions. Anthocyanins and 3-deoxyanthocyaninidins have roles in plants other than as attractants and can act as antioxidants and antibacterial agents. Anthocyanins together with other flavonoids are also important factors for plant resistance to insect attack (Harborne, 1988).

In plants phenylpropanoids such as flavonols, flavonoids such as anthocyanin and 3-deoxyanthocyanidin, and phytoalexins are produced in response to various stress situations (Dixon and Paiva, 1995; Winkel-Shirley 2002). For instance, mechanical wounding and wounding by insects and other herbivores, produce an increase in coumestrol, coumarin, psoralens, chlorogenic acid, ferulate esters, lignin, suberin and wall bound phenolic acids whereas pathogen attack produce an increase in pterocarpans, isoflavans, prenylated isoflavonoids, stilbenes, coumarins, furanocoumarins, flavanols, aurones and 3-deoxyanthocyanidins. Low nitrogen induces the production of flavonoids and isoflavonoids and low iron, phenolic acids. Salicylic acid, another product of the phenylpropanoid biosynthetic pathway, is produced in response to numerous biotic and abiotic stressors.

Although flavonoids, which comprise a subset of phenylpropanoids, have common functions in many different plant species, some (functions) are more limited and appear to have evolved differently or independently in certain lineages. For instance, isoflavonoids, which are important in plant defense and as signalling molecules in nitrogen fixation, occur only in legumes and in a few non-legume plants (Winkel-Shirley, 2001). Likewise, 3-deoxyanthocyanins that are involved in pigmentation (Grotewold *et al*., 1994) and defense (Snyder and Nicholson, 1990) are produced only in a few species such as sorghum, maize and gloxinia. Mutations in the first enzyme of the flavonoid biosynthetic pathway in maize and petunia lines, uncovered a role for flavonoids in male fertility. However, a null mutant affecting the same enzyme in *Arabidopsis* was fully fertile indicating that flavonoids are not universally required

for pollen tube formation (Burbulis *et al*., 1996). Despite these anomalies, flavonoids play a number of apparent universal roles in the control of many processes during plant development. These include shading the photosynthetic apparatus for protection against UV radiation and scavenging of free radicals produced during oxidative stress (Harborne and Williams 2000). Esters of hydroxycinnamic acids that comprise the major group of phenylpropanoids, are incorporated into the cell wall where they act as defense compounds or as precursors of wound-inducible lignin and suberin. Some phenolic compounds act as antifungal agents and as internal signal molecules that can alter the pattern of gene expression. Hydroxycinnamic acids and their conjugates accumulate to high levels in all tissues and have a major role in allelopathy, defined as any direct or indirect beneficial or harmful effect of one organism on another, and mediated by the release of allelochemicals into their immediate environment (Ibrahim, 1999). Such allelopathic effects may be enhanced in nutrient-poor soil such as the increased levels of chlorogenic acid in sunflower, and juglone in walnut trees when grown in nitrogen-, phosphorus-, potassium- and/or sulphur-deficient soils. The increased accumulation of anthocyanin in the vegetative tissue of many plant species due to a lack of available phosphate (Winkel-Shirley, 2002), may represent another example of such allelopathic effects. *Arabidopsis* offers a simple example of the complex flavonoid pathway since all but one of the biosynthetic enzymes of the flavonoid pathway is encoded by single genes, unlike the situation in other plants (Winkel-Shirley, 2001). Additionally, the pathway in *Arabidopsis* can be analysed by genetic and molecular resources that may not be readily available for other plant species.

1.2 The molecular biology of anthocyanin synthesis

Anthocyanins are the red and blue pigments that give flowers fruit and leaves colour. They are water-soluble sugar phenols originally derived from the amino acid phenylalanine, a product of the primary shikimic acid pathway. Anthocyanin pigments have varied and numerous biological functions in plants (Harborne, 1976). These include the attraction of insects and animals for pollination and seed dispersal (Kong *et al*., 2003), and the provision of protection against UV irradiation and pathogen attack (Winkel-Shirley, 2002). They have also been implicated in the role of hormone transport regulators (Sablowski *et al*., 1994). Anthocyanin pigments have beneficial properties such as anti-tumorogenic, antimicrobial and antioxidant properties, in man and other animal (Kong *et al*., 2003).

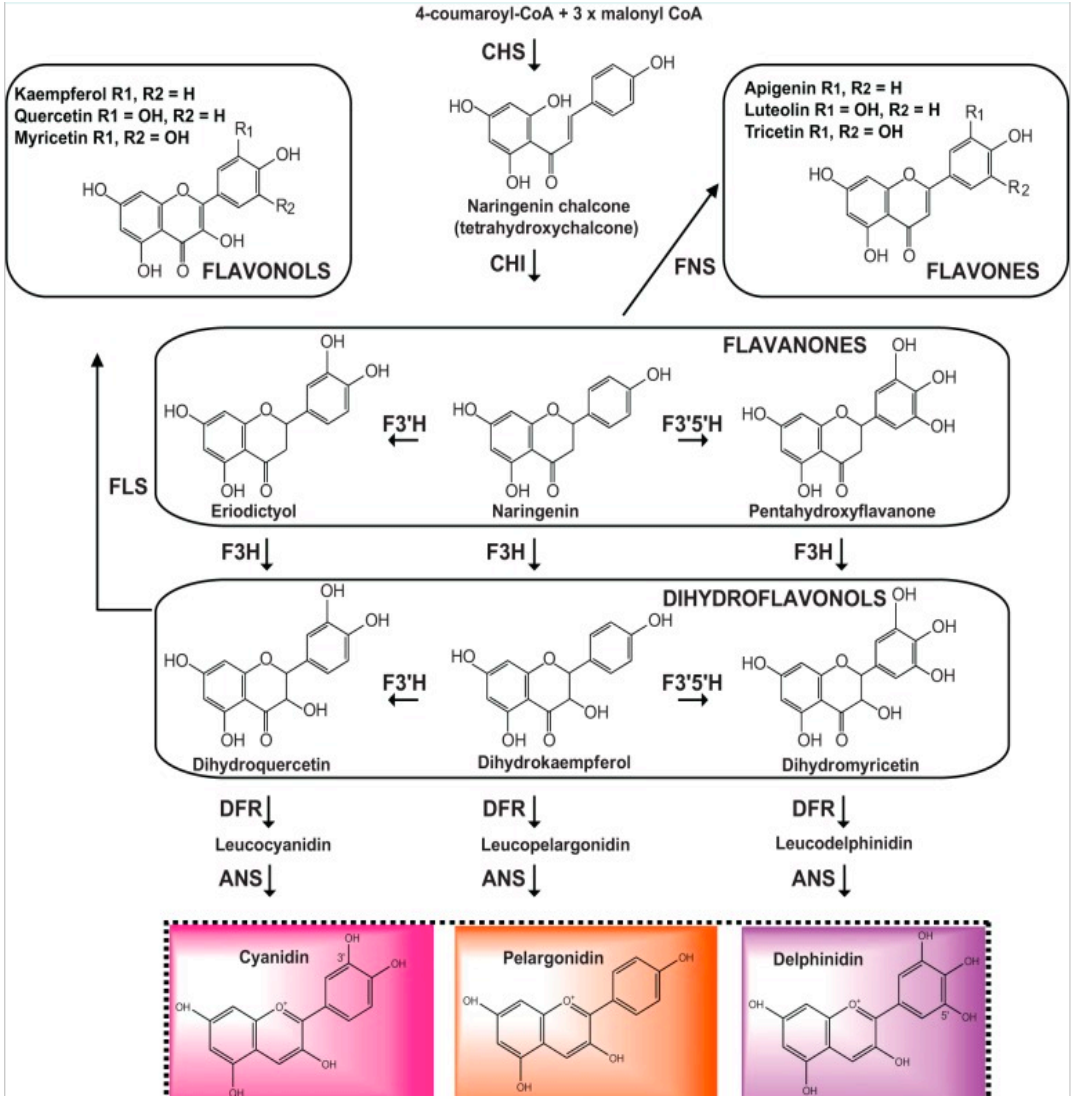


Figure 1.1 Major genes and products of the anthocyanin biosynthetic pathway: cyanidin, pelargonidin and delphinidin. CHS, Chalcone synthase; CHI, chalcone isomerase; FNS, flavone synthase; F3H, flavonoid-3-hydroxylase; F3'H, flavonoid-3'-hydroxylase; F3'5'H, flavonoid 3',5'-hydroxylase; FLS, flavonol synthase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase (from Falcone Ferreyra *et al*., 2012).

Red anthocyanins (anthocyanidin 3-*O*-glucosides) are the first stable coloured metabolites produced in the early-stage reactions of anthocyanin production (see Fig. 1). Violet pigments (anthocyanidin 3, 5-*O*-glucosides) are produced in the late-stage reactions that involve further modification such as acylation, methylation and further glycosylation (Yamazaki *et al*., 1999). Glycosylation at the 5-OH position has been shown to promote stable complex formation in the co-pigmentation of anthocyanins, resulting in a reddish-purple colour (Yamazaki *et al*., 1999) while the late-stage reactions are responsible for the finer adjustment of colour.

The anthocyanin pigment profile varies from one species to the next, for example, each grape variety or species has a unique anthocyanin profile as determined by HPLC (Boss *et al*., 1996b). The same is true for various other plant species that produce anthocyanin pigments. Such profiles can be used for chemotaxonomic identification and classification of grape species and varieties. This uniqueness in pigment production is demonstrated for various plant species. For instance, petunias do not normally produce pelargonidin (brick-red) pigments, whereas snapdragon and maize cannot produce delphinidin (blue) pigments due to the lack of enzymes catalysing these late colour adjustment reactions. Anthocyanins determine the degree of astringency and quality of colour in red wines.

The genetic basis of anthocyanin pigment production began with Mendel's work on flower colour in pea in the 1800s and the study has been expanded to plant species such as maize (*Zea mays*), petunia (*Petunia hybrida*), snapdragon

(*Antirrhinum majus*), perilla (*Perilla frutenscens*) a herb plant from the east used in local traditional medicines and its pigment as food colorant) and grape (*Vitis vinifera*). The number of genes responsible for the production of coloured pigments is large and there are at least 35 in petunia alone (Wiering and de Vlaming, 1984). The pathway is an attractive model system for secondary metabolite production studies in plants and this has been facilitated by the availability of large numbers of non-lethal anthocyanin production mutants and the fact that these pigments are visible markers that can be easily scored visually or by basic biochemical applications.

1.3 Structural genes for anthocyanin production

Anthocyanin production in plants requires two sets of genes – regulatory and structural. Anthocyanin production occurs as part of the general phenylpropanoid biosynthetic pathway that is responsible for the production of polyphenolic compounds such as lignins, tannins and other flavonoids that are produced via branchpoints in the general pathway. The first flavonoid biosynthetic pathway gene isolated was the chalcone synthase (*CHS*) gene from parsley (*Petrosileum crisum*) (Cornish and Holton, 1995). The gene product is also the first committed enzyme in anthocyanin biosynthesis. Dihydroflavonols are the precursors of both anthocyanins and flavonols. Consequently there is potential for competition between flavonoid metabolising enzymes for common substrates (Holton and Cornish, 1995). Dihydroflavonols are predominantly located in the cytosol

whereas anthocyanins are sequestered in the vacuole after modification by widely divergent glutathione S-transferases (GST) (Beld *et al*., 1989). This is to prevent their potential toxic effect in the cytosol. At least eight key enzymes (see Fig. 1) are required for anthocyanin production and most of them have been well characterised by cloning and sequencing in a number of species such as petunia, maize and *Arabidopsis*. Table 1.1a and 1.1b provide a summary of details such as gene loci and structure, and enzyme function. The following section highlights some specifics regarding enzyme function and gene structure of some of the key proteins in the anthocyanin biosynthetic pathway.

1.3.1. Gene copy number and structure

Most of the enzymes involved in the anthocyanin biosynthetic pathway are encoded by multi-gene families (Harborne and Williams 2000). There is speculation in the literature regarding the origin of these multi-gene families. The current consensus supports the hypothesis that they arose from a single ancestral gene-by-gene duplication and re-arrangement. For example, phenylalanine ammonia-lyase (PAL) is a member of a large gene family in grape where least twelve members have been identified (Quan *et al*., 2013), and there are twelve different CHS genes in petunia (Holton and Cornish, 1995). The intron-exon structure of these genes is generally conserved across species. This is exemplified by the dihydroflavonol reductase (DFR) gene of bean which contains three introns in the same positions as the maize DFR gene that corresponds to the first three of

five introns present in the gene from petunia and snapdragon (Kristiansen and Rohde, 1991). The 3-*O*-glucosyl transferase (3-GT) gene contains a single intron whose position appears to be conserved in *Arabidopsis thaliana* and barley (*Hordeum vulgare*). The gene is the least conserved of all the anthocyanin structural genes (Sparvoli *et al*., 1994) but contains a signature motif common to all glycosyl tranferases including 5-*O*-glucosyltransferase (5-GT) (Yamazaki *et al*., 1999).

1.3.2. Genetics and enzyme function of late stage anthocyanin structural genes

Anthocyanidin synthase (ANS) and glycosyltransferases (3-GT, 5-GT and ART) are involved in the late stages of anthocyanin production.

1.3.2.1 Anthocyanidin synthase (ANS)

ANS catalyses the transformation of colourless leucoanthocyanidins to coloured anthocyanidins.. This is proposed to involve two steps: in the first step, ANS removes an hydroxyl group under acidic conditions from the basic ring structure; a dehydratase is involved in the second step which results in the formation of a double bond between C-3 and C-4 on the flavonoid ring (Heller and Forkmann, 1988; Boss *et al*., 1996b). However, *in vitro* studies have shown that the hydration-dehydration step can be catalysed by acid without requirement for other enzymes (Saito *et al*., 1999). ANS has sequence similarity with 2-oxoglutarate-

dependent dioxygenases including flavone-3-hydroxylase (F-3-H), flavonol synthase (FLS), and amino-cyclopropane-1-carboxylate (ACC) oxidase (Holton and Cornish, 1995).

1.3.2.2. UDP-Glucose flavonoid: 3-*O*-glucosyltransferases (3-GT)

UDP-Glucose flavonoid: 3-*O*-glucosyltransferases (3-GT) plays an important role in metabolite accumulation. The latter enzyme specifically glucosylates anthocyanidins, rather than flavonoids as is suggested by its name, at the 3-*O* position during red fruit ripening to produce the first stable red pigment (Ford *et al*., 1998). Glycosylation increases solubility in aqueous solutions and reduces chemical reactivity. Glycosylated compounds are transportable storage compounds that are preferentially sequestered in the vacuole, or waste products that can be removed from the cytosol (Ford *et al*., 1998). It has been proposed that the full activity of ANS and UFGT requires a multi-enzyme complex (Saito *et al*., 1999). UF:3-GT is encoded by *Bronze*-1 (*Bz*1 allele in maize) (Furte*k et al*., 1988). Sequence polymorphisms among three *Bz*1 alleles include deletions/additions, a transposable element insertion upstream of the promoter region and single base pair substitution. Mutable *Bz*1 alleles with transposable elements belonging to the *Ac*/*Ds*, *Spm*/*dSpm*, *Mu*, *Cy*/*rcy* and *Mut* families have been described by Furtek *et al*. (1988). Callis *et al*. (1987) have demonstrated an enhancement of gene expression by the *Bz*1 intron.

1.3.2.3. Anthocyanin rhamnosyl transferase (ART)

Like 3-GT, ART (or 3-RT) is a sugar transferase and catalyses rhamnosylation of anthocyanin. It adds a rhamnose group to the 3-*O*-glucose position to produce anthocyanin 3-rutinosides. This enzyme shows some homology to glucosyltransferases and is encoded by the *rt* or the *difG* gene in petunia (Bugliera *et al*., 1994; Kroon *et al*., 1994). Anti-sense suppression of ART results in varied phenotypes such as wild type, uniform loss of pink colour, or variegation with red or purple sectors. Two mutant *rt* alleles in petunia contain transposable elements. *dTph1*, a transposable element from the *Ac*/*Ds* family, inserted into the promoter region of ART blocks transcription, and *dTph3* insertion into the coding region results in premature polyadenylation of transcript. Both mutations cause a variegated phenotype in somatic tissue and reversions in their progeny.

1.3.2.4. UDP Glucose flavonoid 5-*O*-glucosyltransferases (5-GT)

The production of purple colour in fruits and flowers results from the modification of anthocyanin by 5-GT or UF:5-GT. The enzyme is a sugar transferase that belongs to the glycosyltransferase superfamily implicated in the production of anthocyanin, auxin metabolism and other as yet uncharacterised functions (Gong *et al*., 1999b). The 5-GT gene has been isolated and fully characterised in perilla (Yamazaki *et al*., 1999), and partially characterised for red campion (*Silene diocia*) (Kamsteeg *et al*., 1978), petunia, stocks (*Matthiola incana*) and Chinese aster (*Callistephus chinensis*) (Seyffort, 1982). The enzyme demonstrates broad substrate specificity with respect to the sugars attached to the

hydroxyl groups on the three ring structures. Both 3-GT and 5-GT may have evolved from a common ancestor to define specificity for the glycosylation of a particular hydroxyl group (Yamazaki *et al*., 1999). A single 5-GT RFLP has been detected between red and green forms of perilla by Southern hybridisation (Yamazaki *et al*., 1999) and a number of RFLPs for CHS, DFR and 3-GT (Gong *et al*., 1997). However, despite these differences in gene expression between these two colour phenotypes, no significant difference in genomic organisation of the gene has been detected (Saito *et al*., 1999).

1.3.3. Additional structural genes

1.3.3.1 Flavonoid 3',5'-hydroxylase (F3,5H)

The purple/blue flowers of delphinidins results from the initial hydroxylation of the basic flavonoid ring structure at the 3',5' positions. This hydroxylation is catalysed by the enzyme flavonoid 3',5'-hydroxylase, a cytochrome P450 enzyme that is encoded by the *hf1* and *hf2* loci in petunia. It catalyses the 3',5' NADPH- and oxygen-dependent hydroxylation of dihydroflavonols to produce leucoanthocyanidins which are the precursors of purple anthocyanins. Roses and carnations lack this enzyme and are therefore unable to produce purple or blue flowers naturally. The gene has recently been expressed in rose (*Rosacea*) to produce the characteristic blue coloration. The activity of flavonoid 3',5'-hydroxylase is regulated by the *difF* locus which encodes cytb5, a heme-containing, membrane bound protein that is hypothesised to act as an alternative

electron donor to NADPH cyt P450 reductase which, in turn is required for P450 activity (de Vettten *et al*., 1999). F3',5'H is encoded by the *Pr* locus in maize (Coe *et al*., 1988). The aleurone of *Pr* plants is purple due to the accumulation of mostly cyanidin glucosides, whereas the aleurone of *pr* plants is red due to the accumulation of mostly pelargonidin glucosides (Holton and Cornish, 1995).

1.3.3.2. Anthocyanin methyltransferases (AMT)

Genetic loci in petunia *Mt1*, *Mt2*, *Mf1* and *Mf2* control the activity of four different methyltransferase isoenzymes. *Mt1* and *Mt2* control methylation only at the 3'position, whereas *Mf1* and *Mf2* control methylation at the 3' and 5' positions (Jonsson *et al*., 1983; Wiering, 1984). An additional petunia methyltransferase gene has been cloned, but its locus remains to be identified (Quattrocchio *et al*., 1999).

1.3.3.3. Glutathione S-transferase (GST)

Glutathione S-transferase catalyses which is currently viewed as the last genetically defined step in anthocyanin production (Marrs *et al*., 1995, Passeri *et al*., 2016). The enzyme is responsible for the transport of anthocyanin into the vacuole by catalysing its conjugation to glutathione, a process that facilitates its transport. The *An13* (*difI*) locus encodes glutathione transferase in petunia (Alfineto *et al*., 1998; de Vetten, 1999) while the *bronze2* (*Bz2*) gene of maize (Marrs *et al*., 1995) shares homology with stress-related proteins such as GST (Schmitz and Theres, 1992). Recessive mutations of the *Bz2* gene result in a

change of vegetative colour from purple to bronze/red-brown (Neuffer *et al*., 1968).

1.4 The regulatory genes of the anthocyanin biosynthetic pathway

Anthocyanin biosynthesis is controlled by at least two distinct classes of regulatory genes, each of which comprises multigene families. These regulatory genes encode the Myb- and Myc-like transcription factors (TFs), and were first discovered in maize. The two transcription factors act together to activate the expression of a number of anthocyanin biosynthesis structural genes. Aspects of Myb and Myc transcription factors, their gene targets and gene loci are discussed below.

1.4.1. *Cl*/*Pl* (Myb) gene family transcription factors

The Myb transcription factors represent a large family of genes with a diverse set of functions. Functions for most of these genes are unknown, but some play key roles in the regulation of secondary metabolite production in plants (Passeri *et al*., 2016), control of cell shape, disease resistance and hormone responses. Unlike animals, plants express a large repertoire of Myb proteins. Petunia expresses between 20 and 40 Myb genes (Avila *et al*., 1993) and *Arabidopsis* approximately 100 (Uimari and Strommer, 1997). Myb proteins bind the promoter regions of structural genes and induce conformational changes in the target DNA binding site (Solano *et al*., 1995). Mybs show high affinity for selectively nicked DNA that has enhanced DNA

flexibility (Solano *et al*., 1995). Plant Mybs have two imperfect 51/52 amino acid repeats II and III located at the conserved N-terminal end. Each repeat contains a helix-helix-turn-helix motif and three conserved tryptophan residues spaced 18 or 19 amino acids apart. The repeat motif is involved in sequence-specific DNA binding. Myb-like genes from perilla (Myb-*p1*) (Gong *et al*., 1999b) and potato (*MYB-St1*) (Baranowskij *et al*., 1994) have only one DNA binding motif, repeat III, yet they are fully functional as transcriptional activators. Studies using a yeast two-hybrid system have demonstrated interaction of both Mybs with Myc (Gong *et al*., 1999a).

In *Arabidopsis thaliana*, the R2R3 MYB transcription family is one of the largest families of transcription factors (Romero *et al*., 1998). The MYB domain contain three imperfect repeats, R1, R2 and R3 and analysis of protein-protein interaction specificities revealed a conserved amino acid signature ([DE]Lx_2 [RK]$x_3Lx_6Lx_3R$) for R3 (Zimmerman *et al*., 2004). Each imperfect repeat contains a helix-turn-helix signature of approximately 53 amino acids and the conserved amino acid signature is located on helices 1 and 2 of R3. Specific plant R2R3 MYB transcription factors requires interaction with specific basic helix-loop-helix (bHLH) transcription factor, MYCs, to induce anthocyanin biosynthesis (Goff *et al*., 1992) and certain MYB and MYC proteins can physically interact with each other (Quattrocchio *et al*., 1998) to regulate flavonoid biosynthesis in planta.

Recently, a number of papers reported on the role of MYB transcription factors in colouration of apple skin and flesh, and pear skin (Yang *et al*., 2015). The first

apple R2R3 MYB transcription factor *MdMYB1*, was isolated from *Malus domestica* Borkh. cv Cripps' Pink (Takos *et al*., 2006). Deduced amino acid sequence analysis indicated the gene encodes an ortholog of anthocyanin transcription factors in other plant species. *MdMYB1* is induced by light and regulate anthocyanin biosynthesis in apple fruit skin (Takos *et al*., 2006). In grape cells, MYB1, in combination with the product of the bHLH partner *AtEGL3,* induced transcription from the promoters of *MdDFR* and *MdUFGT*, two key enzymes for flavonoid biosynthesis in apple (Treutter *et al*., 2005). In the red fruit skin cultivar, Cripps' Red, a number of the genes for anthocyanin biosynthesis were coordinately upregulated in response to light (Takos *et al*., 2006). A derived cleaved amplified polymorphic sequence (dCAPS) marker for an allele of MdMYB1 displayed segregation with red fruit skin colour (Takos *et al*., 2006) and this marker will be useful in breeding programmes aimed at the development of red blush cultivars.

A second paper delineating the regulation of red colouration in apple fruit cortex and skin, identified a MYB transcription factor gene, *MdMYB10*, that has sequence similarity to known MYB anthocyanin transcription factors in other plant species (Espley *et al*., 2007). MdMYB10 can induce anthocyanin biosynthesis in both heterologous and homologous systems, generating sectors of pigmented areas in transient assays in tobacco leaves and intensely pigmented apple plants after transformed with constitutively expressed MdMYB10 (Espley *et al*., 2007). Induction of anthocyanin biosynthesis in transient assays by MdMYB10 showed an absolute requirement for co-expression of two distinct bHLH apple transcription

factors, MdbHLH3 and MdbHLH33. In addition, this paper reports on the regulation of transcript levels of anthocyanin biosynthetic genes in a red-fleshed, red-skinned and red-leaved cultivar, 'Red Field' compared with a white-fleshed, red-skinned, green-leaved cultivar, Pacific Rose™. Quantitative PCR analysis for apple biosynthetic genes Chalcone synthase (*CHS*), Chalcone isomerase (*CHI*), Flavone 3β-hydroxylase (*F3H*), Dihydroflavonol 4-reductase (*DFR*), Leucoanthocyanidin dioxygenase (*LDOX*) and Uridine diphosphate (UDP)-glucose:flavonoid 3-*O*-glycosyltransferase (*UFGT*) confirmed induction of all the genes of the anthocyanin biosynthetic pathway in cortex and skin of 'Red Field' compared to that of Pacific Rose™. However, measurement by quantitative PCR indicated that all of the genes of the pathway were not significantly differentially induced in the leaves of the red-leaf "Red Field' or the green-leaf Pacific Rose™.

Candidate gene mapping in a population segregating for the red flesh and foliage phenotype, identified the *Rni* locus as the major genetic determinant of red leaf and core of apple fruit (Chagné *et al*., 2007) and the *MdMYB10* gene (Espley *et al*., 2007) co-segregates with the *Rni* locus. *MdMYB10* could be the gene underlying the Rni locus since there were no recombinants between the marker for this gene and red phenotype in a population of 516 individuals (Chagné *et al*., 2007). Pierantoni *et al*. (2010) have demonstrated that a MYB10 identified from pear, *PcMYB10*, cannot be directly linked to red colour production in an anthocyanin producing mutant, Max Red Bartlett, of the green skinned European pear, Max Bartlett. Recently, Li *et al*., (2012) have reported a positive correlation between anthocyanin biosynthetic gene expression and the PcMYB10 transcription factor in pears, and

Wu *et al*., (2012) reported a R2R3 MYB transcription factor (PcMYB10) that was strongly positively correlated with anthocyanin accumulation in 'Wujiuxiang' pears in response to both developmental and cold-temperature induction. However, anthocyanin biosynthesis and accumulation in pear are more complicated compared to other plants, and are determined by fruit cultivar, maturity and environmental factors (Wu *et al*., 2012).

1.4.2. Myc (*R*/*B*) gene family of transcription factors

The *myc* family of transcription factors contain a basic helix-loop-helix (bHLH) leucine zipper domain that is involved in protein dimerization and DNA binding. In addition, it contains proline-rich, glutamine-rich and C-terminal acidic domains for transcriptional activation. The N-terminal and bHLH domains of these transcription factors are well conserved (Gong *et al*., 1999a) and the regions between these two domains are negatively charged. The N-terminus is believed to be the MYB-binding or interaction domain that is divided into two sub-domains containing three and two conserved tryptophans, respectively. These residues may form part of the hydrophobic core. Snapdragon Myc select C-box (core sequence CACGTG) and G-box (core sequence (C)ACGT(G) that matches the Myb consensus sequence binding site elements (ACE) as homo-dimers (Martinez-Garcia *et al*., 1998) while hetero-dimerization reduces DNA binding affinity.

1.4.3. Myb and Myc gene loci

Table 1.2a and 1.2b summarise the gene loci encoding Myb and Myc and their target genes in different plant species. This table is based on the one presented by Holton and Cornish (1995). The maize *C1* was the first Myb-like gene to be isolated from plants and is required for pigment accumulation in the aleurone, whereas *Pl* is active in vegetative tissue where it performs the same function. *C1* interacts with the MYC B protein, P, Zm 1 and Zm 38 (Cone *et al*., 1993; Franken *et al*., 1994). The *Delila* MYC-like gene of snapdragon controls the red pigmentation pattern (Jackson *et al*., 1991), however, it is required for pigmentation in the petal tube, but not in the petal lobe. MYB-like *Rosea* that interacts with *Delila*, appears to be functionally equivalent to *C1* of maize (Martin and Gerats, 1993).

1.4.4. Additional (non-Myb, non-Myc) regulatory genes

Additional regulatory genes appear to function upstream of the above regulatory genes. *An11* encodes a WD-40 protein (de Vetten *et al*., 1997) that is related to the β subunit of heterotrimeric G proteins. AN11 controls anthocyanin production exclusively in the flower. AN2, a Myb transcription factor, can complement an *an11* mutation suggesting that AN11 regulates *AN2* (Spelt *et al*., 2000). The *TRANSPARENT TESTA GLABRA1* (*TTG1*) gene encode a WD-40 repeat protein that regulates trichome differentiation (leaf hair formation) and anthocyanin biosynthesis in the seed coat of *Arabidopsis* (Walker *et al*., 1999). *ttg1* mutants are blocked at the DFR step (Fig. 1), the same as for the *Delila* mutation in snapdragon. *TTG1* is similar to *An11* of petunia but is not related to Myc transcription factors

although Lloyd *et al*. (1992) have demonstrated complementation of *ttg1* mutants with constitutive expression of the maize *R* gene. This implies that Myb could play a role in signal transduction to downstream Myc transcription factors. This apparently conflicts with evidence that An11 in petunia regulates Myb, but it is possible that TTG1 regulates a Myb-Myc complex (Walker *et al*., 1999). The *ANTHOCYANINLESS 2* (*ANL2*) gene encodes a homeobox protein involved in anthocyanin pigment production in the sub-epidermal layer of the leaf, and root development in *Arabidopsis*. It belongs to the HD-GL2 group of homeodomain proteins and contains a conserved helix-turn-helix motif for DNA binding, and a proline-glutamine stretch and an acidic region implicated in transcriptional activation (Kubo *et al*., 1999).

CHALCONE SYNTHASE (*CHS*) encodes a key enzyme in phenylpropanoid synthesis and is the committing enzyme in anthocyanin biosynthesis. Its activation by low levels of UV-B is well characterised and the activation pathway does not require known photoreceptors but involves ELONGATED HYPOCOTYL5 (HY5) (Ulm *et al*., 2005), a basic domain/leucine zipper (bZIP) transcription factor and a key regulator of photomorphogenesis under all light qualities (Ulm and Nagy, 2005) including DNA damaging UV-B. In *Arabidopsis*, UVR8 regulates a range of genes essential for UV-B protection including the transcription factor HY5, a key effector of the UV-B signalling pathway. Association of UVR8 with chromatin in the *HY5* promoter region supports its function in regulating *HY5* transcription (Brown *et al*., 2005) but only upon UV-B exposure (Kaiserli and Jenkins, 2007). Nuclear

localisation studies with UVR8 confirmed the nuclear localisation of UVR8, but this was insufficient to induce *HY5* transcription (Kaiserli and Jenkins, 2007). This suggests that there must be additional triggers to transcription of *HY5* and other genes regulated by UVR8. Additional targets for UVR8 include genes associated with protection against oxidative stress such as glutathione peroxidases (Milla *et al*., 2002) and photo-oxidative damage namely Early Light Inducible Proteins (ELIPs) (Hutin *et al*., 2003). UVR8 induces *ELIP1* expression via HY5 (Harari-Steinberg 2001).

1.5 Expression Patterns of anthocyanin genes

The production of anthocyanin pigment is developmentally (temporally) regulated in a tissue specific (spatial) manner. This section aims to describe the timing and location of the expression of the enzymes of the anthocyanin biosynthetic pathway in plants. Many plants display tissue specificity with regard to anthocyanin production. For instance, the eastern medicinal plant, *Perilla frutescens* produces anthocyanin in leaf tissue whereas petunias and snapdragon produce pigment only in flowers, particularly petals. Maize produces anthocyanin in the aleurone and some vegetative organs whereas apple, grape and pear produce pigments in flowers, and fruit flesh and skin.

1.5.1. Temporal control

In *Arabidopsis* (Kubasek, 1992), grape (Boss *et al*., 1996b) and maize (Taylor, 1990) anthocyanin pigment production is under temporal control, and induced by UVB and visible white light. For instance, flavonoid synthesis can be induced in cell and tissue culture by UV light, high irradiance (Wasternack and Hause, 2002), and sugar in *Arabidopsis* (Jeong *et al*., 2010). It can also be induced in tissue of excised leaves incubated under high light in a 2% glucose solution (Brugliera *et al*., 1994). In grape, structural genes are induced within six hours of continuous exposure to light as demonstrated in time-course experiments using Northern blots, and reach peak levels within 12 to 24 hours (Sparvoli *et al*., 1994; Boss *et al*., 1996b). Dark-grown seedlings have been shown to express basal mRNA levels of anthocyanin structural genes while *PAL* is constitutively expressed. Maize and Perilla exhibit similar expression patterns for anthocyanin structural genes under variable light conditions (Gong *et al*., 1997) and expression of structural genes is co-ordinately induced in both species. In Perilla *Myb-p1* is light induced, whereas *Myc-rp* is regulated independently of light conditions (Gong *et al*., 1999b).

In addition to light quality, temperature can also differentially regulate mRNA synthesis of anthocyanin structural genes. Low temperature stress induces some structural and regulatory genes in *Arabidopsis* (Leyva *et al*., 1995). In grape, anthocyanin production is characterised by two peaks, one very briefly early in berry development and again after veraison (berry ripening) when colour development occurs (Boss *et al*., 1996b).

1.5.2. Spatial control

Anthocyanin pigment is produced only in certain plant organs or tissues. Thus, certain sets of regulatory genes must control this spatial regulation of anthocyanin production. Myb and Myc genes, *R* and *C1,* control anthocyanin production in maize aleurone, and paralogous regulatory genes control pigment biosynthesis in other parts of the plant (Moyano *et al*., 1996). The MYC-like regulatory protein encoded by *Delila* is crucial for the expression of genes in the later steps of flavonoid biosynthesis in petal tubes but not in petal lobes (Sablowski *et al*., 1994). A genetic study on anthocyanin production in leaves and stems in Perilla have demonstrated pigment production in stem only (Gong *et al*., 1997).

As mentioned earlier, the structural genes of the anthocyanin biosynthetic pathway can be divided into two halves, "early" and "late" on the basis of their regulation, which can be controlled separately. The designation "early" or "late" differs between species, for instance F3H constitutes an "early" gene in *Arabidopsis*, but a "late" gene in snapdragon. The designation can also differ for various tissues within a particular species (Pelletier *et al*., 1997).

The loss of anthocyanin synthesis can arise by mutation(s) or other alterations in individual structural genes of the biosynthetic pathway or in any of the regulatory genes controlling the expression of a number of structural genes (Boss *et al*., 1996b).

1.6 The stress response and the role of anthocyanins

1.6.1. Introduction

Plants are subjected to a range of variables in environmental conditions. As sessile organisms, they are incapable of removing themselves from any undesirable conditions that occur in their immediate environment. Consequently they have developed a host of defense responses that allow them to counter the full range of negative environmental effects in order to survive (Chrispeels *et al*., 1999). This adaptive response to various biotic and abiotic stress conditions is dependent on the particular type of stress or the combination of such stresses. To survive the daily and seasonal variations in diverse environmental conditions such as temperature, light and water availability, they have to respond in an appropriate manner. Superimposed on this seasonal and daily environmental variation are numerous biotic and abiotic stress conditions. Biotic stress results from bacterial and fungal pathogen attack, and herbivore and insect feeding. Abiotic stress results from factors such as high levels of visible and UV light, drought, and salt and nutrient stress. For their survival, plants have to respond to these two sets of variables in a specific and coordinated manner. This response requires the fine integration of a network of signalling pathways, leading to the elicitation of the appropriate response. The stress/defense response usually involves a set of components/responses such as the production of reactive oxygen species (ROS) and H_2O_2, membrane and/or membrane channel modulators, proteasome activation for the degradation of transcription factors and heat shock

damaged proteins, together with an increase in the production of heat shock proteins (Chrispeels *et al*., 1999). An investigation of the transcriptome resulting from the induction of various biotic and abiotic stress conditions has provided insight into the different classes of genes that are differentially regulated during the defense response in *Arabidopsis* (Kong *et al*., 2003).

1.6.2. Biotic stress responses

Among the biotic factors influencing plant life are pathogens and herbivores. Plants respond to invasion by pathogens, such as bacteria and fungi, and to herbivore and insect feeding by producing an array of biochemical changes that include the production of reactive oxygen species, antimicrobial compounds, antioxidants and signalling molecules such as salicylic acid (SA), and jasmonic acid (JA) and methyl jasmonate (MeJA) (Mahalingam *et al*., 2003). In addition, they respond by the localised activation of a cell-death programme called the hypersensitive response (HR), and by systemic activation of cellular and molecular defenses referred to as systemic acquired resistance (SAR) (Cao, 1997). This systemic acquired resistance is mediated by second messengers that include reactive oxygen species (ROS), SA, MeJA and ethylene (Penninckx, 1996). There is significant overlap between plant responses to pathogens and environmental stresses (Pastori, 2003). For instance, the plant's response to the abiotic stressor, ozone, shows extensive overlap at the biochemical and molecular level with the pathogen defense response and include the production of ROS as well as induction of HR and SAR (Sharma, 1996).

Notwithstanding the numerous commonalities between the defense response and the stress response, a plant's response to each environmental stress is uniquely tailored to increase its ability to survive the specific environmental challenge (Reymond et. al, 2000). To understand the integration of the comprehensive network of all the genes, proteins and molecules, including signalling molecules, that mediate plant stress and defense responses requires identification and characterisation of the molecular components involved during the initial response and at the physiological level to a particular stress or pathogen (Mahalingam *et al*., 2003).

1.6.3. Abiotic stress responses

The perception of abiotic stresses and signal transduction to initiate adaptive responses are critical steps in determining the survival and reproduction of plants exposed to adverse environments. Plants have stress-specific adaptive responses as well as responses that protect plants from more than one environmental stress. There are multiple stress perception and signalling pathways present in plants. Some are highly specific while others may cross-talk at various steps (Chinnusamy *et al*., 2004). Molecular and biochemical studies suggest that abiotic stress signalling in plants involves receptor-coupled phosphorelays, phosphoinositol-induced Ca^{2+} changes, mitogen-activated protein kinase cascades and transcriptional activation of stress-responsive genes. In addition, protein posttranslational modifications and adapter or scaffold-mediated protein-protein interactions are also important in abiotic stress signal transduction (Xiong and Zhu, 2001).

Among the numerous abiotic factors that impact negatively on plant productivity are low temperature, frost, heat, high light conditions, ultraviolet light, darkness, oxidation stress, hypoxia, wind, touch, nutrient imbalance, salt stress, osmotic adjustment, water deficit, and desiccation (Wasternack and Hause, 2002). Among these, low temperature, drought and salinity are major adverse environmental factors that limit plant productivity (Xiong and Zhu, 2001). Analysis of salt overly sensitive (sos) *Arabidopsis* mutants revealed a novel calcium-regulated protein kinase pathway for response to the ionic aspect of salt stress. In-gel kinase assays identified several SOS-independent protein kinases that are either activated specifically by osmotic stress or by multiple abiotic and biotic stresses. Genetic analysis has defined the Salt-Overly-Sensitive (SOS) pathway, in which a salt stress-induced calcium signal is probably sensed by the calcium-binding protein SOS3 which then activates the protein kinase SOS2. Both ABA-dependent and -independent signalling pathways appear to be involved in osmotic stress tolerance. Components of mitogen-activated protein kinase (MAPK) cascades may act as converging points of multiple abiotic as well as biotic stress signalling pathways (Chinnusamy *et al*., 2004). Calcium ions represent both an integrative signal and an important convergence point of many disparate signalling pathways. Calcium-binding proteins, like calcineurin B-like (CBL) proteins, have been implicated as important relays in calcium signalling. *In vivo* studies in *Arabidopsis* indicate that the calcium sensor protein CBL1 may constitute an integrative node in plant responses to abiotic stimuli and contribute to the regulation of early stress-related transcription factors of the C-Repeat-Binding Factor/dehydration-responsive element (CBF/DREB) type (Albrecht *et al*., 2003).

1.6.4. Anthocyanins in the stress response in plants

The production of anthocyanin in the vegetative tissue is the hallmark of the stress response in plants (Winkel-Shirley, 2002). The stress response is defined as the combined response to various abiotic stresses whereas the defense response refers to the same action against pathogens and herbivores (Chrispeels *et al*., 1999; Pastori *et al*., 2003). Continued progress is being made in understanding the role that flavonoids play in the stress response in plants, as well as the mechanisms that control the amounts and types of flavonoids that are being produced in this response (Chalker-Scott, 2004). The ultra-violet (UV)-absorbing characteristics of flavonoids have long been considered to be evidence for their role in UV protection. Flavonoids and flavanoids are often present in the epidermis of leaves and tissue that are susceptible to UV light, such as pollen and the apical meristem. Evidence to support the role of flavanoids in UV protection came from experiments in *Arabidopsis* mutants for anthocyanin biosynthesis (Li *et al*., 1993). The transparent testa-4 mutant (*tt4*), that has reduced flavonoid but normal sinapate ester levels, is more sensitive to UV-B than wild type when grown under high UV-B irradiance. The *tt5* and *tt6* mutants characterised by reduced levels of UV-absorptive leaf flavonoids and the monocyclic sinapate ester phenolic compounds are extremely sensitive to UV-B damage (Winkel-Shirley *et al*., 1995). This demonstrates that both flavonoids and other phenolic compounds are important for UV-B protection in plants in vivo. Mutations in chalcone synthase or chalcone isomerase resulted in UV-hypersensitive

phenotypes in these mutants (Li *et al*., 1993; Landry *et al*., 1995; Booij-James *et al*., 2000). Chalcone isomerase mutants corresponding to decreased sinapate ester production showed the highest UV sensitivity (Li *et al*., 1993). Studies by Ryan *et al*. (2002) in petunia and *Arabidopsis* have provided new evidence that UV light induces the synthesis of flavonols with higher hydroxylation levels. Because flavonoid hydroxylation levels does not affect the UV absorbing properties of these compounds but does affect their antioxidant capacity, it was suggested that flavonoids play as yet uncharacterised roles in UV stress responses (Winkel-Shirley, 2002). However, numerous other studies have implicated various flavonoid compounds in UV protection (Li *et al*., 1993; Landry *et al*., 1995; Booij-James *et al*., 2000; Bieza and Lois, 2001; Ryan *et al*., 2001; Ryan *et al*., 2002). A role for flavonoids in UV protection was further supported by the isolation of an Arabidopsis mutant that is tolerant of extremely high levels of UV-B (Bieza and Lois, 2001). This mutant shows constitutively high levels of a number of phenolics, including flavonoids, and upregulation of the CHS gene. The accumulation of anthocyanin in leaves of many woody species prior to being shed has long been debated amongst biologists because it is unclear what effects anthocyanins may have on leaf function during senescence. Recent evidence reported by Feild *et al*. (2001) for red-osier dogwood (*Cornus stolonifera*) supports the hypothesis that anthocyanins form a pigment layer in the palisade mesophyll layer that decreases light capture by chloroplasts. Measurements of leaf absorbance demonstrated that red-senescing leaves absorbed more light of blue-green to orange wavelengths (495-644 nm) compared with yellow-senescing leaves. Using chlorophyll a fluorescence measurements, they observed that

maximum photosystem II (PSII) photon yield of red-senescing leaves recovered from a high-light stress treatment, whereas yellow-senescing leaves failed to recover after 6 hours of dark adaptation, which suggests photo-oxidative damage. A role of anthocyanins as screening pigments was further investigated by measuring the responses PSII photon yield to blue light, which is preferentially absorbed by anthocyanins, versus red light, which is poorly absorbed. They found that dark-adapted PSII photon yield of red-senescing leaves recovered rapidly following illumination with blue light. However, red light induced a similar, prolonged decrease in PSII photon yield in both red- and yellow-senescing leaves. This supports a role for anthocyanins in optical masking of chlorophyll that reduces the risk of photo-oxidative damage to leaf cells as they senesce. This optical masking increases the efficiency of nutrient retrieval from senescing autumn leaves.

The production of anthocyanin pigments in vegetative tissue is characteristic of the stress response in plants (Winkel-Shirley, 2002). Anthocyanin production in the testa of seeds enhances fertility (Burbulis *et al*., 1996) whereas biotic and abiotic stress conditions such as phosphate starvation, high light stress and pathogen attack have been shown to result in the production of anthocyanin in plant leaves (Dixon and Paiva, 1995). In addition to the respective biotic and abiotic stress conditions that result in increased anthocyanin production, various other stress situations regulate the level of anthocyanin in leaves often by affecting the stability of the pigment (Steyn *et al*., 2004). Temperature has a major influence on anthocyanin stability. Whereas low temperature increases the levels of anthocyanin, high temperature tends to decrease

the level by bleaching action (Steyn *et al*., 2004). This bleaching effect can be prevented by metals such as magnesium and ferric ion through complex formation with the pigment, thereby increasing the half-life of the pigment. The increase in anthocyanin production by magnesium stress appears not to result from a stress related action as the levels of PAL and CHI, two enzymes involved in the stress response in plants, are unchanged.

The *pho3* mutant that displays a phosphate-deficient phenotype, is characterised by low acid phosphatase (APase) levels. Mutants have increased levels of anthocyanin and starch, but lowered phosphate levels in roots and shoots at normal phosphate supply levels, when compared to wild type (Zakhleniuk *et al*., 2001). It is likely that the *pho3* mutant lacks a phosphate transporter or a phosphate-signalling component that renders it incapable of responding to low phosphate stress. Protein phosphatase inhibitors inhibit CHS induction but induce PAL activity indicating that their effect on CHS is not a general transcriptional inhibitory effect. CHS induction requires protein dephosphorylation as well as protein kinase activity (Christie and Jenkins, 1996).

The role of anthocyanin in shading the thylakoid from photo-oxidative damage during conditions of high light stress is well characterised. Anthocyanin in the palisade mesophyll decreases light absorption by chloroplasts. This optical masking of chlorophyll by anthocyanin reduces photo-oxidative damage to leaf cells during phosphate and other nutrient retrieval from leaves to roots thereby improving the retrieval of phosphate and other nutrients during leaf senescence (Feild *et al*., 2001)

Table 1.1a. Structural enzymes of the 'early' anthocyanin biosynthetic pathway summarising function and motifs, and gene locus, gene product size and copy number. Some of the transparent testa (*tt*) mutant loci of *Arabidopsis* are indicated in parenthesis.

Enzyme name, function & motif	Locus (gene product size)	Copy number
Phenylalanine ammonia-lyse (PAL), de-amination using NADPH as cofactor.	*C2*, *whp*, *Nivea* (700 aa)	15-20 (Grape) 1 (Rice)
Chalcone synthase (CHS) uses 4-coumaroyl-CoA and & malonyl-CoA as substrates.	*DifD* (*tt4*)	3 (Perilla) mutiple in Petunia, Grape
Chalcone isomerase (CHI), stereospecific isomerisation of chalcones (yellow) to flavanones (colourless).	*Po*, (*tt5*)	mutiple in Petunia
Flavonone-3-hydroxylase (F3H), stereo-specific hydroxylation of flavonone to dihydroflavonols at 3-position, 3 His residues implicated in iron binding, the last common step in flavonoid synthesis.	*An3*, *Incolorata* *tt6* (360 aa)	1 (Petunia) 2 (Perilla) multiple in Grape

Table 1.1b. Structural enzymes of the 'late' anthocyanin biosynthetic pathway summarising function and motifs, and gene locus, gene product size and copy number.

Dihydroflavonol reductase (DFR) reduces Dihydroflavonols to leucoanthocyanidins, NADPH-dependent hydrophilic reaction, C-terminal repeat motifs.	*An6*, *ant18*, *A1* *Pallida*, (*tt3*) (350 aa)	2 (Perilla) 3 (Petunia)
Anthocyanidin synthase (ANS) or leucoanthocyanidin dehydroxylase (LDOX), 2-oxoglutarate-dependent oxygenases catalyses coloured anthocyanidin synthesis from colourless leucoanthocyanidin, iron cofactor bound by His-Asp-His residues.	*difA*, *ant17* *candi* (430 aa)	1(*Arabidopsis*) 2 (Perilla)
UDP-glucose: flavonoid 3-*O*-glucosyltransferase (UF3GT) glycosylates anthocyanidin to anthocyanidin 3'-glucose, UDP-glucose-binding domain at the C-terminus.	*difh*, *Bronze-1* (*Bz-1*) (470 aa)	1 (Grape)
UDP-glucose: flavonoid 5-*O*-glucosyltransferase (UF5GT) glycosylates anthocyanidin 3-*O*-glucoside to anthocyanidin 3,5-di-*O*-glucosides, UDP-glucose-binding domain at C-terminus.	(460 aa)	2 (Perilla)

Table 1.2a. Myc loci and homologues in plants

Species	Locus	Target gene	Gene cloning reference
Maize	*R*	CHS, DFR, 3GT	Consonni *et al*., 1988
	R (*S*)	CHS, DFR, 3GT	Perrot and Cone, 1989
	R (*Sn*)	CHS, DFR	Tonelli *et al*., 1991
	R (*Lc*)	CHS, DFR	Ludwig *et al*., 1989
	B	DFR, 3GT	Chandler *et al*., 1989
Snapdragon	*Delila*	F3H, DFR, ANS, 3GT	Goodrich *et al*., 1992
	Eluta	F3H, DFR, ANS, 3GT	Almeida *et al*., 1989
Petunia	*An1*	*chsJ*, DFR, ANS, 3GT	Beld *et al*., 1989;
		3RT, AMT, GST	Jonnson *et al*., 1984
	An4	*chsJ*, DFR, ANS, 3GT	Beld *et al*., 1989;
		3RT, AMT, F3',5'H, GST	Quattrocchio, 1999
	An11	*chsJ*, DFR, ANS, 3GT	de Vetten *et al*., 1997
		5GT, 3RT, AMT, GST	
	Jaf13	DFR, ANS, 3GT, 5GT,	Quattrocchio *et al*., 1998
		3RT, AMT, GST	

Table 1.2b. Myb loci and homologues in plants

Species	Locus	Target gene	Gene cloning reference
Maize	*C1*	DFR, 3GT	Chandler *et al*., 1989
	Pl	CHS, DFR, 3GT	Cone and Burr, 1989
	Vp1	F3H, DFR, ANS, 3GT	McCarty *et al*., 1989
Snapdragon	*Rosea*	F3H, DFR, ANS, 2GT	Holton and Cornish, 1995
Petunia	*An2*	*chsJ*, DFR, ANS, 3GT, 3RT, AMT, GST	Quattrocchio *et al*., 1999

1.7 The phenomenon of 'Bon Rouge'

1.7.1. Introduction

The 'Bon Rouge' pear cultivar was developed from a bud mutation that has been reported to occur only once on a Williams Bon Chretien (Bartlett) pear cultivar. The bud sport was discovered on a Bon Chretien tree by D. Mouton on his farm Ongegund, Simondium (Western Cape Province, South Africa) and 'Bon Rouge' was released by the Stellenbosch Institute for Fruit Technology (INFRUITEC), South Africa in 1993 (Jolly, INFRUITEC INFO, Number 627, 1993). Cloned material from this mutant sport was evaluated during the mid to late 1960s and was found to have good potential as a red pear on the local market. Reversion is a problem with 'Bon Rouge' and the reversion may manifest as reverted branches or as spurs with the rest of the tree producing fruit with good colour, or on the whole tree producing fruit with little or no red cover colour.'Bon Rouge' is

harvested during mid to late January although this is still dependent on area of cultivation and seasonal characteristics. The cultivar stores well making it quite suitable for exporting to overseas markets during the northern hemisphere winters. Further evaluation of this cultivar established that it is suitable for canning but not drying. 'Bon Rouge' is currently not a protected cultivar and is available from commercial nurseries.

1.7.2. Is 'Bon Rouge' chimeric?

The reversion in 'Bon Rouge' fruit skin manifests as single green coloured sectors of various widths on a red background, as multiple green coloured stripes on a red background or completely reverted green skinned fruit. Half-red, half-green leaves have also been observed. This poses the question. Is 'Bon Rouge' chimeric or are the colour differences the result of mixed red and green clonal material i.e. genetic mosaics (single organisms with coexisting cells of different genotypes).

Plant chimeras are specific types of genetic mosaics in which genetically dissimilar cells are present in the shoot apical meristem (Marcotrigiano *et al*., 1995). These cells divide to produce cells that eventually form the body of the plant and their arrangement in the shoot apex determines the stability of the chimeral state and thus the plant's phenotype. Besides cytochimeras, where different apical layers possess different ploidy levels (Satina *et al*., 1940), there are three types of chimeras that are defined by their arrangement of genetically dissimilar cells in the shoot apex.

Sectorial chimeras describe the arrangement of a wedge of genetically dissimilar cells in all three apical cell layers. This chimeral state is most likely to arise when spontaneous or induced mutations occur during the early stages of embryo development when the number of cells and cell layers are minimal.

Mericlinal chimeras arise when part of one or more cell layers of a shoot apical meristem is genetically dissimilar to the rest of the layer(s). Variability in the fraction size and position of genetically dissimilar cells produce endless variation in the arrangement of mericlinal chimeras.

Periclinal chimeras are chimeras in which one or more entire apical cell layer is genetically distinct from another apical cell layer. Of the three chimeral types, periclinal chimeras are the most stable because apical layers remain independent and axillary buds maintain the same apical organisation as the terminal bud from which they were derived (Marcotrigiano, 1991). In grape cultivars, DNA profiling using microsatellite markers has occasionally identified more than two alleles at a locus in certain individuals. This phenomenon has been identified as periclinal chimerism (Franks *et al*., 2005 and references therein) that plays a role in clonal differences and cultivar identification in long-lived clonally propagated crops such as grapevine that are derived from historically ancient cultivars. Periclinal chimera meristem structures were also demonstrated in polymorphic clones observed in a collection of 145 accessions belonging to 'Pinot gris', 'Pinot noir', 'Pinot blanc', 'Pinot meunier' and 'Pinot moure' cultivars (Hocquigny *et al*., 2004).

Investigations of genetic mosaics have identified herbivory (Marcotrigiano 2000) and possibly mechanical wounding or pruning as a means of releasing mutations sequestered in the shoot apical meristem. When mutations occur in the L2 meristem layer of sexually reproducing cells, gametes arising from this layer carry mutations that may or may not have an impact on the genetic make-up of the offspring. Meristem destruction by herbivory or mechanical means often result in the production of two shoot meristems in the leaf axils of some nodes. These are referred to as the primary and secondary axillary meristems (Marcotrigiano 2000). Seeds derived from the secondary axillary meristem were not always descended from the L2 (second) layer of the terminal shoot meristem as is expected for terminal and primary shoot meristems. Further analysis demonstrated that secondary meristems did not maintain the same order of cell layers as in the terminal shoot meristem and that reproductively sequestered cell layers with mutant cells can be repositioned into gamete-forming cell layers and transfer mutant genes to their offspring.

1.7.3. Stress-induced chimeras

Various biotic and abiotic stress responses are associated with increased anthocyanin production. Phosphate starvation (Bariola *et al*., 1999; Franco-Zorrilla *et al*., 2002; Miura *et al*., 2005), nitrogen limitation in maple trees (Schaberg *et al*., 2003), and UV radiation combined with high light stress (Kubasek *et al*., 1992; Stapleton 1992; Logemann *et al*., 2000), result in increased flavonoid accumulation. Alternatively low temperature stress combined with high light stress increase anthocyanin production (Leyva *et al*., 1995) but not in 'Bon Rouge' pears (Steyn *et al*., 2004). Anthocyanin levels

in pear skin decrease towards harvest as a consequence of pigment breakdown under high light stress (Steyn *et al*., 2004), and low nitrogen levels combined with high sucrose levels increase anthocyanin accumulation in *Arabidopsis* (Martin *et al*., 2002).

The *Arabidopsis UV RESISTANCE LOCUS* (*UVR8*) is implicated in responses to UV-B (280 – 320nm). The *UVR8* locus has sequence similarity to the eukaryotic guanine nucleotide exchange factor for Ran, Regulator of Chromatin Condensation 1 (RCC1) (Kliebenstein *et al*., 2002) but has little exchange activity suggesting a different function. Both UVR8 and RCC1 locate mainly to the nucleus and interact with chromatin via histones (Brown *et al*., 2005). UVR8 regulate a UV-B signal transduction pathway that results in production of protective flavonoid and other phenolic compounds that accumulate in the epidermal layers to provide a protective UV-absorbing screen (Logemann and Hahlbrock, 2002). Mutation in *Arabidopsis UVR8* reduces the UV-B induction of flavonoids by repressing chalcone synthase (CHS) mRNA and protein production (Kliebenstein *et al*., 2002). *CHALCONE SYNTHASE* (*CHS*) encodes a key enzyme in phenylpropanoid synthesis and is the committing enzyme in anthocyanin biosynthesis (Wade *et al*., 2001). Its activation by low levels of UV-B is well characterised and the activation pathway does not require known photoreceptors but involves ELONGATED HYPOCOTYL5 (HY5), a basic domain/leucine zipper (bZIP) transcription factor and a key regulator of photomorphogenesis under all light qualities (Ulm and Nagy, 2005) including DNA damaging UV-B. In *Arabidopsis*, UVR8 regulates a range of genes essential for UV-B protection including the transcription factor HY5, a key effector of the UV-B signalling pathway. Association of UVR8 with

chromatin in the *HY5* promoter region supports its function in regulating *HY5* transcription (Brown *et al*., 2005) but only upon UV-B exposure (Kaiserli and Jenkins, 2007). Nuclear localisation studies with UVR8 confirmed the nuclear localisation of UVR8, but this was insufficient to induce *HY5* transcription (Kaiserli and Jenkins, 2007). This suggests that there must be additional triggers to transcription of *HY5* and other genes regulated by UVR8. Additional targets for UVR8 include genes associated with protection against oxidative stress such as glutathione peroxidases (Milla *et al*., 2002) and photoxidative damage namely Early Light Inducible Proteins (ELIPs) (Hutin *et al*., 2003). UVR8 induces *ELIP1* expression via HY5 (Harari-Steinberg 2001).

Finally, post-transcriptional gene silencing (PTGS) is a general term for a variety of mechanism that decreases gene expression via mRNA degradation (Maine, 2000). PTGS was discovered fortuitously in organisms that were virally infected, treated with exogenous RNA or transformed. The first description of PTGS was in Petunia transformed with chalcone synthase transgenes where it was termed co-suppression (Napoli *et al*., 1999; van der Krol *et al*., 1990). Subsequently it was described in the filamentous fungi, *Neurospora crassa* as quelling (Romano, 1992). Viral infection can trigger co-suppression (Ruiz *et al*., 1998) suggesting a biological role for PTGS as an anti-viral defense mechanism. Further experiments in the soil nematode *Caenorhabditis elegans* uncovered another phenomenon, RNA interference (RNAi), which is triggered by double stranded RNA (dsRNA) (Fire *et al*., 1998) and transgene induced co-suppression (Guo *et al*., 1995). In both endogenous and non-endogenous PTGS, the dynamic role of RNA in the genome as opposed to its long-held role as the intermediate for protein production was established with the identification of small RNA species that

corresponded to the transgene transcript. These short RNA sequences, now also known as short interfering RNA (siRNAs) were in the antisense orientation whereas the transgene was in the sense orientation indicating that these siRNAs were derived from active processing of the sense precursor (Matzke and Birchler, 2005). Viral infections engineered to induce PTGS and TGS (transcriptional gene silencing) of endogenous genes or stably integrated expressed transgenes (Jones *et al*., 2001) in *Arabidopsis* mutants, facilitated the identification of plants defective in RNA-triggered silencing (Mourrain *et al*., 2000). These mutant screens have identified a number of the proteins involved in the recognition and processing of the types of RNA responsible for PTGS and TGS (Dalmay *et al*., 2000). Proteins with homology to RNA-dependent RNA polymerases (RdRP) such as *Neurospora* QDE-1, *C. elegans* EGO-1 (Smardon *et al*., 2000) and *Arabidopsis* SGS-2/SDE-1 (Mourrain *et al*., 2000; Dalmay *et al*., 2000), function in PTGS. RdRP function in the synthesis of double stranded RNA from exogenous or aberrant transcripts (Dalmay *et al*., 2000). *rde-1* and q*de-2* genes are members of the *piwi/sting* family and Piwi/Sting proteins that are related to eIF2C, a proposed translation factor (Catalanotto *et al*., 2000) while *sde3* appears to be an RNA helicase with a RNA processing role in PTGS (Dalmay *et al*., 2001). Proteins from the ARGONAUTE (AGO) family such as AGO1 have a conserved PAZ (PIWI-ARGONAUTE-ZWILLE) domain (Carmell *et al*., 2002). In plants and other eukaryotes, AGO1 forms a physical complex with other proteins to form the RNA-induced silencing complex (RISC). SDE1 catalyse the synthesis of dsRNA from unusual transcripts that are not already double-stranded but the production of siRNAs from longer transcripts that are completely or partially double-stranded requires another enzyme,

RNaseIII Dicer. The first plant Dicer to be identified was *DICER-LIKE 1* (*DCL-1*) (Schauer *et al*., 2002).

miRNAs are encoded in the intragenic regions of eukaryotes and are transcribed to form 'stem-loop' miRNA precursors that are processed by Dicer-like enzymes to form processed miRNA of approximately 22-25 nt in length (Grant-Downton and Dickinson, 2006). These processed stem-loop miRNA are exported from the nucleus and incorporated into the RISC to catalyse degradation of homologous mRNA, and can anneal to their target mRNA to block translation or effect degradation of the target transcription (Grant-Downton and Dickinson, 2006).

1.7.4. 'Bon Rouge' trees

The 'Bon Rouge' cultivar (*Pyrus communis* L.) evaluated in this study was under commercial production in the South Western Cape region of South Africa. 'Bon Rouge' is a 'late' flowering/fruit producing cultivar in bloom during mid to end October, approximately two to three weeks after its parent cultivar, Williams Bon Chretien (Bartlett). At flowering, 'Bon Rouge' trees are characterised by dark red to bronze leaves (Figure 1.4) and fruit skin with a similar colour a few weeks later (Figures 1.2 and 1.3). Both 'Bon Rouge'' and its revertants maintain their respective fruit skin phenotypes throughout the early part of the growing season.

1.7.5. Reversion in 'Bon Rouge' trees

The level of reversion in 'Bon Rouge' trees were estimated by determining the number of trees that show reversion to the parent (green) phenotype, either as whole tree reversions or reversions of branches or parts of branches on 'Bon Rouge' trees, and by phenotypic evaluation of fruit skin colour throughout the growing season. In a commercial orchard planted with approximately 5000 trees, the rate of reversion to the parent (green) phenotype was estimated to be close to 10%.

1.7.6. Pear crosses

Open crosses

Thirty-six F1 progeny developed from seeds obtained from fruit harvested from open pollinated 'Bon Rouge' trees growing in a commercial orchard were evaluated phenotypically for anthocyanin pigment in leaves. The progeny exhibited a segregation ratio of 1:1 for anthocyanin pigment production in leaves, indicating that pigment production is inherited as a simple Mendelian trait.

Closed crosses

A controlled cross with 'Bon Rouge' as the pistillate parent and ‘Packham’s Triumph’ as the staminate parent generated 48 progeny with a segregation ratio of 23:25 for anthocyanin production in leaves indicting that a single locus controls anthocyanin pigment production in leaves.

Figure 1.2 Fruit from a 'Bon Rouge' tree showing the true phenotype (right), the reverted phenotype (left) and a mixed phenotype (middle) of fruit skin.

Figure 1.3 'Bon Rouge' pear (right) and a fully reverted green pear (left) harvested from 'Bon Rouge' trees.

Figure 1.4 'Bon Rouge' tree showing red skinned fruit (below) of the true phenotype and green skinned fruit (above) of the reverted phenotype with a reduced amount of anthocyanin.

Figure 1.5 Seedling from a controlled pollinated cross 'Bon Rouge' x 'Packham's Triumph' grown under greenhouse conditions.

Figure 1.6 Seedling from an open pollinated 'Bon Rouge' cross grown under greenhouse conditions exhibiting the red and green leaf phenotypes.

1.8 Differential gene expression analysis

Anthocyanin production in plant leaves and fruit skin is a complex trait that is determined by genetic variants at numerous loci (Shirley *et al*., 1995). Production of anthocyanin in the leaves and fruit skin of plants is a highly variable trait that is genetically correlated with developmental stage of the plant, in a tissue specific manner. To understand the molecular basis for the genetic differences between the two phenotypic variants, red and green leaves and fruit skin, of the commercial cultivar 'Bon Rouge', we employed a differential gene expression analysis using differential display and mRNAseq.

Differential display (DD) was first described by Liang and Pardee (1992). The technique combined 3′ anchored oligo(dT) primers and short 5′ arbitrary primers in PCR to amplify subsets of the transcriptome with the resulting cDNA fragments separated on denaturing polyacrylamide gels, and visualized by autoradiography. Liang *et al*. (1993) showed that 80 of a total of 240 primer combinations would be sufficient to cover the complete transcript for translations into proteins. Differential display has been used to isolate genes involved in secondary metabolism from different chemically defined phenotypes (Yamazaki *et al*., 1999). The availability of two phenotypic variants (red and green) exhibited by 'Bon Rouge' leaves and fruit skin, and present on the same tree presents a unique model to investigate how a genotype defines a metabolic phenotype, because two chemo-varietal forms regarding anthocyanin, are available. Currently, an automated differential display technique using fluorescently labeled primers and capillary electrophoresis is available for high throughput gene expression analysis. The technique requires the attachment of a sampler with precise collection capabilities for the collection of the differentially expressed fragments in individual tubes, to the capillary electrophoresis system. The differential display technique has been used extensively to study gene expression in organisms that have limited DNA and RNA sequence information. At the time of this study, a whole sequenced pear genome was not available in the public database. Recently, a number of pear genomes have been sequenced and made available to the public. These include *Pyrus communis* (in 2010) and *Pyrus pyrifolia* (in 2011).

Individual differential display fingerprints were generated for red and green phenotypes of Bon Rouge. Bands that displayed at least a two-fold increase in density as visualised between the red and green sets of fingerprints were excised from gels, re-amplified, cloned and sequenced. Re-amplification of cDNA from excised bands was performed under the same PCR conditions and with the same primer combination that was used in the initial amplification for differential display. Cloned sequences were characterised by sequence similarity search.

One revolutionary aspect of RNAseq is its ability to provide information regarding transcript abundances in species whose genomes have not yet been sequenced. In such situations, transcript quantification involves a combination of de novo assembly and abundance estimation, both of which are challenging tasks when a reference genome is not publicly available. Despite the recent development of several computational tools and approaches for data assembly and analysis de novo assembly of short reads without a reference genome remains a significant challenge (Deng *et al.*, 2018). Contigs assembled from reads from the two phenotypic variants of 'Bon Rouge', respectively, were mapped to an assembly of all the reads that were generated during six runs for the two phenotypes on the Illumina GAII. To obtain contig sequences for the two phenotypes, the best-assembled transcript sets from different available tools (Velvet and CLC Bio) at different k-mers ranging from 21 to 31 (for Velvet) were selected. The parameters considered were: transcripts having assembly length higher than 100 bp, average coverage, average transcript size, percentage of transcripts having length higher than 1000 bp, N50 value and highest transcript length. In our case CLC Bio emerged as the best choices for

performing assembly as this analysis displayed the best balance between transcripts number, coverage, maximum and average transcript length. Assembly and subsequent annotation of mRNAseq data is highly dependent on the differences in analytical methods used. Using Blast2GO for functional annotation of the differentially expressed contigs for 'Bon Rouge' and its reverted phenotype was especially useful as only a draft pear genome was publicly available at the time of the analyses. In such instances, GO terms are well suited for inferring functions for genes for which such have not yet been assigned.

Both mRNAseq and differential display allow the survey of the transcriptome in an unbiased manner. However, compared to the mRNAseq, the differential display method provides a low throughput analysis of the transcriptome, whereas the mRNAseq method provides a comprehensive analysis. Compared to differential display, mRNAseq has been shown to be more extensive in its survey of the transcriptome (Lievens *et al*., 2001).

CHAPTER 2

MATERIALS AND METHODS

2.1 LIST OF REAGENTS

ABI PRISM 310 Genetic analyser capillaries	Applied Biosystems, CA, USA
ABI PRISM® BigDye™ Terminator v3.0	
Cycle Sequencing Ready Reaction Kit	Applied Biosystems, CA, USA
Acetate filter	GE Life Science, UK
Acrylamide	BIO-RAD, WI, USA
Agar	Whitehead Scientific, SA
Agarose (DL1E)	Whitehead Scientific, SA
Agilent Bioanalyzer DNA 1000 chip	Agilent, CA, USA
Ammonium persulphate	Sigma Aldrich, MO, USA
Ampicillin	Promega, Madison, WI, USA
Amplitaq® DNA Polymerase	Applied Biosystems, CA, USA
Amplitaq® DNA Polymerase	Applied Biosystems, CA, USA
Amplitaq® DNA Polymerase	Applied Biosystems, CA, USA).

Binding Buffer	Invitrogen, CA, USA
BIORAD Experion™RNA StdSens Analysis Kit	BIO-RAD, WI, USA
Boric acid	Sigma Aldrich, MO, USA
Bovine serum albumin	Sigma Aldrich, MO, USA
Bromophenol blue	Sigma Aldrich, MO, USA
Chloroform	Sigma Aldrich, MO, USA
DEPC	Sigma Aldrich, MO, USA
DNA ladder (100 bp)	Invitrogen, CA, USA
DNA Master^{PLUS} SYBR Green I Kit	Roche, Switzerland
DNA Pol I (10 U/μL)	Invitrogen, CA, USA
DNA reference standard	Invitrogen, CA, USA
DNAse1 (10 Units/μL)	GeneHunter Corporation, TN, USA
dNTP (25 μM)	GeneHunter Corporation, TN, USA
dNTP (250 μM)	GeneHunter Corporation, TN, USA
dNTP mix (10 mM)	GeneHunter Corporation, TN, USA
DTT (100mM)	Invitrogen, CA, USA
DTT (10mM)	Invitrogen, CA, USA
Dynabeads mRNA Purification Kit	Invitrogen, CA, USA
Dynabeads mRNA purification Kit	Invitrogen, CA, USA
E. coli MC1061 strain containing the *lac*1^{q}ZM15	Invitrogen, CA, USA
EDTA	Sigma Aldrich, MO, USA
Elution Buffer (EB)	Illumina, CA, USA
Ethanol	Sigma Aldrich, MO, USA

Ethidium bromide	Sigma Aldrich, MO, USA
Experion™RNA StdSens Analysis Kit	BIO-RAD, WI, USA
First strand buffer (5X)	Invitrogen, CA, USA
Formaldehyde	Sigma Aldrich, MO, USA
Formic acid	Sigma Aldrich, MO, USA
Fragmentation Buffer (10X)	Ambion, CA, USA
GENE CLEAN Kit	Promega, WI, USA
GeneScan™-500 LIZ™ Size Standard	Applied Biosystems, CA, USA
Genetic analyser 10X running buffer with EDTA	Applied Biosystems, CA, USA
Genomic DNA Sample Prep Oligo Only kit	Illumina, CA, USA
Genomic DNA Sequencing Sample Prep Kit	Illumina, CA, USA
Glass plates (35cm x 45 cm)	Whitehead Scientific, CT, SA
Glucose	Sigma Aldrich, MO, USA
Glycerol	Sigma Aldrich, MO, USA
Glycogen (10 mg/mL)	Ambion, CA, USA
Glycogen (5μg/μL)	Ambion, CA, USA
H-AP primer (2 μM)	GeneHunter Corporation, TN, USA
H-T_{11}M (2 μM)	GeneHunter Corporation, TN, USA
HCl	Sigma Aldrich, MO, USA
Idaein chloride (cyanidin 3-galactoside)	Carl Roth, Germany
Industrial bleach	Sigma Aldrich, MO, USA
IPTG	Sigma Aldrich, MO, USA
Isopropanol	Sigma Aldrich, MO, USA

Kodak Biomax MR film	Kodak, Framingham, USA
Ligation buffer (2X)	Promega, WI, USA
Liquid nitrogen	Afrox, CT, SA
Lithium Chloride (LiCl)	NEB, MA, USA
M13 Forward Control Primer	Applied Biosystems, CA, USA
M13 Reverse Control Primer	Applied Biosystems, CA, USA
MessageClean® Kit	GeneHunter Corporation, TN, USA
Methanol	Sigma Aldrich, MO, USA
MMLV reverse transcriptase	GeneHunter Corporation, TN, USA
N,N′-dimethyl-formamide	Promega, Madison, WI, USA
NaCl	Sigma Aldrich, MO, USA
NaOH	Sigma Aldrich, MO, USA
pGEM®-T Easy Vector System I	Promega, Madison, WI, USA
pGEM®-T Easy Vector System II	Promega, Madison, WI, USA
pGEM® - 3Zf(+) (0.2 μg/μL)	Applied Biosystems, CA, USA
Phenol:Chloroform (1:1)	Sigma Aldrich, MO, USA
Phenol/Chloroform (3:1)	Sigma Aldrich, MO, USA
Phusion Buffer (5X)	NEB, MA, USA
Phusion polymerase	NEB, MA, USA
Plant DNAZOL®	GIBCO BRL®, CA, USA
Polyacrylamide	Sigma Aldrich, MO, USA
POP-4 (Performance Optimized Polymer-4)	Applied Biosystems, CA, USA
Potassium Acetate	Sigma Aldrich, MO, USA

QIAquick gel extraction kit	QIAGEN, MD, USA
QIAquick MinElute column	QIAGEN, MD, USA
QIAquick PCR spin column	QIAGEN, MD, USA
Random Hexamer Primers (3μg/μL)	Invitrogen, CA, USA
Reaction buffer (10X)	GeneHunter Corporation, TN, USA
RNAimage® Kits	GeneHunter Corporation, TN, USA
RNAse A (100 μg/μL)	Sigma Aldrich, MO, USA
RNaseH (2 U/μL)	Invitrogen, CA, USA
RNaseOUT (40 U/μL)	Invitrogen, CA, USA
RNeasy® Plant Mini kit	QIAGEN, MD, USA
Roche DNA Master^{PLUS} SYBR Green I kit	Roche Diagnostics, USA
RT buffer (5X)	GeneHunter Corporation, TN, USA
SDS	BIO-RAD, WI, USA
Second strand buffer (10X)	Invitrogen, CA, USA
Silane	Sigma Aldrich, MO, USA
Sodium Acetate	Sigma Aldrich, MO, USA
Stop Buffer	Ambion, CA, USA
SuperScript II (200 U/μL)	Invitrogen, CA, USA
T4 DNA ligase (3Weiss units/μL)	Promega, WI, USA
Taq DNA polymerase	QIAGEN, MD, USA
TEMED	Sigma Aldrich, MO, USA
Tris-Base	Sigma Aldrich, MO, USA
Tris-HCl	Sigma Aldrich, MO, USA

Tris-HCl (10mM)	Invitrogen, CA, USA
Tryptone	Sigma Aldrich, MO, USA
Uncut Plasmid, pGEM®-3Zf(+) (0.2 ng/μL)	Promega, Madison, WI, USA
Washing Buffer B	Invitrogen, CA, USA
Whatman No. 1 filter paper	Whatman, CA, USA
X-Gal (5-bromo-4-chloro-3-indolyl-B-D-galactose)	Promega, WI, USA
Xylene cyanol	Sigma Aldrich, MO, USA
Yeast extract	Sigma Aldrich, MO, USA
α-[^{33}P] dATP (2000 Ci/mmole)	PerkinElmer, MA, USA

2.2 PIGMENT EXTRACTION AND CHARACTERISATION

2.2.1 Sample material collection

Leaves from Bon Rouge and Bon Rouge reverted branches on field grown trees were collected and immediately frozen in liquid nitrogen. Sample materials were transported to the laboratory and stored at -20°C until pigment extraction. Skins from red and green reverted Bon Rouge pears were collected by peeling directly into liquid nitrogen in the orchard, transported to the laboratory and stored at -20°C until pigment extraction.

2.2.2 Anthocyanin pigments extraction

Frozen leaves and skin were ground to a fine powder in liquid nitrogen using a mortar and pestle. A 5.0 g sample of ground powder from red and green phenotypes respectively, were extracted in 10 mL acidified methanol (1% HCl in methanol) at -20°C for 1 hour. After extraction, leaf tissue was pelleted by centrifugation at 10 300 x g for 15 minutes at 4°C. The supernatant was rotary evaporated and re-dissolved in 2 mL of acidified methanol, filtered through sterile 0.45μm acetate filter (GE Infrastructure Water and Process Technologies Life Science Microseparations) and retained for HPLC analysis.

2.2.3 Anthocyanin quantification

Anthocyanin concentration was determined by measuring absorbance at 530 nm and calculated with reference to the standard, idaein (cyanidin 3-galactoside) chloride (Carl Roth, Germany).

2.2.4 HPLC analysis

A 5.0 μL aliquot of the filtered sample was diluted into 1.0 mL methanol containing 1% (v/v) HCl. Pigments were quantified by reverse-phase high performance liquid chromatography. A 5μm C18 column (250 x 4.6 mm (Phenomenex, CA, USA) and a 5 μm guard column (12.5 mm) (Phenomenex, CA, USA) was maintained in a 30°C water bath throughout the chromatographic run. The mobile phase consisted of 5% formic acid in water (A) and 5% formic acid in methanol (B) with a linear gradient of 25% to 65% in the first 18 minutes

and from 65% to 100% during the last three minutes. The flow rate was maintained at 1.0 mL per minute and the injection volume varied between 5 and 20 μL. Measurements were performed at 280 and 350 nm for anthocyanins and other phenolics, respectively, and at 530 nm for anthocyanins. Anthocyanin standards idaein (cyanidin 3-galactoside) chloride (Carl Roth, Karlsruhe, Germany) and flavonols standards rutin (quercitin 3-rutinoside), hyperoside (quercitin 3-D-galactoside) and quercitrin (quercitin 3-rutinoside) (Sigma Aldrich, MO, USA) were used to establish standard elution profiles.

2.2.5 Liquid Chromatography Mass Spectrometry (LC-MS) and MALDITOF MS analysis

LC-MS and MALDITOF MS was performed at the Central Analytical Facility of the University of Stellenbosch, South Africa using standard protocols.

2.3 DIFFERENTIAL DISPLAY ANALYSIS

Differential Display analysis was performed with the RNAimage kits 1, 4 and 7 (GeneHunter Corporation, Nashville, TN, USA).

2.3.1 Sampling of plant material for differential display analysis

Leaves from Bon Rouge and Bon Rouge reverted branches growing in a commercial orchard were collected and frozen in liquid nitrogen in the field.

Mature leaves were collected during the early stages of the growing season to confirm identification of the appropriate phenotype. Sample materials were transported to the laboratory and stored at -70°C until pigment extraction.

2.3.2 RNA extraction and analysis

RNA was extracted according to the method described by Levi *et al*. (1992). Extraction procedures were performed at 4°C except where indicated. Using several additions of liquid nitrogen, leaf tissue was ground to a fine powder with a mortar and pestle. Lyophilised tissue (0.1 g) was extracted with 300 µL homogenisation buffer in a 2 mL microfuge tube to which 300 µL of chloroform was added. The tubes were inverted for 5 minutes until a white emulsion formed. A further 5 mL of chloroform was added to improve subsequent phase separation and the tubes mixed by inversion for a further 5 minutes. Emulsified homogenate was centrifuged at 2 500 x g for 15 minutes in an Eppendorf bench top centrifuge. The upper aqueous phase was transferred to a new 2 mL microfuge tube and re-extracted with 300 µL chloroform for 5 minutes. After centrifugation at 2 500 x g for 15 minutes, the upper aqueous phase was transferred to a new microfuge tube, mixed with 30 µL of 3 M NaCl and 600 µL methanol, and stored at -20°C for at least 1 hour. Tubes were centrifuged for 15 minutes at 4 000 x g. The supernatant was discarded and the pellet dissolved in 500 µL 1X TE followed by clarification by centrifugation at 12 000 x g for 10 minutes. The supernatant was transferred to a new microfuge tube and 50 µL 3 M NaCl and 275 µL isopropanol was added. After mixing, the tubes were stored at –20°C for 1 hour then

centrifuged at 14 000 x g for 10 minutes. Isopropanol was removed and the pellet washed with 400 μL 70% ethanol. After removal of the ethanol wash the pellet was dissolved in 300 μL TE and the LiCl concentration raised to 2 M by adding 100 μL 8 M LiCl. The mixture was incubated overnight at 4°C followed by centrifugation at 14 000 x g for 10 minutes. The resultant pellet was dissolved in 300 μL TE and 450 μL 1.5 M potassium acetate (1.5 M KAc). The mixture was incubated at 4°C for 5 hours then centrifuged at 14 000 x g for 10 minutes. The pellet was dissolved in 300 μL TE followed by the addition of 30 μL 3 M NaCl and 660 μL ethanol. The mixture was incubated at 4°C for 1 hour then centrifuged as before. The supernatant was discarded, the pellet carefully washed with 400 μL 70% ethanol and centrifuged as before. The final pellet was vacuum dried and dissolved in 30 μL TE. The final RNA solution was clarified by centrifugation at 12 000 x g for 5 minutes. The supernatant containing RNA was retrieved and stored at -80°C for downstream applications. RNA quality and yield was established by formaldehyde-agarose gel electrophoresis as described below.

2.3.3 DNAse1 treatment of RNA

Following extraction and quantification, total RNA was treated for complete removal of DNA from the RNA sample. This cleanup step was performed with the MessageClean® Kit (GeneHunter Corporation, Nashville, TN, USA) as described in the manufacturer's instructions. Total RNA extracted from red and

green leaves was treated as described below. Indicated amount of reagents were added to a sterile 1.5 mL microfuge tube in the following order:

Total RNA	50.0 μL (10 – 50 μg)
10X Reaction buffer	5.7 μL
DNAse1 (10 Units/μL)	1.0 μL
Total volume	**56.7 μL**

Reagents were mixed by gentle tapping and incubated at 37°C for 30 minutes. Removal of protein contaminants and DNAse1 from the RNA sample was effected by phenol/chloroform extraction followed by ethanol precipitation of RNA. A 30 μL volume of phenol/chloroform (3:1) was added to the RNA/DNAse1 mix and incubated on ice for 10 minutes. The mixture was centrifuged at 16 100 x g for 5 minutes at 4°C. The supernatant was removed to a sterile microfuge tube and the RNA precipitated by the addition of 5 μL of 3 M sodium acetate (NaOAc) and 200 μL of 100% ethanol. Mixtures were incubated at -80°C for at least 1 hour followed by centrifugation at 16 100 x g for 10 minutes at 4°C. The supernatant was carefully removed and the pellet rinsed with 500 μL of 70 % ethanol in DEPC water. Rinse solution was centrifuged at 16 100 x g for 5 minutes at 4°C and the ethanol removed with a micropipette. This step was repeated to ensure removal of all traces of ethanol. Total extracted RNA was re-dissolved in 20 μL of DEPC treated water and between 1 and 2 μg aliquots was stored at -70°C. RNA integrity was established by electrophoresis of 2 - 3 μg of the sample on a 6% formaldehyde-agarose gel as described below. The RNA

concentration was determined by OD_{260} on a 1:1000 dilution of the sample material. Total RNA was diluted to a final concentration of 0.1 μg/μl in DEPC water for differential display analysis.

2.3.4 RNA quantification and integrity

RNA quantification and integrity was established by agarose-formaldehyde gel electrophoresis as described by Sambrook et al. (1989).

2.3.5 cDNA production by reverse transcription of mRNA

Differential display analysis was performed using the RNAimage® Kit (GeneHunter Corporation, Nashville, TN, USA) as described in the manufacturer's instructions. Due to financial considerations, kit numbers 1, 4 and 7 were used. Each kit contains three one-base-anchored H-T_{11}M primers (where M may be A, C or G) that can be used in combination with any of the 8 supplied arbitrary primers specifically designed for plant materials with large genomes, and other reagents required for the differential display technique (see Reagents, Appendix 1). Essentially, three reverse transcription reactions, each containing one of the three different one-base-anchored H-T_{11}M primers (where M may be A, C or G), were performed using RNA isolated from either red or green leaves, in thin-walled 0.5 mL PCR tubes. All reagents were thawed on ice and subsequent preparations performed at 4°C. Each reaction tube contained the following reagents in a total volume of 20 μL:

dH_2O	9.4 μL

5X RT buffer	4.0 μL
dNTP (250 μM)	1.6 μL
Total RNA (DNA-free)	2.0 μL (0.1 μg/μL, freshly diluted)
H-T_{11}M (2.0 μM)	2.0 μL
Total volume	19.0 μL

The components were gently mixed and kept on ice until amplification. To facilitate preparation and to minimize pipetting error, a core mix containing all reagents except total RNA, was prepared for each one-base-anchored H-T_{11}M primer RT reaction. PCR conditions for RT reactions was as follows:

65°C for 5 minutes, 37°C for 60 minutes and a final step for MMLV reverse transcriptase inactivation (without denaturation of the mRNA/cDNA duplex) at 75°C for 5 minutes. A holding step at 4°C was included for all amplifications. After incubation at 37°C for 10 minutes, the thermal cycler was paused and 1 μL of MMLV reverse transcriptase was added to each tube, quickly mixed by finger tapping and the programme restarted. After amplification, the tubes were spun briefly to collect the condensate and stored at -20°C until cDNA amplification with radioactive labelled primers.

2.3.6 cDNA amplification using radiolabelled nucleotides

All components for cDNA amplification were thawed on ice and PCR reactions were set up in thin walled 0.2 μL tubes at room temperature. To avoid pipetting error, a core mix was prepared for each subpopulation of cDNA generated from

each individual H-T_{11}M primer. The RT-mix and the H-AP primers were pipetted individually and the core mix scaled up to the volume required for the number of primer combinations tested for comparative analysis of a pair of RNAs. This sample set comprised the cDNA generated from RNA prepared from red and green pear tree leaves, respectively. The following reagents were added to a final volume of 20 μL into sterile thin walled 0.2 μL PCR tubes.

dH_2O	10.0 μL
10X PCR buffer	2.0 μL
dNTP (25 μM)	1.6 μL
H-AP primer (2 μM)	2.0 μL
H-T_{11}M (2.0 μM)	2.0 μL
RT-mix (generated with the H-T_{11}M primer)	2.0 μL
α-[^{33}P] dATP (2000 Ci/mmole)	0.2 μL
Taq DNA polymerase (Qiagen)	0.2 μL
Total volume	20.0 μL

Reagents were mixed by pipetting and kept on ice until amplification. PCR conditions were as follows:

94°C for 30 seconds, 40°C for 2 minutes, 72°C for 30 seconds for 40 cycles. This was followed by one cycle at 72°C for 5 minutes and a holding step at 4°C.

Reactions were stored at -20°C for denaturing poly-acrylamide gel electrophoresis the following day.

2.3.7 Denaturing PAGE and autoradiography

Radioactive labelled cDNA amplification products were electrophoresed through 6% denaturing poly-acrylamide gels essentially as described by Maniatis *et al.*, (1989). Glass plates (35cm x 45 cm) were thoroughly cleaned with soap and water, and the shorter plate (with ears) treated with silane to prevent the gel from sticking to both glass surfaces and to facilitate transfer of the gel from the plate to Whatman No. 1 paper following electrophoresis. Plates were sealed with duct tape to prevent leakage of the gel solution. The following solution was prepared for a set of gels.

dH_2O	75 mL
10X TBE	10 mL
Acrylamide solution	15 mL
TEMED	0.07 mL
Ammonium persulphate	0.3 mL

The contents were mixed by gentle swirling to prevent the formation of bubbles, and poured between the plates with the use of a 50 mL syringe. A well-forming comb producing 48 spaces was inserted immediately after pouring the gel. Gels were allowed to set at room temperature and left for at least two hours to ensure

complete polymerisation. After polymerisation, the sealing tape was removed and the gels pre-electrophoresed in 1 X TBE for 30 minutes at 75 Watts, 30 Amps and a maximum of 2000 V/40 cm. All wells were flushed with a 5 mL syringe prior to sample loading to remove residual urea that normally interferes with electrophoresis. Radioactive labelled cDNA samples were prepared by adding 4 μL of loading buffer to each tube followed by denaturation at 95°C for 5 minutes. Samples were stored on ice during the loading step. An aliquot of 10 μL from each sample was loaded onto the gel. The cDNAs generated by 24 primer combinations (the total number of combinations per kit) for a sample pair, could be analysed on a single 48 well containing gel. Samples were electrophoresed at room temperature in 1 X TBE at 75 Watts, 30 Amps and a maximum of 2000 V/40 cm until the leading dye, bromophenol blue, reached the bottom of the gel. After electrophoresis, the gel was allowed to cool and then transferred to Whatman No.1 paper cut to the size of the gel. To facilitate transfer of the gel from the glass plate, a glass pipette was rolled over the surface of the paper before it was removed from the plate. Gels were dried on a Speed Gel SG 200 drier (Savant, Farmingdale, NY USA), then exposed to Kodak Biomax MR film (35 x 43 cm) for a period of between 12 and 72 hours depending on the efficiency of the radio-labeling step.

2.4 CLONING AND SEQUENCING OF cDNAs

2.4.1 Excision of cDNA fragments from differential display gels

Band displaying a doubling in intensities as visualised on the autoradiogram was selected for re-amplification and subsequent applications (see cloning and sequencing, section 2.4.4.). The autoradiogram and the gel was aligned with the use of marker spots containing radioisotope mixed with a dye applied to two opposite corners of the gel, and by alignment with the indent on the Kodak Biomax MR film that was marked with a pen. Selected bands were located by punching through the film with a sterile needle then excised from the gel with a sterile blade.

2.4.2 DNA precipitation from gel slices

Paper containing gel slices were soaked for 15 minutes in 1.5 mL Eppendorf tubes containing 100 μL dH_2O then boiled with the tubes tightly sealed with parafilm for the same time period. Samples were centrifuged in a microfuge for 2 minutes at maximum speed to collect the condensation and to pellet the paper debris and the gel slice. The supernatant was transferred to a fresh, sterile Eppendorf tube and the DNA precipitated with 10 μL 3M sodium acetate, 5 μL of glycogen (10 mg/mL) and 450 μL of 100% ethanol. Samples were incubated at -80°C for 30 minutes after which the DNA was pelleted in a microfuge at maximum speed for 10 minutes at 4°C. After removal of the supernatant, the DNA pellet was washed with 200 μL of ice-cold 85% ethanol, re-pelleted in a microfuge and the residual

ethanol removed. Pellets were dried briefly at room temperature and the DNA dissolved in 10 μL of dH_2O. A 4 μL aliquot was used in the re-amplification procedure and the rest stored at -20°C for subsequent applications.

2.4.3 Re-amplification of cDNA fragments

Re-amplification was performed under the same PCR conditions and with the same primer set used in the initial amplification. However the dNTP concentration was raised from 2 μM to 20 μM and radio-isotopes omitted. The reaction was performed in a total volume of 40 μL using the following reagents.

dH_2O	20.4 μL
10 X PCR buffer	4.0 μL
dNTP (250 μM)	3.2 μL
H-AP primer (2 μM)	4.0 μL
H-T_{11}M (2 μM)	4.0 μL
cDNA template	4.0 μL
Taq DNA polymerase (Qiagen)	0.4 μL
Total volume	40.0 μL

PCR conditions were as follows:

94°C for 30 seconds, 40°C for 2 minutes, 72°C for 30 seconds for 40 cycles. This was followed by one cycle at 72°C for 5 minutes and a holding step at 4°C.

After reamplification, 30 μL of the amplicon was electrophoresed through a 1.5% agarose gel at 90 V until the leading dye reached the bottom of the gel. Ethidium bromide was added to the gel mix to facilitate visualisation of the PCR product under ultraviolet light. Successfully reamplified product sizes were verified by comparison with their size on the denaturing poly-acrylamide gel. Unsuccessful reamplifications were repeated with a 1:100 dilution of the original PCR product using the same PCR conditions and reagents concentrations. Over 90% of the excised bands were successfully reamplified. This was within the success rate for this procedure. PCR products were cloned using pGEM®-T Easy Vector System II (Promega, Madison WI, USA). Product was stored at -20°C between applications.

2.4.4 Cloning and sequencing of amplified cDNA fragments

2.4.4.1 Preparation of competent *E. coli* cells

Competent *E. coli* strain MC1061 cells (containing the *lac*1^{q}ZM15) was prepared from a 40% glycerol stock maintained at -80°C. Dedicated glassware for competent cell preparation was baked at 220°C overnight or heat sterilised twice before use. An aliquot from the glycerol stock of *E. coli* strain MC1061 cells was aseptically streaked on a sterile nutrient agar plate and incubated at 37°C overnight. A single colony was picked aseptically from the plate and inoculated into 10 mL of sterile LB (Luria Bertani; 10 gm Tryptone, 5 gm Yeast extract and 5 gm of NaCl in 1 L of dH_2O, heat sterilised for 20 minutes) medium. The inoculated medium was incubated overnight at

37°C with continous shaking (300 rpm). Simultaneously, four 250 mL polypropylene bottles and four 50 mL Oak Ridge tubes filled with 100% industrial bleach were shaken at 37°C overnight. After overnight shaking the bottles and tubes were rinsed with DEPC treated water to remove all traces of bleach. Additionally, four 1L Erlenmeyer flasks were capped with foil and baked overnight at 220°C. For preparation of competent cells, 100 mL LB medium was added to the sterilised Erlenmeyer flasks and heat sterilised for 20 minutes. After the LB medium had cooled to 37°C, the medium in each flask was inoculated aseptically with 2 mL of the overnight *E. coli* culture and incubated overnight at 37°C with constant shaking (300 rpm) until the OD_{600} reached 0.4. Cultures were transferred into the sterile 250 mL polypropylene bottles and centrifuged at 4 000 x g for 10 minutes at 4°C. After centrifugation, the supernatant was discarded and the cell precipitate placed on ice. All subsequent procedures were performed strictly at 4°C. The pellet was rapidly resuspended by gentle swirling in 200 mL ice-cold sterile dH_2O, then centrifuged at 4 000 x g for 10 minutes at 4°C. This last procedure was repeated and after removal of supernatant, the pellet rapidly resuspended by gentle swirling in 50 mL ice-cold glycerol. Resuspended cells were stored in 1.0mL aliquots at -70°C. The transformation efficiency of competent cells was determined by transformation with an uncut plasmid, pGEM®-3Zf(+) (0.2 ng/μL) and calculating the cfu/μg DNA. Transformation efficiency above 1 x 10^8 cfu/μg DNA was deemed satisfactory. To calculate the transformation efficiency, 100 μL of competent cells was transformed with 0.1 ng uncut

plasmid DNA in a 1.5 mL microfuge tube. The contents were mixed by gently tapping the tube then left on ice for 20 minutes. Competent cells were heat-shocked for 45 - 50 seconds in a water bath at exactly 42°C without shaking and immediately returned to ice for 2 minutes. A 900 μL aliquot of LB broth was added to the tubes and incubated at 37°C with shaking at 150 rpm. After a 1.5hour incubation, 100 μL of a 1/10 dilution of the transformation culture was plated on duplicate LB/ampicillin/X-Gal/IPTG plates and incubated at 37°C for 16 – 24 hours. To facilitate colony selection, plates were incubated for a longer period or stored at 4°C after the 37°C incubation. Competent cells with a 1 x 10^8 cfu/μg DNA produced approximately 100 colonies per plate when 100 μL of the transformation culture was plated. When a higher colony number was desired, the total volume of the transformation culture was centrifuged at 1 000 x g for 10 minutes and the pelleted cells re-suspended in 200 μL LB broth. A 100 μL aliquot was plated on duplicate plates and treated as above. White colonies generally contained inserts and these were screened (for inserts) by colony PCR (see Section 2.4.4.4)

2.4.4.2 Ligation of cDNA fragments into pGEM®-T Easy vector

PCR products were cloned using pGEM®-T Easy Vector System II (Promega, Madison, WI, USA). PCR product purity was verified by 1.5% agarose gel electrophoresis before cloning. Only one band was visible on agarose gel electrophoresis for each re-amplification reactions. To ensure absolute

integrity of the amplified fragment, bands cut from agarose gels were purified with the GENE CLEAN Kit (Promega, Madison, WI, USA) before cloning. In some instances, the PCR product from re-amplification was used but only when a single band corresponding to the size of the band cut from the differential display gel, was visualised after agarose gel electrophoresis. Prior to vector ligation, DNA concentration was determined by comparison with a standard DNA sample electrophoresed simultaneously through 1.5% agarose gel containing ethidium bromide. A vector: insert ratio between 1: 3 and 3: 1 was determined for each insert by using the following calculation.

For a 500 bp insert:

10 ng vector/kb size of vector x insert: vector molar ratio = ng of insert

Ligations protocols included both positive and background controls. Tubes containing the pGEM®-T Easy vector and the control insert DNA were centrifuged briefly to collect the contents at the bottom of the tube. The 2X ligation buffer was mixed vigorously by vortexing then stored on ice before use. All ligations were set up in 1.5 mL sterile microfuge tubes at room temperature as follows:

	Standard reaction	Positive control	Background
2X Rapid Ligation Buffer	5.0 μL	5.0 μL	5.0 μL
pGEM-T Easy vector (10 ng)	1.0 μL	1.0 μL	1.0 μL
PCR product (1 ng/μL)	5.0 μL	----	----

Control insert DNA	----	2.0 μL	----
T4 DNA ligase (3Weiss units/μL)	1.0 μL	1.0 μL	1.0 μL
Sterile dH_2O to a final volume of	10.0 μL	10.0 μL	10.0 μL

Reagents were mixed by pipetting and incubated at 20°C for 2 hours.

2.4.4.3 Transformation of *E. coli* cells

Competent *E. coli* cells were thawed in an ice water bath to preserve their competency to a large degree. Duplicate LB/ampicillin/IPTG/X-Gal plates were prepared for each ligation and for the determination of the transformation efficiency of competent cells. Tubes containing ligation reactions were centrifuged briefly in a microfuge to collect the contents in the bottom of the tube. For transformations, 2 μL from each ligation reaction was transferred to a sterile 1.5 mL microfuge tube. For determination of the transformation efficiency of competent cells, the same volume containing 0.1 ng uncut plasmid DNA, was added to a sterile 1.5 mL microfuge tube. The required volume of competent cells was thawed in an ice water bath for approximately 5 minutes. The tube was tapped carefully to mix the competent cells and 50 μL transferred to each tube containing ligation reactions and uncut plasmid. The content was mixed by gently tapping the tube, then left on ice for 20 minutes. Competent cells were heat-shocked for 45 - 50 seconds in a water bath at exactly 42°C without shaking, then immediately returned to ice for 2 minutes. A 950 μL aliquot of LB broth was added to each ligation

reaction containing tube (900 μL was added to the tube containing cells transformed with uncut plasmid) and incubated at 37°C with shaking at 150 rpm. After 1.5 hours, 100 μL of the transformation culture was plated on duplicate LB/ampicillin/X-Gal/IPTG plates and incubated at 37°C for 16 – 24 hours. For determination of transformation efficiency, a 100 μL of a 1/10 dilution of the transformation mix was plated and similarly incubated. To facilitate blue/white colony selection, plates were incubated for a longer period or stored at 4°C after the 37°C incubation. Competent cells with a 1 x 10^8 cfu/μg DNA produced approximately 100 colonies per plate when 100 μL of the transformation culture was plated. When a higher colony number was desired, the total volume of the transformation culture was centrifuged at 1 000 x g for 10 minutes and the cells pellets re-suspended in 200 μL LB broth. A 100 μL aliquot was plated on duplicate plates and treated as above. White colonies generally contained inserts and these were screened for inserts by colony PCR (see Section 2.4.4.4).

2.4.4.4 Colony PCR with M13 Primers (to confirm ligations)

To verify the validity of the insert, colony PCR was performed on positive clones as described below essentially as described by Sambrook et al. (1989).

2.4.4.5 PCR products were resolved by agarose gel electrophoresis as described by Sambrook et al. (1989).

2.4.4.6. Plasmid DNA preparation

A single colony containing the vector with the desired insert as determined by colony PCR, was picked aseptically from the LB-Ampicillin plate and inoculated into 10 mL of LB medium contained in 15 mL sterile Greiner tubes. Inoculates were incubated at 37°C with shaking (150 rpm) overnight. After incubation, inoculates were centrifuged at 4 000 x g for 10 minutes at 4°C. The supernatant was discarded and the pellet resuspended in 300 µL of GTE buffer (50 mM Glucose, 50 mM Tris-HCl at pH 8.0, 10 mM EDTA) containing RNAse A (100 µg/µL). Resuspended pellets were transferred to sterile 1.5 mL microfuge tubes. After incubation for five minute at room temperature, 300 µL of Lysis solution (200 mM NaOH, 1% SDS) was added and the samples mixed by gentle inversion of the tubes. After an additional 5 minute incubation at room temperature, 300 µL of Neutralisation solution (1.5 M potassium acetate pH 4.8) was added and the samples mixed by gentle inversion of the tubes. Samples were centrifuged at 12 000 x g for 20 minutes at room temperature. A 600 µL aliquot of the supernatant was transferred to a fresh sterile microfuge tube, mixed thoroughly by shaking with 360 µL of isopropanol and incubated at -20°C for at least one hour to precipitate plasmid DNA. Following precipitation, samples were centrifuged at 12 000 x g for 10 minutes at room temperature to pellet the DNA. The supernatant was discarded and the pelleted DNA suspended in 500 µL 70% ethanol, centrifuged at 12 000 x g for 10 minutes at room temperature after which the supernatant was discarded. This ethanol wash was repeated and the resultant

DNA pellet dissolved in 500 µL 1X TE (50 mM Tris-HCL, 10 mM EDTA) pH 8.0. Plasmid DNA was further purified by the following procedure. An equal volume of phenol:chloroform was added to the sample and the contents thoroughly mixed by vortexing for 30 seconds. Phase separation was effected by centrifugation in a microfuge at maximum speed for 1 minute and the top phase transferred to a fresh tube. The phenol:chloroform extraction was repeated and 400 µL of the top phase transferred to a fresh microfuge tube. Plasmid DNA was precipitated from this extract by addition of 1/10 the volume of 3 M sodium acetate (pH 5.5) and 1 mL of cold absolute (at least 96%) ethanol. Tube contents were mixed by vortexing for five seconds then incubated at -20°C for at least one hour. After ethanol precipitation, the DNA was pelleted by centrifugation at 12 000 x g for 10 minutes at room temperature and the supernatant discarded. The resultant pellet containing plasmid DNA was washed with 250 µL 70% ethanol by brief vortexing, then centrifuged at 12 000 x g for 5 minutes at room temperature. The ethanol wash was repeated and the last traces of ethanol removed with a pipette. The pellet was air-dried for 5 minutes at room temperature and dissolved in 50 µL of sterile dH_2O, vortexed briefly then clarified by pulse spin. DNA quality and concentration was determined by agarose gel elctrophoresis by comparison with a known DNA standard. Samples were stored at -20°C between applications.

2.4.4.7 Re-amplification of cloned fragments with M13 primers

Inserts were amplified in the GeneAmp 9700 thermal cycler (Applied Biosystems, Foster City, CA, USA) for automated sequencing in the ABI PRISM 310 Genetic Analyser (Applied Biosystems, Foster City, CA, USA) using M13 forward and reverse primers supplied with the ABI PRISM® BigDye™ Terminator v3.0 Cycle Sequencing Ready Reaction Kit with Amplitaq® DNA Polymerase, FS. DNA concentration was determined by measuring absorbance at a wavelength of 260 nm. Between 5 and 20 ng of DNA template was used in each reaction. Then following components were added to a 1.5 mL eppendorf tube:

DNA template	5 - 20 ng
M13 primer (forward or reverse)	1 μL (10 pM)
Ready reaction mix	1 μL (10 pM)
H_2O	x μL
Total volume	20.0 μL

The following parameters were used in a thermal cycler set for 25 cycles:

Rapid thermal ramp (1°C/sec) to 96°C, 96°C for 10 seconds, rapid thermal ramp to 50°C, 50°C for 10 seconds, rapid thermal ramp to 60°C, 60°C for 4 minutes. Following amplification, the extension product was concentrated in a benchtop centrifuge at 14 000 x g for 2 minutes. The total volume of each extension reaction was pipetted into a 1.5 mL microcentrifuge tube followed by addition of 16 μL of deionised water and 64 μL of non-denatured 95%

ethanol to a final concentration of 60 ±3%. Each mixture was vortexed briefly and left at room temperature to precipitate the extension products. Extensions were pelleted in a microfuge at maximum speed for 20 minutes. Immediately following centrifugation, the supernatant was removed. Tubes were rinsed with 250 μL 75% ethanol and vortexed briefly to re-precipitate the pellets, centrifuged at maximum speed for 10 minutes followed by careful aspiration of the supernatant. This ethanol wash was repeated and the pellets dried in a heat block at 90°C for 1 minute. Extension products were sequenced immediately or stored at -20°C for not longer than 24 hours before sequencing in an ABI PRISM 310 Genetic Analyser (Applied Biosystems, Foster City, CA, USA).

2.4.4.8 Automated sequencing of amplified cloned cDNAs

Cloned DNA fragments were sequenced in an ABI PRISM 310 genetic analyser (Applied Biosystems, Foster City, CA, USA) using Big Dye Cycle Sequencing kits (Applied Biosystems, Foster City, CA, USA) essentially as described by the manufacturer.

2.5 QUANTITATIVE RT-PCR

2.5.1 Reagents, glass and plastic ware

All glass materials were baked at 220°C for 12 hours after thorough washing and a final rinse with DEPC treated water. Plastic ware was sterilised twice and all reagent used was dedicated to RNA extraction. All solutions were prepared in DEPC water.

2.5.2 Plant material

Leaves for RNA extraction was collected from trees growing in a commercial orchard early in the growing season, frozen in liquid nitrogen and transported to the laboratory. Sampling material was stored at -80°C until RNA extraction.

2.5.3 RNA extraction and analysis

RNA extraction was performed using the RNeasy® Plant Mini kit (Qiagen, MD, USA) according to the manufacturer's instructions.

2.5.4 RNA quantification

RNA quantification was performed as described in section 2.3.4)

2.5.5 Primer design for specific cDNA fragment amplification

Primer3 software was used for primer design with default parameters settings at (http://www-genome.wi.mit.edu/genome_software/other/primer3.html).

2.5.6 RT-PCR reactions in the Light Cycler (Roche Biochemicals)

For quantitative RT-PCR, the SYBR green DNA kit was used essentially as described by the manufacturer and the products analysed by gel electrophoresis.

PCR cycling conditions:

Reverse Transcription: 42-44 °C for 30 min

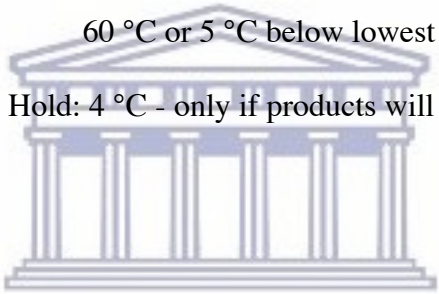
Initial denaturation: 94 °C for 2 min

Then 40 cycles: Denaturation: 94 °C for 15 sec

Annealing, extension, and read fluorescence:

60 °C or 5 °C below lowest primer T_m, 1 min

Hold: 4 °C - only if products will be run out on a gel



2.6 cDNA LIBRARY PREPARATION FOR TRANSCRIPTOME SEQUENCING

2.6.1 Reagents and plastic ware

All reagents were certified DNAse and RNAse free. Plastic ware was sterile and certified DNAse and RNAse free. All solutions, where necessary, were prepared in ultrapure water.

2.6.2 Plant material

Leaves for RNA extraction was collected from trees growing in a commercial orchard early in the growing season, frozen in liquid nitrogen and transported to the laboratory. Sampling material was stored at -80°C until RNA extraction.

2.6.3 RNA extraction

RNA extraction was performed using the RNeasy® Plant Mini kit (Qiagen, MD, USA) as described by the manufacturer. Buffer RLC was used for RNA extraction.

2.6.4 RNA quantification and analysis

RNA quality was determined on the BIO-RAD Experion™ Platform using the protocol for RNA analysis with BIO-RAD Experion™ RNA StdSens Analysis Kit essentially as describe by the manufacturer.

2.6.5 cDNA library preparation for transcriptome sequencing

High Throughput transcriptome sequencing was performed on the Illumina GAII platform using the Illumina mRNA Seq V2 protocol. This protocol incorporates several steps designed to convert total RNA into a library of short template cDNA molecules for high throughput DNA sequencing. Initially poly-A containing mRNA molecules were purified using poly-T oligo-attached magnetic beads from the Dynabeads mRNA Purification Kit (Invitrogen, #610-06). The purified mRNA was fragmented into shorter sequences using divalent cations under

elevated temperature. After fragmentation, the cleaved RNA fragments were copied to first strand cDNA using reverse transcriptase and a high concentration of random hexamer primers. This was followed by second strand cDNA synthesis using DNA Polymerase I and RNaseH. Subsequently, short cDNA fragments were prepared for sequencing on the Illumina Genome Analyzer II.

The protocol was performed using reagents obtained from Invitrogen, NEB, Qiagen, and Ambion in conjunction with Illumina's Genomic DNA Sample Prep Oligo Only kit (#FC-102-1003). Alternatively, reagents for steps 1-4 were purchased from other vendors and used in conjunction with reagents provided in the Genomic DNA Sequencing Sample Prep Kit (#FC-102-1001) for the remainder of the protocol. The complete Genomic DNA Sample Prep kit reagents for the PCR step are in a 2X Master Mix format containing Phusion polymerase, dNTPs, and Phusion Buffer. The protocol was modified for this adjustment. The detailed protocol was as follows:

2.6.5.1 mRNA purification from Total RNA

1. A total of 10 μg of total RNA was diluted with nuclease-free H_2O to 50 μL in a 1.5 mL RNase free Eppendorf tube (Ambion).
2. The sample was heated at 65°C for 5 minutes to disrupt the secondary structures, and placed on ice.

3. An aliquot of 100 μL of Dynal oligo(dT) beads (from Dynabeads mRNA Purification Kit, Invitrogen, #610-06) was transferred to a 1.5 mL RNase free Eppendorf tube.

4. The beads were washed twice with 100 μL of Binding Buffer (20 mM Tris-HCl pH 7.5, 1.0 M LiCl and 2 mM EDTA; Invitrogen, #610-06). The supernatant was removed after the second wash and the beads resuspended in 50 μL of Binding Buffer,

5. The 50 μL aliquot of total RNA from step 2 was added to the beads contained in 50 μL of Binding Buffer (total volume = 100 μL. The tube was rotated at RT for 5 minutes and the supernatant removed.

6. The beads with RNA attached was washed twice with 100 μL of Washing Buffer B (10 mM Tris-HCl PH 7.5, 0.15 M LiCl, 1 mM EDTA; Invitrogen, #610-06).

7. To prepare for a second round of oligo-dT purification, 80 μL of Binding Buffer was aliquoted to a fresh 1.5 mL RNase free Eppendorf tube.

8. The supernatant was removed from the beads of step 6, to which 20 μL of 10mM Tris-HCl (Invitrogen, #610-06) was added. The mixture with the beads was heated at 80^0C for 2 minutes to elute mRNA. Immediately after the heating step, the tube was placed on the magnet stand the supernatant containing the mRNA was transferred to the tube from step 7. A 100 μL of Washing Buffer B was added to the remaining beads.

9. The mRNA sample from step 8 was heated at 65°C for 5 minutes to disrupt secondary structures and the tube placed on ice.

10. The beads from step 8 was washed twice with 100 μL of Washing Buffer B and the final supernatant removed.

11. A 100 μL of the mRNA sample from step 9 was added to the beads and the tube rotated at room temperature for 5 minutes.

12. The supernatant was removed and the beads washed twice with 100 μL of Washing Buffer B.

13. The supernatant was removed from the beads, and 10 μL of 10 mM Tris-HCl (Invitrogen, #610-06) was added followed by a heating step at 80^0C for 2 minutes to elute mRNA. The beads were immediately placed on the magnet stand and the supernatant containing the mRNA transferred to a fresh 200 μL thin wall PCR tube. The final volume containing the mRNA was ~9 μL.

2.6.5.2 Fragmentation of mRNA

For mRNA fragmentation, the following reaction mix was assembled:

10 X Fragmentation Buffer (Ambion, #AM8740)	1 μL
mRNA	9 μL

1. The tube was incubated in a PCR thermocycler at 70°C for exactly 5 minutes. The reaction was stopped by the addition of 1 μL of Stop Buffer (Ambion, #AM8740) and the tube placed on ice.

2. The solution was transferred to a 1.5 ml microcentrifuge tube to which 1 μL of 3 M sodium acetate, pH 5.2, 2 μL of glycogen (5 μg/μL, Ambion,

#AM9510) and 30 µL of 100% ethanol was added. The tube was incubated at -80°C for 30 minutes.

3. The tube was centrifuged at 14000 x g for 25 minutes at 4°C in a microcentrifuge and the pellet washed with 70% ethanol. The ethanol was removed and the pellet and air-dried at room temperature. The fragmented RNA was suspended in 10.5 µL of RNase free water.

2.6.5.3 First strand cDNA synthesis

1. The following reaction was assembled in a 200 µL thin wall PCR tube:
   - Random Hexamer Primers (3 µg/µL, Invitrogen, #48190-011) 1 µL
   - Fragmented mRNA (from 2.6.5.2 above) 10.5 µL

2. The sample was incubated in a PCR thermocycler at 65°C for 5 minutes and the tube placed on ice.

3. The following reaction mix was prepared in the following order (10% extra reagent was prepared for multiple samples):
   - 5X first strand buffer (Invitrogen, #18064-014) 4 µL
   - 100 mM DTT (Invitrogen, #18064-014) 2 µL
   - dNTP mix (10 mM) 1 µL
   - RNaseOUT (40 U/µL) (Invitrogen, #10777-019) 0.5 µL

4. The total volume (7.5 µL) of the mixture was added to the tube, mixed, and the tube heated at 25°C in a thermocycler for 2 minutes.

5. 1 µL SuperScript II (200 U/µL, Invitrogen, #18064-014) was added to the sample followed by incubation in a thermocycler with the following program:

- Step 1 25°C 10 min
- Step 2 42°C 50 min
- Step 3 70°C 15 min
- Step 4 4 °C Hold

2.6.5.4 Second strand cDNA synthesis

1. The tube from section 2.6.5.3, step 5 was placed on ice and 61 µL of H_2O was added to the first strand cDNA synthesis mix.

2. The following reagents were added in the indicated order:

- 10X second strand buffer (500 mM Tris-HCl pH7.8, 50 mM $MgCl_2$, 10 mM DTT) 10 µL
- dNTP mix (10 mM) 3 µL

3. After thorough mixing, the tube was incubated on ice for 5 minutes followed by addition of the following reagents in the indicated order:

- RNaseH (2 U/µL, Invitrogen, #18021-014) 1 µL
- DNA Pol I (10 U/µL, Invitrogen, #18010-025) 5 µL

4. After thorough mixing, the tube was incubated at 16°C in a thermocycler for 2.5 hours.

5. The DNA was purified with a QIAquick PCR spin column (Qiagen, #28106) and eluted in 30 μL of EB solution.

For Steps 2.6.5.5 to 2.6.5.9 reagents that are part of the Illumina Genomic DNA Sample Prep Kit (#FC-102-1001) was used, unless otherwise stated.

2.6.5.5 End repair

1. The following reaction mix was prepared in the order indicated:

   - Eluted DNA 30 μL
   - H_2O 45 μL
   - T4 DNA ligase buffer with 10 mM ATP 10 μL
   - dNTP mix (10 mM) 4 μL
   - T4 DNA polymerase (3 U/μL) 5 μL
   - Klenow DNA polymerase (5 U/μL) 1 μL
   - T4 PNK (10 U/μL) 5 μL

2. The sample was incubated at 20°C for 30 minutes and the DNA purified with a QIAquick PCR spin column (QIAGEN, #28106).

3. End repaired cDNA was eluted from the column in 32 μL of EB solution.

2.6.5.6 Addition of a single A base

1. The following reaction mix was prepared by addition of the indicated reagents in the specified order:

   - Eluted DNA 32 µL
   - Klenow buffer 5 µL
   - dATP (1 mM) 10 µL
   - Klenow 3' to 5' exo- (5 U/µL) 3 µL

2. The sample was incubated at 37°C in for 30min. and the DNA purified with a QIAquick MinElute column (QIAGEN, #28006).

3. The DNA was eluted in 19 µL of EB solution.

2.6.5.7 Adapter ligation

1. The following reaction mixture was prepared for adaptor ligation:

   - Eluted DNA 19 µL
   - DNA Ligase buffer 25 µL
   - Adaptor oligo mix 1 µL
   - DNA Ligase (1 U/µL) 5 µL

2. The sample was incubated at room temperature for 15 minutes and the DNA purified with a QIAquick MinElute column (Qiagen, #28006).

3. cDNA was eluted in 10 µL of EB solution.

2.6.5.8 Gel purification of cDNA templates

1. A 2% agarose gel in a final volume of 50 mL 1x TAE buffer was prepared for template purification.
2. The total sample volume of 10 μL was loaded into one well, and 1 μL of 100 bp DNA ladder (Invitrogen 15628-019) into another (for multiple sample loads, at least two empty wells were spaced between samples to prevent cross contamination).
3. The sample was electrophoresed through the gel at 120V for 45~60min until sufficient separation of the 100 bp and 200 bp bands of the DNA ladder was achieved.
4. A gel slice at 200 bp (+/- 25bp) was excised with a sterile blade and the DNA purified with QIAquick gel extraction kit (Qiagen, #28706).
5. The DNA was eluted into 30 μL of EB solution.

2.6.5.9 PCR enrichment of purified cDNA templates

1. The following PCR master mix was prepared, with 10% extra reagent for multiple samples, and 20 μL dispensed into each PCR tube:
   - 5 X cloned Phusion Buffer (NEB, #F-530) 10 μL
   - PCR primer 1.1 1 μL
   - PCR primer 2.1 1 μL
   - 25 mM dNTP mix 0.5 μL

- Phusion polymerase (NEB, #F-530) 0.5 μL
- H2O 7 μL

2. The total volume of 30 μL of purified ligation mix from section 2.6.5.8, step 5 was added to the PCR tube.

3. **The following PCR cycling parameters was used for amplification:**

- 98°C 30 sec
- 98°C 10 sec
- 65°C 30 sec 15 X
- 72°C 30 sec
- 72°C 5 min
- 4°C hold

4. cDNA was purified with a QIAquick column (Qiagen, #28106) and eluted in 30 μL of EB solution.

5. An aliquot of 1 μL of the purified product from step 4 was analysed on an Agilent Bioanalyzer DNA 1000 chip (Agilent, #5067-1504) to check the quality of the final product and quantify the DNA concentration. A distinct band at ~200 bp. was observed for all samples used for cluster generation on the Illumina Cluster Station using the standard protocols.

2.7 QUANTIFICATION OF cDNA LIBRARIES

cDNA libraries were quantified on the Roche LightCycler™ using the Roche DNA Master^{PLUS} SYBR Green I kit. For quantification, cDNA libraries were diluted to the same range as the control template for quantitative PCR (qPCR). Prior to qPCR, a fresh dilution of each of the libraries was prepared as the cDNA libraries do not store well at low concentrations.

2.7.1 Standard curve generation using a control template

A 2 µL aliquot of the control template was added to 198 µL of water solution to dilutions was prepared to give 7 control template dilutions in the range 100 pM to 1.6 pM. After each serial dilution, samples were vortex mixed to ensure complete dispersal of sample material in the water solution. Three independent sets of serial dilutions were prepared for control template.

2.7.2 Preparation of the qPCR reaction mix

For all PCR reagent preparation, and wherever possible, master mixes were used to minimize pipetting error.

The following reaction mix was prepared immediately before amplification:

	1x mix	Final Conc.
LightCycler SYBR Master Mix (5X)	4 µL	1X
qPCR Primer 1.1 (10 µM)	1 µL	0.5 µM
qPCR Primer 2.1 (10 µM)	1 µL	0.5 µM
Ultrapure water	12 µL	
	18 µL	

The sample was thoroughly mixed by gently tapping the tube.

The tube was covered with foil to protect the sample from light and placed on ice.

An aliquot of 18 µL of the mix was added 2 µL of the cDNA library sample to be quantified.

2.7.3. Quantification by qPCR on the Roche LightCycler™

The sample carousel was placed in the qPCR machine in the correct orientation and the optical lids thoroughly cleaned with lens tissue to remove any dust before the lid of the LightCycler qPCR machine was closed.

The following thermal profile was used in conjunction with the LightCycler DNA Master^{PLUS} SYBR Green I kit:

Hot start (Initial denaturation) - 1 cycle

95ºC 10 minutes

Cycling (Amplification) - 40 cycles

95ºC	10 seconds
61ºC	10 seconds
72ºC	20 seconds

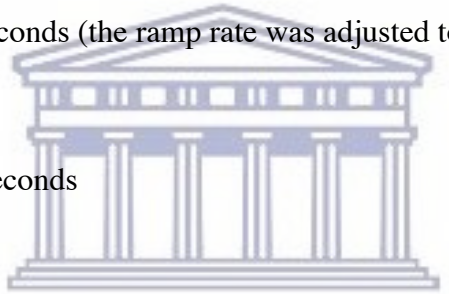
Extension time depended on the size of amplified fragment and was calculated at 25bp/s.

Melting curve analysis: - 1 cycle

95ºC	10 seconds
65ºC	60 seconds
95ºC	0 seconds (the ramp rate was adjusted to 0.1ºC/s)

Cooling

40ºC	30 seconds



The quantified library was diluted to a standard concentration for clustering. The cluster manual suggests starting the cluster generation process with a DNA library concentration of 10 nM.

The following reagent mix was prepared for denaturation of the template:

EB (Elution Buffer)	18 μL
10 nM Template DNA	1 μL
2 N NaOH	1 μL

2.8 CLUSTER GENERATION AND mRNA SEQUENCING ON THE ILLUMINA GAII PLATFORM

The protocols for cluster generation on the Illumina GAII Cluster Generation station was performed essentially as described by the manufacturer in the Cluster Generation Protocol Manual.

The protocols for mRNA sequencing on the Illumina GAII Sequencing Station was performed essentially as described by the manufacturer in the Sequencing Protocol Manual. A full protocol is supplied as a supplement in Appendix 6.

CHAPTER 3

ANTHOCYANIN CHARACTERISATION IN 'BON ROUGE'

Introduction

It is a well-established fact that anthocyanin pigment production causes the red fruit skin and leaves associated with 'Bon Rouge' and other similar phenotypes of pear and apple cultivars. Observation over a number of years indicated the occurrence of many green-skinned fruit, and green leaves on branches of 'Bon Rouge' trees. It is widely accepted that a reduction in anthocyanin concentration causes the reversion to the green-skinned phenotype. To investigate the difference in pigment production in the two phenotypic variants of 'Bon Rouge', anthocyanin pigment quantification was performed by spectrophotometry using a commercial anthocyanin standard (cyanidin 3-galactoside also known commercially as ideain chloride, Roth Biochemicals) as reference. Spectrophotometric analyses were followed by HPLC and ESI-MS to confirm the presence of cyanidin 3-galactoside in the red phenotype.

For anthocyanin identification, HPLC followed by ESI-MS is the method of choice for the majority of investigators because of its unique advantages. HPLC with tandem ESI-MS provides the intact molecular ion as well as fragment ions by collision-induced decomposition (CID) technology in the same run. Cyanidin 3-galactoside identification

and peak assignment was based on comparison of its retention times and mass spectral data with that of the idaein standard and published data. Cyanidin 3-galactoside was detected according to the respective m/z values of its parent (287 m/z) and product ions 449 m/z. Fragment peaks generated by HPLC were further characterised by NMR.

Equal weights of fresh leaves from red or green forms were used for extraction in acidic methanol. Extracts were filtered through a 0.45 mm nylon filter and subjected to HPLC followed by LC-ESI-MS.

3.1 Quantification of anthocyanin in red and green phenotypes of 'Bon Rouge'

Anthocyanin concentration was determined by measuring absorbance at 530 nm and calculated with reference to the standard, idaein (cyanidin 3-galactoside) chloride. The red leaf extract had a significantly higher concentration of cyanidin 3-galactoside at 50 mg/L compared to 1.0 mg/L for the green leaf extract (table 3.1).

Table 3.1 Quantification of anthocyanin pigment concentration in red and green leaf extracts of 'Bon Rouge'

Phenotype	Absorbance units (Au)	Concentration
Idaein standard	1.0	100.0 mg/L
Red leaf extract	0.5	50.0 mg/L
Green leaf extract	0.01	1.0 mg/L

3.2 HPLC analysis of leaf pigment extracts

HPLC analyses of red and green leaf extracts displayed differences with respect to their pigment profiles when compared to the standard, cyanidin 3-galactoside. At 530 nm the red leaf extract in acidified methanol displayed a single peak with a retention time of 10 minutes (fig. 3.2.2.1) similar to that for cyanidin 3-galactoside (fig. 3.2.1.1). At 280 nm, the chromatogram for cyanidin 3-galactoside displayed two major peaks at retention times of 10 and 12 minutes respectively (fig.3.2.1.2) while the red leaf extract displayed three major peaks at retention times of 10, 12 and 15 minutes respectively (fig. 3.2.2.2). Chromatograms for red leaf extracts in acetone displayed a single major peak at a retention time of 10 minutes when collected at 530 nm (fig. 3.2.2.3) and 280 nm (fig. 3.2.2.4) respectively, while the green leaf acetone extract displayed a single peak at a retention time of 25 minutes when collected at 530 nm (fig. 3.2.3.1).

3.2.1 HPLC analysis of cyanidin 3-galactoside (idaein) (standard)

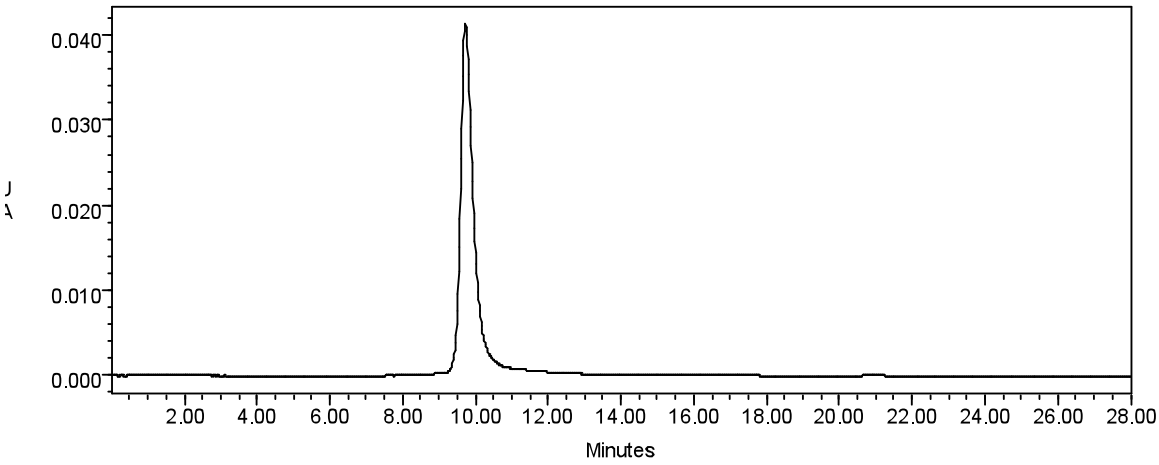


Figure 3.2.1.1 Chromatogram of idaein in 1% HCl methanol collected at 530 nm. Injection volume = 100 μL.

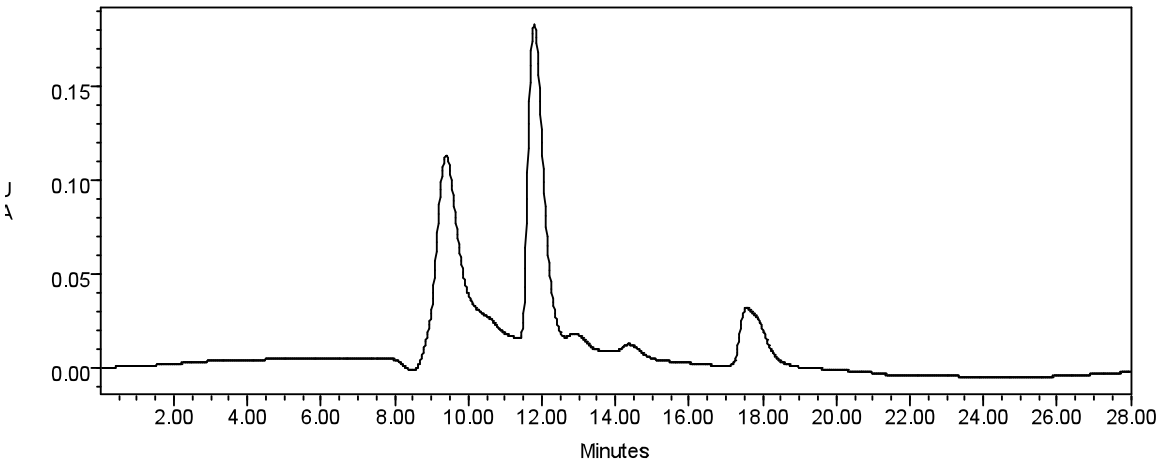


Figure 3.2.1.2 Chromatogram of idaein in 1% HCl methanol collected at 280 nm. Injection volume = 100 μL.

3.2.2 HPLC analysis of red 'Bon Rouge' leaf pigment extract

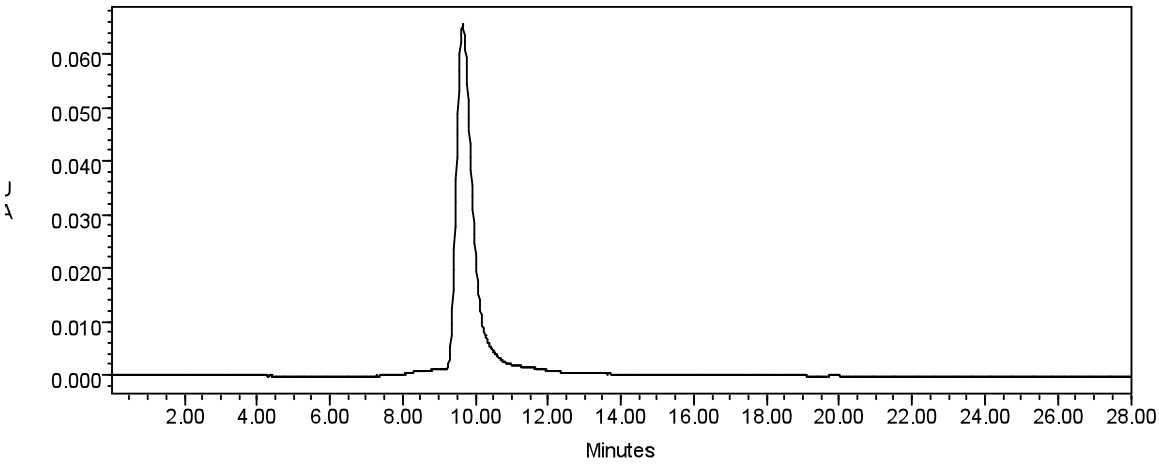


Figure 3.2.2.1 Chromatogram of red leaf pigment extract in 1% HCl methanol collected at 530 nm. Injection volume = 100 μL.

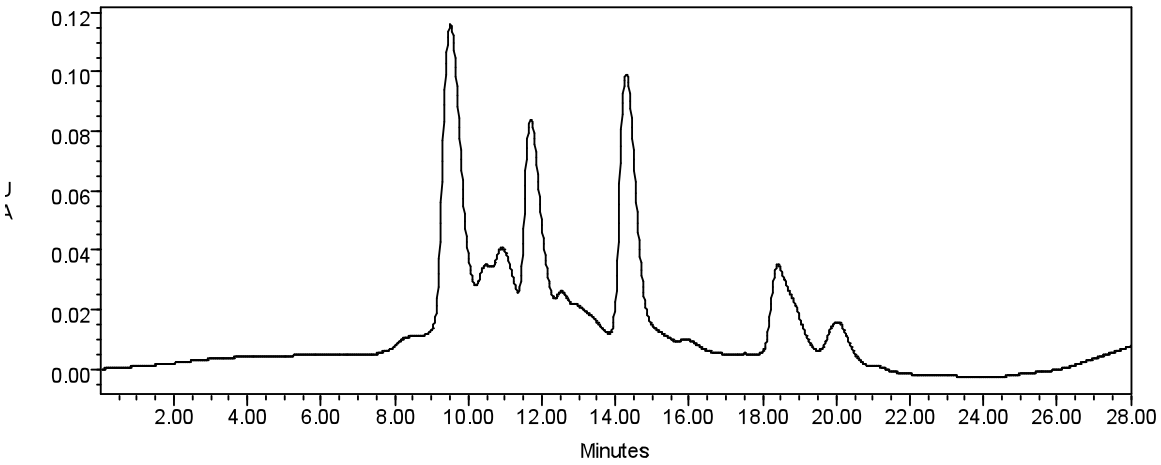


Figure 3.2.2.2 Chromatogram of red leaf pigment extract in 1% HCl methanol collected at 280 nm. Injection volume = 100 μL.

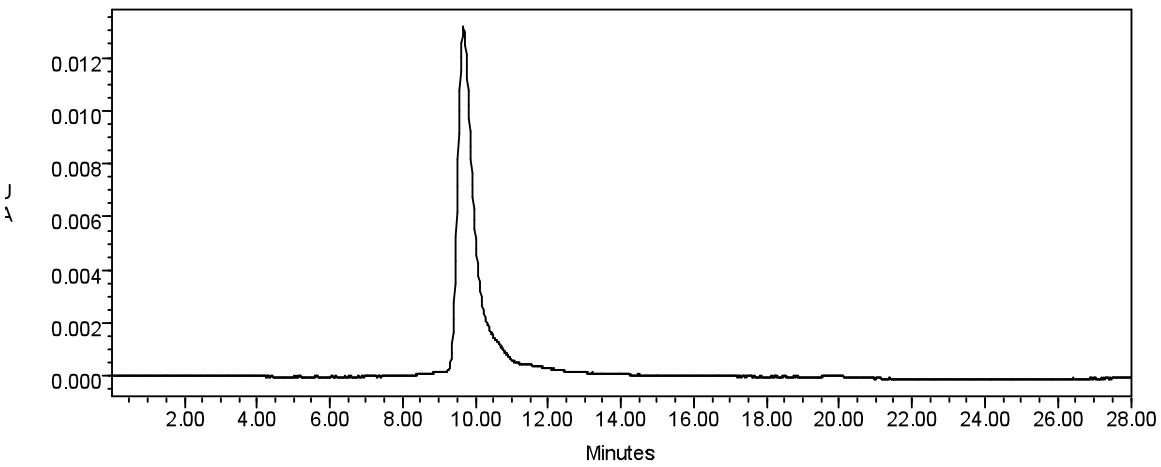


Figure 3.2.2.3 Chromatogram of red leaf pigment extract in acetone collected at 530 nm. Injection volume = 100 μL.

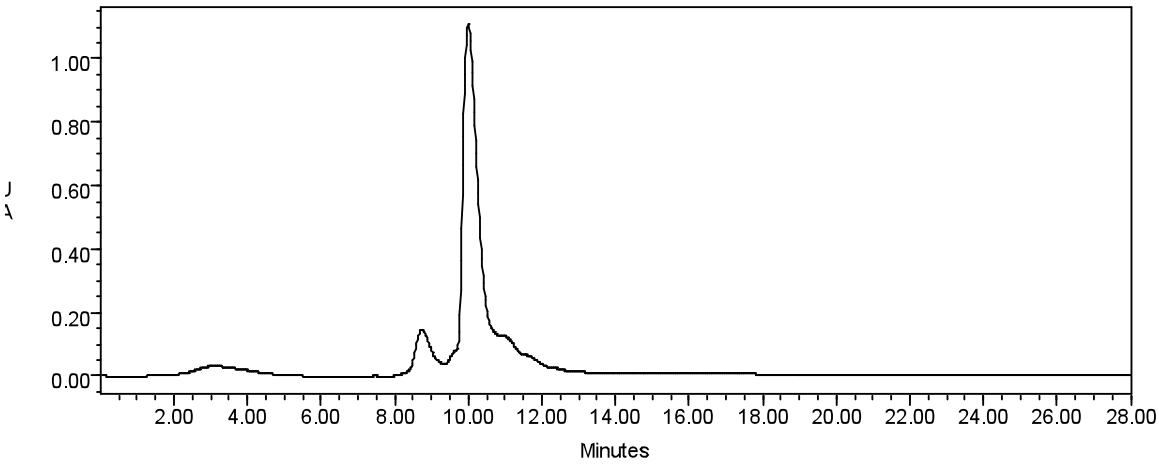


Figure 3.2.2.4 Chromatogram of red leaf pigment extract in acetone collected at 280 nm. Injection volume = 100 μL.

3.2.3 HPLC analysis of reverted 'Bon Rouge' leaf pigment extracts

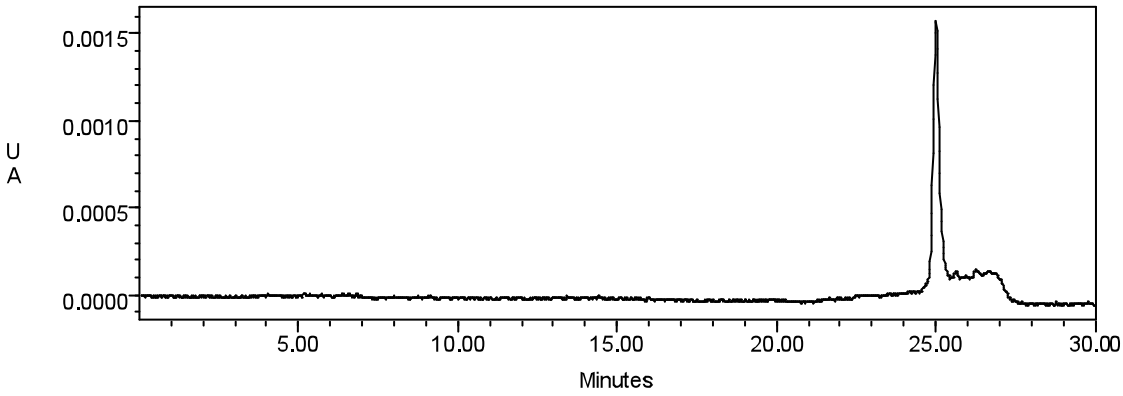


Figure 3.2.3.1 Chromatogram of green (reverted) leaf pigment extract in acetone collected at 530 nm. Injection volume = 100 μL.

3.3 LC-MS of HPLC peaks

The expanded LC-MS spectra of idaein standard display a single 447 *m/z* ion (fig. 3.3.1). The red leaf extract displayed the major 449 *m/z* ion but also the 447 *m/z* ion which may have resulted from the reduction of anthocyanins to epicatechins by anthocyanidin reductase (ANR), and another minor peak at 447 *m/z*. A minor peak corresponding to a 449 *m/z* ion is also present in the spectrum for red leaf extract.

The ES^- tandem mass spectrum of the 449 *m/z* ion (fig. 3.4.1) selected in the LC-MS spectrum of red leaf extract (fig. 3.3.1) show a major 287 *m/z* ion that correspond to the cyanidin core of cyanidin 3-galactoside. This 287 *m/z* ion resulted from the loss of the sugar moiety (449 *m/z* ion – 162 *m/z* ion).

The ES^- tandem mass spectrum of 447 *m/z* ion (fig. 3.4.2) identified in the LC-MS spectrum of red leaf extract (fig. 3.3.1) show a major 284 *m/z* ion and the original 447 *m/z* ion. This 284 *m/z* ion may represents the flavanol core of a flavanol glycoside that resulted from the loss of a sugar moiety from the 447 *m/z*.

Because the idaein standard displayed a 447 *m/z* ion instead of the expected 449 *m/z* ion in its LC-MS spectrum (fig. 3.3.1) and a major 284 *m/z* ion (fig. 3.4.2) instead of the expected 287 *m/z* ion, the idaein standard was subjected to Nuclear Magnetic Resonance (NMR) analysis to confirm its structure as cyanidin 3-galactoside.

3.3.1 LC-MS of idaein standard and red leaf pigment extracts

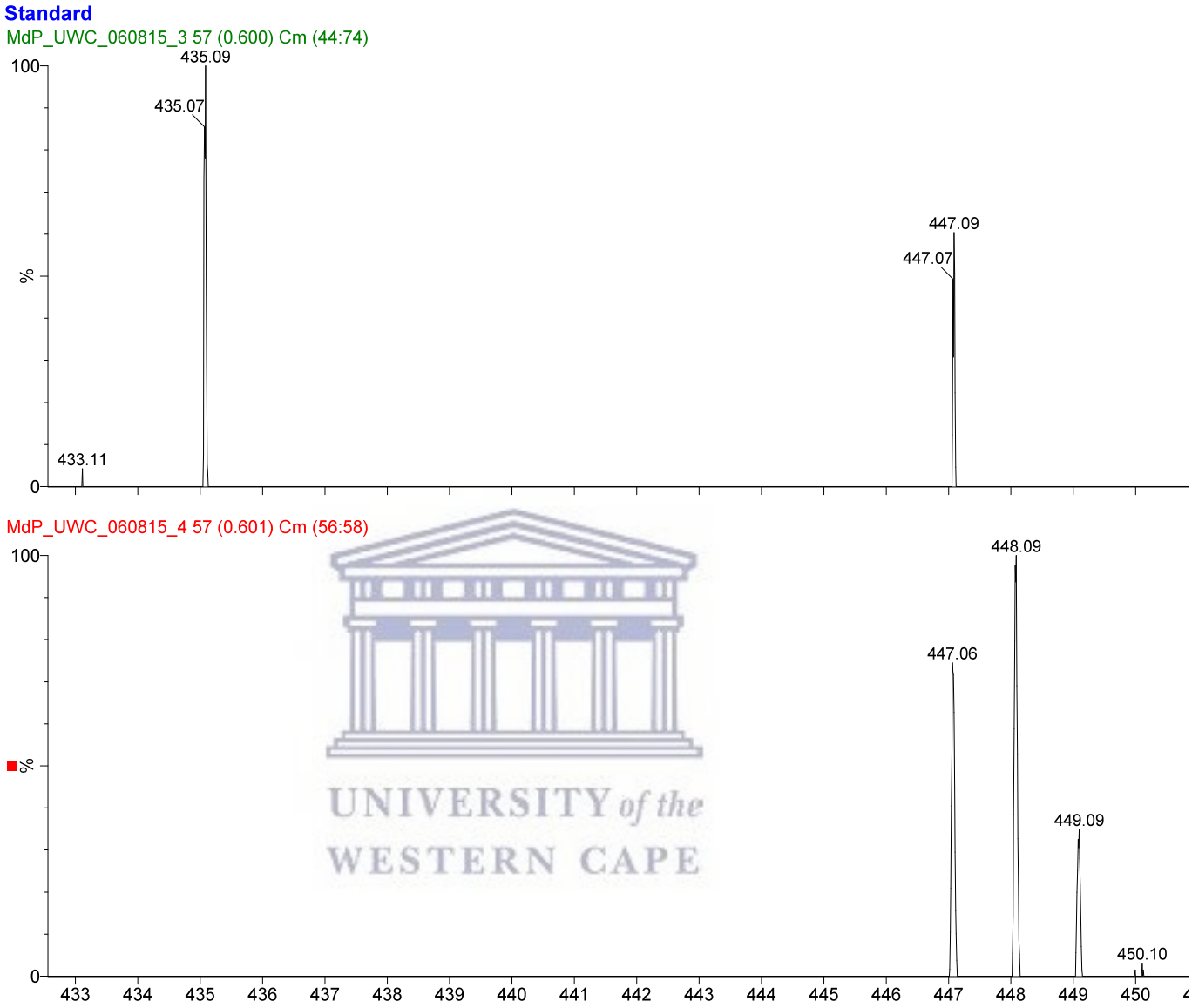


Figure 3.3.1 Expanded LC-MS spectra of idaein standard (above) and red leaf extract (below).

3.4 MS-MS analysis of ions generated during LC-MS

3.4.1 MS-MS analysis of 449 *m/z* ion generated during LC-MS

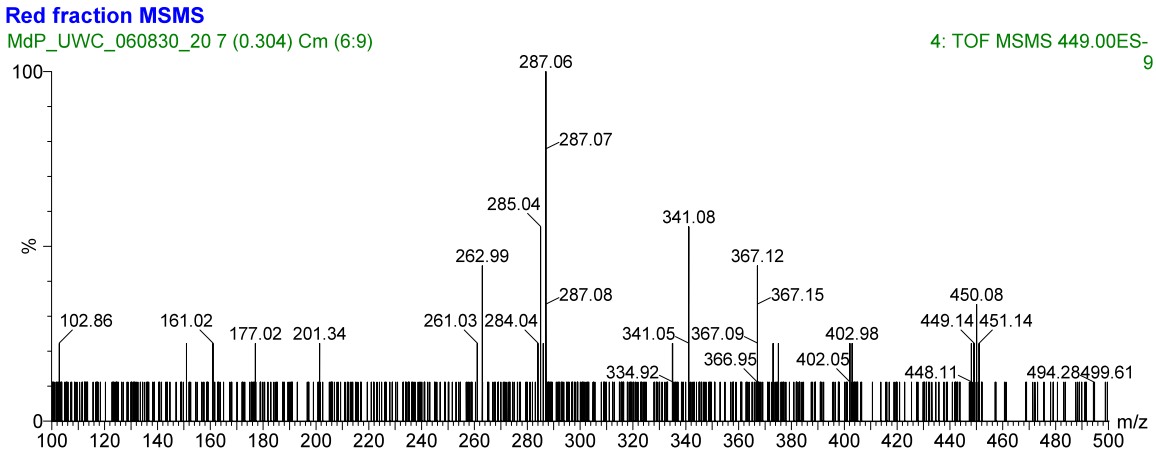


Figure 3.4.1 ES⁻ tandem mass spectrum of 449 *m/z* ion selected in the LC-MS spectrum of red phenotype leaf extract.

3.4.2 MS-MS analysis of 447 *m/z* ion generated during LC-MS

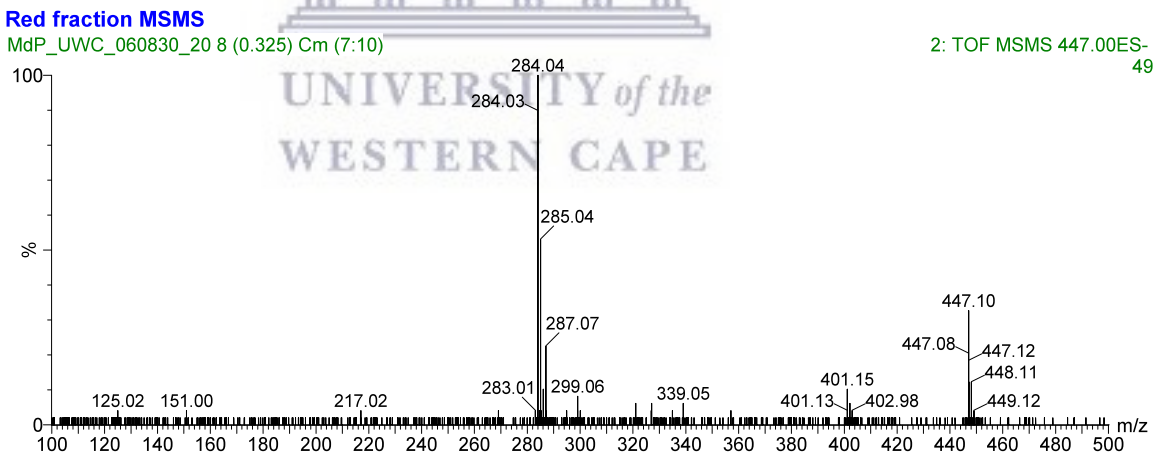


Figure 3.4.2 ES⁻ tandem mass spectrum of 447 *m/z* ion identified in the LC-MS spectrum of red phenotype leaf extract.

3.5 NMR analysis of HPLC fractions

According to the ^{1}H NMR spectrum for the idaein standard (table 3.5.1.1 and fig. 3.5.1.1), all the protons for cyanidin 3-galactoside could be assigned. No O*H* protons were visible in the ^{1}H NMR spectrum of the idaein standard due to exchange with the alcohol deuterium of CD_3O*D*. Similarly all the carbons for cyanidin 3-galactoside could be assigned in the ^{13}C NMR spectrum of the idaein standard (table 3.5.1.2 and fig. 3.5.1.2).

Further analysis by COSY spectrum (fig 3.5.1.3), ghmqc spectrum (fig. 3.5.1.4) and ghsqc spectrum (fig. 3.5.1.5) confirmed the structure of the idaein standard as cyanidin 3-galactoside (fig. 3.5.3.1).

The LC-MS spectrum of the red leaf extract displayed a fragment ion at 435 m/z (fig. 3.3.1). This fraction was sampled, designated RZ, and analysed by NMR. All protons could be assigned for the minor pigment, RZ, extracted from red leaves according to the ^{1}H NMR spectrum (table 3.5.2.1 and fig. 3.5.2.1).

Carbon were assigned according to the ^{13}C NMR spectrum for the minor pigment, RZ, extracted from red leaves (table 3.5.2.2 and fig. 3.5.2.2). With the ^{1}H spectrum assigned for the phenyl ring and the double bond, the corresponding ^{13}C resonances could be identified using the ghsqc spectrum (fig. 3.5.2.4). Further analysis by COSY spectrum (fig. 3.5.2.3), ghmqc spectrum (fig. 3.5.2.4) and ghsqc spectrum (fig. 3.5.2.5) indicated the structure of the minor pigment, RZ, in the red leaf extract as 2,3-dihydroxycyclopentyl (2*E*)-3-(3,4-dihydroxyphenyl) acrylate (fig. 3.5.3.2).

3.5.1 NMR analysis of cyanidin 3-galactoside (idaein standard)

Table 3.5.1.1 Proton assignment for cyanidin 3-galactoside. No O*H* protons are visible in the ^{1}H NMR spectrum due to exchange with the alcohol deuterium of CD_3OD.)

^{1}H NMR Spectrum:

Chemical Shift (ppm)	No. of protons	Multiplicity	Coupling constant/s	Assignment
9.02	1	s	-	H4
8.26	1	dd	2.3, 8.7	H6'
8.06	1	d	2.3	H2'
7.01	1	d	8.7	H5'
6.89	1	d	1.8	H8
6.65	1	d	1.8	H6
5.26	1	d	7.7	H1"
3.99	1	dd	7.7, 9.7	H2"
3.96	1	d	3.3	H4"
3.76 to 3.80	3	m	-	H5" and 2H6"
3.68	1	dd	3.3, 9.7	H3"

Table 3.5.1.2 Carbon atom assignment for cyanidin 3-galactoside (idaein standard)

^{13}C NMR spectrum:

Chemical shift (ppm)	Assignment
170.29	C7
164.33	C2
159.14	C5
157.60	C9
155.78	C4'
147.37	C3'
145.71	C3
136.88	C4
128.26	C6'
121.22	C1'
118.43	C2'
117.40	C5'
113.33	C10
104.40	C1"
103.28	C6
95.09	C8
77.78	C5"
74.90	C3"
72.07	C2"
70.09	C4"
62.33	C6"

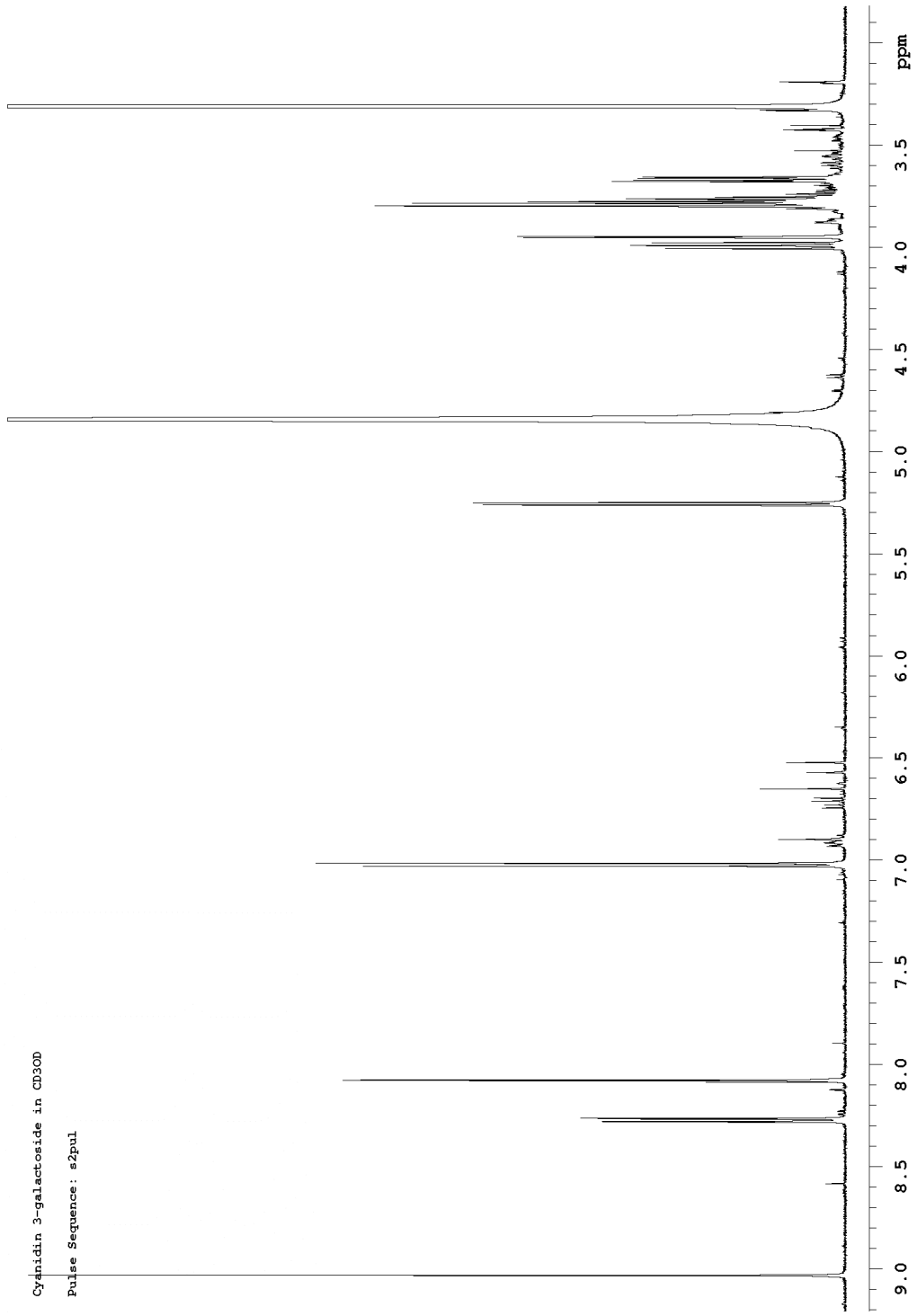


Figure 3.5.1.1 ^{1}H spectrum of cyanidin 3-galactoside

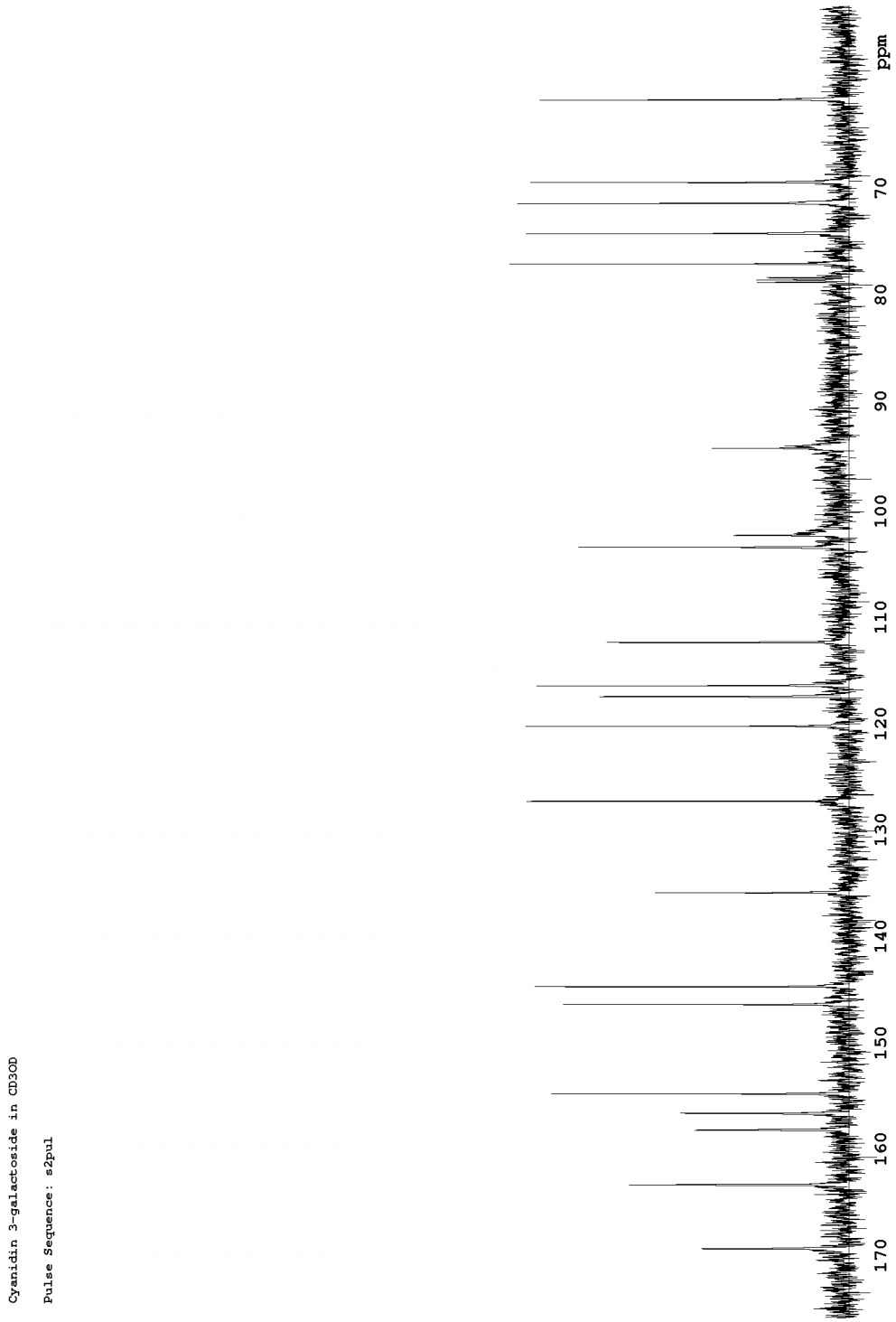


Figure 3.5.1.2 ^{13}C spectrum of cyanidin 3-galactoside

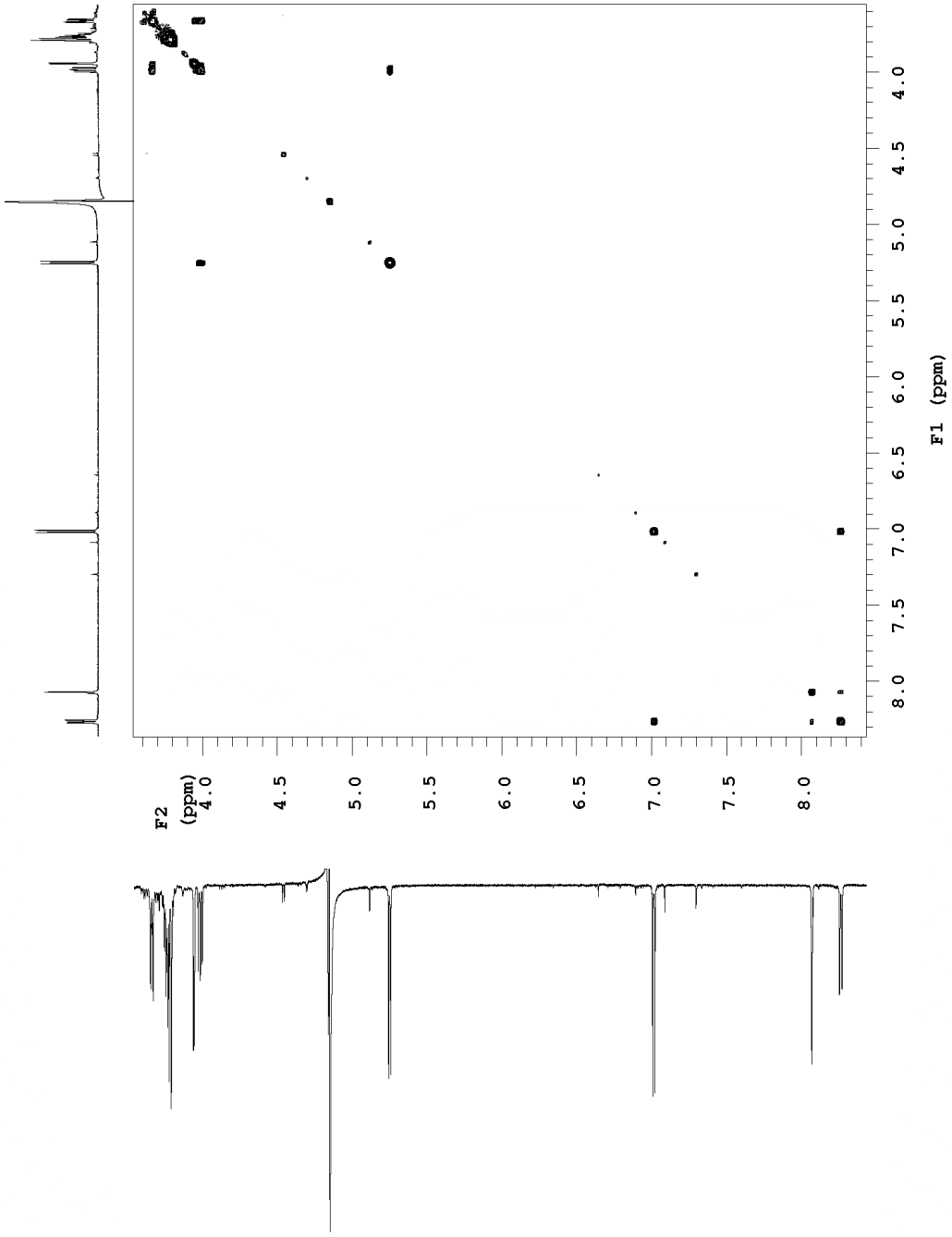


Figure 3.5.1.3 COSY spectrum of cyanidin 3-galactoside

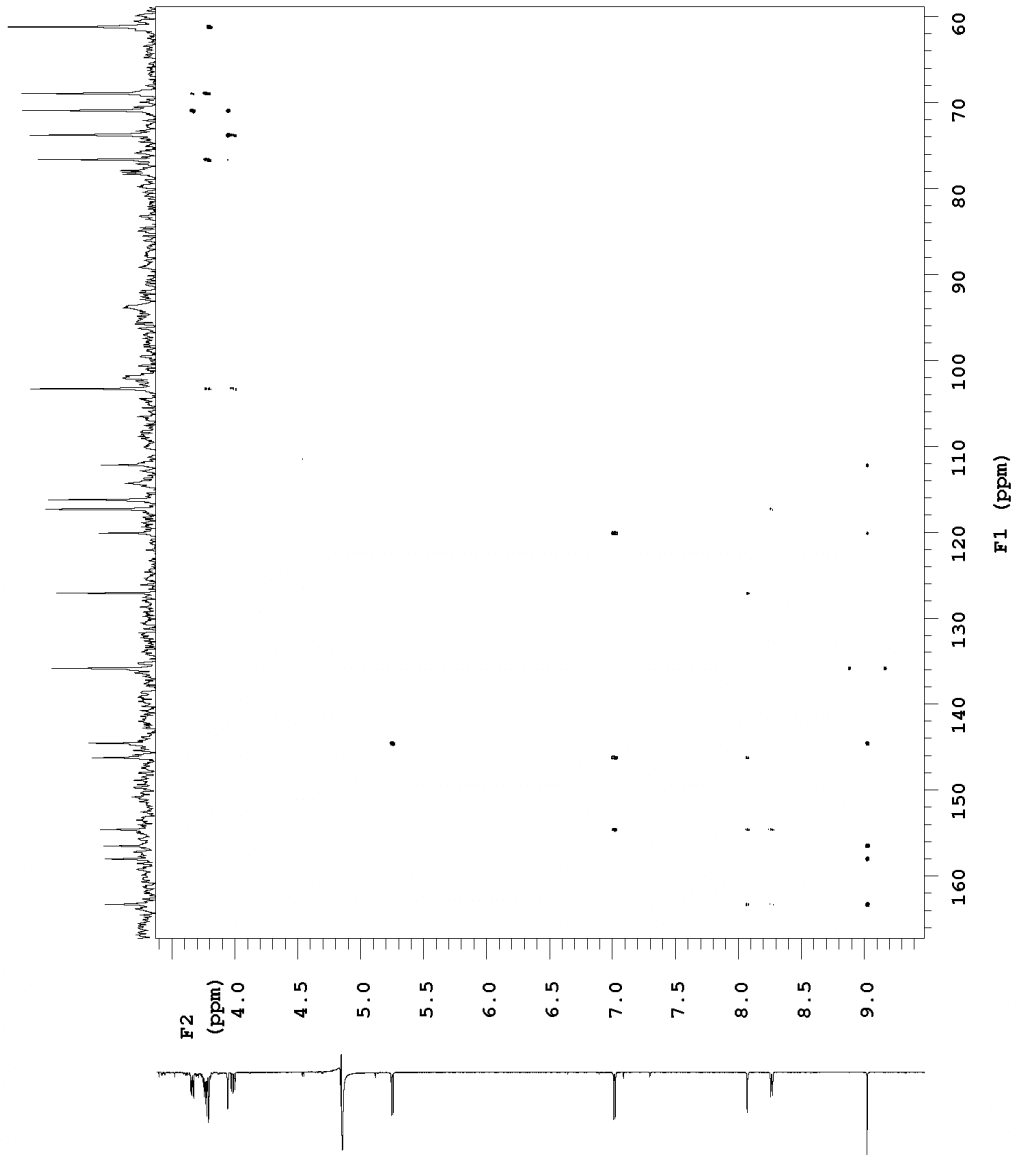


Figure 3.5.1.4 ghmqc spectrum of cyanidin 3-galactoside

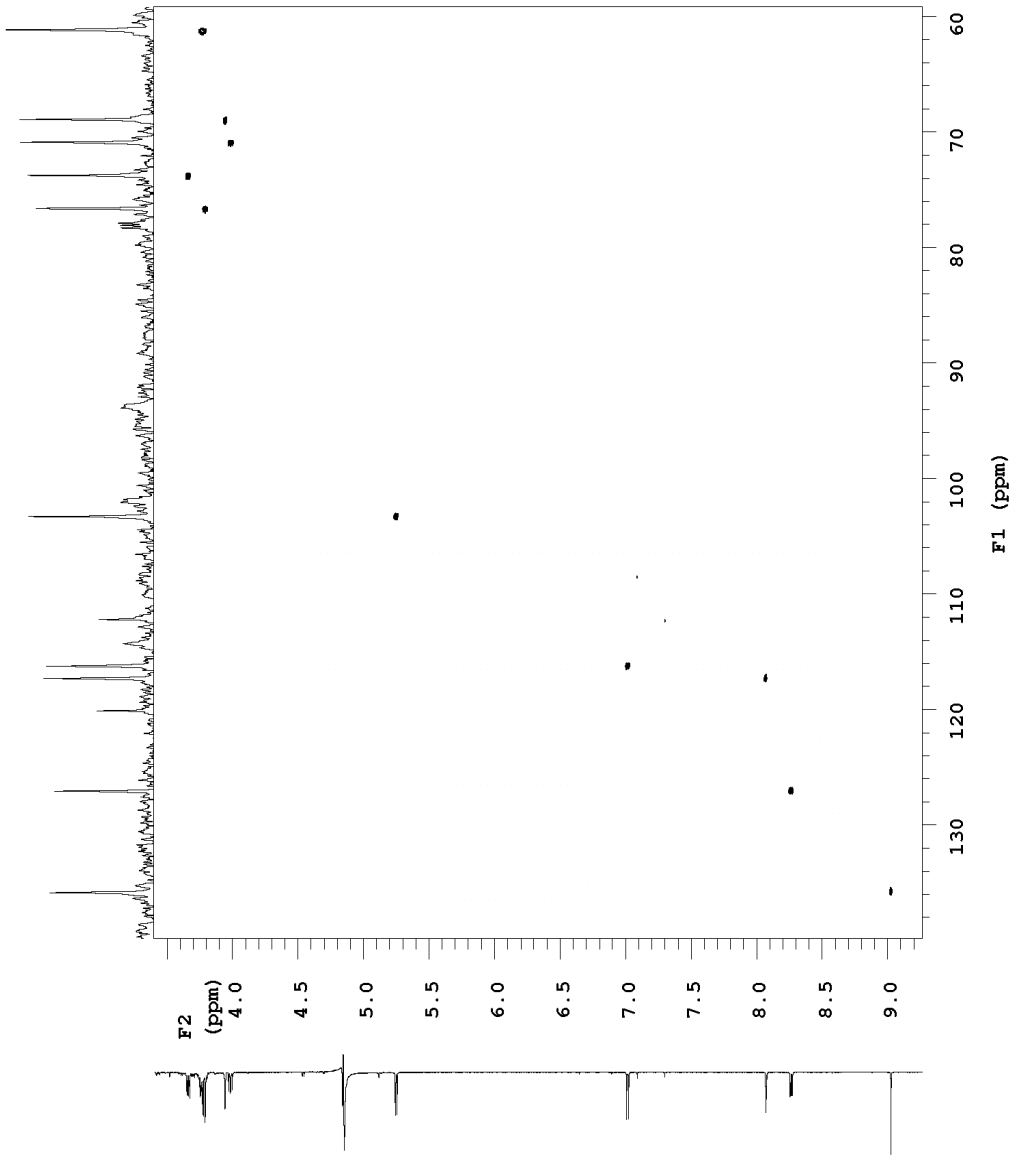


Figure 3.5.1.5 ghsqc spectrum of cyanidin 3-galactoside

3.5.2 NMR analysis of a minor red leaf extract pigment, RZ

Table 3.5.2.1 Proton assignment with the ^{1}H NMR spectrum for the minor red pigment, RZ, extracted from red leaves

^{1}H NMR Spectrum:

Chemical Shift (ppm)	No. of protons	Multiplicity	Coupling constant/s	Assignment
8.20	1H	s	-	H6
7.56	1H	d	15.9	H12
7.05	1H	d	2.0	H14
6.95	1H	dd	2.0,8.2	H18
6.78	1H	d	8.2	H17
6.27	1H	d	15.9	H11
5.35	1H	multiplet	-	H2
4.15	1H	multiplet	-	H5
3.71	1H	dd	3.1,8.9	H1
2.16 to 2.21	2H	multiplet	-	$H3_a$ and $H4_a$
2.01 to 2.08	2H	multiplet	-	$H3_b$ and $H4_b$

Table 3.5.2.2 Carbon assigment using the ^{13}C NMR spectrum for the minor red pigment, RZ, extracted from red leaves. With the ^{1}H spectrum assigned for the phenyl ring and the double bond, the corresponding ^{13}C resonances could be identified using the ghsqc spectrum

^{13}C NMR spectrum:

Chemical shift (ppm)	Assignment
168.79	C8
149.56	C16
146.98	C12
146.81	C15
127.83	C13
122.94	C18
116.48	C17
115.38	C11
115.16	C14
73.97	C1
72.20	C2
71.85	C5
39.34	C3
38.45	C4

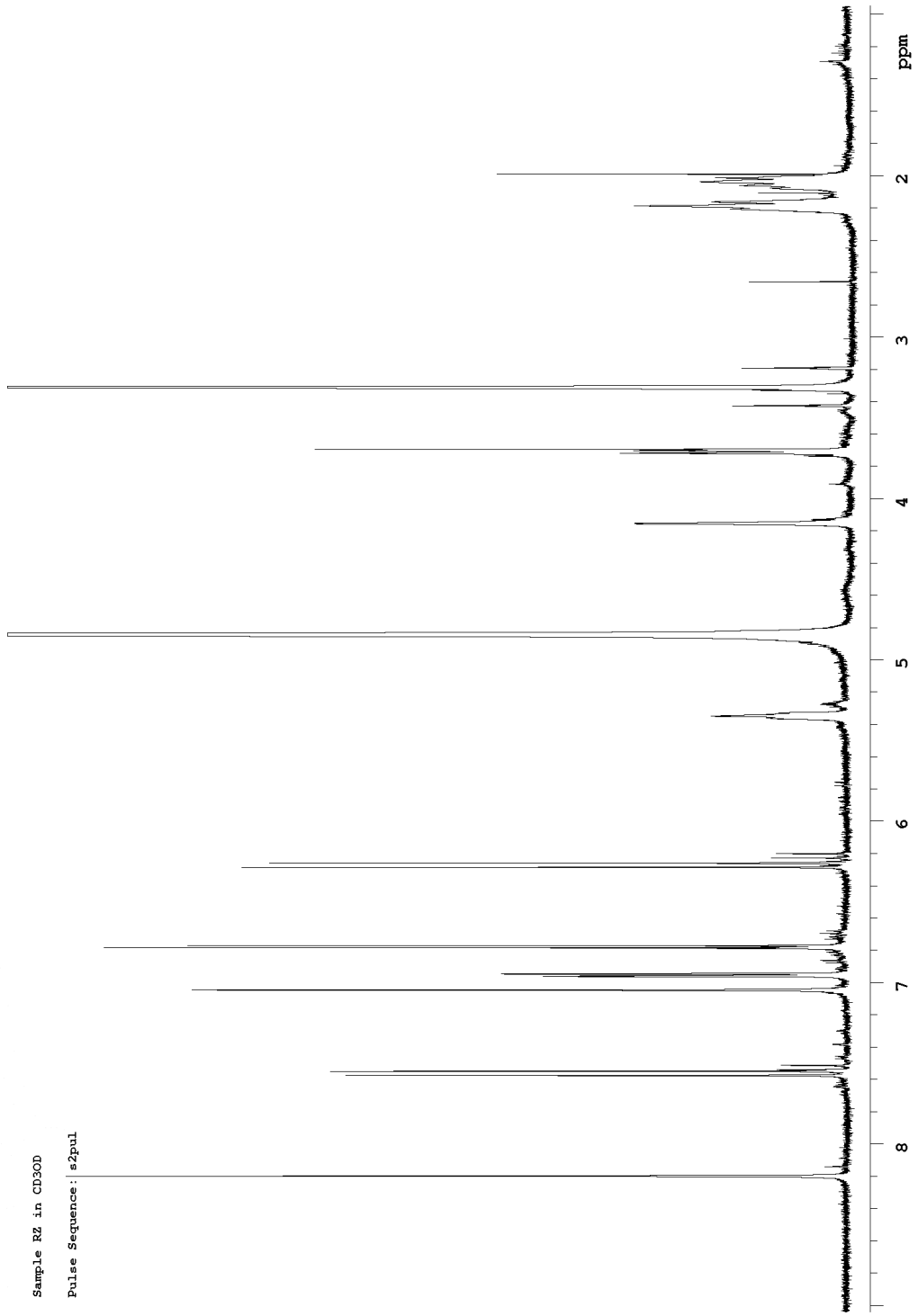


Figure 3.5.2.1 ^{1}H spectrum of RZ

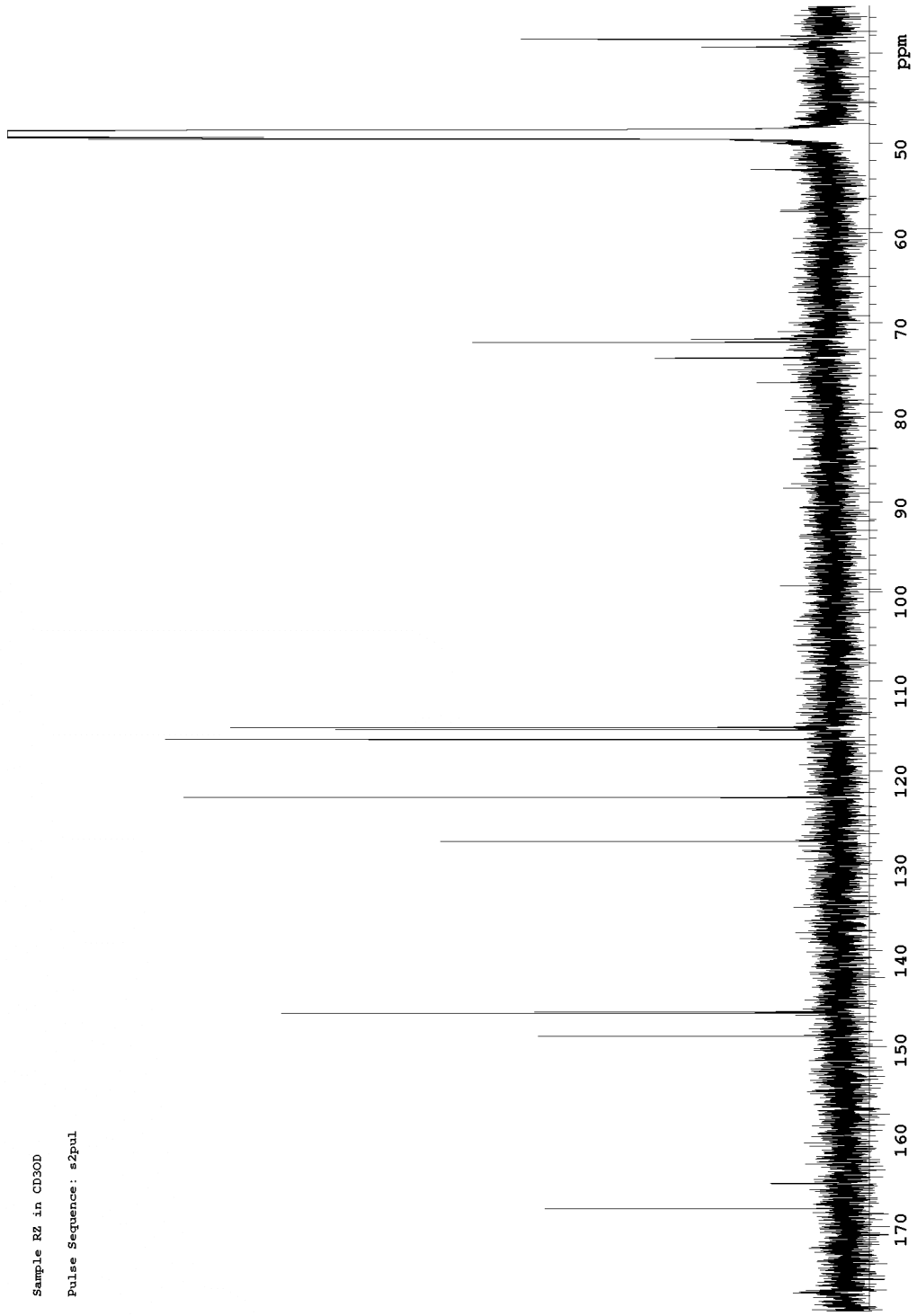


Figure 3.5.2.2 ^{13}C spectrum of RZ

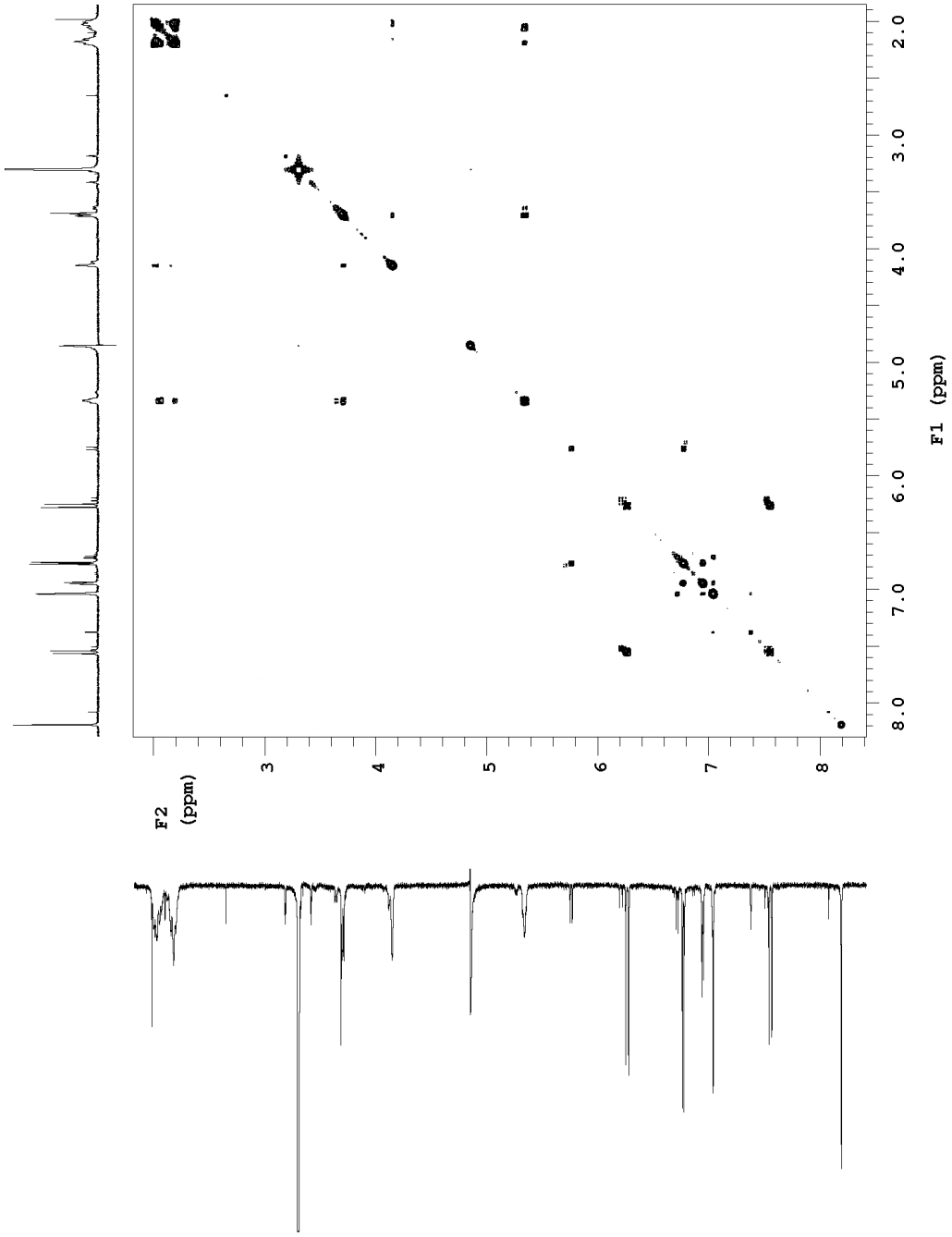


Figure 3.5.2.3 COSY spectrum of RZ

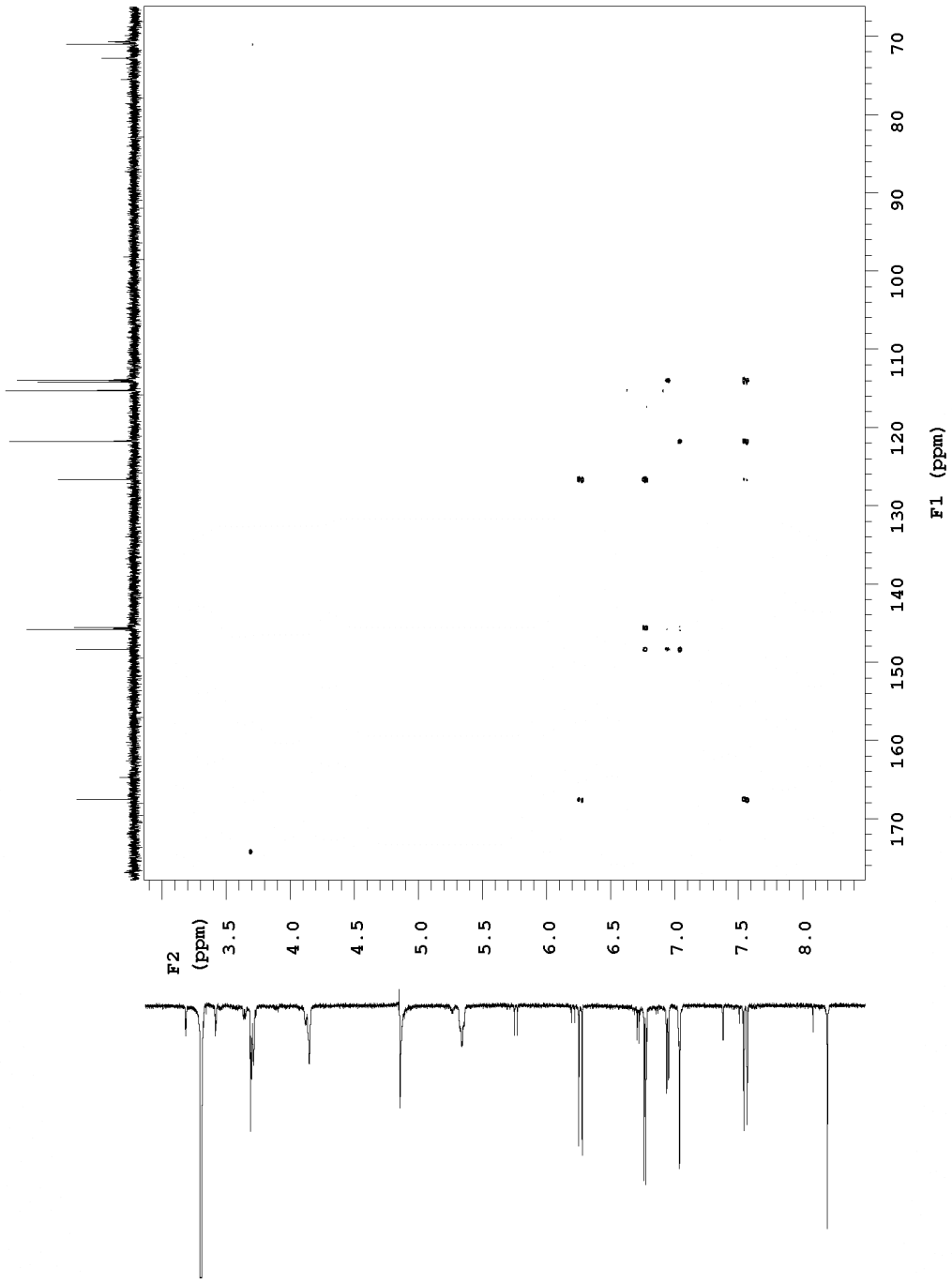


Figure 3.5.2.4 ghmqc spectrum of RZ

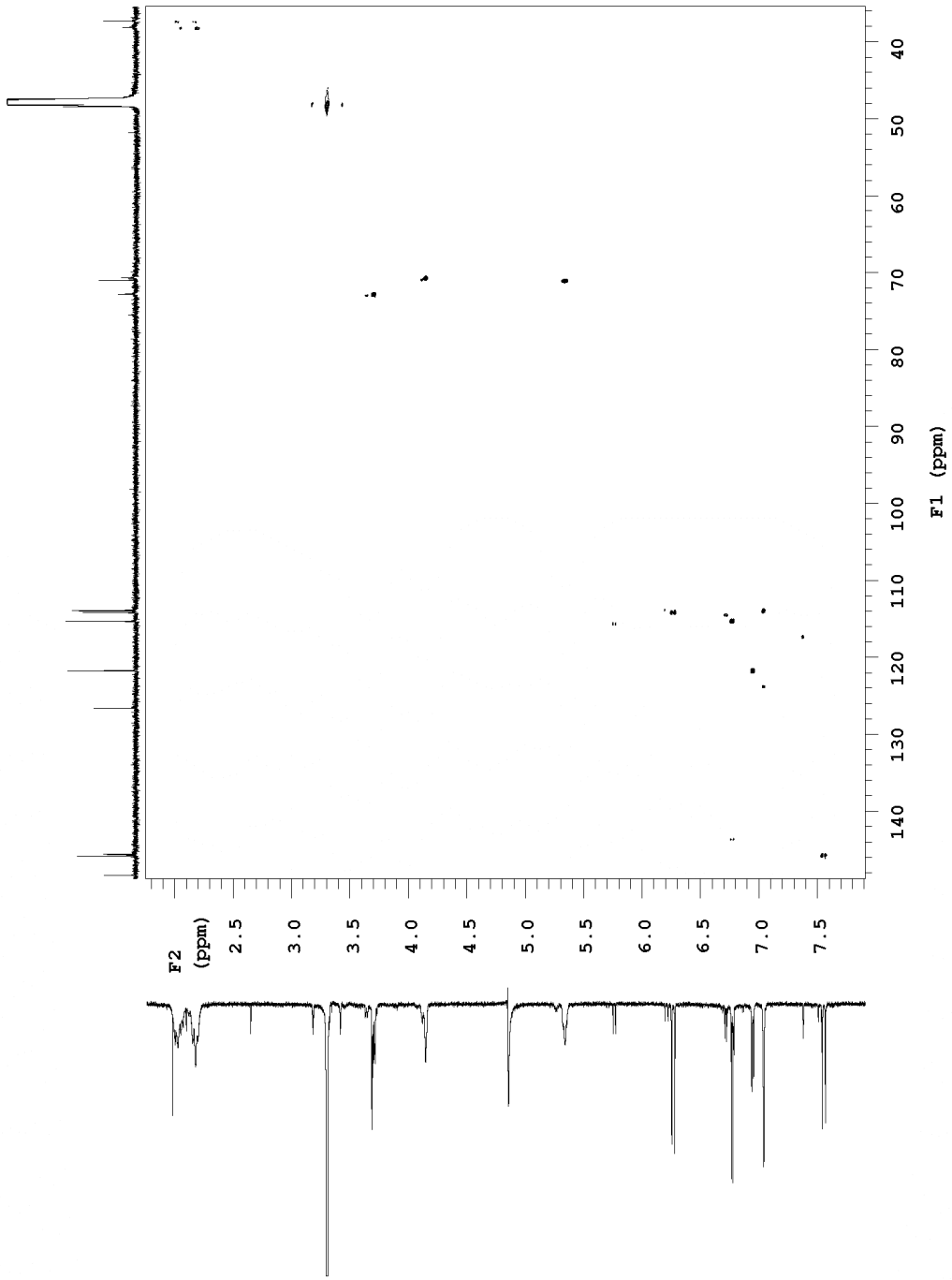


Figure 3.5.2.5 ghsqc spectrum of RZ

3.5.3 Pigment structure identification by NMR

3.5.3.1 Structure of cyanidin 3-galactoside standard

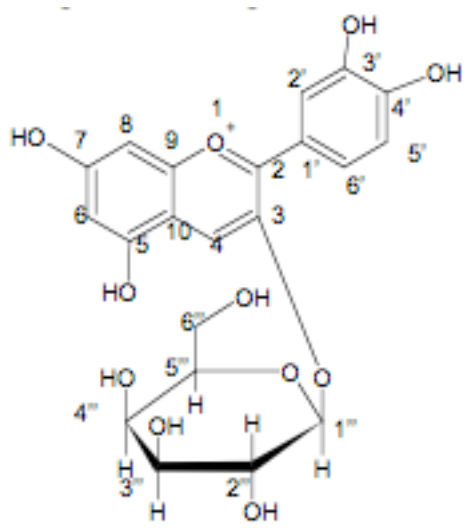


Figure 3.5.3.1 cyanidin 3-galactoside (chloride ion)

3.5.3.2 Structure of pigment RZ isolated from red leaves

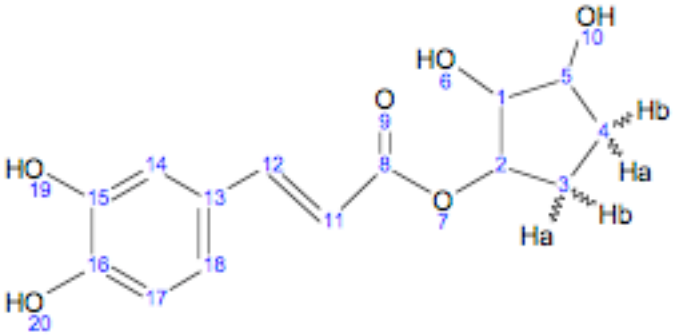


Figure 3.5.3.2 2,3-dihydroxycyclopentyl (2*E*)-3-(3,4-dihydroxyphenyl) acrylate (NB: numbering is arbitrary).

3.6 Results and Discussion

3.6.1 Quantification of anthocyanin concentration

Anthocyanin pigment concentration was measured in 0.1% HCl methanol extracts for both phenotypes in a Nanodrop spectrophotometer™ and by comparison with the commercial standard, idaein chloride. Red leaf extracts contain 50.0 mg/L cyanidin 3-galactoside compared to 1.0 mg/L for green leaf extracts (table 3.1). HPLC analyses at 530 nm confirmed the presence of a single pigment peak at 10 minutes (fig. 3.2.2.1) in the red leaf extract corresponding to cyanidin 3-galactoside when compared to the commercial standard, idaein (fig. 3.2.1.1). A single peak at 25 minutes was visible in the HPLC spectrum of reverted (green) leaf extract collected at 530 nm (fig. 3.2.3.1). The presence of a pigment peak in the visible range (530 nm) and the low concentration of pigment measured in the Nanodrop™ is consistent with the production of some colour pigment in reverted leaves but at a level significantly lower than in 'Bon Rouge' leaves. The pigment in green leaf extracts has not been fully characterised.

3.6.2 Characterisation of cyanidin 3-galactoside by LC-MS

To confirm the presence of the major pigment, cyanidin 3-galactoside, we analysed the peak collected during HPLC of the red leaf extract and on the standard cyanidin 3-galactoside, idaein chloride, by liquid chromatography mass spectrometry (LC-MS) and tandem mass spectrometry (MS-MS) on. LC-MS and MS-MS spectra were collected in negative ion mode and the spectrum for red leaf extract displayed three major ions (fig. 3.3.1 bottom). In addition to the expected

449 *m/z* ion representing cyanidin 3-galactoside, two additional ions at 448 *m/z* and 447 *m/z* respectively, were visible in the spectrum. These two ions most likely result from the reduction around the C2-C3 bond of the C-ring of the anthocyanidin core to epicatechin by anthocyanidin reductases. Reduction of the cyanidin core to epicatechin has been reported for a number of species, including apple (Pfeiffer *et al.,* 2006). Unexpectedly, the commercial standard displayed an ion at 447 *m/z* instead of the expected 449 *m/z* (fig. 3.3.1 top). The 447 *m/z* ion could result from the loss of protons from the cyanidin core. ES^- MS-MS analyses of the 449 *m/z* ion and the 447 *m/z* ion identified a cyanidin core for both ions. The cyanidin core of cyanidin 3-galactoside is represented by the 287 *m/z* ion that result from the loss of the sugar moiety (449 *m/z* – 162 *m/z*). Similarly the flavanol core of the reduced cyanidin 3-galactoside is represented by the 285 *m/z* ion that result from the loss of the sugar moiety (447 *m/z* – 162 *m/z*). The presence of a major ion, 284 *m/z* in the mass spectrum in addition to the expected 285 *m/z* most likely results from the loss of an additional H^+ from the reduced cyanidin core.

3.6.3 NMR Structure determination cyanidin 3-galactoside

Standard cyanidin 3-galactoside was dissolved in CD_3OD and spectra were collected in a Varian ^{Unity}*Inova* 600 NMR spectrometer with a ^1H frequency of 600 MHz and a ^{13}C frequency of 150 MHz. The ^1H spectrum was referenced to the CD_2*H*OD peak at 3.31 ppm and the ^{13}C spectrum to the *C*D_3OD peak at 49.5 ppm . In addition to the 1-dimensional spectra, 2-dimensional ^1H-^1H COSY (fig.

3.5.1.3), ghsqc (fig. 3.5.1.5) and ghmqc (fig. 3.5.1.4) spectra were also carried out to assist in the assignments of the ^{1}H (fig. 3.5.1.1) and ^{13}C (fig. 3.5.1.2) spectra. The numbering of the cyanidin backbone follows convention, as does that of the galactoside ring. No O*H* protons were visible in the ^{1}H NMR spectrum due to the exchange with the alcohol deuterium of CD_3O*D*.

Based on the chemical shifts, multiplicities and coupling constants almost all non-exchangable protons for the cyanidin backbone in the idaein chloride sample could be unambiguously assigned from the ^{1}H NMR spectrum. The downfield region of the ^{1}H NMR spectrum of idaein chloride showed a singlet at δ 9.02 (H-4), a 3H AMX system at δ 8.09 (dd, 8.7 Hz, 2.3 Hz, H-6'), 7.86 (d, 2.3 Hz, H- 2') and 6.88 (d, 8.7 Hz, H-5'), and an AB system at δ 6.76 (δ, 1.8 Hz, H-8) and 6.61 (δ, 1.8 Hz, H-6), which is in accordance with a cyanidin derivative. The sugar region showed the presence of only one sugar unit. The anomeric coupling constant (7.7 Hz) and the six ^{13}C resonances in the sugar region of the ^{13}C spectrum of idaein chloride were in accordance with ß-glucopyranose, the sugar moiety in cyaniding 3-galactoside. The pigment isolated from the red leaf extract co-chromatographed (HPLC) with authentic cyanidin 3-galactoside standard, idaein chloride. The ^{1}H-^{1}H COSY spectrum was able to confirm the assignments. The only protons where assignment was difficult was for those of H6 and H8. Initial observations indicated the H8 and H6 protons are not typical exchangeable protons such as the -OH's, however when the tautomeric forms of the ring concerned are considered the exchange can be understood:

Figure 3.6.1 Tautomeric forms of the ring structure (see Appendix 7)

This exchange was confirmed in the ^{13}C spectrum where the C8 and C6 carbons are small multiplets which appear to show some kind of coupling as would be expected if connected to deuterium. The deuterium exchange of H6 and H8 has also been observed by other workers (Wang *et al.,* 1999). In general the assignments of the cyandin backbone agree well with those for similar compounds found in the literature (Wang *et al.,* 1999). For the galactoside ring protons the anomeric proton (H1") can be easily assigned based on its down field chemical shift of 5.26 ppm (table 3.5.1.1) relative to the other remaining protons and the typical anomeric proton coupling constant of 7.7 Hz. Based only on coupling constants obtained from the ^{1}H NMR spectrum the protons H2", H4" and H3" could subsequently easily be assigned. The assignments agreed with the correlations in the ^{1}H-^{1}H COSY spectrum (fig. 3.5.1.3) and the remaining multiplet integrating for three protons could then be assigned to the remaining protons, H5" and the two H6" protons.

3.6.4 NMR Structure determination of a minor pigment RZ

The sample was dissolved in CD_3OD and spectra collected on a Varian ^{Unity}*Inova* 600 NMR spectrometer with a ^{1}H frequency of 600 MHz and a ^{13}C frequency of

150 MHz. The ^{1}H spectrum was referenced to the CD_2HOD peak at 3.31 ppm and the ^{13}C spectrum to the CD_3OD peak at 49.0 ppm. In addition to the 1-dimensional spectra, 2-dimensional ^{1}H-^{1}H COSY, ghsqc and ghmqc spectra were also generated to assist in the assignments of the ^{1}H and ^{13}C spectra.

Because the fragment peak appeared in the visible range of the spectrum it was expected that this compound was a cyanidin derivative. However, the notable absence of certain key peaks of the cyanidin backbone, as was shown for cyanidin 3-galactoside above, indicated that this was not the case. Analysis of the ^{1}H (table and fig. 3.5.2.1), ^{13}C (table and fig. 3.5.2.2) and 2-dimensional spectra revealed the structure of the isolated compound to be that of 2,3-dihydroxycyclopentyl (2*E*)-3-(3,4-dihydroxyphenyl) acrylate (numbering is arbitrary). Compound RZ was obtained from the red leaf extract as a colorless oily product. The ^{1}H NMR spectrum of RZ revealed a 3,4-dihydroxylcinnamoyl moiety. However, the two multiplets in RZ, appearing at δ 2.16 to 2.21 and δ 2.01 to 2.08, were assigned to two methylene groups, respectively. The COSY experiment showed that two CH_2 protons in RZ were correlated and adjacent to each other and also coupled to other hydrogens. The ^{13}C NMR spectrum of this compound revealed that there were only one carbonyl carbon, eight methine carbons, and two methylene carbons. Three of the methine carbons at δ 73.97, 72.20, and 71.85 were oxygenated and showed correlations to three methine protons at δ 3.71, 5.35, and 4.15, respectively, as evident from the ghmqc spectrum. Also, five other methine carbons at δ 115.16, 116.48, 122.94, 146.81, and 115.38 showed correlations to three aromatic protons appearing at δ 7.05, 6.78, and 6.95 and two olefinic

protons at 7.56 and 6.27 ppm, respectively. Therefore, compound RZ was assigned as 1-(3″,4″-dihydroxy- l-cinnamoyl)-cyclopenta-2,3-diol.

This compound is a known natural product with antioxidant properties and has been previously isolated from tart cherries (Wang *et al.,* 1999). The most obvious features in the ^{1}H NMR spectrum were the phenyl ring and the trans double bond (with a large $^{3}J_{H-H}$ coupling constant of 15.9 Hz). The peak at 5.35 ppm initially appeared to be an anomeric proton of a sugar due to its shift, but the fact that it is a broad multiplet and that its attached carbon has a shift of 72.20 ppm, rather than approximately 100 ppm as would be expected, indicated that this was not the case. The peaks due to the phenyl and double bond protons were sharp while the peaks further upfield between 5.5 and 2.0 ppm were broad, and based on the shifts of these broad peaks, integrals, and on analysis of the cosy spectrum (fig. 3.5.2.3), the cyclopentyl ring could be identified. The diastereotopic nature of the CH_2's in the ring is clearly seen from the ghsqc spectrum (fig. 3.5.2.4). The sample was run in CD_3OD and thus due to deuterium exchange it is expected that no O*H* peaks will be seen in the ^{1}H spectrum. However the presence of a singlet at 8.20 ppm in the ^{1}H spectrum that has no directly bonded carbon atom, according to the ghsqc spectrum, is interesting. This peak integrates for one proton and can be assigned to one of the OH's on the cyclopentyl moiety. Due to the close proximity of the H6 proton and the carbonyl oxygen the formation of a hydrogen bond between the two atoms is highly probably (fig. 3.6.1), which will result in the prevention of any deuterium exchange at this particular OH group giving rise to the signal in the ^{1}H spectrum. With the ^{1}H spectrum assigned for the phenyl

ring and the double bond, the corresponding ^{13}C resonances could be identified using the ghsqc spectrum. The quaternary carbons could then also be identified using the ghmqc spectrum which had been optimised for a 7 Hz J_{H-C} coupling constant, allowing the $^{3}J_{H-C}$ couplings in the saturated system to be seen. The ghmqc spectrum also showed a weak correlation between the signal at 5.35 ppm in the ^{1}H spectrum and the carbonyl carbon (distinctive at 168.79 ppm in the ^{13}C spectrum), which indicated how the two parts of the molecule were joined. As the ^{1}H-^{1}H COSY had allowed the assignments of the protons of the cyclopentyl group the corresponding carbons could be identified from the ghsqc spectrum. All the peaks assigned correspond well with those in the literature (Wang *et al.,* 1999), although the hydrogen bonded OH was not identified and its existence is most likely pH dependent.

CHAPTER 4

DIFFERENTIAL GENE EXPRESSION

TRANSCRIPTOME ANALYSIS: DIFFERENTIAL DISPLAY

In an attempt to identify the molecular determinant for anthocyanin production in 'Bon Rouge', or reduction in anthocyanin production in the reverted phenotype, differential gene expression was measured by differential display using RNAimage kits (GeneHunter Corp.), and whole transcriptome sequencing by mRNAseq on the Illumina GAII platform. mRNAseq was carried out to survey comprehensive changes in gene expression between the two phenotypic variants of 'Bon Rouge' while the differential display technique allowed investigation of only a selected subset of the expressed genes. Differential display analysis was performed using three of the ten available RNAimage® Kit (GeneHunter Corporation, Nashville, TN, USA). Each kit provided a total of 24 primer combinations for differential display analysis supplied as three one-base-anchored H-T_{11}M primers (where M may be A, C or G) used in combination with eight arbitrary primers (see table 4.1). A total of 72 primer combinations produced on average 50 bands per fingerprint resulting in an approximate 3600 bands for all generated fingerprints from red and green phenotypes. A total of 30 bands that displayed at least a two-fold increase in density as visualised between the red and green sets of fingerprints (see fig. 4.1.2.)

were excised from gels, re-amplified, cloned and sequenced. Re-amplification of cDNA from excised bands was performed under the same PCR conditions and with the same primer combination that was used in the initial amplification for differential display. Re-amplified PCR products were cloned into the pGEM®-T Easy Vector System II (Promega, Madison, WI, USA) and ligations protocols included both positive and background controls. Transformation was effected in competent *E. coli* strain MC1061 cells containing the selectable marker *lac*1^{q}ZM15. Cloned inserts were amplified in the GeneAmp 9700 thermal cycler (Applied Biosystems, Foster City, CA, USA) for automated sequencing in the ABI PRISM 310 Genetic Analyser (Applied Biosystems, Foster City, CA, USA) using M13 forward and reverse primers supplied with the ABI PRISM® BigDye™ Terminator v3.0 Cycle Sequencing Ready Reaction Kit with Amplitaq® DNA Polymerase, FS. Extension products were sequenced immediately or stored at -20°C for not longer than 24 hours before sequencing. Twenty-seven of the thirty bands that were excised from the sequencing gels were successfully re-amplified, cloned and sequenced (see Appendix 1). Sequences were characterised by BLAST similarity search at the NCBI site against the non-redundant nucleotide database. The results of the BLAST are listed in table 4.3.

4.1 RNA isolation and quantification

The quality and integrity of total RNA extracts from red and green leaves were established by electrophoresis on a 6% formaldehyde-agarose gel. The ethidium bromide stained rRNA bands were sharp and showed a two-fold increase in intensity of 28S over

18S rRNA characteristic of high quality RNA (fig. 4.1.1). Total RNA concentration was determined by spectrophotometry and integrity established by a OD_{260}/OD_{280} ratio > 1.8 and a OD_{260}/OD_{230} ratio > 2.0.

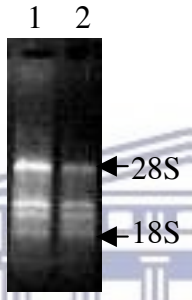
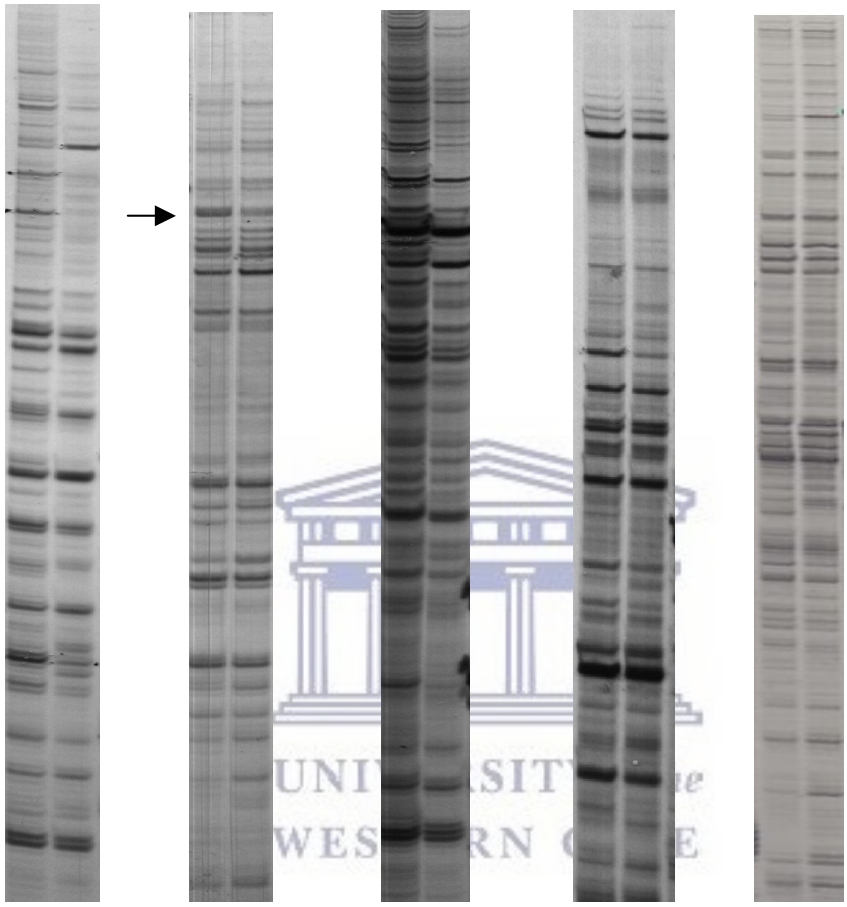


Figure 4.1.1 RNA extracts from red (lane 1) and green (lane 2) leaves resolved by agarose-formaldehyde gel electrophoresis. Banding patterns for both phenotypes show the characteristic 28S and 16S bands indicative of high quality, intact RNA.

4.2 cDNA fingerprints from differential display gels



A **B** **C** **D** **E**

Figure 4.1.2 cDNA fingerprints from differential display gels generated with the following primer combination: A (H-T_{11}G/H-AP1-55), B (H-T_{11}G/H-AP1-26), C (H-T_{11}A/H-AP1-25), D (H-T_{11}A/H-AP1-25) and E (H-T_{11}A/H-AP1-27). The bands aligned with the arrow were excised from gel, re-amplified and sequenced.

Table 4.1 Primer combinations used to produce cDNA amplicons

Designated band number	Primer combination used in cDNA synthesis
M1	H-T_{11}A/H-AP1-3
M2	H-T_{11}A/H-AP1-4
M3	H-T_{11}A/H-AP1-5
M4	H-T_{11}A/H-AP1-5
M5	H-T_{11}A/H-AP1-5
M6	H-T_{11}G/H-AP1-7
M7	H-T_{11}A/H-AP1-25
M8	H-T_{11}A/H-AP1-25
M9	H-T_{11}A/H-AP1-25
M11	H-T_{11}G/H-AP1-26
M12	H-T_{11}G/H-AP1-26
M13	H-T_{11}G/H-AP1-26
M14	H-T_{11}G/H-AP1-27
M15	H-T_{11}A/H-AP1-27
M16	H-T_{11}A/H-AP1-55
M18	H-T_{11}C/H-AP1-51
M20	H-T_{11}C/H-AP1-54
M21	H-T_{11}C/H-AP1-55
M22	H-T_{11}C/H-AP1-55
M23	H-T_{11}C/H-AP1-55
M24	H-T_{11}G/H-AP1-49
M25	H-T_{11}G/H-AP1-51
M27	H-T_{11}G/H-AP1-55
M28	H-T_{11}G/H-AP1-55
M29	H-T_{11}G/H-AP1-55
M30	H-T_{11}G/H-AP1-55
M31	H-T_{11}G/H-AP1-55

4.3 Results for similarity search by BLAST (October 2017)

Cloned cDNAs that were differentially regulated between the transcriptomes of red and green phenotypes according to the differential display analyses, were characterised by similarity search with the BLAST tool (Altschul *et al.,* 1990) at the NCBI (www.ncbi.nlm.nih.gov) against non-redundant nucleotide sequence databases. The similarities of transcripts from the two phenotypes to genes with the closest similarity as measured by the lowest E value, are reported in Table 4.3.

Table 4.3 Summary of BLAST results

Seq ID,	E value,	%ID/Seq Alignment,	Bit score,	GI #,	Similar to
Over-expressed in red leaves:					
M3	1e-24	95/80	124	XM_009362734	PAPS reductase
M7	9e-151	98/310	544	HM044853	*Pp* Defensin
M9	1e-34	98/91	158	AB0930930	*Pc* Expansin
M11	2e-156	94/372	562	XM_009345692	*Pp* ELIP
M14	3e-150	99/305	542	XM_009340901	OMTranslocator
M21	2e-18	98/315	203	DR997344	*Md* PGIP-like
M25	1e-128	99/261	470	XM_009336517	*Pp* VHA c"
M30	1e-49	100/112	207	XM_009357974	QRS N-terminus
M31	1e-49	100/112	207	XM_009357974	QRS N-terminus
Over-expressed in green leaves:					
M15	0.0	99/377	658	XM_009360581	SQD2
M20	2e-68	94/177	270	XM_009378121	mtCarrier protein
M22	6e-93	98/201	351	XM_009380595	Thiazole synthesis
M23	1e-60	95/157	244	XM_009339047	Starch phosphoryl
M27	5e-138	99/283	501	XM_009380595	Thiazole synthesis

4.4 Quantitative RT-PCR of differentially expressed sequences

To confirm differential expression of the partial genes as indicated by differential display, we selected a sub-set of the partial genes that could possibly be linked to anthocaynin sythesis, for quantitative RT-PCR on the Roche Light Cycler (Roche Biochemicals) using the Fast Start SYBR green DNA kit. For quantitative RT-PCR total RNA was extraction with the RNeasy® Plant Mini kit (Qiagen, MD, USA). RNA integrity and quantification was established as described in section 2.3.4. For amplification of the cloned cDNA fragments primers were designed using the Primer 3.0 software (http://www-genome.wi.mit.edu/genome_software/other/primer3.html) with default parameters settings (see Appendix 2). A sub-set of the cDNA sequences was selected for quantitative RT-PCR analysis. Data analysis was performed with Light Cycler Software (Roche Diagnostics, Bern, Switzerland). For cDNA fragments M7 and M11, quantitative RT-PCR confirmed over-expression of these two cDNA fragment in the red compared to the green phenotype (see fig. 4.4 and subsequent description). Similarly for M15 the quantitative RT-PCR results confirmed over-expression of this cDNA fragment in the green compared to the red phenotype (see fig. 4.4 and subsequent description).

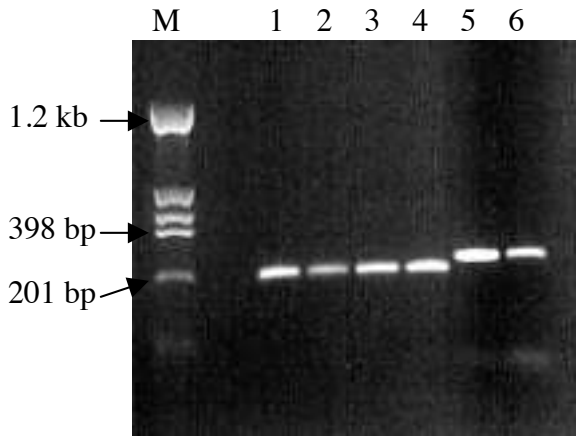


Figure 4.1.3 Gel electrophoresis of quantitative RT-PCR amplicons derived for cDNA sequences M7 (lanes 1, 2), M15 (lanes 3, 4) and M11 (lanes 5, 6) from red (lanes 1, 3, 5) and green (lane 2, 4, 6) leaves, resolved by agarose gel electrophoresis

4.5 Annotation of differentially expressed cDNAs

Partial cDNAs representing expressed genes with homology to Early Light Inducible Proteins (ELIPs), a *Prunus persica* defensin, expansin, an oxoglutarate-malate translocator, glutaminyl tRNA synthetase (QRS), a polygalacturonase inhibitor protein-like peptide and the vacuolar H^+-ATPase pump subunit c" were up-regulated in the red phenotype, whereas those with homology to a chloroplast biosynthetic enzyme, *Citrus sinensis* thi1, a mitochondrial carrier protein, a starch phosphorylase H and sulfolipid synthase (SQD2) were up-regulated in the green form (see table 4.3).

Annotation by BLAST of genes over-expresseded in the green 'Bon Rouge' phenotype

Partial cDNA M15

The partial cDNA sequence completely matched a number of EST clones from the apple nucleotide sequence database (E value = 0, Bit score = 685). The second hit with the same bit score and E value was to sulfoquinovosyl synthase (SQD2) (XM_009360581) or the sulfolipid synthase gene, the equivalent of the bacterial *SqdX*. Sulfoquinovosyl synthase is responsible for catalysis of the second step in the production of sulfoquinovosyl diacylglycerol (Yu *et al.,* 2002), an anionic nonphosphorous glycolipid at physiological pH found in the photosynthetic membranes of seed plants (Mulichak *et al.,* 1999) and photosynthetic bacteria (Güler *et al.,* 1996). Production of SQDG increases in leaves and chloroplast at high temperatures in drought resistant plants whereas drought sensitive plants showed a decrease in SQDG levels (Taran *et al.,* 2000). Plants with mutations in *SQD2* show limited growth under conditions of phosphate starvation (Yu *et al.,* 2002) that suggests a role for SQDG in photosynthetic membranes when glycolipid synthesis is impaired due to phosphate limitation. The anionic nature of the head group, sulfoquinovose, maintains the negatively charged lipid-water interface that is essential for proper functioning of photosynthetic membranes (Yu *et al.,* 2002). Plants that carry mutations in SQD2 show limited growth under conditions of phosphate starvation (Yu *et al.,* 2002) that suggest a compensatory role for sulfoquinovosyl diacylglycerol (SQDG) as components of photosynthetic membranes when glycolipid synthesis is impaired due to phosphate limitation during growth. The increased production of sulfoquinovosyl

diacylglycerol in leaves and chloroplast of plants in indicative of an adaptive response to high temperatures in drought resistant plants whereas drought sensitive plants showed a decrease in SQDG levels (Taran *et al*., 2000). Phosphate starvation induces an increase in the production of anthocyanin via a cytokinin-mediated action, and mutations in CRE1, the cytokinin receptor suppress the increased production of anthocyanin under phosphate-limited growth (Franco-Zorilla *et al*., 2002). The production of anthocynin in vegetative tissue of plants is characteristic of the phosphate starvation response (Winkel-Shirley, 2002) but the molecular mechanism by which this anthocyanin increase is mediated is still not clear. According to Nikiforova et al. (2003), a decrease in sulpholipid causes a decrease in photosynthetic capacity under already limiting phospholipid conditions that results from phosphate limitation. This in turn results in oxidative stress as a consequence of high light stress. Glutathione, a crucial metabolite in the scavenging of reactive oxygen species, is reduced due to limited sulphur availability and consequently anthocyanin increases to protect against high light stress (Nikiforova *et al*., 2003). If this scheme holds true for the plants in this study, then decreased sulpholipid synthesis associated with a decreased expression of SQD2 in the red 'Bon Rouge' phenotype would be consistent with increased production of anthocyanin and increased sulpholipid synthesis would be consistent with decreased anthocyanin synthesis in the green 'Bon Rouge' phenotype.

Partial cDNA M20

Partial cDNA M20 displays significant homology (E = 2e-68, Bit score = 270) to a mitochondrial carrier protein (mtCP). All mtCPs characterised to date have a high pI and

a molecular weight of approximately 30 kD (Laloi, 1999). mtCPs represent a superfamily of related transporters in eukaryotes that are characterised by a stretch of 300 amino acids arranged in three domains. Each domain consists of about 100 amino acids that are arranged in two membrane spanning α-helices separated by a hydrophilic extra-membrane loop. These proteins function as homodimers with 12 trans-membrane domains (Saraste and Walker, 1982) in the specific transport of a vast array of metabolites and solutes between the mitochondrion and cytosol. These include adenine nucleotide transporter (ANT), uncoupling protein (UP), oxoglutarate-malate translocator (OMT), the phosphate transporter (PiC), the dicarboxylate and tricarboxylate transporters, and amino acid and cofactor transporters (Palmieri *et al.,* 1996). All mitochondrial carrier proteins contain up to three copies of the 10 amino acid consensus sequence referred to as the mitochondrial energy transfer signature (METS; Prosite PDOC00189), P-x-[DE]-x-[LIVAT]-[RK]-x-[LRH]-[LIVMFY]-[QGAIVM] (Laloi, 1999). The partial cDNA identified in this gene screen has limited sequence information to allow assignment for function and we aim to obtain the full-length sequence or a significant part of the complete gene in an attempt to identify its specific function in mitochondrial metabolite transport. A partial cDNA (M14) upregulated in the red 'Bon Rouge' phenotype displays significant similarity to the oxoglutarate-malate translocator (OMT) and it would be interesting to determine the function of this transporter upregulated in the red 'Bon Rouge' phenotype.

Partial cDNAs M22 and M27

Two partial cDNAs upregulated in the green phenotype showed similarity to the chloroplast thiazole biosynthesis precursor protein gene, *Citrus sinensis thi1* encoding the enzyme involved in the biosynthesis of the thiazole ring of thiamine pyrophosphate (Vitamin B_1) in plants. Partial cDNA M27 (similar to XM_009380595, E value = 5e-138, Bit score = 501) displayed more significant homology than partial cDNA M22 (similar to XM_009380595, E value = 6e-93, Bit score = 351) to *Citrus sinensis thi1*. Thiamine pyrophosphate is an essential coenzyme for all organisms that depend on fermentation, respiration or photosynthesis for ATP synthesis and is synthesised via two independent routes (Godoi *et al*., 2006). One route is responsible synthesis of 2-methyl-4-amino-5-hydroxymethylpyrimidine pyrophosphate (pyrimidine moiety) and the other for 4-methyl-5-(β-hydroxyethyl) thiazole phosphate (thiazole moiety). An *Arabidopsis thi1* showed high homology to the gene product of THI4 involved in thiazole biosynthesis in *Saccharomyces cerevisiae* that was able to complement the UV-sensitive phenotype of *E. Coli* (AB1886), uvrA (Machado *et al*., 1996). THI1 from *Arabidopsis* is encoded by a single gene and is directed simultaneously to chloroplasts and mitochondria from a single nuclear transcript that produces two products by differential usage of two in-frame translational start codons (Chabregas *et al*., 2003). In maize, *thi1* expression is differentially regulated and in citrus, *thi* genes are developmentally regulated in response to ethylene (Jacob-Wilk *et al*., 1997). Research in eukaryotic thiamine biosynthesis is at an early stage and although the three-dimensional structure of *Arabidopsis thi1* has been solved at a 1.6A resolution, details concerning its function are not known (Godoi *et al*., 2006). Two key proteins for yeast thiazole biosynthesis, THI1 and THI4, have been

identified but their structure and function cannot be predicted using bioinformatic analyses (Settembre *et al.,* 2003).

Partial cDNA M23

The partial cDNA fragment displays significant homology (E = 1e-60, Bit score = 244) to the cytosolic H isoform of starch phosphorylase (EC 2.4.1.1). Starch phosphorylase, also referred to as α-glucan phosphorylase, is a key enzyme in glucan catabolism in animals, prokarytes and fungi (Newgard *et al.,* 1989). It is considered to be a critical enzyme for starch degradation in plants where it catalyses the reversible phosphorolysis of linear α-1,4-glycosidic linkages in glucan substrates to yield glucose-1-phosphate (Smith *et al.,* 2005). The enzyme is reversible under physiological conditions and was previously associated with starch synthesis, a role now assigned to starch synthases (Smith *et al.,* 2005). The enzyme exists in multiforms in animals, microorganisms and plant, and higher plant starch phosphorylase isozymes are classified into types L or H, according to their low and high affinities for glucans, respectively (Mori *et al.,* 1993). The L and H isozyme sequences are highly conserved except for a unique 78 amino acid insertion in the middle of the L isozyme that specifies striking differences in substrate affinities (Mori *et al.,* 1991). Type L isozymes have high affinity for small linear amylose, maltodextrin and highly branched amylopectins whereas type H isozymes prefer various glucans, including glycogen, as substrates (Gerbrandy, 1974). All higher plants studied thus far have plastidial and cytosolic isoforms of starch phosphorylase, which are encoded by separate genes (Mori *et al.,* 1991), and since starch is synthesised exclusively in the plastids, only plastidial isoforms are implicated in its metabolism (Zeeman *et al.,* 2004).

Plastidial starch phosphorylase isoforms are not critical for starch degradation in *Arabidopsis* leaves but has a role in abiotic stress tolerance (Zeeman *et al.,* 2004), while cytosolic starch phosphorylase H isoforms have not yet been assigned a critical role in starch metabolism (Kossmann and Lloyd, 2000). The *pal1* and *pal2* mutants of *Arabidopsis* are T-DNA insertion mutants defective in phenylalanine ammonia-lyse (PAL; EC 4.3.1.25) expression that display differential regulation of starch phosphorylase H (Rohde *et al.,* 2004). PAL catalyses the first reaction in the phenylpropanoid pathway that produces numerous flavonoids, including anthocyanins. The *pal1* and *pal2* mutants display no visible phenotypic alteration but have far-reaching effects on phenylpropanoid, amino acid and carbohydrate metabolism (Rohde *et al.,* 2004).

Annotation by BLAST of genes up-regulated in the red 'Bon Rouge' phenotype

Partial cDNA M3

Partial cDNA fragment M3 showed similarity (E = 1e-24, Bit score = 124) to a sequence encoding a PAPS reductase-like enzyme (XM_009362734). PAPS reductases function in the sulphate assimilation pathway in plants and microorganisms (Asahi, 1960) that produce reduced sulphur from inorganic sulphate for the synthesis of sulphur-containing amino acids, methionine and cysteine (reviewed in Wilson, 1962), coenzymes and iron-sulphur clusters of enzymes (Schmidt and Jäger, 1992), and a variety of S-containing secondary metabolites such as glucosinolates that play an important role in plant defences against herbivores and pathogens (Kopriva, 2006; Kopriva *et al.,* 2007). Plant sulphur

assimilation is the major source of reduced sulphur for animal and human diets and the pathway is stimulated by light (Asahi, 1960) while the biochemical and genetic regulation is affected by oxidative stress, heavy metal (cadmium) exposure and sulphur deficiency (Mendoza-Cózatl *et al.,* 2005). The first step of assimilatory sulphate reduction is sulphate activation catalysed by ATP sulphurylase (EC 2.7.7.4) (Suter *et al.,* 2000). Adenosine 5'-phosphosulphate (APS) acts as the substrate for APS kinase (EC 2.7.1.2.5), which forms adenosine 3'-phosphate 5'-phosphosulphate (PAPS) in the subsequent activation step (Schmidt and Jäger, 1992). PAPS is reduced by PAPS reductase (EC 1.8.99) when the enzyme reacts initially with reduced thioredoxin then with PAPS to form SO_3^{2-}, oxidised thioredoxin and adenosine 3'-phosphate 5'-phosphate (PAP). The SO_3^{2-} is reduced to sulphide by sulphite reductase (EC 1.8.7.1), which is subsequently incorporated into O-acetyl-L-serine via O-acetyl-L-serine thiollyase (EC 4.2.99.8) to form cysteine (Schmidt and Jäger, 1992). For higher plants sulphate reduction, there is debate whether APS sulphotransferase catalyses reductive transfer of sulphate from APS to reduced glutathione, or whether PAPS reductase carries out thioredoxin-dependent reduction of PAPS (Setya *et al.,* 1996). In the moss, *Physcomitrella patens*, orthologues of both APS reductase and PAPS reductase genes exist but the APS reductase does not contain the FeS cluster associated with APS reductases from higher plants (Kopriva *et al.,* 2007). The PAPS reductase-like partial cDNA identified in this differential gene screen needs to be further characterised using the full-length cDNA for this gene in an attempt to elucidate its function in sulphur assimilation in red 'Bon Rouge' phenotypes. The molecular mechanism for the regulation of sulphur uptake and assimilation in higher plants has been reported (Maruyama-

Nakashita *et al*., 2005) but the biological questions regarding sulphur metabolism in higher plants cannot be answered by research on *Arabidopis* alone (Kopriva, 2006). Research on sulphur metabolism in other species will be required to answer some of these questions.

Partial cDNA M7 (sequence similarity to Pyrus defensin protein 1)

Expression of this mRNA-representing cDNA is upregulated in the red phenotype and is significantly homologous (E = 9e-151, Bit score = 544) to a *Pyrus pyrifolia* mRNA, complete cds, accession number HM044853, that has been characterised as a defensin (DFN1). Defensins are low molecular weight antimicrobial, cysteine-rich, proteins found in plants and animals that inhibit the growth of fungi, oomycetes and gram-positive bacteria (Broekaert *et al*., 1995). They constitute an important part of innate immunity in plants (Thomma *et al*., 2002) with some defensins expressed specifically in response to infection from pathogens like fungi while other are constitutively expressed. In *Arabidopsis,* antifungal proteins, PDF1.1 and PDF1.2 expression require both jasmonic acid and ethylene (Penninckx *et al*., 1996). Defensins are small basic proteins of between 45 and 54 amino acids are characterised by two to six disulphide bridges required for structural stabilisation of these molecules (Thomma *et al*., 2002). Plant defensins contain eight cysteines in four disulphide bridges that are critical for the stabilisation of their particular globular structure (Meyer *et al*., 1996), and they are structurally and functionally related to insect and mammalian defensins (Thomma *et al*., 2002). All plant defensin pre-proteins identified to date have a signal peptide that targets the protein for extracellular secretion, a mature functional domain of between 45 and 54 amino acids and

in the case of petunia PhD1 and PhD2 (Lay *et al.,* 2003) and tobacco NaD1 (Nielson *et al.,* 1996), an acidic and hydrophobic amino acid rich C-terminal prodomain for vacuolar targeting (Lay *et al.,* 2003). PhD1 and PhD2 also contain an additional disulphide bridge although this does not change the activity of the protein but appears to confer additional rigidity and thermostability to the proteins (Lay *et al.,* 2003). The deduced amino acid sequence for the partial M7 cDNA contains three of four conserved cysteines found in the C-terminal half of the conceptual translation for the functional domain for *Pyrus pyrifolia* defensin gene, including the acidic, hydrophobic amino acid rich C-terminal prodomain that targets the protein to the vacuole. Expression of defensin is associated with biotic stress responses and increased expression of M7 in the red phenotype may suggest an enhanced response to pathogen attack. Increased expression of anthocyanin during biotic stress responses is well documented but the molecular mechanisms for many of these responses remains to be detailed (Dixon and Paiva, 1995 and references therein).

Partial cDNA M9

The partial cDNAs represented by M9 displayed weak homology to *Pyrus communis* PcExp3 (ABO93030) and *Pyrus communis* PcExp2 (AJ11692) mRNA for expansin, as determined by BLAST in the nucleotide database (nr) at NCBI (E = 1e-34, Bit score 158). Expansins are extracellular proteins that function in cell wall loosening during cellular growth but exhibit no detectable hydrolase or transglycosylase activity (McQueen-Mason and Cosgrove, 1995). Furthermore, they are regulated by extracellular pH and are developmentally regulated in a tissue specific manner (Cosgrove, 2000). Two gene subfamilies code for α-expansins and β-expansins that display limited amino acid

sequence conservation characteristics. The HFD motif and an FRRV signature (or a closely related sequence) are key recognition elements for the canonical α- and β-expansin family (Cosgrove, 2000). These molecules have been implicated in 'acid growth', the enlargement of plant cells in response to low extracellular pH (Sampedro and Cosgrove, 2005), increased cell wall extensibility during dehydration and rehydration in *Craterostigma plantagineum* (Jones and McQueen-Mason, 2004), soybean cyst nematode infection (Tucker *et al.,* 2007), pollen tube invasion of stigma in grasses, wall disassembly during fruit ripening, petiole abscission and numerous other cell separation processes (Cosgrove, 2000). In addition to the two reported α- and β-expansin subfamilies, two additional types, δ- and γ-expansin that are truncated forms of the canonical expansins, have been isolated (Li *et al.,* 2003). Expansin-like molecules such as plant natriuretic peptide share evolutionary and functional relationships with expansins and other cell wall loosening molecules but do not display cell wall loosening activity (Gehring and Irving, 2003). Although we can infer much from gene expression patterns and sequence information for species whose genomes are well represented in databases, functional tests like *in vitro* assays and gene knockout experiments are necessary to test these inferences and assign specific roles for proteins like expansins (Cosgrove, 2000).

In *Pyrus*, a number of expansins were shown to be differentially expressed during growth development and ripening of fruit (Hiwasa *et al.,* 2003). Treatment with propylene, an inducer of ethylene production, induced accumulation of *PcExp2*, *-3*, *-5* and *-6* mRNA and fruit softening after late stage fruit development whereas *PcExp1* mRNA increased at the late expanding stage of fruit development and during ripening. *PcExp4* was constitutively expressed while *PcExp7* was expressed only in young fruit. *PcExp4* and

PcExp6 mRNA were detected in flowers and *PcExp4*, *PcExp5*, *PcExp6* and *PcExp7* mRNA accumulated more abundantly in young growing tissue than in fully expanded mature tissue. The results demonstrate the differential expression of multiple expansin genes in response to hormonal regulation during pear fruit ripening and tissue development (Hiwasa *et al*., 2003). Increased expression of partial cDNA homologues of *PcExp2* and *PcExp3* in the red 'Bon Rouge' phenotype most likely suggests a function in cell expansion of leaf tissue.

Partial cDNA M11

Partial cDNA M11 display similarity (E = 2e-156, Bit score = 562) to a *Pyrus pyrifolia* Early Light Inducible Proteins (ELIP1) accession number XM_009345692. ELIPs (ELIP1, ELIP2, ELIP3 and ELIP4) are nuclear encoded, stress-regulated chloroplast proteins that are related to the chlorophyll *a*/*b*-binding (Cab) proteins (Heddad *et al*., 2006), stress enhanced proteins (SEPs) and the one-helix proteins (OHPs). They are members of a group of low and high molecular weight proteins characterised by three membrane spanning α-helices with helix I and III showing high homology to the corresponding helices from Cab proteins (Grimm *et al*., 1989) that contain four putative chlorophyll-binding residues. ELIPs are synthesised in the cytoplasm and transported to the chloroplast where they are inserted in the thylakoid via a pathway that involves chloroplastic signal recognition particle 43 (cpSRP43) (Hutin *et al*., 2003). The reported function for ELIPs is the binding of chlorophyll pigments for protection under conditions of high light stress (Heddad *et al*., 2006) with release of intact pigment upon removal of the light stress with a more rapid recovery from high light stress-induced photoinhibitory

conditions (Steyn *et al.,* 2004). The photoprotective role of ELIPs includes the binding of free chlorophyll to prevent the formation of reactive oxygen species (Hutin *et al.,* 2003) or as xanthophyll chelators for dissipation of excess absorbed light energy (Krol *et al.,* 1999). This function has been challenged by the demonstration that suppression of ELIP1 and ELIP2 in *Arabidopsis* does not affect tolerance to photoinhibition or photoxidative stress although no alternative function for ELIPs have been put forward (Rossini *et al.,* 2006). In mature plants, photoinhibitory conditions that produce induction of ELIPs include high light stress (Pötter and Kloppstech 1993), high light and cold (Montané *et al.,* 2000) salinity (Sävenstrand *et al.,* 2004), UV-B (Adamska *et al.,* 1992; Valledor *et al.,* 2012) and dessication stress (Bartels *et al.,* 1992).

The transcription factor for ELIP1 is the constitutively expressed LONG HYPOCOTYL5 (HY5) (Harari-Steinberg *et al.,* 2001). HY5 levels are regulated during photomorphogenesis, or development in light, and skotomorphogenesis (development in the dark), by interaction with the CONSTITUTIVELY PHOTOMORPHOGENIC1 (COP1) (Hardtke *et al.,* 2000). COP1 is an E3-ubiquitin ligase that targets HY5 for degradation during the dark diurnal phase. Skotomorphogenesis requires the physical interaction of COP1, a negative regulator of photomorphogenesis, with HY5, a constitutively expressed nuclear localised positive regulator of photomorphogenesis, which ultimately results in the degradation of HY5 by the 26S proteasome in the nucleus. To exert its negative effect on HY5, COP1 has to locate to the nucleus in the dark. Alternatively, photomorphogenesis requires the inactivation of COP1 with concomitant decreased degradation of HY5. The resultant increase in HY5 mediates the typical response of wild type plants to light including reduced hypocotyl growth. HY5 has also

been shown to be the transcriptional activator for chalcone synthase (CHS), the first committed enzyme in the anthocyanin biosynthetic pathway, and a number of downstream enzymes required for anthocyanin synthesis (Vandenbussche *et al.,* 2007). The degradation of HY5 by the 26S proteasome is regulated by phosphorylation/dephosphorylation (Hardtke *et al.,* 2000), by a yet unknown enzyme system. The dephosphorylated form of HY5 is preferentially degraded by the 26S proteasome with the phosphorylated form showing a greater stability toward proteasomal degradation (Hardtke *et al.,* 2000). This allows for further modulation and /or control of the level of available transcription factor, HY5 beyond transcriptional and translational control. For chalcone synthase under the control of the transcription factor, HY5, and the first committed step in anthocyanin production, this could result in enhanced gene expression and protein levels with subsequent increase in anthocyanin production. Expression of ELIP under the control of constitutively expressed HY5, which is also the transcription factor for CHS, may coincidently result in the production of anthocyanin during high light stress. Sequestration in an acidic vacuole would then be required to stabilise anthocyanin pigments and the upregulation of the VHAc" subunit in the red phenotype may play a role in pigment stabilisation.

Partial cDNA M14

The partial cDNA sequence for M14 shared significant similarity (E = 3e-150, Bit score = 542) with a 2-oxoglutarate/malate translocator (OMT) (GI number XM_009340901), a member of the plastidic dicarboxylic acid transporter (DCT) family. The plastidic DCTs of higher plants consist of two distinct groups that share overlapping substrate

specificities (Kore-eda *et al.,* 2005). The first group represented by OMTs transports dicarboxylates such as malate, succinate, fumarate, glutarate and 2-oxoglutarate while the second group represented by general dicarboxylic acid transporters (GDTs) and also referred to as glutamate/malate transporters use glutamate and aspartate in addition to those transported by OMTs, as substrates (Hinz *et al.,* 1988). The 2-oxoglutarate/malate translocator in tandem with the glutamate/malate transporter participate in the two translocator model (Woo *et al.,* 1987) where cytosolic malate is imported in exchange for stromal glutamate by the glutamate/malate transporter while uptake of 2-oxoglutarate facilitates the direct re-export of malate via the OMT. Exported malate is subsequently used for glutamine synthesis (Woo *et al.,* 1987). The stromal 2-oxoglutarate is used in combination with glutamine for the synthesis of glutamate by coupling of the stromal enzymes glutamine synthase and glutamine/2-oxoglutarate aminotransferase (GOGAT). Elucidation of the coupling of these two transporters has provided a better understanding of ammonia assimilation during nitrite reduction and photorespiration (Woo *et al.,* 1987). Characterisation of the 2-oxoglutarate/malate translocator from spinach (Menzlaff and Flügge, 1993) identified the first plastidic transporter with 12 predicted transmembrane domains that does not conform to the 2 x 6 α-helices forming dimers typical of plastid transporters.

In *Panicum miliaceum*, a NAD-malic enzyme-type C4 plant, the expression of the mitochondrial 2-oxoglutarate/malate translocator is regulated by light and development (Taniguchi and Sugiyama, 1997). The steady-state level of mRNA was higher in leaves than in nonphotosynthetic tissues and its expression was restricted to bundle sheath cells (BSC) but not mesophyll cells. Specific 2-oxoglutarate/malate translocator of *P.*

miliaceum in BSC mitochondria is expressed in concert with C4 enzymes during the differentiation of BSC and parallels the capacity of C4 photosynthesis (Radchuk *et al.,* 2006). Tobacco DiT1, a 2-oxoglutarate/malate translocator, antisense-repressed in intact transgenic plants, caused reduced transport capacity for 2-oxoglutarate across plastid envelope membranes. The reduction in 2-oxoglutarate transport resulted in impaired allocation of carbon precursors for amino acid synthesis, accumulation of organic acids, and a significant decrease in protein content, photosynthetic capacity and sugar pools in leaves. The phenotype was consistent with a role of DiT1 in both, primary ammonia assimilation and the re-assimilation of ammonia resulting from the photorespiratory carbon cycle (Schneidereit *et al.,* 2006). Decreased photosynthetic capacity due to photoinhibition as a consequence of high light stress is associated with anthocyanin production in leaves and fruit skin of blush pear cultivars such as Rosemarie and Forelle (Steyn *et al.,* 2004).

Partial cDNA M21

Partial M16 cDNA matched a *Malus* EST, DR997344 (E = 2e-18, Bit score = 203), that displays similarity to the mRNA for (PGIP)-like protein from *Fragaria x ananassa* (AAP33475). Polygalacturonase-inhibiting proteins (PGIPs) inhibit endopolygalacturonases (PGs) (EC 3.2.1.15) produced by fungal pathogens during plant infection, which degrade the host cell wall by cleavage of α-(1 → 4) linkages between the D-galacturonic acid residues in polygalacturonan. The cleavage results in cell separation and softening of host tissue (D'Ovidio *et al.,* 2006) and this critical process in pathogenesis has been described for the fungal pathogens, *Botrytis cinerea* (ten Have *et*

al., 1998), *Alternaria citri* (Isshiki *et al.,* 2001) and *Claviceps purpurea* (Oeser *et al.,* 2002) among others. PGIPs are members of the large superfamily of leucine-rich repeat (LRR) defense proteins (Toubart *et al.,* 1992), are present in the cell walls of all plants investigated to date and inhibit PGs of fungal pathogens but not those of plants and bacteria (Ferrari *et al.,* 2003). The LRRs are of the extracytoplasmic type and characterised by the consensus sequence GxIPxxLGxLxxLxxLxLxxNxLx (Kajava, 1998). In *Malus x domestica* cv. Florina, the consensus sequence for a leucine-rich-repeat receptor-like protein kinase m1 (LRPKm1) that is somewhat similar to PGIPs is given as LxxLxxLxxLdLxNxLSGxIPxx (Komjanc *et al.,* 1999). Interaction between PGIPs and endopolygalacturonases (PGs) results in the production of polygalacturonides (OGs) that produces a wide range of defense responses (Cervone *et al.,* 1989). These responses result in a reduction in symptoms in transgenic tobacco and tomato plants when challenged with *Botrytis cinerea* (Ferrari *et al.,* 2003). Transgenic tobacco transformed with grapevine *PGIP* (*VvPGIP1*), showed reduced susceptibility to *Botrytis cinerea* and differentially inhibited fungal PGs (Joubert *et al.,* 2006).

PGIPs are regulated during development, and biotic and abiotic stress responses such as wounding and pathogen infection or treatment with elicitors like salicylate and cold produce PGIP synthesis (Ferrari *et al.,* 2003; Li *et al.,* 2003). In *Vitis* ('Riesling' and 'Gloire de Montpellier') PGIP production and genes encoding anthocyanin biosynthetic enzymes, including phenylalanine ammonia-lyase, chalcone synthase and chalcone isomerase, are differentially regulated in response to fungal pathogens *Plasmopara viticola* and *Pseuperonospora cubensis* (Kortekamp, 2006).

Partial cDNA M25

The partial cDNA sequence for M25 shared significantly high similarity (E = 1e-128, Bit score = 470) with the vacuolar membrane H^+-ATPase subunit c" from pear (*Pyrus pyrifolia*) (GI number XM_009336517). The vacuolar-type H^+-transporting adenosine triphosphatase (V-ATPase; EC 3.6.1.34) is a membrane-bound, primary active transport protein active in the vacuolar tonoplast and various other sites in the plant endomembrane system (Sze, 1985). In plant vacuoles, V-ATPase are responsible for energising ion and metabolite transport (Ratajczak, 2000), acidification of the intracellular vacuolar compartment (Forgac, 1998) and is essential for the maintenance of ion homeostasis, secondary activated transport and adaptation to environmental stress (Seidel *et al.,* 2004). Physiological and biochemical studies have demonstrated that acidification of the vacuole by this proton pump drives the uptake and release of ions and metabolites across the tonoplast membrane (Sze *et al.,* 2002). Tonoplast V-ATPAse is highly abundant comprising approximately 6.5 – 35% of the total tonoplast protein of various plant species (Fischer-Schliebs *et al.,* 1997). V-ATPases are conserved in all eukaryotes (Lüttge and Ratajczak, 1997) and consists of at least 12 distinct subunits arranged in two large subcomplexes: the cytosolicV_0 and the membrane bound V_1 (Domgall *et al.,* 2002). The cytosolic V_1 complex comprises subunits A to H and catalyses the hydrolysis of ATP that is coupled to proton pumping into a compartment via the membrane-bound V_0 complex. The V_0 complex comprises three integral membrane proteins namely subunits a, c, c", and one hydrophilic subunit d (Arata *et al.,* 2002). Most V_1 subunits are encoded by single genes, whereas V_0 subunits are encoded by multiple genes found in duplicated segments of the genome of *Arabidopsis* (Sze *et al.,* 2002). *Arabidopsis* (Sze *et al.,* 2002)

and *Mesembryanthemun crystallinum* (Kluge *et al.,* 2003) contain an additional V_0 complex subunit e that has yet to be assigned a location in the multimeric complex.

In *Arabidopsis* 13 subunits are encoded by 27 genes, indicating a significant level of complexity for this proton pump (Sze *et al.,* 2002). The vacuolar H^+-ATPase c'' subunit, designated VHA c'' is a homolog of the yeast *VMA16* proteolipid 23kD subunit (Sze *et al.,* 2002).

In mature plant cells, the vacuole is the largest intracellular compartment occupying approximately 90% of the cell volume and functions as a store for salts, metabolites, plant pigments such as anthocyanins, sugars, organic acids, including malate, and numerous other solutes (Sze, 1992) under various environmental conditions. Malate is a central metabolite in plants and night-time vacuolar malate accumulation is energised by the primary vacuolar H^+ pumps, V-ATPase and vacuolar H^+pyrophosphatase (V-PP_iase) (Hafke *et al.,* 2003). Both pump H^+ into the vacuole to establish an H^+ electrochemical gradient that drives malate thermodynamically downhill from the cytosol into the vacuole across the tonoplast (Hafke *et al.,* 2003). Day-time photo-inhibited C4 plants that display crassulacean acid metabolism (CAM) synthesise large amounts of malate in their chloroplast-containing cell vacuoles at night as a result of dark-fixation of CO_2 via phosphoenolpyruvate carboxylase (Kluge *et al.,* 2003). During the following light period, malate is transported out of the vacuole, decarboxylated in the cytosol with the resultant CO_2 used for the synthesis of organic sugars in the Calvin cycle (Lüttge *et al.,* 2002). This adaptation to high light stress is a demonstration of the crucial role for vacuolar acidification and proton translocation as a function of vacuolar proton pumps (Kluge *et al.,* 2003).

Under conditions of heat, osmotic, cold and salinity stress, distinct coordinated changes in transcript levels were observed in roots and leaves for all VHA subunits in the halophyte *Mesembryanthemun crystallinum* (Kluge *et al.,* 2003). Subunits A, c and E were shown to be upregulated in salinity stress, and VHA c expression was developmentally regulated in *Arabidopsis* (Perera *et al.,* 1995) and cotton (Hasenfratz *et al.,* 1995). Vacuolar ATPase c from *Pennisetum glaucum* (the *Arabidopsis* VHA c homolog) is regulated by abiotic stress (Tyagi *et al.,* 2005) and different isoforms are differentially regulated not only developmentally but also under different environmental conditions (Tyagi *et al.,* 2006). Isoform III of *P. glaucum* vacuolar ATPase c is constitutively expressed in roots and shoots and does not respond to stress whereas isoform I is upregulated under stress. Isoform II is expressed primarily in roots but under salinity stress, its expression is induced in shoots and downregulated in roots (Tyagi *et al.,* 2006). There is also increasing evidence that modulation of the holoenzyme structure might influence V-ATPase activity (Ratajczak, 2000) and antibodies against VHA c" from *Arabidopsis* and *C. limona* inhibit proton pumping activity (Aviezer-Hagai *et al.,* 2003). The monocotyledonous dessication-tolerant plant, *Xerophyta viscosa* Baker (family Velloziacaea) belonging to a small group of angiosperms collectively known as resurrection plants, can tolerate exposure to high light and extreme temperatures (Gaff *et al.,* 1971). *Xerophyta viscosa* subjected to dehydration, extreme temperatures (-20°C) and salinity (NaCl shock), display increased steady state mRNA levels of *XvVHA c"1* and a stress responsive-expression pattern (Mundree and Farrant, 2000). According to the authors, *XvVHA c"1* is the subunit whose expression has been demonstrated to be critical for plant stress responses. Southern hybridisation analysis has identified a small gene

family of *XvVHA c''1* homologues (Mundree and Farrant, 2000) that might aid 'molecular slip' control or regulation of proton translocation for this pump.

The partial cDNA identified in this screen for differentially expressed genes, also displays significant homology to the *C. limona VHA c''*. This subunit is critical for *C. limona* vacuolar acidification resulting in a vacuolar pH = 2.0 (Aviezer-Hagai *et al.,* 2003) and increased expression of this subunit may suggest a role in vacuolar acidification of pear tissue vacuoles. In morning glory (*Ipomoea tricolor*) petals increased vacuolar pH has been associated with petunidin (an anthocyanidin) colour change from red to blue (Yoshida *et al.,* 2005) and in *Petunia hybrida* (Quattrocchio *et al.,* 2006) an acidic vacuole has been implicated in red pigment colour stabilisation and intensity for cyanidins (Spelt *et al.,* 2002).

Partial cDNAs M30 and M31

Two partial cDNAs displayed similarity (M30, E = 1e-49, Bit score = 207 and M31, E = 1e-49, Max score = 207) to the N-terminal sequence of *Fragaria vesca* glutaminyl tRNA synthetase (QRS). Glutaminyl tRNA synthetase catalyses the covalent linkage of glutamine to its cognate tRNA during protein translation. However, homology searches have indicated similarity for both partial cDNAs with the N-terminal extension of QRS suggesting misspriming of the olig-dT primer used in differential display analysis to the nuclear localisation signal found in the N-terminal part of the gene. The N-terminal extension of QRS has no critical role in the catalytic reaction.

CHAPTER 5

DIFFERENTIAL GENE EXPRESSION

TRANSCRIPTOME ANALYSIS: mRNA SEQUENCING

Introduction

In an attempt to identify the molecular determinant for anthocyanin production in 'Bon Rouge', or reduction in anthocyanin production in the reverted phenotype, differential gene expression was measured by differential display using RNAimage kits (GeneHunter Corp.), and whole transcriptome sequencing by mRNAseq on the Illumina GAII platform. mRNAseq was carried out to survey comprehensive changes in gene expression between the two phenotypic variants of 'Bon Rouge' since the differential display technique allowed investigation of only a subset of the expressed genes due to the fact that only one-third of the possible 240 primer combinations was used in our analysis. RNA was extracted from both red and green leaves of 'Bon Rouge' using an RNeasy plant mini kit (Qiagen) according to the manufacturer's instructions. Following total RNA extraction, samples were treated with DNase and purified by the RNA clean-up procedure using the RNeasy kit. RNA quantity, quality and integrity were determined by spectrophotometry for differential display. For mRNAseq, RNA quality and integrity, as well as quantity were determined by RNA chip based electrophoresis

on the BIO-RAD platform. Polyadenylated RNA was selected using oligo(dT) purification and reverse-transcribed to cDNA. The cDNA was fragmented, blunt-ended, and ligated to the Illumina TruSeq Adaptors (Illumina Inc., http://www. illumina.com). These libraries were size-selected for an insert size of 200 bp, and quantified using the SyberGreen DNA Master kit (Applied Biosystems) on the LightCycler system (Roche). cDNA libraries were sequenced on the Illumina GAII platform. Paired-end (PE) sequencing runs of 50 cycles and single-end (SE) sequencing runs of 72 cycles was performed on the Illumina GAII for the transcriptome of red and green phenotypes of 'Bon Rouge', respectively. Read pairs had a mean insert size of 100 bp. The raw Paired-end 50 bp sequencing data was transformed into Single-end 50 bp reads, using GERALD base-calling (a CASAVA package tool provided by Illumina). The resulting sequence reads were stored in FASTQ format. Read trimming was performed with Velvet (Zerbino *et al.,* 2008) to exclude reads with quality below Q30. Three lanes were sequenced for each phenotypes resulting in a total of 17 million reads. The filtered reads were de novo assembled into contigs using the CLC Genomics Workbench de novo assembly tool, and Velvet. For the latter, a series of independent assemblies were performed to analyse the effects of varying de novo assembly parameters.

The maximum length of transcripts are 5188 bases (CLC Bio) and 5088 bases (with Velvet) respectively, and average length 646 bases for CLC Bio. The number of reads used in the assembly were relatively low (7847794 for Velvet and 4414967 for CLC Bio) compared to what was commonly recorded in the literature at the time of the analysis, but could be due the low initial number of total reads produced from mRNAseq. However, many studies reported the incorporation of only 50 % of the total

reads into contig assembly for plant mRNAseq data at that time.

Contigs over-expressed in either phenotype were annotated by similarity search with the BLAST tool at NCBI against a number of nucleotide databases. In addition we used Blast2GO to annotate the top 100 contigs over-expressed in the red phenotype.

5.1. RNA isolation and quantification

RNA extraction was performed with the RNeasy® Plant Mini kit (Qiagen, MD, USA) as prescribed by the manufacturer. RNA quality was determined on the BIO-RAD Experion Platform using the protocol for RNA analysis with BIO-RAD RNA chips essentially as described by the manufacturer. RNA samples with a minimum Q value of 8.0 were used in cDNA library construction.

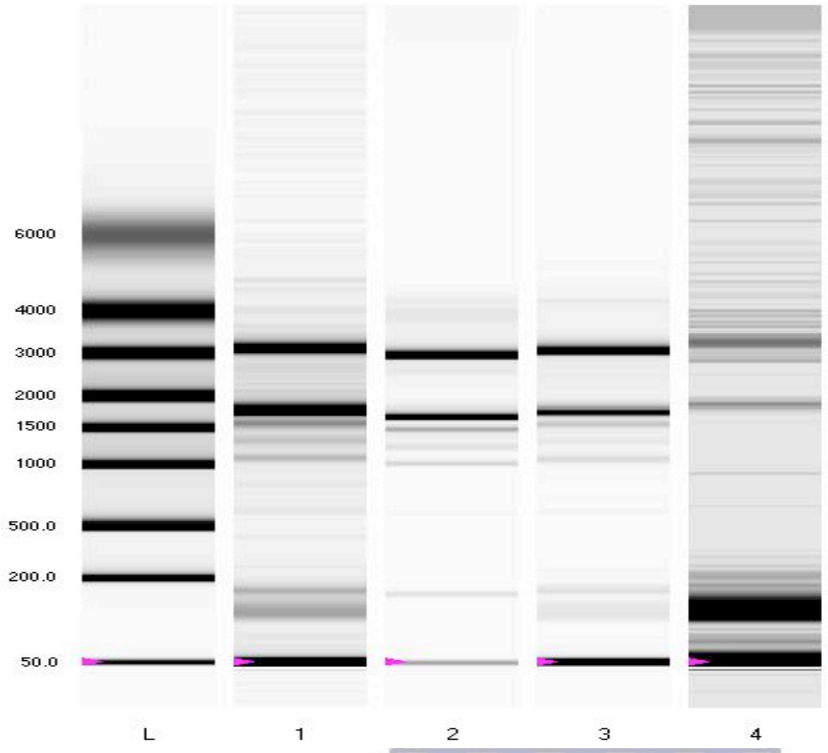


Figure 5.1 Gel image of total RNA quantification on the BIO-RAD Experion System.

L RNA molecular weight marker, lane 1 RNA extracted from red leaf, lane 2 RNA extracted from red leaf, lane 3 RNA extracted from green leaf, lane 4 RNA extracted from green leaf. RNA depicted in lanes 2 (with a Q value of 10) and 3 (with a Q value of 8) were used for cDNA library construction.

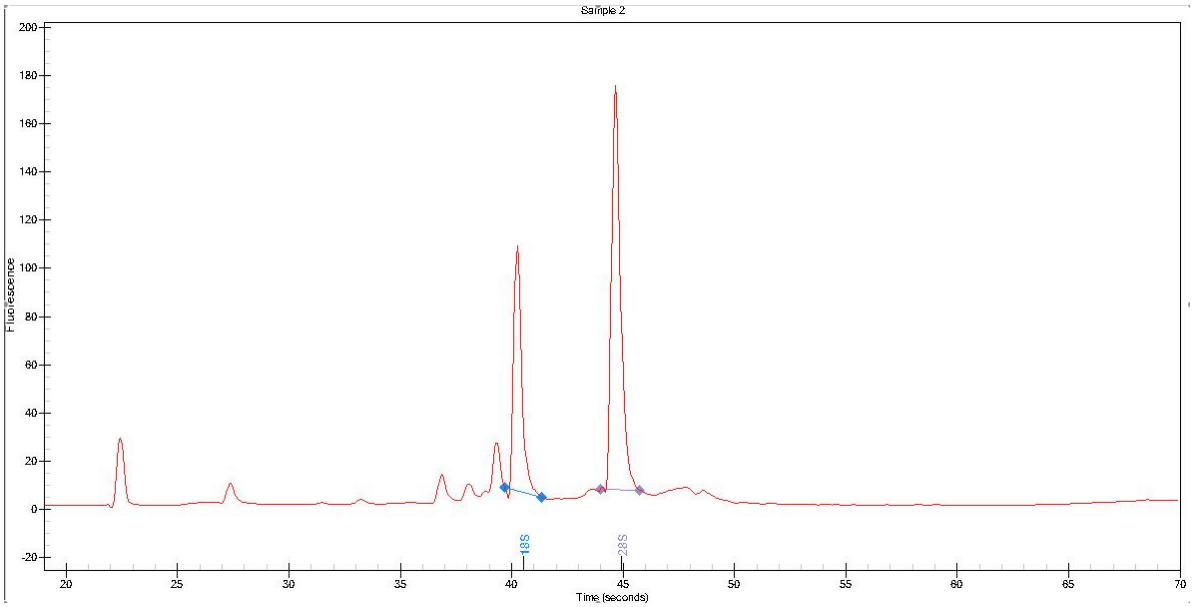


Figure 5.2 Total RNA quantification for the red phenotype from lane 2 in figure 5.1 on the BIO-RAD Experion System indicating the 18S RNA and 28S RNA peaks

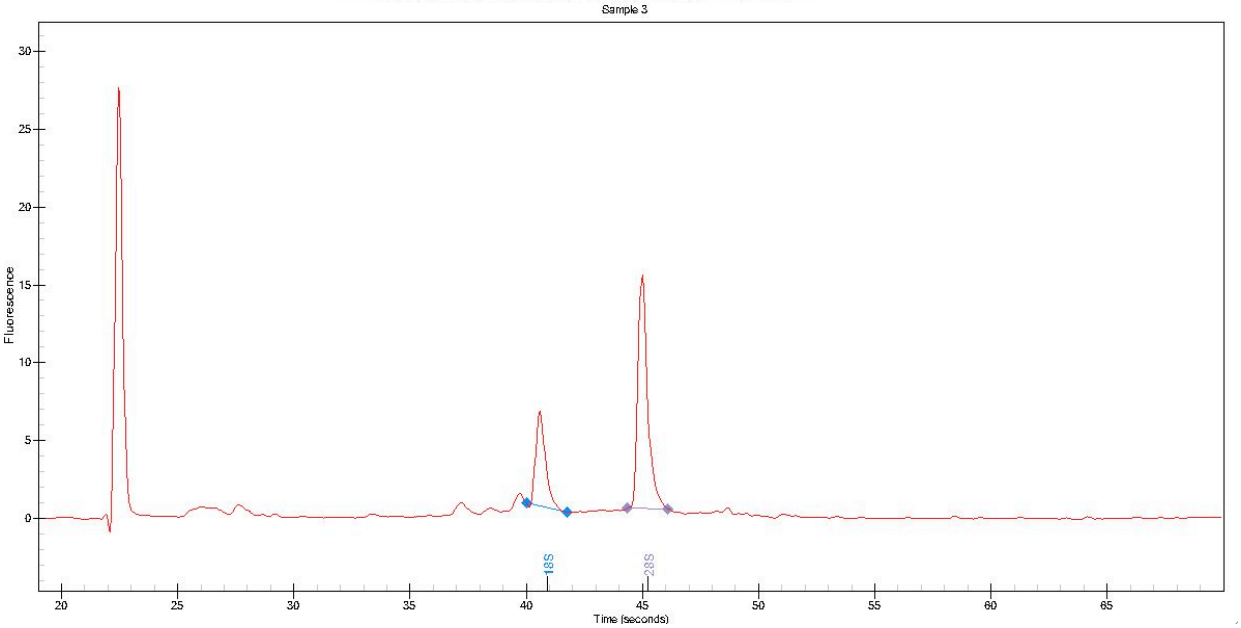


Figure 5.3 Total RNA quantification for the green phenotype from lane 3 in figure 5.1 on the BIO-RAD Experion System indicating the 18S RNA and 28S RNA peaks

5.2 Assembly of reads into contigs

Table 5.1 Velvet assembly output for stated parameters

k-mer	cov cutoff	exp cov	n50	max length	no. contigs	reads used
25	3	100	120	4543	13736	4941291
		1000	124	4543	13708	5100668
25	30	100	392	4482	2746	7489491
		1000	596	5088	2225	7847794

Table 5.2 CLC Bio's Genomic Workbench assembly output at two settings (100 bp and 200bp) for the minimum contig length.

min length	max length	no. contigs	av length	reads used
100	5191	24581	302	5057975
200	5191	8191	646	4414967

5.3 Differential contig expression ratios (RI), red (R) versus green (G)

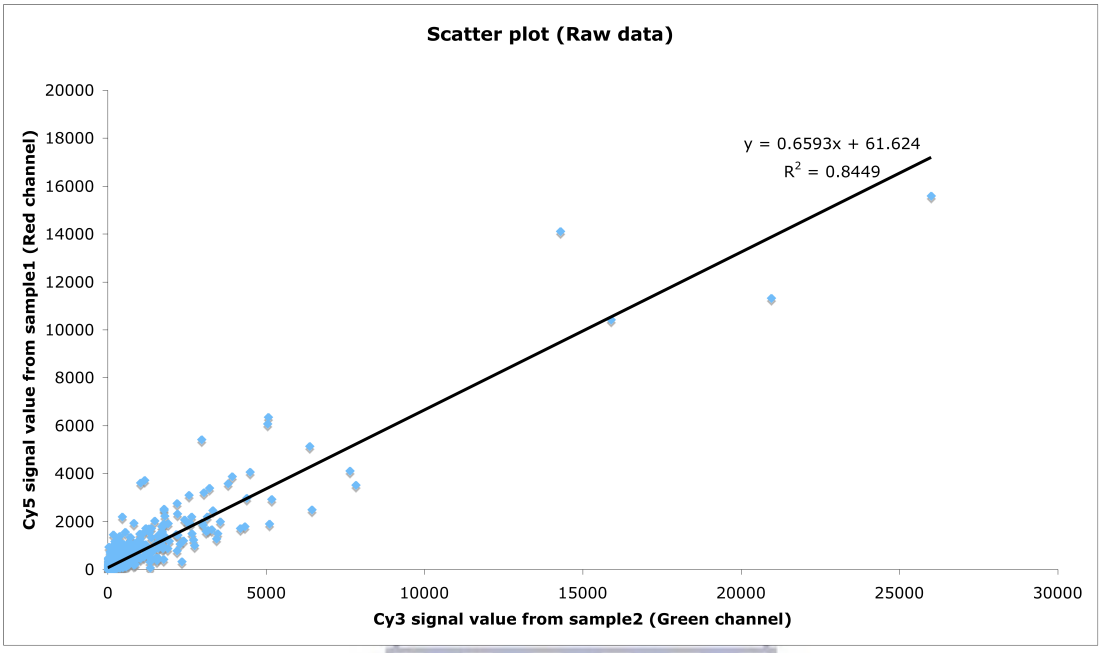


Figure 5.4 Relative intensity (RI) plot for the Log ratio (raw data) of RPKM for red contig versus RPKM for the corresponding green contig.

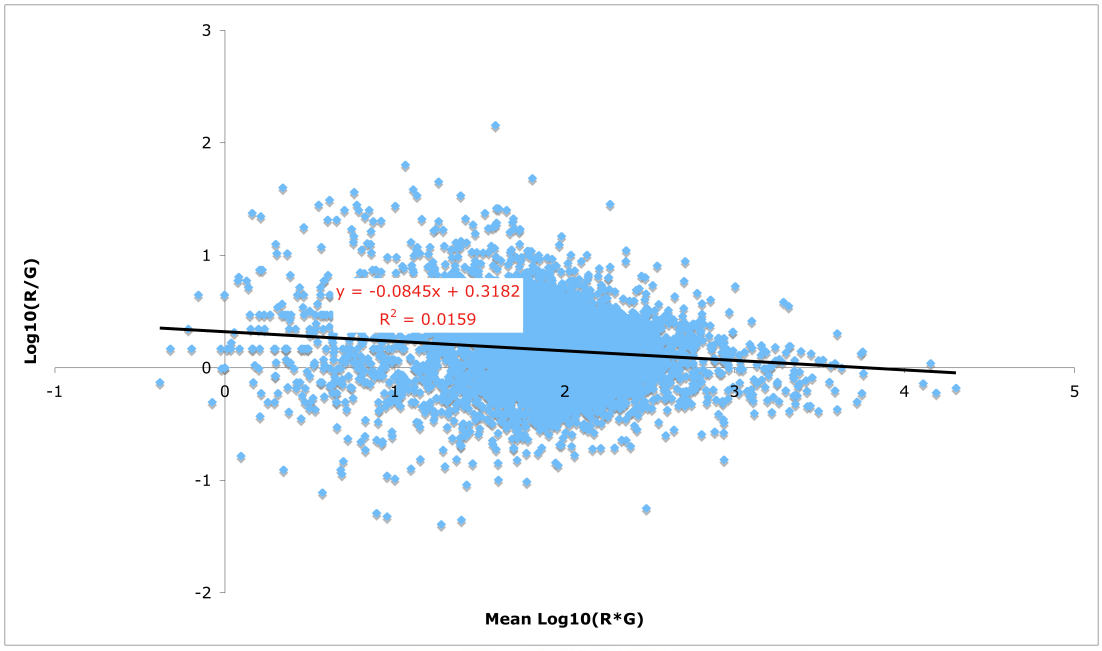


Figure 5.5 Relative intensity (RI) plot for the Log Ratio (Logbase10) of RPKM for red contig versus RPKM for the corresponding green contig.

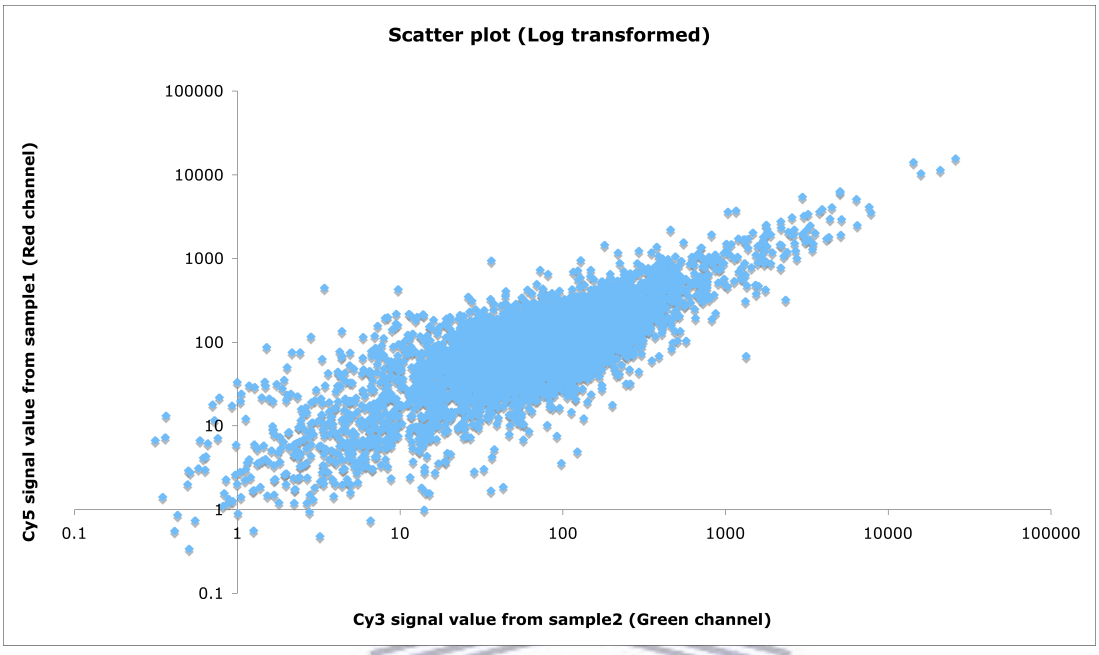


Figure 5.6 Relative intensity (RI) plot for the Log Ratio (Logbase2) of RPKM for red contig versus RPKM for the corresponding green contig.

Table 5.3 Interpretation for RI plot expression levels.

Log ratios of gene expression values are often easier to interpret than raw ratios			
Time (t)	Behavior of Gene	Raw ratio value	Log_2 Ratio Value
0	Basal level of Expression	1.0	0.0
1	No change	1.0	0.0
2	2X UP	2.0	1.0
3	2X Down	0.5	-1.0

5.4 Annotation of transcripts over-expressed in the cDNA library from the red and green phenotype

Table 5.4.1 Over-expressed genes identified from the red phenotype cDNA library. The first twenty-one transcripts were expressed only in the red phenotype library whereas the last ten were significantly up-regulated in the red phenotype compared to the green phenotype library. (na = not applicable, nf = not found).

Contig	GI number	Gene name	E value	Query cov	Max ID
1121	XM_008361364.1	AAA ATPase	0.0	100%	98%
1974	XM_008362989.1	nc	0.0	99%	95%
2559	XM_008383847.1	nc	4e-163	98%	95%
2572	XM_008228123	RHM2	4e-79	100%	93%
2597	XM_008344705	EH3	1e-103	100%	94%
2598	FJ18513	Ej ROP1.1	2e-127	100%	99%
3102	XM_008343661	RAD23d	0.0	99%	97%
3148	KC855732	D-Xylosidase1	1e-118	100%	98%
3233	GU732446	AP2 TF	0.0	99%	91%
3273	XM_008359690	mtPeptidase	1e-134	100%	94%
3338	DQ222994	Mt Lipocalin	2e-59	81%	75%
3655	NM122006	RPT6A	7e-104	74%	86%
4276	XM_004294961	DTC transporter	1e-58	84%	88%
4736	XM_004307038	NADH dehydrogenase	6e-66	93%	88%
4959	nf	na	na	na	na
5046	AK285688	Gm clone	8e-54	35%	79%

5107	nf	na	na	na	na
5209	nf	na	na	na	na
5799	XM_004307190	50S ribosomal protein L1	1e-103	60%	86%%
6069	nf	na	na	na	na
699	XM_004307091	ATP Phospho-ribosyltransferase	3e-49	32%	86%
71 (148)	JN573599	MdGST	0.0	100%	97%
7522	XM_004290937	LCAT	5e-49	77%	81%
4825	EU310513	PsABP2	5e-180	87%	88%
4700	XM_002521974	RcZFP1	6e-162	73%	77%
482	XM_004287856	Succinate DH	0.0	97%	92%
1847	XM_004290182	3-ketoacyl-CoA thiolase	0.0	91%	87%
587	XM_009374226.1	RHD3-like	0.0	62%	98%
3298	NM001250070	Clathrin Hc	3e-58	99%	85%
3457	KC154001	Cell number regulator 20	2e-63	46%	91%
4077	XM002528265	Clathrin Hc	0.0	90%	87%

Table 5.4.2 Transcripts expressed exclusively in the green phenoptype. (na = not applicable, nf = not found).

Contig	GI number	Gene name	E value	Query cov	Max ID
273	XM_008379382	ATP/ADP tranlocon	0.0	100%	98%
276	XM_008341797	CCD4	0.0	98%	98%
620	XM_008368934	Ω-3-desaturase	0.0	97%	91%
1668	XM_008390596	calreticulin-3-like	2e-142	99%	97%
1770	XM_008359435	MdEIN3	0.0	99%	97%
1818	XM_008346833	Tryptophan synthase	0.0	99%	99%
1925	XM_008339745	Aconitase hydratase	0.0	99%	97%
2050	XM_008387233	Glucomannan-synthase	2e-114	100%	99%
2241	XR_523847	unknown locus	0.0	99%	93%
2303	XM_008381163	unknown protein	4e-141	94%	96%
2632	XM_008340716	eIF-3	2e-123	100%	94%
3110	XM_008345463	Nuceolin-like	2e-129	100%	99%
3740	XM_003521448.1	BEL1-like	1e-166	68%	75%
4317	nf	na	na	na	na
4590	EF150643.1	SHMT	0.0	98%	86%
5340	NM_118909.2	Rhodanese-like	3e-43	72%	77%
5772	XM_002328254.1	unknown	0.0	77%	82%
7346	GU732486.1	MdEIN3	4e-121	100%	90%

5.5 Mapping of assembled contigs ratio (RED:GREEN) to anthocyanin structural genes

Table 5.5.1 Expression ratios of anthocyanin genes in red versus green phenotypes mapped to contigs assembled with reads from both phenotypes. Expression ratios are measured as the RPKM ratio of RED:GREEN (R:G) mapped to the specific pear gene indicated in the first column and involved with anthocyanin biosynthesis. *Expression units in RPKM (reads per kilobase exon per million mapped reads). Phenylalanine ammonia lyase (PAL), Chalcone synthase (CHS), Chalcone isomerase (CHI), Flavonone 3 hydoxylase (FHT), Flavonol synthase (FLS), Dihydroflavonol reductase (DFR), Anthocyanidin synthase (ANS), UDP-Glucose flavonoid 3-glucosyl transferase (UFGT), Leucoanthocyanidin reductase (LAR), Anthocyanidin reductase (ANR).

Gene	Accession number (Gene length in bp)	RPKM* (R:G)	Total reads (mapped to the gene)
PAL	DQ230992 (2163)	2.6	1523/1527
CHS	AY786998 (1176)	1.9	407/931
CHI	EF446163 (748)	1.2	58/58
FHT	AY965342 (1095)	5.5	0/30
FLS	DQ230993 (1014)	2.2	1025/1025
DFR	AY227730 (1044)	7.4	73/396
DFR	AY227731 (1044)	4.5	0/20
DFR	AY227732 (1044)	9.8	0/11
ANS	DQ230994 (1074)	6.7	326/852
UF3GT	GQ325589 (1404)	2.4	402/408
LAR1	DQ251190 (1069)	2.1	712/1109
LAR2	DQ251191 (1345)	0.6	12/19
ANR	DQ251189 (1102)	1.6	79/130
MYB10	EU153757 (1545)	80/0	0/4

5.6 DISCUSSION

Transcriptome analysis of red and green phenotypes by mRNAseq

To identify differentially expressed genes between the red and green phenotypes of 'Bon Rouge', and associated with anthocyanin production in the red phenotype I performed mRNAseq on the Illumina GAII platform. Reads were assembled with Velvet (Zerbino *et al*., 2008) and CLC Bio Genomic Workbench. The representative genes associated with the assembled contigs were identified in similarity searches using the BLAST tool at the NCBI's non-redundant nucleotide database (Altschul *et al*., 1990).

5.6.1 RNA isolation and quantification

Total RNA was prepared from the leaves of 'Bon Rouge' and the reverted phenotype with the use of the Qiagen RNeasy kit according to the manufacturers instructions. Total RNA was quantified with BIO-RAD RNA quantification kits on the BIO-RAD Experion Instrument. Samples with a minimum total concentration of 10 μg and an OD260/280 ratio of 1.8 to 2.0 was used for next generation high throughput sequencing. Samples with acceptable RNA quality showed a 28S rRNA band at 4.5kb that was twice the intensity of the 18S rRNA band at 1.9kb. RNA samples with a minimum Q value of 8.0 were used for cDNA library construction.

5.6.2 cDNA library construction and quantification

cDNA libraries for both phenotypes were constructed using Illumina's RNAseq adapted protocol (see methods cDNA library quantification). Libraries were quantified on the Roche LighCycler® according to the manufacturers instructions.

5.6.3 Sequencing of cDNA libraries

Seventeen million single and paired end (SE and PE) reads were generated in six lanes of an Illumina mRNAseq flow cell during three runs on the Illumina GA11 platform. A set of reads for the libraries from both red and green phenotypes of 'Bon Rouge' was always generated in the same run. Reads were pre-processed on the Illumina pipeline and included base calling, quality scoring and preliminary cleanup. Reads with acceptable quality scores were used for assembly. In the first two runs, 51 bp were sequenced for each end of a 200 bp fragment and in the third run, a single end of a 72 bp sequence was generated for each 200 bp fragment. For the second run the quality of the last 15 bp from the reverse read was found to be of low quality and was subsequently removed from all the sequence for both phenotypes.

5.6.4 Assembly of reads into contigs

Reads were assembled with Velvet (Zerbino *et al.,* 2008) and CLC Bio Genomic Workbench. For the Velvet assembly a range of parameters such as k-mer length,

coverage cut off and expected coverage were tested for optimisation of the assembly. K-mer values between 21 and 27, a coverage cut-off range between 3 and 30 and an expected coverage between 30 and 1000 were tested. For assembly with CLC Bio Genomic Workbench, default parameters were used at two different settings of minimum contig length namely 100bp and 200 bp. Although the Velvet assembly produced a N50 of 596 with 7847794 reads at an expected coverage of 1000, the CLC BIO assembly was used in downstream analyses because it produced a comparable average contig length of 646 at a minimum contig length setting of 200bp even though only 4414967 reads were assembled into 8191 contigs (table 5.2). At the specified settings, Velvet assembly produced only 2225 contigs (table 5.1). The assembled contigs were used as a matrix to map subsets of reads from each phenotype to identify contigs that were significantly over-expressed in the red compared to the green phenotype.

5.6.5 Mapping of reads to the assembled contigs

A subset of the reads from run 3 on the Illumina GAII was mapped to the assembled contigs generated using the default parameters in CLC Bio Genomic Workbench with a minimum contig length setting of 200 bps. In an attempt to identify the genes that are differentially expressed between the red and green phenotypes, and associated with anthocyanin production in the red phenotype, I mapped a subset of the reads against the contigs assembled with Velvet (Zerbino *et al.,* 2008) and CLC Bio Genomic Workbench in similarity searches by BLAST (Altschul *et al.,* 1990). The optimum assembly used in downstream analyses was based on the highest N50 value and was achieved with CLC

Bio Genomic Workbench. In order to annotate the transcripts representing the assembled contigs, I used the CLC Bio Genomic workbench assembly with a minimum contig length setting of 200 bp.

5.6.6 Determination of differential expression of contigs in red and green phenotypes

To measure expression in RPKM for each contig in either or both red and green phenotypes, a subset of reads generated from either red or green phenotypes were mapped to the contigs assembled from all the reads generated from both red and green phenotypes (8191 in total). This assembly was generated using the default parameters in CLC Bio Genomic Workbench with a minimum contig length setting of 200 bps. The RPKM ratio for a particular contig was calculated by measuring the RPKM for that contig in the red phenotype to the RPKM for that same contig in the green phenotype. RPKM ratios were log transformed (logbase10 and logbase 2) to calculate the relative expresssion ratio (RI) for each contig in red versus green phenotype. (The logbase10 and logbase2 ratios are to be supplied on a storage device).

5.6.7 Representative genes expressed only in the red phenotype

A set of twenty one contigs were expressed only in the red when compared to those expressed in the green phenotype and similarity searches by BLAST identified the representative genes corresponding to these transcripts as listed in table 5.4. The rest of a

total of 3627 contigs that were over expressed in the red compared to the green phenotype ranged in expression values for RPKM RED:GREEN ratios of 130 for a *Malus domestica* Glutathione transferase followed by a ratio of 44 for an auxin binding protein and 41 for a Zinc finger protein for the most highly expressed transcripts in the red compared to the green phenotype.

Valosin/p97 AAA ATPase (yeast homolog CDC48)

Many critical cellular functions such as membrane fusion, gene transcription, and DNA replication and repair are controlled by the covalent linkages of ubiquitin (Ub) to substrate proteins (Maric *et al*., 2014). Ubiquitinated proteins can be delivered directly to the proteasome for degradation or via p97/VCP (valosin-containing protein). The chaperone p97 or valosin-containing protein (p97/VCP) has been recognized as key player within the ubiquitin/proteasome system (Cayli *et al*., 2009). Whereas the proteasome degrades ubiquitinated proteins, the homohexameric ATPase p97/VCP appears to control the ubiquitination status of targeted substrates. p97/VCP extracts mono- or oligo-ubiquitinated substrates from complexes and present them to the ubiquitin/proteasome system for degradation (Bègue *et al*., 2017). It is a member of the family of ATPases associated with various cellular activities (AAA^+) and forms a homohexamer. Two AAA^+ cassettes of each monomer build two consecutive stacked rings (termed D1 and D2 domains) in the hexamer underneath a ring of flexible N-terminal domains. Its structure is reminiscent of the proteasome base complex that contains a heterohexameric ring of the AAA^+ ATPases Rpt1-6 that resembles the lid of the proteasome.

The COP9 signalosome (CSN), an important mediator of light responses in plants, is also involved in the ubiquitin/proteasome system by controlling the neddylation of ubiquitin E3 ligases. p97/VCP colocalizes and directly interacts with subunit 5 of the CSN (CSN5) *in vivo* and is associated with the entire CSN complex in an ATP-dependent manner. Therefore, CSN and p97/VCP could form an ATP-dependent complex that resembles the 19 S proteasome regulatory particle and serves as a key mediator between ubiquitination and degradation pathways (Cayli *et al.,* 2009). This analogy is based on pairwise sequence homologies between all CSN and RP lid subunits, and structural homologies between the homohexameric AAA^+ ATPase p97/VCP and the AAA^+ ATPases Rp1-6 of the regulatory particle base that form a heterohexamer (Stone, 2014). It has been proposed that p97/VCP and the CSN RP plays a global regulatory role in protein turnover, a proposal that lead to p97/VCP being termed a molecular 'gearbox' that regulates the ubiquitination status of substrates. Cayli *et al.,* (2009) demonstrated that the deneddylase activity of CSN5 and the deubiquitinase USP15 are involved in regulating the ubiquitination status of proteins recruited to p97/VCP and that both activities are crucial to the proper functioning of the ubiquitin/proteasome system.

Ruggiano *et al.,* (2014) proposes a role for the CDC48/p97 ATPase as the driving force for substrate location from the ER membrane to the cytosol in ER associated degradation (ERAD). After release from the membrane, substrates are kept soluble and transferred to the proteasome by cytosolic chaperones or shuttle factors like RAD23 and Dsk2. The final degradation step for ERAD substrates is facilitated by the 26S proteasome.

CDC48, also named p97 or valosin-containing protein (VCP) in animals, is a member of the AAA^+ ATPase (ATPase associated with various cellular activities) which assembles

as a homohexameric complex (Bègue *et al*., 2017). In plants, the role of CDC48 remains poorly understood but available evidence indicate that it displays similar functions to its animal and yeast counterparts, such as endoplasmic reticulum associated protein degradation (ERAD), cell expansion and differentiation, cytokinesis and membrane fusion (Bègue *et al*., 2017, Gallois et al. 2013). Previous studies have demonstrated that it contributes to development, cell division, the ubiquitin-proteasome degradation system, low-temperature-induced freezing tolerance and centromere disassembly (Gallois et al. 2013). Recently the contribution of CDC48 in plant immunity has emerged (Niehl et al. 2012). In Arabidopsis thaliana, CDC48 isoforms localize in different subcellular compartments including the nucleus and the cytoplasm (Gallois et al. 2013), and also in association with the endoplasmic reticulum (ER) and the plasma membrane.

Clone similar to LIL3 and ELIP (Unigene database)

This contig has similarity to a clone with reasonable similarity to a Light inducible-like protein (Lil3) and low similarity to ELIP (via a Unigene analysis). Unlike ELIP, LIL3 may not be involved in light protection since its expression is not dependent on strong illumination according to the Nottingham *Arabidopsis* Stock Center microarray database (Craigon *et al.,* 2004). Instead, LIL3 is proposed to transfer de novo synthesized chlorophyll to the photosystems because it is associated with pigment-binding proteins that appear temporally at the greening stage of barley seedlings (Reisinger *et al.,* 2008). This suggests that LIL3 has a unique biological function distinct from those of other light-harvesting chlorophyll-binding-like proteins (Takahashi *et al.,* 2014). To elucidate the function of LIL3 proteins, Tanaka *et al.,* (2010) analysed *A. thaliana* transposon

mutants lacking one or both isoforms of LIL3 proteins. Both mutants are impaired in the synthesis of α-tocopherol and phytylated chlorophyll as a result of a reduced content of geranylgeranyl reductase. Geranylgeranyl reductase enzyme is responsible for the reduction step of geranylgeranyl pyrophosphate to phytyl pyrophosphate (phytyl-PP) (Tanaka *et al.,* 1999), which is required for chlorophyll, tocopherol, and phyloquinone biosynthesis. The authors demonstrated a physical interaction of LIL3 and geranylgeranyl reductase, suggesting that LIL3 stabilizes geranylgeranyl reductase in plastid membranes.

Rhamnose synthase

L-Rhamnose is a component of the plant cell wall pectic polysaccharides rhamnogalacturonan I (RG-I) and rhamnogalacturonan II (RG-II) (Ridley *et al*., 2001). It is also present in diverse secondary metabolites including anthocyanins, flavonoids, and triterpenoids (Markham *et al*., 2000), in certain types of plant glycoproteins (Haruko and Haruko, 1999), and in arabinogalactan proteins (Pellerin *et al*., 1995). The specific enzymes that attach rhamnose to acceptor molecules are known as rhamnosyltransferases. The small number of plant rhamnosyltransferases that have been investigated thus far are all involved in flavonoid rhamnosylation.

MUCILAGE MODIFIED 4 (MUM4) encodes a rhamnose synthase that is required for L-Rhamnose synthesis, a component of plant cell wall pectic polysaccharides, various secondary metabolites including anthocyanins and flavonoids, some glycoproteins and for seed mucilage biosynthesis. UDP-L-rhamnose is one of the major components of the

plant cell wall skeleton and rhamnose synthase plays a key role in the conversion of UDP-D-glucose into UDP-L-rhamnose (Wang *et al.,* 2008).

Mutations in MUM4 lead to a decrease in seed coat mucilage and incomplete cytoplasmic rearrangement. Western *et al.,* (2004) demonstrated that MUM4 encodes a putative NDP-l-rhamnose synthase, an enzyme required for the synthesis of the pectin rhamnogalacturonan I, the major component of *Arabidopsis* mucilage.

The cellular phenotype seen in mum4 mutants is similar to that of several transcription factors (AP2 [APETALA2], TTG1 [TRANSPARENT TESTA GLABRA1], TTG2 MYB61, and GL2 [GLABRA2]). The GLABRA2 (GL2) gene in *Arabidopsis thaliana* encodes a transcription factor that is required for the proper differentiation of several epidermal cell type (Shi *et al.,* 2011). Expression studies suggest that MUM4 is developmentally regulated in the seed coat by AP2, TTG1, and GL2 (Western *et al.,* 2004). AP2, TTG1, TTG2, MYB61, and GL2 have all been shown to affect flavonoid biosynthesis in the seed coat and other mature plant organs.

Epoxide hydrolase 3 (EH3)

Epoxide hydrolases (EH, EC 3.3.2.10), also known as epoxide hydratase (EC 3.3.2.3) are present in diverse organisms, such as plants, insects, fungi, bacteria, yeast and mammals. They form part of a sub-category of a broad group of hydrolytic enzymes that include esterases, proteases, dehalogenases, and lipases (Beetham *et al.,* 1995). Epoxide hydrolases catalyse the conversion of epoxide or arene oxides to trans-dihydrodiol by adding a water molecule to the cyclic ether. The rice genome contains more than 10 EH like coding genes (Wang *et al.,* 2010) and an epoxy hydrolase has also been identified in

apple. Currently there is no comparative pear sequence in rhe public database. Two types, microsomal epoxide hydrolase (mEH from mammalian, fungal, insect and bacterial species), and soluble epoxide hydrolase (sEH from from mammals, plants, and bacteria) have broad and complementary substrate selectivity and are involved with detoxification of mutagenic, toxic and carcinogenic xenobiotic epoxides. Plant seeds are known to contain oxylipins such as vernolic acid, a monoepoxide of linoleic acid. The hydrolysis of oxylipins to the corresponding diol by epoxide hydrolases produces important intermediates for cutin synthesis, for the production of aromatic components, or for antifungal defenses in plants (Morisseau *et al.,* 2000). In potato the biological function of epoxide hydrolase has not been established but the expression in leaves suggests a role in the production of the protective cutin layer. The enzyme has been demonstrated to be developmentally regulated, as well as by environmental cues. Epoxide-containing fatty acids are the preferred endogenous substrate for the potato EH and this appears to the situation for most plant EHs (Morisseau *et al*. 2000). The potato EH is also suggested to play a role defending against exogenous epoxides.

Rho GTPases of plant (ROP1)

Rho family small GTPases are signaling switches controlling many eukaryotic cellular processes such as the control of cell polarity in eukaryotic cells, and the control of polar growth in pollen tubes (Gu *et al.,* 2006). Plant Rho GTPase (ROP1) displays a unique desiccation-associated ABA signaling transduction through which the ROP1 gene is regulated during the different stage of pollen maturation. In lily, the Rop1 gene was spatially and temporally regulated during anther development (Hsu *et al.,* 2010). The

ROP1 complex is activated by the binding of GTP, and conversion to the inactive GDP form is catlysed by GTPase. Conversion from the GDP- to GTP-bound form is catalyzed by guanine nucleotide exchange factors (GEFs). In *Arabidopsis* the Rho GEFs, named RopGEFs consists of a family of 14 members that contains a conserved central domain, the domain of unknown function 315 (DUF315), and variable N- and C-terminal regions (Gu *et al.,* 2006). Disruption in F-actin produce changes in ROP1 activity indicating a role for calcium in the negative feedback regulation of the ROP1 activity (Yan *et al.,* 2009). During pollen maturation, the ROP1 gene is regulated via a unique desiccation-associated ABA signaling transduction (Hsu *et al.,* 2010).

RADIATION SENSITIVE23 (RAD23)

The nuclear-enriched RADIATION SENSITIVE23 (RAD23) family proteins bind ubiquitin (Ub) conjugates and flag these for degradation by the proteasome (Derrien and Genschik, 2014). Transcriptional roles have been postulated for the ubiquitin receptor RAD23 suggesting they may be key components of proteasome transcriptional specificity (Wade *et al.,* 2010). RAD23 and cell division cycle protein 48 (CDC48) are two key regulators of post-ubiquitylation events that act on distinct and overlapping sets of substrates (Baek *et al.,* 2011).

RAD23 is an adaptor protein that binds to both ubiquitylated substrates and to the proteasome. It serves as a link for ubiquitylated substrates to the proteasome for degradatiojn by the proteasome. However, even though it is linked to the proteasome, RAD23 escapes degradation because it lacks an effective initiation region at which the proteasome can attack the protein and unfold it (Fishbain *et al.,* 2011). The RAD23

proteins bind Ub conjugates, especially those linked internally through Lys-48, via their UBA domains, and associate with the 26S proteasome Ub receptor RPN10 via their N-terminal UBL domains (Farmer *et al.,* 2010).

The ubiquitin/26S proteasome system (UPS) directs the turnover of misfolded and numerous regulatory proteins, thereby controlling many aspects of plant growth, development, and survival. RAD23 proteins appear to play an essential role in the cell cycle, and morphology and fertility of plants through their delivery of ubiquitin/26S proteasome system substrates to the 26S proteasome (Stone, 2014).

Arabinosidase

Alpha-N-arabinofuranosidase (EC 3.2.1.55) also known as arabinosidase, alpha arabinosidase and alpha-L-arabinosidase form part of the hydrolases and glycosylases glycosidases classes. These enzymes hydrolyse O- and S-glycosyl compounds and hydrolyse alpha-L-arabinofuranosides, alpha-L-arabinans containing (1,3)- and/or (1,5)-linkages, arabinoxylans and arabinogalactans. Some beta-galactosidases (EC 3.2.1.23) and beta-D-fucosidases (EC 3.2.1.38) also hydrolyse alpha-L-arabinosides. Softening of fleshy fruits and cell walls during ripening is associated with the catabolism of cell wall components such as alpha-L-arabinofuranosides, alpha-L-arabinans. In strawberry, pectin degradation combined with the loss of neutral sugars such as arabinose, increase during ripening, and probably contributes to fruit softening (Rosli *et al.,* 2009).

An alpha-L-Arabinofuranosidase (alpha-L-arafase) has been characterized in Japanese pear (*Pyrus pyrifolia*). The enzyme comprises a single 62-kD polypeptide as determined

on SDS-PAGE (Tateishi *et al.,* 2005) and is related to a cDNA clone, PpARF2 whose transcript and related protein were detected only in the ripening fruit. The increase in alpha-L-arafase activity was closely associated with the increase in transcript and related protein in the ripening fruit. Transcripts of PpARF2 were not detected in buds, leaves, roots, or shoots of the Japanese pear. The deduced amino acid sequences of PpARF2 displayed minimal identity with those of other plants or bacteria.

Apetala 2 (AP2) transcription factor

In *Arabidopsis*, the Apetala 2/ ETHYLENE RESPONSE FACTOR1-like (AP2/ERF-like) genes are represented by a large transcription factor family of 147 genes (Nakano *et al.,* 2006), several of which are upregulated by ethylene (Alonso and Stepanova, 2004). ERF1 is bound and activated by the transcription factor ETHYLENE INSENSITIVE3 (EIN3) which increases during apple fruit development and ripening (Solano *et al.,* 1998). This developmental process is associated with an increase in the expression of cell wall hydrolase genes such gene as POLYGALACTURONASE1 (PG1) whose increase in expression correlates with both an increase in ethylene, and following cold treatment. The cold-treatment induced increase in PG1 is mediated via a COLD BINDING FACTOR (CBF) gene that transactivates the PG1 promoter. The addition of exogenous ethylene significantly increases the transactivation (Stockinger *et al.,* 1997).

Lipocalin

Lipocalins represent an ancient protein family in bacteria, protoctists, plants, arthropods, and chordates (Wong *et al.*, 2017). These small ligand-binding proteins display a simple

tertiary structure that allows them to bind small, generally hydrophobic molecules. Currently, very little is known about plant lipocalins. Previously, Charron *et al.,* (2005) have reported the cloning of the first true plant lipocalins and later identified and characterized a plant lipocalins and lipocalin-like proteins using an integrated approach of data mining, expression studies, cellular localization, and phylogenetic analyses. Plant lipocalins can be classified into two groups; temperature-induced lipocalins (TILs) and chloroplastic lipocalins (CHLs). Plant temperature-induced lipocalins (TILs) have been shown to be responsive to heat stress (HS), but the nature of this response has not yet been elucidated. It is postulated that AtTIL1 is an essential component for thermotolerance and most likely act against lipid peroxidation induced by acute heat stress. In addition, violaxanthin de-epoxidases (VDEs) and zeaxanthin epoxidases (ZEPs) can be classified as lipocalin-like proteins (Hieber *et al.,* 2000). An *Arabidopsis*, chloroplastidic lipocalin (AtCHL) is involved in the protection of thylakoidal membrane lipids against reactive oxygen species, especially singlet oxygen, produced in excess light.

Regulatory Particle 6A (RPT6A)

RPT6 is a component of the regulatory particle that form a heterohexamer made up of AAA^+ ATPases Rp1-6 (Wei *et al.,* 2008). The regulatory particle is part of the COP9 signalosome (CSN) that is reminiscent of the proteasome base complex that also contains a heterohexameric ring of the AAA^+ ATPases Rpt1-6. The COP9 signalosome (CSN) is an important mediator of light responses in plants and also involved in the ubiquitin/proteasome system by contolling the neddylation of ubiquitin E3 ligases.

p97/VCP colocalizes and directly interacts with subunit 5 of the CSN (CSN5) *in vivo* and is associated with the entire CSN complex in an ATP-dependent manner. CSN is an ATP-dependent complex that resembles the 19 S proteasome regulatory particle and serves as a key mediator between ubiquitination and degradation pathways (Cayli *et al*., 2009) in photomorphogenesis in plants.

Glutathione transferase (GST) (up-regluated in the red phenotype)

The flavonoid transport enzyme Glutathione *S*-transferase (GSTs) is suggested to be the last genetically defined step in flavonoid biosynthesis (Walbot *et al*., 2000). The GST enzymes have a central role in the transport of flavonoids through the cytoplasm to the vacuolar membrane (Walbot *et al*., 2000) where the acidic vacuole preserves the flavonoids and prevent their degradation for recycling into the phenylpropanoid pathway. GSTs may also play a role in the regulation and signaling of the flavonoid biosynthetic pathway (Loyall *et al*., 2000). GST mRNA transcripts were found to be abundant in immature fruit while the fruit are actively synthesizing flavonoids (Steyn *et al*., 2004).

The plant-specific phi class glutathione transferases (GSTs) are often highly stress-inducible and expressed in a tissue-specific manner. This suggests an important protective role in the regulation of the binding and transport of protective defence-related compounds such as anthocyanins (Dixon *et al*., 2011, Monticolo *et al*., 2017). AN9, a petunia (Petunia hybrida) glutathione S-transferase is required for efficient anthocyanin transport from its biosynthetic site in the cytoplasm to stabilised storage in the vacuole. Mueller *et al*., (2000) demonstrated that AN9 bind flavonoids using three assays: inhibition of the glutathione S-transferase activity of AN9 toward the common substrate

1-chloro 2,4-dinitrobenzene, equilibrium dialysis, and tryptophan quenching. Consequently they concluded that AN9 is a flavonoid-binding protein, and proposed that in vivo it serves as a cytoplasmic flavonoid carrier protein.

Auxin binding protein (ABP) (up-regluated in the red phenotype)

Auxin regulates plant cell division, elongation, and differentiation through signal transduction (Christian *et al.,* 2006). In plants, auxin signals are perceived by proteins referred to as auxin receptors. As a result numerous attempts have been made to purify auxin-binding proteins (ABPs) that specifically bind auxin. In climacteric fruits most aspects of the ripening process are triggered and maintained by ethylene (Lelièvre *et al.,* 1997) and numerous studies showed that the endogenous auxin content significantly increased during fruit ripening, concomitant with the production of climacteric ethylene (Miller *et al.,* 1987). Some ABPs have been identified as proteins with enzymatic activities such as β-glucosidase (Campos *et al.,* 1992), 1,3-β-glucanase (Macdonald *et al.,* 1991), glutathione S-transferase (Bilang *et al.,* 1993, Zettl *et al.,* 1994), manganese superoxide dismutase (Feldwisch *et al.,* 1995), and glutathione dependent formaldehyde dehydrogenase (Sugaya and Sakai 1996).

Zinc finger protein-like (ZFP-like) (up-regluated in the red phenotype)

A number of zinc finger proteins that act as transcription factors have been characterized in plants.

ZPT2 (for petunia zinc-finger protein 2)-related proteins contain two Cys2/His2-type zinc-finger motifs and an ethylene-responsive element binding factor-associated

amphiphilic repression motif. During abiotic stress many plant genes are down-regulated and ZPT2 are thought to function as transcriptional repressors under such conditions (Kodaira *et al*., 2011).

LOW QUANTUM YIELD OF PHOTOSYSTEM II1 (LQY1) is a small Zn finger containing thylakoid membrane protein of *Arabidopsis thaliana* that shows disulfide isomerase activity. LQY1 interacts with the photosystem II (PSII) core complex and may act in repair of photo-damaged PSII complexes. A mutant of LQY1 were found to have a lower quantum yield of PSII photochemistry and reduced PSII electron transport rate following high-light treatment. This suggests an involvement for LQY1 protein in maintaining PSII activity under high light stress by regulating repair and reassembly of PSII complexes (Lu *et al*., 2011).

Stress-associated proteins (SAPs), containing A20/AN1 zinc-finger domains, confer abiotic stress tolerance in different plants. However, the interacting partners and downstream targets have not yet been identified. (Giri *et al*., 2011).

The *Arabidopsis* TT1 gene encodes a WIP-type zinc finger protein that is expressed in the seed coat endothelium. Production of proanthocyanidin pigments (PAs) in the wild type of *Arabidopsis thaliana* produces brown seed coats. The pigmentation requires activation of phenylpropanoid biosynthesis genes and mutations in some of these genes cause a yellow appearance of seeds, termed transparent testa (tt) phenotype. TT1 is also required for the regulation of expression of CHS, the first enzyme of flavonoid biosynthesis in other plant organs. Expression of the genes encoding enzymes of this pathway are controlled by trimeric complexes of MYB and bHLH transcription factors, and the WD40 factor TTG1. Appelhagen *et al*., (2011) have demonstrated the interaction

of TT1 with the R2R3 MYB protein TT2 to produce PA in the seed coat of wild type *Arabidopsis*.

Tandem CCCH Zinc Finger (TZF) proteins can affect gene expression at both transcriptional and post-transcriptional levels. The *Arabidopsis* TZF1 (AtTZF1) has been demonstrated to bind both DNA and RNA *in vitro*, and can exchange between the nucleus and cytoplasm. However, this binding has not been demonstrated in vivo and little is known about the molecular mechanisms underlying AtTZF1's profound effects in plants on development, stress responses and growth (Pomeranz *et al.,* 2011).

Zinc finger proteins are that are characterized by a B-box containing motif play an important role in light signaling in plants. LIGHT-REGULATED ZINC FINGER1 (LZF1)/SALT TOLERANCE HOMOLOG3 (STH3) is a B-box encoding gene that interacts at the gene level with two key regulators of light signaling, ELONGATED HYPOCOTYL5 (HY5) and CONSTITUTIVE PHOTOMORPHOGENIC1 (COP1). Chang *et al.,* (2008) identified LZF1 as a gene regulated at the transcriptiona level by HY5. LZF1 was found to act as a positive regulator of de-etiolation affecting both anthocyanin production and chloroplast biogenesis. However, in this same study it was demonstrated that since hy5 was not epistatic to sth3, a *LZF1* homolog, the activity of LZF1/STH3 can not be regulated by HY5 alone. HY5 binding targets was enriched in light-responsive genes and transcription factor genes (Lee *et al.,* 2007), suggesting that HY5 is an hierarchical regulator of the transcriptional cascades for photomorphogenesis. Mutation in the HY5 gene also causes defects in lateral root formation, secondary thickening in roots, and chlorophyll and anthocyanin production (Oyama *et al.,* 1997; Holm *et al.,* 2002). Furthermore, sth3 suppresses the cop1 hypocotyl phenotype in the

dark as well as anthocyanin accumulation in the light (Datta *et al.,* 2008), indicating that factors additional to HY5 is responsible for the regulation of LZF1/STH3.

Succinate dehydrogenase (up-regluated in the red phenotype)

Succinate dehydrogenase is an inner mitochondrial membrane succinate:ubiquinone oxidoreductase (EC 1.3.5.1) that is also known as succinate-coenzyme Q reductase (SQR). It is an iron-sulfur protein that catalyses an oxidation-reduction reaction during which succinate is oxidized to fumarate and ubiquinone (Q) reduced to ubiquinol (QH_2) with FAD acting as an electron acceptor. The reaction forms part of both the citric acid cycle, and the membrane-associated electron transport system that couples oxidation-reduction reactions to the transfer of protons across a membrane. The succinate dehydrogenase co-factor FAD is covalently attached to the enzyme via a histidine side chain and is part of the short electron transfer chain from succinate to QH_2. $FADH_2$, the electron acceptor, never dissociates from the enzyme so the final product of this reaction is QH_2 (Cheng *et al.,* 2008). The reduction of Q facilitates the transfer of two electrons from the cytosol to the inner mitochondrial membrane. The resulting proton motive force is used to drive the synthesis of ATP. Mitochondrial succinate dehydrogenases are composed of two hydrophilic and two hydrophobic subunits. Two phospholipids, cardiolipin and phosphatidylethanolamine, are associated with the hydrophobic subunits of the complex.

RHD3-like (Root Hair Defective 3-like) ((up-regluated in the red phenotype)

In a study on nitrogen starvation in *Arabidopsis*, Wang *et al.*, (2015) identified a new

allele of *ROOT HAIR DEFECTIVE3* (*RHD3*) showing an anthocyanin overaccumulation phenotype under nitrogen starvation conditions. The study demonstrated a close relationship among RHD3, anthocyanin biosynthesis, and ethylene signaling. The anthocyanin overaccumulation phenotype observed in *rhd3* mutants indicated a negative role of RHD3 in anthocyanin biosynthesis. It is known that ethylene negatively regulates anthocyanin under sucrose stress, primarily through the ETR1-, EIN2-, and EIN3/EIL1-mediated signaling pathways (Jeong *et al*., 2010). The exact molecular mechanisms underlying these interactions are not fully understood. In *Arabidopsis* seedlings, it is known that low nitrate induces a transient rise in ethylene production and enhances the expression of the ethylene signaling components CTR1, EIN2, and EIL1, and that these form a negative feedback pathway to reduce high-affinity nitrate uptake by suppressing *NRT2.1* expression (Zheng *et al*., 2013). This study consistently showed that ethylene negatively regulates the anthocyanin biosynthesis that is induced by low nitrogen conditions in the hypocotyls, in which ETR1, EIN2, and EIN3/EIL1 play essential roles, and the contribution from other ethylene receptors was modest.

This working model by Wang et al., (2015) proposes that nitrogen starvation activates both stimulating and inhibitory signaling pathways to fine-tune anthocyanin biosynthesis. The inhibitory signaling pathway is primarily mediated by the ET-dependent pathway, which in turn has partial dependence on RHD3 function. The precise molecular and cellular mechanism of RHD3 in ethylene signal transduction requires further investigation.

5.6.8 Representative genes expressed only in the green phenotype

A set of eighteen contigs were expressed only in the green when compared to the red phenotype and similarity searches by BLAST identified two of these representative genes corresponding to these transcripts that may function at a hierarchy in the regulation of anthocyanin biosynthesis. These are EIN3 and calreticulin3 (CRT3). A sub-set of the over-expressed contigs for the green phenotypes is listed in (table 5.4.2).

ETHYLENE-INSENSITIVE3 (EIN3)

Of the contigs that were expressed in the green phenotype only, two could be identified by similarity searches using the BLAST tool at the NCBI's non-redundant database as transcription factors involved with ethylene signalling. Contig 273 show similarity to an ETHYLENE-INSENSITIVE3 (EIN3) transcription factor (GU732486.1, E value = 0, maximum score = 2139, query coverage = 99% and maximum identity = 96%). A second contig (contig 7346) that was expressed in the green phenotype only, display similarity to an ETHYLENE-INSENSITIVE3-like (EIL1) transcription factor but with lower scores and shorter contig length. Ethylene (ET) is a major plant hormone that regulates plant development and tolerance to necrotrophic fungi (Zhu *et al.,* 2011) via the EIN3 transcription factor. Ethylene also modulates sucrose and glucose sensitivity during *Arabidopsis* seedling development and controls anthocyanin biosynthesis (Gibson *et al.,* 2001). Anthocyanin accumulation is suppressed by ethylene signaling and activated by sugar and light signaling. Jeong *et al.,* (2010) reported the presence of an anthocyanin induction pathway that is independent of HY5 but dependent upon photosynthetic electron transport in acyanic mesophyll cells. EIN3 is a short-lived

protein whose degradation is mediated by two F-box proteins, EBF1 and EBF2 via the ubiquitin/26S proteasome pathway in the absence of ethylene (Tacken *et al.,* 2012). Ethylene treatment reduces the levels of CONSTITUTIVE TRIPLE RESPONSE 1 (CTR1) and EIN2 that facilitates stabilization of the EIN3 protein due to decreased interaction with EBF1 and EBF2 with subsequent accumulation of EIN3 in the nucleus.

In *Arabidopsis* the ethylene response pathway is mediated by the nuclear proteins ETHYLENE-INSENSITIVE3 (EIN3) and ETHYLENE-INSENSITIVE3-like (EIL3) (Solano *et al.,* 1998, Yu *et al*., 2017), and related proteins that result in a triple enhanced response mediated by the negative regulator, CTR1 (Chao *et al.,* 1997). EIN3 and EIL1 are signal transduction transcription factors in ethylene signaling and expression of EIN3 or EIL1 promote the binding and activation of secondary transcription factors, including ETHYLENE RESPONSE FACTOR1 (ERF1) (Kendrick and Chang, 2008).

Calreticulin3 (CRT3)

Contig 1668 shows homology to the endoplasmic reticulum (ER) localized protein calreticulin 3 (CRT3) (XM_00838560, E value = 2e-142, maximum score = 514, query coverage = 99% and maximum identity = 97%), an important component for protein folding, endoplasmic reticulum associated degradation (ERAD) of misfolded proteins and Ca^{2+} homeostasis in the ER of animal cells (Michalak *et al.,* 1999). In addition, animal CRTs have been implicated in more than 40 other cellular functions, highlighting their versatility (Michalak *et al.,* 1999). Compared to their functions in animal cells, the role of plant CRTs is less clear but is believed to have chaperone-like functions in plants (Michalak *et al.,* 2009). The best evidence for such a function is CRT facilitated

formation of stress-induced, i.e. heat-shock-induced, protein complexes in tobacco leaves, suggesting that it may bind to unfolded proteins and therefore possibly function as a molecular chaperone (Michalak *et al.,* 2009).

CRT appears to reside mainly in the ER, and the Golgi apparatus, but has also been shown to localize to the nuclear envelope in plant cells (Michalak *et al.,* 2009). Plants and mammals appear to contain two subgroups of CRT proteins and *Arabidopsis* contain three CRT family members that are classified into an AtCRT1a/1b, and an AtCRT3 group, on the basis of sequence homology (Christensen *et al.,* 2010). Several recent studies propose that AtCRT3 is necessary for the folding of the bacterial epitope elf18 responsive EF-Tu receptor (EFR) associated with Pathogen-Associated Molecular Patterns (PAMPs)-triggered immunity in plants (Christensen *et al.,* 2010). Pattern recognition receptors in eukaryotes initiate defense responses on detection of microbe-associated molecular patterns shared by many microbial species.

To further characterise the role of CRT3 in plants, Saijo *et al.,* (2009) have taken a genetic approach to identify and isolate *Arabidopsis* mutants that are insensitive to PAMPs by focusing on the two best-studied PAMP/PRR (pathogenesis response receptor) pairs in plants flg22/FLS2 and elf18/EFR. Leu-rich repeat receptor-like kinases FLS2 and EFR recognize the bacterial epitopes flg22 and elf18, derived from flagellin and elongation factor-Tu, respectively. Saijo *et al.,* (2009) described *Arabidopsis* 'priority in sweet life' (*psl*) mutants that display de-repression of sucrose-induced anthocyanin accumulation in the presence of elf18. EFR but not FLS2 accumulation and signalling, are impaired in *psl1*, *psl2*, and *stt3a* plants. *PSL1* and *PSL2*, encode calreticulin3 (CRT3) and UDP-glucose:glycoprotein glycosyltransferase, respectively, that act in concert with

STT3A-containing oligosaccharyltransferase complex in an N-glycosylation pathway in the ER. The up-regulation of calreticulin3 in the green phenotype may suggest a role for the repression of anthocyanin production in the presence of the bacterial epitope elf18. Further functional characterisation of calreticulin in the green phenotype of 'Bon Rouge' may explain/elucidate/highlight the role of this protein in the suppression of anthocyanin production in the green phenotype, in response to elf18.

Carotenoid cleavage dioxygenase (CCD4)

A second contig (276) that was expressed in the green phenotype only, and identified by similarity searches using the BLAST tool at the NCBI's non-redundant nucleotide database was similar to an apple Carotenoid cleavage dioxygenase (CCD4) (EU327777.1, E value = 0, maximum score = 1051, query coverage = 97% and maximum identity = 92%). Carotenoid cleavage dioxygenases (CCDs) are non-heme iron oxygenases that cleave carotenes and xanthophylls to colourless apocarotenoids (Rubio *et al*., 2008). This oxidative cleavage of carotenoids occurs in plants, animals, and microorganisms. Apocarotenoids are abundant in the thylakoid membranes of plants and cyanobacteria where they act as photoprotective pigments, accessory pigments in thylakoid membrane (Markwell *et al*., 1992) and signaling molecules with diverse functions, including the plant hormone abscisic acid. The first identified gene encoding a carotenoid cleavage dioxygenase was the maize Vp14 gene that is required for abscisic acid (ABA) biosynthesis (Tan *et al*., 1997).

Although carotenoid cleavage dioxygenase (CCD) genes have been functionally characterized in a number of plant species, the biochemical role and enzymatic function

of the subclass 4 (CCD4) members remained largely unknown until recently. Ahrazem *et al*., (2010) characterized the plastoglobule-targeted enzyme (CCD4) that catalyses the formation of volatile C_{13} ketones, such as β-ionone, by cleavage of the C_9-C_{10} and C_9'-C_{10}' double bonds of cyclic carotenoids. The apple MdFS2 (MdCCD4) gene product was demonstrated to catalyse the same reaction on the substrate β-carotene to produce β-ionone (Huang *et al*., 2009).

5.6.9 Blast2GO analysis

To characterise and functionally annotate the transcript expressed only in the red phenotype, or those over-expressed in the red phenotype with a red:green logbase2 ratio of 2 and above, open source software Blast2GO was employed. Thirty of the top one hundred and five contigs subjected to this analysis show similarity to enzymes with established Enzyme Commission (EC) numbers in pathways represented in the KEGG database. Of these, only five could be directly associated with anthocyanin production, i.e. displayed enzyme codes similar to those involved with catalytic steps in the anthocyanin biosynthesis pathway. These include phenylalanine ammonia-lyase (PAL) and UDP-galactose transporter and glutathione transferases (GST). In addition to the Blast2GO analysis, a BLAST search with the one hundred and five contig sequences over-expressed in the red phenotype was used as queries against the non-redundant nucleotide database.

Anthocyanin production in plants is facilitated by a complex of biosynthetic enzymes localized to the cytosolic membranes of the endoplasmic reticulum. For stabilisation of

pigment colour and to prevent its toxic effects in the cytosol, anthocyanin pigments have to be transported to, and sequestered in the acidic vacuole (Zhao, 2010). A member of the glutathione S-transferase (GST) family was reportedly involved in vacuolar anthocyanin transport (Kitamura, 2006), and its molecular function is currently being elucidated (Monticolo *et al*., 2017).

Glutathione S -transferases (GSTs; E.C. 2.5.1.18) form a superfamily of multifunctional, dimeric enzymes, best known for their role in enzymatic detoxification of xenobiotics (Moons, 2005). In plants GSTs appear to be involved in plant growth and development and have been shown to bind hormones such as auxin and cytokinin, while it can also be induced by a wide variety of phytohormones such as ethylene, auxin, methyl jasmonate, and salicylic and abscisic acid (Moons, 2005). Such interaction imply that these hormones regulate many aspects of plant development and that plant GSTs are essential to plant growth and development. However, evidence to substantiate this role has been limited (Jiang *et al.,* 2010)

Recently Sun *et al.,* (2012) demonstrated that the *Arabidopsis Transparent Testa 19* (*TT19*), a glutathione S-transferase, functions as a carrier in the transport of cyanidin and/or anthocyanins to the tonoplast for sequestration in the vacuole. TT19 is localised both in the cytoplasm and on the tonoplast, while conjugated to cyanidin and cyanidin 3-galactoside. However, the researchers demonstrated that TT19 does not conjugate these pigments with glutathione, thus supporting the hypothesis that TT19 is a carrier protein for anthocyanins to facilitate its sequestration into the acidic vacuole for pigment stabilisation.

Two contigs over-expressed in the red compared to the green phenotype displayed sequence similarity to glutathione transferases as determine by Blast2GO analysis. Contig 71, over-expressed in the red phenotype with a logbase2 ratio of 7.018 for red to green shows sequence similarity to a glutathione transferase (EC:2.5.1.18). Contig 3681, over-expressed in the red phenotype with a logbase2 ratio of 3.733 for red to green shows sequence similarity to a glutathione transferase, *AtGSTu17*. However, in the Blast2GO annotation, the latter enzyme was assigned the enzyme commission number, EC:4.4.1.5, which is classified in the KEGG database as a lactoylglutathione lyase (ketone-aldehyde mutase) that catayses an aldo-keto isomerisation. *AtGSTu17* is a member of the GST N family, Class Tau (U) subfamily. Plant Tau GSTs are cytosolic dimeric proteins involved in cellular detoxification by catalyzing the conjugation of glutathione (GSH) with a wide range of endogenous and xenobiotic alkylating agents, environmental toxins such as herbicides and products of oxidative stress. In rare instances it catalyses the bio-activating isomerization of secondary metabolites (Cummins *et al.,* 2011). In addition plant Tau GSTs play important roles in intracellular signalling, biosynthesis of anthocyanin, responses to soil stresses and responses to auxin and cytokinin hormones (Licciardello *et al.,* 2014). Tau GSTs, are also involved in responses to different environmental stresses including heat, cold and drought, and chemical compounds such as hydrogen peroxide, salicylate, DTT, CuSO4, and also in herbicide metabolism.

According to a study by Jiang *et al.,* (2010), AtGSTU17 is mainly involved in FR light signaling and is regulated by various photoreceptors, in particular phyA, under all light conditions. Its function appears to affect anthocyanin accumulation, and flowering time. The double mutant *atgstu17phyA* showed defects in physiological responses, including

slightly longer hypocotyls in low FR, reduced anthocyanin levels, and insensitivity in FR-mediated inhibition of greening. Auxin transport proteins like PIN7, and other auxin responsive-genes are affected in abscisic acid treated atgstu17 mutants (Jiang *et al.,* 2010). A number of auxin efflux carriers like PIN1, PIN2 and PIN7 are reduced in ROOT MERISTEMLESS 1 (RML1) mutants that lack the RML1 gene that encodes the first enzyme of glutathione synthesis (Kopriva *et al.,* 2010). The reduced glutathione levels result in altered auxin transport. Four contigs were annotated by Blast2GO as auxin effluc proteins or auxin binding proteins.

Contig 4825 over-expressed in the red phenotype with a logbase2 ration of 5.454 was annotated by Blast2GO analysis as an auxin binding protein abp19a (GO:0045735).

Contig 4939 and 7893 over-expressed in the red phenotype with a logbase2 ration of 3.892 and 3.320, respectively, were similarly annotated by Blast2GO analysis as auxin binding protein abp19a (GO:0045735).

Contig 74 over-expressed in the red phenotype with a logbase2 ratio of 3.341 was annotated by Blast2GO analysis as an auxin efflux carrier component (GO:0055085) while a sequence similarity search by BLAST against the non-redundant nucleotide database with this query sequence annotated this contig as PIN1. Further investigation into the interaction of these various auxin-binding proteins with glutathione transferases should facilitate an understanding of their role in anthocyanin production in the red phenotype.

Contig 6805 over-expressed in the red phenotype with a logbase2 ratio of 3.376 shows sequence similarity to a phenylalanine ammonia-lyase (PAL). Phenylalanine ammonia-lyase [EC:4.3.1.25] is the first committed step for phenylpropanoid biosynthesis in plants.

It catalyzes the deamination of phenylalanine to *trans*-cinnamic acid and ammonia. PAL is a member of a large gene family in *Arabidopsis* and grape as demonstrated by complex hybridisation patterns (Sparvoli *et al.,* 1994). In *Arabidopsis*, *PAL1* and *PAL2* genes encode the principal PAL enzymes of phenylpropanoid metabolism (Raes *et al.,* 2003).

PAL, like tyrosine ammonia-lyase (TAL, EC:4.3.1.24), catalyzes the non-oxidative deamination of phenylalanine to *trans*-cinnamate and direct the carbon flow from the shikimate pathway, a primary metabolic pathway, to numerous branches of the general phenylpropanoid pathway, which is deemed a secondary metabolic pathway in plants. The unusual deamination reaction is independent of any co-factor while the required electrophilic prosthetic group, 4-methylidene-imidazole-5-one, is auto-catalytically formed.

Contig 3155 over-expressed in the red phenotype with a logbase2 ratio of 4.176 show sequence similarity to an UDP-galactose/UDP-glucose transporter. The role of this transporter in the supply of galactose or glucose for conjugation to the anthocayanidin core in the reaction catalysed by UFGT, could provide further insight into the production of anthocyanin in the red leaf phenotype of 'Bon Rouge'.

Other contigs over-expressed in the red phenotype with similarity to genes coding for enzymes with EC numbers in the KEGG database pathways, include contig 261, a 3-ketoacyl- thiolase peroxisomal-like enzyme EC:2.3.1.16 (logbase2 ratio 5.045), contig 6124, a beta-galactosidase 3-like EC: 3.2.1.23 (logbase2 ratio 3.550), contig 6678, a plasma membrane H^+ATPase EC:3.6.3.6 (logbase2 ratio 3.265), contig 482, succinate dehydrogenase EC:1.3.5.1 (logbase2 ratio 5.176), contig 701, 1-aminocyclopropane-1-carboxylate oxidase EC:1.14.17.4 and EC:1.14.11.0 (logbase2 ratio 3.558), contig 1938,

berberine bridge enzyme EC:1.21.3.3 and MurB reductase EC:1.1.1.158 (logbase2 ratio 4.630), contig 3957, omega-3 fatty acid desaturase EC:1.14.99.33 and EC:1.14.19.0 (logbase2 ratio 3.456), contig 2597, an epoxide hydrolase 2-like, EC:2.7.10.2 (logbase2 ratio not applicable, present only in the red leaf cDNA library), contig 699, an ATP phosphoribosyltransferase-like EC:2.4.2.17 (logbase2 ratio not applicable, present only in red the leaf cDNA library).

Contig 261 over-expressed in the red phenotype with a logbase2 ratio of 5.045 show sequence similarity to a C-acyl transferase (EC:2.3.1.16) that is involved with numerous metabolic pathways listed in the KEGG database. C-acyl transferase catalyse reaction for biosynthesis of unsaturated fatty acids, fatty acid elongation, Ethylbenzene degradation, Benzoate degradation, alpha-Linolenic acid metabolism, Geraniol degradation, Valine, leucine and isoleucine degradation,

Additional contigs only, or over-expressed in the cDNA library prepared from the red phenotype include contig 596, a subtilisin-like protease-like EC:3.4.21.0 (logbase2 ratio not applicable, present only in the red leaf cDNA library); contig 80, a subtilisin-like protease-like EC:3.4.21.0 (logbase2 ratio 3.295); contig 7450, a alpha-xylosidase 1-like EC:3.2.1.0 (logbase2 ratio 3.305); contig 3655, a 26S protease regulatory subunit 8 homolog a-like EC:3.6.4.3 (logbase2 ratio not applicable, present only in the red leaf cDNA library); and contig 1121, a cell division cycle protein 48 homolog EC:3.6.4.3 (logbase2 ratio not applicable, present only in the red leaf cDNA library).

5.6.10 Mapping of Anthocyanin biosynthesis genes over-expressed in the red phenotype

To measure anthocyanin gene expression differences between the two phenotypes, a subset of the reads from each phenotype was mapped to a number of *Pyrus communis* anthocyanin biosynthesis genes available in public databases (see table 5.5). The following genes for anthocyanin production could be mapped for both phenotypes and thus provided a measure of expression difference for these genes.

Phenylalanine ammonia-lyase (PAL)

Phenylalanine ammonia-lyase (EC:4.3.1.25) catalyses the first committed step in phenylpropanoid biosynthesis in plants. It is responsible for the deamination of phenylalanine to *trans*-cinnamic acid and ammonia. Phenylalanine ammonia-lyase (PAL) is a member of a large gene family in *Arabidopsis* and grape as demonstrated by complex hybridisation patterns (Sparvoli *et al.*, 1994). In *Arabidopsis*, *PAL1* and *PAL2* genes encode the principal PAL enzymes of phenylpropanoid metabolism (Raes *et al.,* 2003). PAL, like tyrosine ammonia-lyase (TAL), catalyzes the non-oxidative deamination of phenylalanine to trans-cinnamate and direct the carbon flow from the shikimate pathway, a primary metabolic pathway, to numerous branches of the general phenypropanoid pathway, which is deemed a secondary metabolic pathway in plants. The unusual deamination reaction is independent of any co-factor while the required electrophilic prosthetic group, 4-methylidene-imidazole-5-one, is auto-catalytically formed.

Several copies of the *PAL* genes are found in all plant species, comprising four in *Arabidopsis*, to five in poplar and nine in rice (Hamberger *et al.,* 2007). The individual genes may respond differentially to biotic and abiotic stressors and their expression is developmentally regulated and spatially controlled in a tissue specific manner (Bhuiyan *et al.,* 2009; Lillo *et al.,* 2008). In contrast to *Arabidopsis* and poplar, the gene for catalysis of the first step for phenylpropanoid biosynthesis in the Solanaceae is represented by a remarkable set of an estimated 20 putative *PAL* genes, as demonstrated for *Lycopersicon esculentum* (Chang *et al.,* 2008). However, only a single gene appears to be strongly expressed in all tissues, while the rest appear to be effectively silenced.

Chalcone synthase (CHS)

Twelve different CHS genes have been identified in petunia, although only some are currently known to be expressed (Holton and Cornish, 1995). Chalcone synthase catalyses the first committed step for flavonoid, and subsequently, anthocyanin production in plants. The enzyme is responsible for the sequential condensation of three malonyl-CoA molecules with *p*-coumaroyl-CoA to yield 4,2',4',6'-tetrahydroxychalcone or naringenin (Holton and Cornisch, 1995). The products of the CHS reaction, namely flavonoids, are yellow coloured chalcones that usually do not accumulate to a significant level in planta. This is due to the fact that in most plants species, chalcones are not the final products of the CHS reaction since the pathway proceeds with several enzymatic steps to produce numerous other classes of flavonoids, such as flavanones, dihydroflavonols and eventually, anthocyanins. The latter constitute the major water-soluble pigments in flowers and fruits.

Chalcone Isomerase (CH1)

Chalcone isomerase catalyses the production of naringenin from naringenin chalcone. As is the case for the yellow coloured chalcones, most plants do not accumulate colourless chalcones. After its formation, naringenin chalcone is rapidly isomerized by the enzyme chalcone isomerase (CHI) to form the flavanone naringenin. Even in the absence of CHI, naringenin chalcone spontaneously isomerises to form naringenin (Holton and Cornisch, 1995). To date two forms of CHI have been identified: one form can isomerise 6'-hydroxyl- as well as 6'-deoxy chalcones, while the other converts only 6'-hydroxychalcones to flavanones.

Flavonoid 3-hydroxylase (F3H)

Flavanone 3-hydroxylases, also known as Naringenin 3-dioxygenase catalyses the conversion of naringenin to dihydrokaempferol in the following reaction:

Naringenin + 2-Oxoglutarate + Oxygen <=> Dihydrokaempferol + Succinate + CO_2

Hydroxylation in position C-3 of flavanones to dihydroflavonols has been demonstrated for a wide variety of plant species such as *Arabidopsis*, petunia, snapdragon, apple, pear and maize. The flavanone-3-hydroxylase (F3H) is a member of the 2-oxoglutarate-dependent dioxygenase family that is highly conserved among widely divergent plant species as shown by sequence comparison (Britsch *et al.,* 1992).

Dihydroflavonol reducatase (DFR)

DFR, also known as dihydrokaempferol 4-reductase, acts in the reverse direction, on (+)-dihydroquercetin and (+)-dihydromyricetin. Each dihydroflavonol is reduced to the corresponding cis-flavan-3,4-diol in a stereospecific manner with NAD^+ as the preferred reducing agent instead of $NADP^+$. However, the latter reaction proceeds more slowly. The enzyme catalyses the following reaction:

Dihydrokaempferol + NADPH + H^+ <=> cis-3,4-Leucoanthocyanidin + NADP+

The intron-exon structure of these anthocyanin biosynthesis genes is generally conserved across species. This is exemplified by the dihydroflavonol reductase (DFR) gene of bean which contains three introns in the same positions as the maize DFR gene that corresponds to the first three of five introns present in the gene from petunia and snapdragon (Kristiansen and Rohde, 1991). The 3-*O*-glucosyl transferase (3-GT) gene contains a single intron whose position appears to be conserved in *Arabidopsis thaliana* and barley (*Hordeum vulgare*). The gene is the least conserved of all the anthocyanin structural genes (Sparvoli *et al.,* 1994) but contains a signature motif common to all glycosyl tranferases including 5-*O*-glucosyltransferase (5-GT) (Yamazaki *et al.,* 1999).

Anthocyanidin synthase (ANS)

ANS, also known as Leucoanthocyanidin dioxygenase (LDOX), catalyses the transformation of colourless leucoanthocyanidins to coloured anthocyanidins in the following reaction:

cis-3,4-Leucoanthocyanidin + 2-Oxoglutarate + Oxygen <=> Anthocyanidin + Succinate + CO_2 + 2 H_2O

This is proposed to involve two steps: in the first step, ANS removes an hydroxyl group under acidic conditions from the basic ring structure; a dehydratase is involved in the second step which results in the formation of a double bond between C-3 and C-4 on the flavonoid ring (Heller and Forkmann, 1988; Boss *et al.,* 1996b). However, *in vitro* studies have shown that the hydration-dehydration step can be catalysed by acid without requirement for other enzymes (Saito *et al.,* 1999). ANS has similarity with 2-oxoglutarate-dependent dioxygenases including flavone-3-hydroxylase (F-3-H), flavonol synthase (FLS), and amino-cyclopropane-1-carboxylate (ACC) oxidase (Holton and Cornish, 1995).

UDP-Glucose flavonoid: 3-*O*-glucosyltransferases (3-UFGT)

UDP-Glucose flavonoid: 3-*O*-glucosyltransferases (3-GT) plays and important role in metabolite accumulation. The latter enzyme specifically glucosylates anthocyanidins, rather than flavonoids as is suggested by its name, at the 3-*O* position during red fruit ripening to produce the first stable red pigment (Ford *et al.,* 1998). The enzyme UDP glucose:flavonoid 3-*O*-glucosyltransferase (3UFGT) is responsible for the transfer of the glucose moiety from UDP-glucose to the hydroxyl group in position 3 of the C ring. This essential reaction stabilises anthocyanidins for accumulation as water-soluble pigments in the vacuoles. As such, 3UFGT is regarded as an indispensable enzyme of the main

biosynthetic pathway for anthocyanin biosynthesis and accumulation. Glycosylation of anthocyanidins increases their solubility in aqueous solutions and reduces chemical reactivity under physiological conditions. Glycosylated compounds are transportable storage compounds that are preferentially sequestered in the vacuole, or waste products that can be removed from the cytosol (Ford *et al.,* 1998). It has been proposed that the full activity of ANS and UFGT requires a multi-enzyme complex (Saito *et al.,* 1999). UF:3-GT is encoded by *Bronze*-1 (*Bz*1 allele in maize) (Furte*k et al.,* 1988). Sequence polymorphisms among three *Bz*1 alleles include deletions/additions, a transposable element insertion upstream of the promoter region and single base pair substitution. Mutable *Bz*1 alleles with transposable elements belonging to the *Ac*/*Ds*, *Spm*/*dSpm*, *Mu*, *Cy*/*rcy* and *Mut* families have been described by Furtek et al. (1988) while Callis *et al*. (1987) have demonstrated an enhancement of *Bz*1 gene expression by the *Bz*1 intron.

CHAPTER 6

CONCLUSION

To gain an understanding of the molecular mechanism underlying the colour variation in red and green leaf phenotypes of 'Bon Rouge', differential gene expression analysis was perform with Differential Display and mRNAseq. Both protocols facilitate the investigation of gene expression for organisms without a publicly available genome sequence. Additionally, the colour pigment differences between the red and green fruit skin phenotypes were investigated by HPLC and ESI-MS-MS.

6.1 Pigment characterisation and identification

Anthocyanin pigment concentration was measured in 0.1% HCl methanol extracts for both phenotypes in a Nanodrop spectrophotometer™ and by comparison with the commercial standard, idaein chloride. Red leaf extracts contain 50.0 mg/L cyanidin 3-galactoside compared to 1.0 mg/L for green leaf extracts (table 3.1). HPLC analyses at 530 nm confirmed the presence of a single pigment peak at 10 minutes (fig. 3.2.2.1) in the red leaf extract corresponding to cyanidin 3-galactoside when compared to the commercial standard, idaein (fig. 3.2.1.1). A single peak at 25 minutes was visible in the HPLC spectrum of reverted (green) leaf extract collected at 530 nm (fig. 3.2.3.1). The presence of a pigment peak in the visible range (530 nm) and the low concentration of

pigment measured in the Nanodrop™ is consistent with the production of some colour pigment in reverted leaves but at a level significantly lower than in 'Bon Rouge' leaves. The pigment in green leaf extracts has not been fully characterised.

To confirm the presence of the major pigment, cyanidin 3-galactoside, I analysed the peak collected during HPLC of the red leaf extract and the standard, idaein chloride by liquid chromatography mass spectrometry (LC-MS) and tandem mass spectrometry (MS-MS). LC-MS and MS-MS spectra were collected in negative ion mode and the spectrum for red leaf extract displayed three major ions (fig. 3.3.1 bottom). In addition to the expected 449 *m/z* ion representing cyanidin 3-galactoside, two additional ions at 448 *m/z* and 447 *m/z* respectively, were visible in the spectrum. These two ions most likely result from the reduction around the C2-C3 bond of the C-ring of the anthocyanidin core to epicatechin by anthocyanidin reductases. Reduction of the cyanidin core to epicatechin has been reported for a number of species, including apple (Pfeiffer *et al.,* 2006). Unexpectedly, the commercial standard displayed an ion at 447 *m/z* instead of the expected 449 *m/z* (fig. 3.3.1 top). The 447 *m/z* ion could result from the loss of protons from the cyanidin core. ESI-MS-MS analyses of the 449 *m/z* ion and the 447 *m/z* ion identified a cyanidin core for both ions. The cyanidin core of cyanidin 3-galactoside is represented by the 287 *m/z* ion that result from the loss of the sugar moiety (449 *m/z* – 162 *m/z*). Similarly the flavanol core of the reduced cyanidin 3-galactoside is represented by the 285 *m/z* ion that result from the loss of the sugar moiety (447 *m/z* – 162 *m/z*). The presence of a major ion, 284 *m/z* in the mass spectrum in addition to the expected 285 *m/z* most likely results from the loss of an additional H^+ from the reduced cyanidin core.

NMR Structure determination of Idaein chloride (Cyanidin 3-galactoside)

Idaein chloride was dissolved in CD_3OD and spectra were colllected in a Varian ^{Unity}*Inova* 600 NMR spectrometer with a ^1H frequency of 600 MHz and a ^{13}C frequency of 150 MHz. The ^1H spectrum was referenced to the CD_2*H*OD peak at 3.31 ppm and the ^{13}C spectrum to the *C*D_3OD peak at 49.5 ppm. In addition to the 1-dimensional spectra, 2-dimensional ^1H-^1H COSY (fig. 3.5.1.3), ghsqc (fig. 3.5.1.5) and ghmqc (fig. 3.5.1.4) spectra were also carried out to assist in the assignments of the ^1H (fig. 3.5.1.1) and ^{13}C (fig. 3.5.1.2) spectra. The obtained ^1H NMR spectrum was very useful in confirming the presence of the carbohydrate moiety besides the aromatic part in this compound, and to rule out other glucosides such as flavonoids. The identification of glucose and the determination of its substitution pattern and anomeric configuration could also be obtained from the ^{13}C NMR spectrum.

The downfield region of the ^1H NMR spectrum showed a singlet at δ 8.85 (H-4), a 3H AMX system at δ 8.09 (δδ, 8.7 Hz, 2.3 Hz, H-6'), 7.86 (δ, 2.3 Hz, H-2') and 6.88 (δ, 8.7 Hz, H-5'), and an AB system at δ 6.76 (δ, 1.8 Hz, H-8) and 6.61 (δ, 1.8 Hz, H-6), which is in accordance with a cyanidin derivative. The sugar region showed the presence of only one sugar unit. The anomeric coupling constant (7.7 Hz) and the six ^{13}C resonances in the sugar region of the ^{13}C NMR spectrum were in accordance with β-glucopyranose. The numbering of the cyanidin backbone follows convention, as does that of the galactoside ring. No O*H* protons were visible in the ^1H NMR spectrum due to the exchange with the alcohol deuterium of CD_3O*D*. Based on the chemical shifts, multiplicities and coupling constants almost all non-exchangable protons for the cyanidin

backbone in the idaein chloride sample could be unambiguously assigned from the ^{1}H NMR spectrum. In general the assignments of the cyandin backbone agree well with those for similar compounds found in the literature.

NMR Structure determination of a minor pigment, RZ, from 'Bon Rouge' extract

The sample was dissolved in CD_3OD and spectra collected on a Varian ^{Unity}*Inova* 600 NMR spectrometer with a ^{1}H frequency of 600 MHz and a ^{13}C frequency of 150 MHz. The ^{1}H spectrum was referenced to the CD_2*H*OD peak at 3.31 ppm and the ^{13}C spectrum to the *C*D_3OD peak at 49.0 ppm. In addition to the 1-dimensional spectra, 2-dimensional ^{1}H-^{1}H COSY, ghsqc and ghmqc spectra were also generated to assist in the assignments of the ^{1}H and ^{13}C spectra. Because the fragment peak appeared in the visible range of the spectrum it was expected that this compound was a cyanidin derivative. However, the notable absence of certain key peaks of the cyanidin backbone indicated that this was not the case. Analysis of the ^{1}H (table and fig. 3.5.2.1), ^{13}C (table and fig. 3.5.2.2) and 2-dimensional spectra revealed the structure of the isolated compound to be that of 2,3-dihydroxycyclopentyl (2*E*)-3-(3,4-dihydroxyphenyl) acrylate (numbering is arbitrary). This compound is a known natural product which antioxidant properties and has been previously isolated from tart cherries (Wang *et al.,* 1999). The most obvious features in the ^{1}H NMR spectrum were the phenyl ring and the trans double bond (with a large $^{3}J_{H-H}$ coupling constant of 15.9 Hz). The ^{1}H NMR spectrum of RZ revealed a 3,4-dihydroxyphenyl moiety. However, the two multiplets in RZ, appearing at δ 2.15 and 1.95, were assigned to two methylene groups, respectively. The COSY experiment showed that two CH2 protons in RZ were correlated and adjacent to each other and also

coupled to other hydrogens. The ^{13}C NMR spectrum of this compound revealed that there were only one carbonyl carbon, eight methine carbons, and two methylene carbons. Three of the methine carbons at δ 74.8, 73.0, and 68.3 were oxygenated and showed correlations to three methine protons at δ 3.64, 5.35, and 4.14, respectively, as evident from the hmqc spectrum. Also, five other methine carbons at δ 115.1, 116.5, 122.9, 146.8, and 115.8 showed correlations to three aromatic protons appearing at δ 7.04, 6.76, and 6.93 and two olefinic protons at δ 7.58 and 6.30 ppm, respectively. Based on this data, compound RZ was assigned as 2,3-dihydroxycyclopentyl (2*E*)-3-(3,4-dihydroxyphenyl) acrylate. All the peaks assigned correspond well with those in the literature (Wang *et al.,* 1999), although the hydrogen bonded OH was not identified and its existence is most likely pH dependent.

6.2 Differential gene expression analysis by Differential Display

To investigate the molecular mechanism underlying fruit skin colour development in the two phenotypic variants of 'Bon Rouge', differential gene expression was analysed by Differential Display and mRNAseq. Diffential Display allows analysis of a subset of the expressed genes between the two phenotypic variants while mRNAseq facilitates whole transcriptome analysis at expression levels several orders of magnitude greater than that determined by Differential Display. Both protocols facilitate the investigation of gene expression in organisms for which a genome sequence is not publicly available and mRNAseq has become the gold standard for such investigations. A *Pyrus* genome

(Chagné *et al*., 2014) and genetic linkage map (Li *et al*., 2017) has recently been published.

Partial cDNAs isolated by the differential display technique that was exclusively expressed in either the red or the green phenotype of 'Bon Rouge' were further characterised by BLAST. A number of these could be annotated by sequence similarity searches using the BLAST tool at the NCBI's non-redundant nucleotide database. Partial cDNAs annotated in this manner, and expressed only in the red phenotype include a PAPS reductase-like enzyme (GI 56431066), a *Pyrus pyrifolia* mRNA, complete cds, accession number HM044853 that has been characterised as a defensin (DFN1), a *Pyrus communis* PcExp3 (ABO93030) and *Pyrus communis* PcExp2 (AJ11692) mRNA for expansin, a *Fragaria vesca ssp vesca* Early Light Inducible Proteins (ELIP1) accession number XM_004296533, a 2-oxoglutarate/malate translocator (OMT) (GI number 49631985), an mRNA for (PGIP)-like protein from *Fragaria x ananassa* (AAP33475), a polygalacturonase-inhibiting protein (PGIP), a vacuolar membrane H^+-ATPase subunit c" from citrus (*Citrus limon*) (GI number 71923687) and a glutaminyl tRNA synthetase (QRS) were up-regulated in the red phenotype, while cDNAs with homology to a chloroplast biosynthetic enzyme, *Citrus sinensis* thi1, a mitochondrial carrier protein, a starch phosphorylase H and sulfolipid synthase (SQD2) were up-regulated in the green form (see table 4.3).

Annotation by BLAST of partial genes upregulated in the green 'Bon Rouge' phenotype

Sulfolipid synthase (SQD2) is responsible for catalysis of the second step in the production of sulfoquinovosyl diacylglycerol (SQDG) (Yu *et al.,* 2002), an anionic nonphosphorous glycolipid at physiological pH found in the photosynthetic membranes of seed plants (Mulichak *et al.,* 1999) and photosynthetic bacteria (Güler *et al.,* 1996). Plants that carry mutations in SQD2 show limited growth under conditions of phosphate starvation (Yu *et al.,* 2002) that suggest a compensatory role for sulfoquinovosyl diacylglycerol (SQDG) as components of photosynthetic membranes when glycolipid synthesis is impaired due to phosphate limitation during growth. Phosphate starvation induces an increase in the production of anthocyanin via a cytokinin-mediated action. The production of anthocyanin in vegetative tissue of plants is characteristic of the phosphate starvation response (Winkel-Shirley, 2002) but the molecular mechanism by which this anthocyanin increase is mediated is still not clear. Decreased expression of SQD2 in the red 'Bon Rouge' phenotype may be responsible for the increased production of anthocyanin, and increased sulpholipid synthesis would be consistent with decreased anthocyanin synthesis in the green 'Bon Rouge' phenotype. Two partial cDNAs upregulated in the green phenotype showed similarity to the chloroplast thiazole biosynthesis precursor protein gene, *Citrus sinensis thi1* encoding the enzyme involved in the biosynthesis of the thiazole ring of thiamine pyrophosphate (Vitamin B_1) in plants. Its over-expression in the green phenotype may suggest a more efficient pathway for Vitamin B_1 production. The *Vicia faba* cytosolic H isoform of starch phosphorylase (EC

2.4.1.1) also referred to as α-glucan phosphorylase, is a key enzyme in glucan catabolism in animals, prokarytes and fungi (Newgard *et al.,* 1989). It is considered to be a critical enzyme for starch degradation in plants where it catalyses the reversible phosphorolysis of linear α-1,4-glycosidic linkages in glucan substrates to yield glucose-1-phosphate (Smith *et al.,* 2005). The *pal1* and *pal2* mutants of *Arabidopsis* are T-DNA insertion mutants defective in phenylalanine ammonia-lyse (PAL; EC 4.3.1.25) expression that display differential regulation of starch phosphorylase H (Rohde *et al.,* 2004). PAL catalyses the first reaction in the phenylpropanoid pathway that produces numerous flavonoids, including anthocyanins. The *pal1* and *pal2* mutants display no visible phenotypic alteration but have far-reaching effects on phenylpropanoid, amino acid and carbohydrate metabolism (Rohde *et al.,* 2004).

Annotation by BLAST of genes upregulated in the red 'Bon Rouge' phenotype

Partial cDNA fragment M3 showed similarity to a sequence encoding a PAPS reductase-like enzyme (GI 56431066). PAPS reductases function in the sulphate assimilation pathway in plants and microorganisms (Asahi, 1960) that produce reduced sulphur from inorganic sulphate for the synthesis of sulphur-containing amino acids, methionine and cysteine (reviewed in Wilson, 1962), coenzymes and iron-sulphur clusters of enzymes (Schmidt and Jäger, 1992), and a variety of S-containing secondary metabolites such as glucosinolates that play an important role in plant defences against herbivores and pathogens (Kopriva, 2006; Kopriva *et al.,* 2007). A *Pyrus pyrifolia* mRNA, complete cds, accession number HM044853 has been characterised as a defensin (DFN1) and was

over-expressed in the red phenotype. Defensins are low molecular weight antimicrobial, cysteine-rich, proteins found in plants and animals that inhibit the growth of fungi, oomycetes and gram-positive bacteria (Broekaert *et al.,* 1995). In *Arabidopsis,* antifungal proteins, PDF1. 1 and PDF 1. 2 expression require both jasmonic acid and ethylene (Penninckx *et al.,* 1996). Expression of defensin is associated with biotic stress responses and increased expression of this cDNA in the red phenotype may suggest an enhanced response to pathogen attack. Increased expression of anthocyanin during biotic stress responses is well documented but the molecular mechanisms for many of these responses remains to be detailed (Dixon and Paiva, 1995 and references therein). Two partial cDNAs represented by M8 and M9 displayed weak but equal homology to *Pyrus communis* PcExp3 (ABO93030) and *Pyrus communis* PcExp2 (AJ11692) mRNA for expansin, respectively, as determine by BLAST in the nucleotide database (nr) at NCBI. Expansins are extracellular proteins that function in cell wall loosening during cellular growth but exhibit no detectable hydrolase or transglycosylase activity (McQueen-Mason and Cosgrove, 1995). Partial cDNA M11 display sequence similarity to a *Fragaria vesca ssp vesca* Early Light Inducible Proteins (ELIP1) accession number XM_004296533. ELIPs (ELIP1, ELIP2, ELIP3 and ELIP4) are nuclear encoded, stress-regulated chloroplast proteins that are related to the chlorophyll *a*/*b*-binding (Cab) proteins (Heddad *et al.,* 2006), stress enhanced proteins (SEPs) and the one-helix proteins (OHPs). The transcription factor for ELIP1 is the constitutively expressed LONG HYPOCOTYL5 (HY5) (Harari-Steinberg *et al.,* 2001). HY5 levels are regulated during photomorphogenesis, or development in light, and skotomorphogenesis (development in the dark), by interaction with the CONSTITUTIVELY PHOTOMORPHOGENIC1

(COP1) (Hardtke *et al.,* 2000), an E3-ubiquitin ligase that targets HY5 for degradation during the dark diurnal phase. HY5 has also been shown to be the transcriptional activator for chalcone synthase (CHS), the first committed enzyme in the anthocyanin biosynthetic pathway, and a number of downstream enzymes required for anthocyanin synthesis (Vandenbussche *et al.,* 2007). For chalcone synthase under the control of the transcription factor HY5, this could result in enhanced gene expression and protein levels with subsequent increase in anthocyanin production. Expression of ELIP under the control of constitutively expressed HY5, which is also the transcription factor for CHS, may coincidently result in the production of anthocyanin during high light stress. Sequestration in an acidic vacuole would then be required to stabilise anthocyanin pigments and the upregulation of the VHAc" subunit in the red phenotype may play a role in anthocyanin transport into the vacuole for pigment stabilisation. The partial cDNA sequence for M14 shared significant sequence similarity with a 2-oxoglutarate/malate translocator (OMT) (GI number 49631985), a member of the plastidic dicarboxylic acid transporter (DCT) family. Tobacco DiT1, a 2-oxoglutarate/malate translocator, antisense-repressed in intact transgenic plants, causes reduced transport capacity for 2-oxoglutarate across plastid envelope membranes. This reduction in 2-oxoglutarate transport resulted in impaired allocation of carbon precursors for amino acid synthesis, accumulation of organic acids, and a significant decrease in protein content, photosynthetic capacity and sugar pools in leaves. Decreased photosynthetic capacity due to photoinhibition as a consequence of high light stress is associated with anthocyanin production in leaves and fruit skin of blush pear cultivars such as Rosemarie and Forelle (Steyn *et al.,* 2004). Partial M16 cDNA matched a *Malus*

EST, DR997344 that displays sequence similarity with the mRNA for (PGIP)-like protein from *Fragaria x ananassa* (AAP33475). Polygalacturonase-inhibiting proteins (PGIPs) inhibit endopolygalacturonases (PGs) (EC 3.2.1.15) produced by fungal pathogens during plant infection. PGIPs are regulated during development, and biotic and abiotic stress responses such as wounding and pathogen infection or treatment with elicitors like salicylate and cold produce PGIP synthesis (Ferrari *et al.,* 2003; Li *et al.,* 2003). In *Vitis* ('Riesling' and 'Gloire de Montpellier') PGIP production and genes encoding anthocyanin biosynthetic enzymes, including phenylalanine ammonia-lyase, chalcone synthase and chalcone isomerase, are differentially regulated in response to fungal pathogens *Plasmopara viticola* and *Pseuperonospora cubensis* (Kortekamp, 2006). The vacuolar membrane H^+-ATPase subunit c" from citrus (*Citrus limon*) (GI number 71923687). The vacuolar-type H^+-transporting adenosine triphosphatase (V-ATPase; EC 3.6.1.34) is a membrane-bound, primary active transport protein active in the vacuolar tonoplast and various other sites in the plant endomembrane system (Sze, 1985). In plant vacuoles, V-ATPase are responsible for energising ion and metabolite transport (Ratajczak, 2000), acidification of the intracellular vacuolar compartment (Forgac, 1998) and is essential for the maintenance of ion homeostasis, secondary activated transport and adaptation to environmental stress (Seidel *et al.,* 2004). In mature plant cells, the vacuole is the largest intracellular compartment occupying approximately 90% of the cell volume and functions as a store for salts, metabolites, plant pigments such as anthocyanins, sugars, organic acids, including malate, and numerous other solutes (Sze, 1992) under various environmental conditions. In morning glory (*Ipomoea tricolor*) petals increased vacuolar pH has been associated with petunidin (an anthocyanidin) colour change from

red to blue (Yoshida *et al*., 2005) and in *Petunia hybrida* (Quattrocchio *et al*., 2006) an acidic vacuole has been implicated in red pigment colour stabilisation and intensity for cyanidins (Spelt *et al*., 2002).

6.3 Differential gene expression analysis by transcriptome analysis of red and green phenotypes by mRNAseq (whole transcriptome analysis)

To identify differentially expressed genes between the red and green phenotypes of 'Bon Rouge' that are associated with anthocyanin production, I measured differential gene expression with the mRNAseq protocol on the Illumina GAII platform. Reads were assembled with Velvet (Zerbino *et al*., 2008) and CLC Bio Genomic Workbench. The representative genes associated with the assembled contigs were annotated by similarity searches with the BLAST tool at the NCBI's non-redundant nucleotide database (Altschul *et al*., 1990).

cDNA libraries for both phenotypes were constructed using Illumina's RNAseq adapted protocol. Seventeen million single and paired end (SE and PE) reads were generated in six lanes of an Illumina mRNAseq flow cell during three runs on the Illumina GA11 platform. A set of reads for the libraries from both red and green phenotypes of 'Bon Rouge' was always generated in the same run. Reads were assembled with Velvet (Zerbino *et al*., 2008) and CLC Bio Genomic Workbench. For the Velvet assembly a range of parameters such as k-mer length, coverage cut off and expected coverage were tested for optimisation of the assembly. For assembly with CLC Bio Genomic Workbench, default parameters were used at two different settings of minimum contig

length namely 100bp and 200 bp. The assembled contigs were used as a matrix to map subsets of reads from each phenotype to identify contigs that were significantly over-expressed in the red compared to the green phenotype.

To calculate the RPKM ratio (Red phenotype:Green phenotype) for each contig, the RNAseq analysis protocol in CLC Bio's Genomic Work Bench was used to map a subset of the reads against the contigs assembled with CLC Bio Genomic Workbench with a minimum contig lenth setting of 200 bp. In order to annotate the contig-representing transcripts that were only expressed in either the red or green phenotype according to the RPKM ratio, similarity searches were performed with BLAST searches against the non-redundant nucleotide database, while the top 105 contigs over-expressed in the cDNA library prepared from the red phenotype of 'Bon Rouge' was analysed with Blast2GO. The RPKM ratio or relative intensity (RI) (Red phenotype:Green phenotype) for each contig was log transformed to generate logbase10 and logbase2 ratios that provide a more accurate expression difference between the two phenotypes.

6.3.1 Mapping mRNAseq reads to anthocyanin biosynthesis genes

To measure gene expression in RPKM for anthocyanin biosynthesis genes in either or both red and green phenotypes, a subset of reads generated from either red or green phenotypes were mapped to genes that encode *Pyrus communis* anthocyanin biosynthesis genes. These include PAL, CHS, CHI, F3H, and ANS, as well as UDP-glucose: flavonoid 7-O-glucosyltransferase (F7GT) that were isolated and characterised from 'Conference' and 'Pyrodwarf' pears (Fischer *et al.,* 2007).

Anthocyanin-related genes include the genes that encode the enzymes at the beginning of the general phenylpropanoid biosynthetic pathway (*PAL*), at the first reaction specifically leading to anthocyanins (*CHS*), and the consecutive steps (*CHI* and *F3H*), in the last branch point that leads to both colored anthocyanins and colorless flavonols (catechins and proanthocyanidins) (*DFR*), and at the consecutive steps involving anthocyanidin biosynthesis and glycosylation (*ANS* and *UFGT*, respectively). In this mapping strategy, anthocyanin accumulation was positively correlated with the expression of three anthocyanin biosynthetic genes namely *PcDFR*, *PcFHT* and *PcANS*, during anthocyanin accumulation in the red pear phenotype. All the structural genes for anthocyanin biosynthesis in pear and currently listed in the NCBI database, could be mapped in both phenotypes. The highest expression difference was seen for Dihydroflavonol reductase (DFR, DQ227732). Other genes that displayed significant difference expression ratios in the red compared to the green phenotype include a second (DQ227730) and a third DFR gene (AY227731), Anthocyanidin synthase (ANS, DQ230994) and Flavanone 3-hydroxylase (FHT, AY965342). However, for the second and third DFR genes, few reads could be uniquely mapped to the sequence in the data set. Genes with expression ratios below 3.0 RPKM include PAL (DQ230992), CHS (AY786998), CHI (EF446163), FHT (AY965342), UF3GT (GQ325589) and FLS (DQ230993). None of the sequenced reads obtained from the green phenotype could be mapped to MYB10 (EU153575) but for the red phenotype, 80.0 RPKM mapped to this gene. ANS catalyses the critical reaction for the production of the red colour characteristic of cyanidins. This reaction is dependent on the previous reaction catalysed by DFR that was shown to be significantly up-regulated in the red phenotype suggesting successful competition for red anthocyanin

production over colourless flavonol biosynthesis. UF3GT stabilises the cyanindin by addition of a sugar molecule that facilitates transport into the vacuole for red colour preservation.

6.3.2 Annotation of mRNAseq derived contigs by BLAST

A relatively small number of the contigs were exclusively expressed in either the red or the green phenotype of 'Bon Rouge'. In both instances, a number of these could be identified by sequence similarity searches using the BLAST tool at the NCBI's non-redundant nucleotide database. A set of twenty one contigs were expressed only in the red when compared to those expressed in the green phenotype and sequence similarity searches by BLAST identified the representative genes corresponding to these transcripts as listed in table 5.4.1. These include a Valosin/p97 AAA^{+} ATPase (VCP), also known as yeast homolog CDC48, a clone similar to LIL3 and ELIP, a rhamnose synthase, an epoxide hydrolase 3 (EH3), RADIATION SENSITIVE23 (RAD23), Rho GTPases of plant (ROP1), an arabinosidase, Apetala 2 (AP2) transcription factor, a lipocalin, Regulatory Particle 6A (RPT6A), a Glutathione Transferase (GST), an auxin binding protein (ABP), a Zinc Finger Protein-like (ZFP-like) and succinate dehydrogenase. The rest of a total of 3627 contigs that were over expressed in the red compared to the green phenotype ranged in expression values for RPKM RED:GREEN ratios of 130 for a *Malus domestica* glutathione transferase followed by a ratio of 44 for an auxin binding protein and 41 for a Zinc finger protein for the most highly expressed transcripts in the red compared to the green phenotype. A subset of the expression ratios of the top 105

over-expressed contigs in the red phenotype library is listed in Appendix 3. A complete set of over- and under-expressed contigs is to be supplied in storage disk format as the complete list is over 400 pages in length.

The chaperone p97 or p97/VCP has been recognized as key player within the ubiquitin/proteasome system (Cayli *et al.,* 2009) and is a member of the family of ATPases associated with various cellular activities (AAA). It has been proposed that p97/VCP and the CSN play a global regulatory role in protein turnover, a proposal that lead to p97/VCP being termed a molecular 'gearbox' that regulates the ubiquitination status of substrates. The COP9 signalosome (CSN), an important mediator of light responses in plants, is also involved in the ubiquitin/proteasome system by controlling the neddylation of ubiquitin E3 ligases. A role for CDC48 in the turnover of proteins related to the hypersensitive response is supported by the finding that it interacts in vivo with the ADP Ribosylation factor (ARF) ARF1, that is directly involved in plant resistance in Nicotiana benthamiana and in rice. However the mechanism of the interaction of ARF1 with CDC48 is currently unknown (Rosnoblet *et al*., 2016). The interaction between CDC48, Sec 61 and TOM20 warrants further investigation. TOM20 forms part of the TOM (translocase of the outer membrane) complex that serves as the entry gate for almost all protein precursors destined for the mitochondrion (Schulz and Rehling, 2014), and like CDC48, is exclusively expressed in the red phenotype. Sec 61 is involved in endoplasmic reticulum associated degradation (ERAD), a process that requires CDC48 (Römisch, 2016), and is one of the top fifty transcripts over-expressed in the red phenotype.

Contig 1974 has sequence similarity to a clone with reasonable similarity to a Light inducible-like protein (Lil3) and low similarity to ELIP. A cDNA identified by differential display show sequence similarity to an ELIP and although the two sequences do not display similarity to each other, further analysis would provide insight into the relationship between these two expressed sequences.

MUCILAGE MODIFIED 4 (MUM4) encodes a rhamnose synthase that is required for L-Rhamnose synthesis, a component of plant cell wall pectic polysaccharides, various secondary metabolites including anthocyanins and flavonoids, some glycoproteins and for seed mucilage biosynthesis. UDP-L-rhamnose is one of the major components of the plant cell wall skeleton and rhamnose synthase plays a key role in the conversion of UDP-D-glucose into UDP-L-rhamnose (Wang *et al.,* 2008). The cellular phenotype of mum4 mutants is similar to that for several transcription factors (AP2 [APETALA2], TTG1 [TRANSPARENT TESTA GLABRA1], TTG2 MYB61, and GL2 [GLABRA2]), all of which are regulators of anthocyanin biosynthesis in plants. The GLABRA2 (GL2) gene in *Arabidopsis thaliana* encodes a transcription factor that is required for the proper differentiation of several epidermal cell type (Shi *et al.,* 2011). Expression studies suggest that MUM4 is developmentally regulated in the seed coat by AP2, TTG1, and GL2 (Western *et al.,* 2004). AP2, TTG1, TTG2, MYB61, and GL2 have all been shown to affect flavonoid biosynthesis in the seed coat and other mature plant organs.

Epoxide hydrolases (EH, EC 3.3.2.10), also known as epoxide hydratase (EC 3.3.2.3) are present in diverse organisms, such as plants, insects, fungi, bacteria, yeast and mammals. Rho family small GTPases are signaling switches controlling many eukaryotic cellular processes such as the control of cell polarity in eukaryotic cells, and the control of polar

growth in pollen tubes (Gu *et al*., 2006). Plant Rho GTPase (ROP1) displays a unique desiccation-associated ABA signaling transduction through which the ROP1 gene is regulated during the different stage of pollen maturation. The nuclear-enriched RADIATION SENSITIVE23 (RAD23) family proteins bind ubiquitin (Ub) conjugates and flag these for degradation by the proteasome. Transcriptional roles have been postulated for the ubiquitin receptor RAD23 suggesting they may be key components of proteasome transcriptional specificity (Wade *et al*., 2010). RAD23 and cell division cycle protein 48 (CDC48) are two key regulators of post-ubiquitylation events that act on distinct and overlapping sets of substrates (Baek *et al*., 2011). RAD23 is an adaptor protein that binds to both ubiquitylated substrates and to the proteasome. It serves as a link for ubiquitylated substrates to the proteasome for degradation by the proteasome. However, even though it is linked to the proteasome, RAD23 escapes degradation because it lacks an effective initiation region at which the proteasome can attack the protein and unfold it (Fishbain *et al*., 2011). The ubiquitin/26S proteasome system directs the turnover of misfolded and numerous regulatory proteins, thereby controlling many aspects of plant growth, development, and survival. RAD23 proteins appear to play an essential role in the cell cycle, and morphology and fertility of plants through their delivery of ubiquitin/26S proteasome system substrates to the 26S proteasome.

Alpha-N-arabinofuranosidase (EC 3.2.1.55) also known as arabinosidase, alpha arabinosidase and alpha-L-arabinosidase form part of the hydrolases and glycosylases glycosidases classes. These enzymes hydrolyse O- and S-glycosyl compounds and hydrolyse alpha-L-arabinofuranosides, alpha-L-arabinans containing (1,3)- and/or (1,5)-linkages, arabinoxylans and arabinogalactans. Some beta-

galactosidases (EC 3.2.1.23) and beta-D-fucosidases (EC 3.2.1.38) also hydrolyse alpha-L-arabinosides. Softening of fleshy fruits and cell walls during ripening is associated with the catabolism of cell wall components such as alpha-L-arabinofuranosides, alpha-L-arabinans. An alpha-L-Arabinofuranosidase (alpha-L-arafase) has been characterized in Japanese pear (*Pyrus pyrifolia*). The enzyme comprises a single 62-kD polypeptide as determined on SDS-PAGE (Tateishi *et al.,* 2005) and is related to a cDNA clone, PpARF2 whose transcript and related protein were detected only in the ripening fruit. The increase in alpha-L-arafase activity was closely associated with the increase in transcript and related protein in the ripening fruit. Transcripts of PpARF2 were not detected in buds, leaves, roots, or shoots of the Japanese pear. The deduced amino acid sequences of PpARF2 displayed minimal identity with those of other plants or bacteria.

In *Arabidopsis*, the Apetala 2/ ETHYLENE RESPONSE FACTOR1-like (AP2/ERF-like) genes are represented by a large transcription factor family encoded by 147 genes (Nakano *et al.,* 2006), several of which are upregulated by ethylene (Alonso and Stepanova, 2004). ERF1 is bound and activated by the transcription factor ETHYLENE INSENSITIVE3 (EIN3) which increases during apple fruit development and ripening (Solano *et al.,* 1998). This developmental process is associated with an increase in the expression of cell wall hydrolase genes such as POLYGALACTURONASE1 (PG1) whose increase in expression correlates with both an increase in ethylene, and following cold treatment.

Plant lipocalins can be classified into two groups; temperature-induced lipocalins (TILs) and chloroplastic lipocalins (CHLs). Plant temperature-induced lipocalins (TILs) have been shown to be responsive to heat stress (HS), but the nature of this response has not

yet been elucidated. It is postulated that AtTIL1 is an essential component for thermotolerance and most likely act against lipid peroxidation induced by acute heat stress. An *Arabidopsis*, chloroplastidic lipocalin (AtCHL) is involved in the protection of thylakoidal membrane lipids against reactive oxygen species, especially singlet oxygen, produced in excess light. The role of this gene in association with Lil3 and ELIP during high light stress may point to co-operation in the alleviation of this abiotic stressor in plants.

RPT6 is a component of the regulatory particle that form a heterohexamer made up of AAA^+ ATPases Rp1-6 (Wei *et al.,* 2008). The regulatory particle is part of the COP9 signalosome (CSN) that is reminiscent of the proteasome base complex that also contains a heterohexameric ring of the AAA^+ ATPases Rpt1-6. The COP9 signalosome (CSN) is an important mediator of light responses in plants and also involved in the ubiquitin/proteasome system by contolling the neddylation of ubiquitin E3 ligases. p97/VCP colocalizes and directly interacts with subunit 5 of the CSN (CSN5) *in vivo* and is associated with the entire CSN complex in an ATP-dependent manner. CSN is an ATP-dependent complex that resembles the 19 S proteasome regulatory particle and serves as a key mediator between ubiquitination and degradation pathways (Cayli *et al.,* 2009) in photomorphogenesis in plants.

The flavonoid transport enzyme Glutathione *S*-transferase (GSTs) is suggested to be the last genetically defined step in flavonoid biosynthesis (Walbot *et al.,* 2000). The GST enzymes have a central role in the transport of flavonoids through the cytoplasm to the vacuolar membrane (Walbot *et al.,* 2000, Licciardello *et al.,* 2014) where the acidic vacuole preserves the flavonoids and prevent their degradation for recycling into the

phenylpropanoid pathway. UV light induced GSTs play a role in the regulation and signaling to chalcone synthase, an enzyme of the flavonoid biosynthetic pathway, in cell cultures (Loyall *et al.,* 2000). GST mRNA transcripts were found to be abundant in immature fruit while the fruit are actively synthesizing flavonoids (Steyn *et al.,* 2004). Plant-specific phi class glutathione transferases (GSTFs) are often highly stress-inducible and expressed in a tissue-specific manner. AN9, a petunia glutathione S-transferase is required for efficient anthocyanin transport from its biosynthetic site in the cytoplasm to stabilised storage in the vacuole.

Auxin regulates plant cell division, elongation, and differentiation through signal transduction (Christian *et al.,* 2006). Some ABPs have been identified as proteins with enzymatic activities such as β-glucosidase (Campos *et al.,* 1992), 1,3-β-glucanase (Macdonald *et al.,* 1991), glutathione S-transferase (Bilang *et al.,* 1993, Zettl *et al.,* 1994) and glutathione dependent formaldehyde dehydrogenase (Sugaya and Sakai 1996).

A number of zinc finger proteins that act as transcription factors have been characterized in plants. ZPT2 (for petunia zinc-finger protein 2)-related proteins contain two Cys2/His2-type zinc-finger motifs and an ethylene-responsive element binding factor-associated amphiphilic repression motif. During abiotic stress many plant genes are down-regulated and ZPT2 are thought to function as transcriptional repressors under such conditions (Kodaira *et al.,* 2011). Stress-associated proteins (SAPs), containing A20/AN1 zinc-finger domains, confer abiotic stress tolerance in different plants. However, the interacting partners and downstream targets have not yet been identified. (Giri *et al.,* 2011). Expression of the genes encoding enzymes of the phenylpropanoid biosynthesis pathway are controlled by trimeric complexes of MYB and bHLH

transcription factors, and the WD40 factor TTG1. Appelhagen *et al.,* (2011) have demonstrated the interaction of TT1 with the R2R3 MYB protein TT2 to produce PA in the seed coat of wild type *Arabidopsis*.

Succinate dehydrogenase is an inner mitochondrial membrane succinate:ubiquinone oxidoreductase (EC 1.3.5.1) that is also known as succinate-coenzyme Q reductase (SQR). It is an iron-sulfur protein that catalyses an oxidation-reduction reaction during which succinate is oxidized to fumarate and ubiquinone (Q) reduced to ubiquinol (QH_2) with FAD acting as an electron acceptor. The reaction forms part of both the citric acid cycle, and the membrane-associated electron transport system that couples oxidation-reduction reactions to the transfer of protons across a membrane.

Only three of the contigs that were exclusively expressed in the green phenotype, and possibly involved with anthocyanin biosynthesis, could be identified by similarity searches using the BLAST tool at the NCBI's non-redundant database. Contig 1770 show similarity to an ETHYLENE-INSENSITIVE3 (EIN3) transcription factor (GU732486.1). A second contig (contig 7346) that was expressed in the green phenotype only, display similarity to an ETHYLENE-INSENSITIVE3-like (EIL1) transcription factor but with lower a score and shorter contig length. The third, contig 1668 show similarity to calreticulin 3. The protein is encoded by *PSL1* with mutants displaying derepressed sugar induced anthocyanin production in *Arabidopisis* (Saijo *et al.,* 2009).

Ethylene (ET) is a major plant hormone that regulates plant development and tolerance to necrotrophic fungi (Zhu *et al.,* 2011) via the EIN3 transcription factor. Ethylene also modulates sucrose and glucose sensitivity during *Arabidopsis* seedling development and

controls anthocyanin biosynthesis (Gibson *et al*., 2001). Anthocyanin accumulation is suppressed by ethylene signaling and activated by sugar and light signaling. Jeong *et al*., (2010) reported the presence of an anthocyanin induction pathway that is independent of HY5 but dependent upon photosynthetic electron transport in acyanic mesophyll cells. EIN3 is a short-lived protein whose degradation is mediated by two F-box proteins, EBF1 and EBF2 via the ubiquitin/26S proteasome pathway in the absence of ethylene (Tacken *et al.,* 2012).

A second contig (contig 276) that was expressed in the green phenotype only, and identified by similarity searches using the BLAST tool at the NCBI's non-redundant nucleotide database was similar to an apple Carotenoid cleavage dioxygenase (CCD4). Carotenoid cleavage dioxygenases (CCDs) are non-heme iron oxygenases that cleave carotenes and xanthophylls to colourless apocarotenoids (Rubio *et al*., 2008). This oxidative cleavage of carotenoids occurs in plants, animals, and microorganisms. Apocarotenoids are abundant in the thylakoid membranes of plants and cyanobacteria where they act as photoprotective pigments, accessory pigments in thylakoid membrane (Markwell *et al*., 1992) and signaling molecules with diverse functions, including the plant hormone abscisic acid.

The third contig that was expressed in the green phenotype only, and identified by similarity searches using the BLAST tool at the NCBI's non-redundant nucleotide database was similar to calreticulin3. Saijo *et al*., (2009) described *Arabidopsis* 'priority in sweet life' (*psl*) mutants that display de-repression of sucrose-induced anthocyanin accumulation in the presence of elf18. *PSL1* and *PSL2*, encode calreticulin3 (CRT3) and UDP-glucose:glycoprotein glycosyltransferase, respectively, that act in concert with

STT3A-containing oligosaccharyltransferase complex in an N-glycosylation pathway in the ER. The up-regulation of calreticulin3 in the green phenotype may suggest a role for the repression of anthocyanin production in the presence of the bacterial epitope elf18. Further functional characterisation of calreticulin in the green phenotype of 'Bon Rouge' may elucidate the role of this protein in the suppression of anthocyanin production in the green phenotype, in response to elf18.

The differences in expressed representative genes in the transcriptome of the red and the green phenotype appears to result from differences in responses to environmental cues. The representative genes that were expressed in the red phenotype only, are associated with stress responses, in particular heat and light stress. Of these, RAD23 has been shown to be up-regulated in plants during high irradiation responses, and lipocalin expression has been reported during temperature stress induction. A number of the transcripts expressed only in the green phenotype that could be identified in this manner are associated with environmental cues. In particular, the EIN3 transcript that is involved with the triple response to ethylene, and identified as an inhibitor of the sugar and light induced anthocyanin biosynthesis in plants (Jeong *et al.,* 2010). This suggests that the difference in anthocyanin production between the red and green phenotypes of 'Bon Rouge' results not from differences in anthocyanin biosynthesis gene expression but from inhibition of sucrose and light induced anthocyanin production via EIN3 in the green phenotype. The significant up-regulation of the transcript in the red phenotype that displayed similarity to a glutathione transferase (contig 71 in table 5.4.1) and associated with transport of anthocyanin into the plant vacuole for pigment stabilisation, and the up-

regulation of DFR in the red phenotype (table 5.5.1), may play significant roles for anthocyanin production in the red phenotype.

Functional characterisation of the transcripts that were differentially expressed between the two phenotypes is necessary to fully explore the difference in anthocyanin biosynthesis gene expression, and anthocyanin production in the red phenotype, or anthocyanin inhibition in the green phenotype. However, a good starting point for this exploration, as suggested by the results of transcriptome sequencing, would be an investigation of a number of the transcription factors that was differentially regulated between the two phenotypes that control the molecular switch for anthocyanin production. In particular, the AP2 and the zinc finger transcription factors that were exclusively induced in the red phenotype (contig 3233 in table 5.4.1) in conjunction with the regulation of *EIN3* and *PSL1* that were exclusively expressed in the green phenotype (contigs 1770, 7346 and 1668 in table 5.4.2).

6.3.3 Summary: Annotation of mRNAseq derived contigs by Blast2GO

In an attempt to link the over-expressed transcripts in the red phenotype to a common regulator for anthocyanin biosynthesis, the 105 most highly expressed contigs was subjected to a Blast2GO analysis. Thirty of the one hundred and five contigs subjected to this analysis show similarity to enzymes with established Enzyme Commission (EC) numbers in pathways represented in the KEGG database. Of these, only three could be directly associated with anthocyanin production, i.e. displayed enzyme codes similar to

those involved with catalytic steps in the anthocyanin biosynthesis pathway. These include phenylalanine ammonia-lyase (PAL), UDP-galactose transporter and glutathione transferases (GST).

Anthocyanin production in plants is facilitated by a complex of biosynthetic enzymes localized to the cytosolic membranes of the endoplasmic reticulum. For stabilisation of pigment colour and to prevent its toxic effects in the cytosol, anthocyanin pigments have to be transported to, and sequestered in the acidic vacuole (Zhao and Dixon, 2009). Recently Sun *et al.,* (2012) demonstrated that the *Arabidopsis Transparent Testa 19* (*TT19*), a glutathione S-transferase, functions as a carrier in the transport of cyanidin and/or anthocyanins to the tonoplast for sequestration in the vacuole. TT19 is localised both in the cytoplasm and on the tonoplast, while conjugated to cyanidin and cyanidin 3-galactoside. However, the researchers demonstrated that TT19 does not conjugate these pigments with glutathione, thus supporting the hypothesis that TT19 is a carrier protein for anthocyanins to facilitate its sequestration into the acidic vacuole for pigment stabilisation. Two contigs over-expressed in the red compared to the green phenotype displayed sequence similarity to glutathione transferases as determine by Blast2GO analysis. Contig 71 shows sequence similarity to a glutathione transferase (EC:2.5.1.18) while contig 3681 displays sequence similarity to a glutathione transferase, AtGSTu17. However, in the Blast2GO annotation, the latter enzyme was assigned the enzyme commission number, EC:4.4.1.5, which is classified in the KEGG database as a lactoylglutathione lyase (ketone-aldehyde mutase) that catalyses an aldo-keto isomerisation. AtGSTu17 is a member of the GST N family, Class Tau (U) subfamily cytosolic dimeric proteins involved in cellular detoxification by catalyzing the

conjugation of glutathione (GSH) with a wide range of endogenous and xenobiotic alkylating agents, environmental toxins such as herbicides and products of oxidative stress. In rare instances it catalyses the bio-activating isomerization of secondary metabolites (Cummins *et al.,* 2011). In addition, plant Tau GSTs play important roles in intracellular signalling, biosynthesis of anthocyanin, responses to soil stresses and responses to auxin and cytokinin hormones. Tau GSTs, are also involved in responses to different environmental stresses including heat, cold and drought, and chemical compounds such as hydrogen peroxide, salicylate, DTT, CuSO4, and also in herbicide metabolism. According to a study by Jiang *et al.,* (2010), AtGSTU17 is mainly involved in FR light signaling and is regulated by various photoreceptors, in particular phyA, under all light conditions. Its function appears to affect anthocyanin accumulation, and flowering time. The double mutant *atgstu17phyA* showed defects in physiological responses, including slightly longer hypocotyls in low FR, reduced anthocyanin levels, and insensitivity in FR-mediated inhibition of greening.

Four of the contigs over-expressed in the red relative to green phenotype, were annotated by Blast2GO as auxin efflux proteins or auxin binding proteins. These include contig 4825 that was annotated as an auxin binding protein abp19a (GO:0045735). Contigs 4939 and 7893 were similarly annotated as auxin binding protein abp19a (GO:0045735). Contig 74 was annotated as an auxin efflux carrier component (GO:0055085) while a sequence similarity search by BLAST against the non-redundant nucleotide database at NCBI with this query sequence annotated this contig as PIN1. Further investigation into the interaction of these various auxin-binding proteins with glutathione transferases should facilitate an understanding of their role in anthocyanin production in the red

phenotype.

Contig 6805 over-expressed in the red phenotype shows sequence similarity to a phenylalanine ammonia-lyase (PAL). Phenylalanine ammonia-lyase [EC:4.3.1.25] is the first committed step for phenylpropanoid biosynthesis in plants that catalyses the deamination of phenylalanine to *trans*-cinnamic acid and ammonia in a reaction that direct the carbon flow from the primary metabolic shikimate pathway to numerous branches of the general secondary metabolic phenylpropanoid pathway in plants. Contig 3155 over-expressed in the red phenotype show sequence similarity to an UDP-galactose/UDP-glucose transporter. The role of this transporter in the supply of galactose or glucose for conjugation to the anthocyanidin core in the reaction catalysed by UFGT, could provide further insight into the production of anthocyanin in the red leaf phenotype of 'Bon Rouge'.

Other contigs over-expressed in the red phenotype with similarity to genes coding for enzymes with EC numbers in the KEGG database pathways, include contig 261, a 3-ketoacyl- thiolase peroxisomal-like enzyme (EC:2.3.1.16), contig 6124, a beta-galactosidase 3-like (EC: 3.2.1.23), contig 6678, a plasma membrane H^+ATPase EC:3.6.3.6 (logbase2 ratio 3.265), contig 482, succinate dehydrogenase (EC:1.3.5.1), contig 701, 1-aminocyclopropane-1-carboxylate oxidase (EC:1.14.17.4 and EC:1.14.11.0), contig 1938, berberine bridge enzyme (EC:1.21.3.3) and MurB reductase (EC:1.1.1.158), contig 3957, omega-3 fatty acid desaturase (EC:1.14.99.33 and EC:1.14.19.0), contig 2597, an epoxide hydrolase 2-like, (EC:2.7.10.2) shown to be present only in the red leaf cDNA library and contig 699, an ATP phosphoribosyltransferase-like (EC:2.4.2.17), present only in red the leaf cDNA library).

Contig 261 over-expressed in the red phenotype show sequence similarity to a C-acyl transferase (EC:2.3.1.16) that is involved in numerous metabolic pathways listed in the KEGG database. C-acyl transferase catalyse reaction for biosynthesis of unsaturated fatty acids, fatty acid elongation, ethylbenzene degradation, benzoate degradation, alpha-Linolenic acid metabolism, geraniol degradation, and valine, leucine and isoleucine degradation,

Additional contigs over- or only expressed in the cDNA library prepared from the red phenotype include contig 596, a subtilisin-like protease-like (EC:3.4.21.0), contig 80, a subtilisin-like protease-like (EC:3.4.21.0), contig 7450, a alpha-xylosidase 1-like (EC:3.2.1.0), contig 3655, a 26S protease regulatory subunit 8 homolog a-like (EC:3.6.4.3) and contig 1121, a cell division cycle protein 48 homolog that was present only in the red leaf cDNA library).

Both BLAST and Blast2GO returned similar annotations for a number of the contigs analysed by these two programmes. These include contigs 3655, a 26S protease regulatory subunit 8 homolog a-like, contig 1121, a cell division cycle protein 48 homolog, contig 6678, a plasma membrane H^+ATPase, contig 482, succinate dehydrogenase, contig 2597, an epoxide hydrolase 2-like and contig 699, an ATP phosphoribosyltransferase-like.

Although a common regulator for anthocyanin biosynthesis was not identified in the Blast2GO analysis, two enzymes, PAL and GST, was shown to be up-regulated in the cDNA library prepared from the red phenotype. PAL catalyses the reaction at the entry

point for anthocyanin biosynthesis, while GST catalyses what is currently proposed to be last genetically defined step for anthocyanin biosynthesis. A number of stress response genes were up-regulated in the red phenotype library suggesting that various biotic and abiotic stresses are indirectly responsible for the production of anthocyanin. Induction of anthocyanin synthesis in response to various stress responses has been well documented in the literature. The up-regulation of genes for ELIP and DFN in the red phenotype cDNA library could well function as an indirect switch that could induce anthocyanin production in the red phenotype, by increasing the expression of PAL and GST.

6.4 Concluding remarks

Taken together, our data obtained from differential display and mRNAseq point to differential gene expression of stress response genes between the two phenotypic variants of 'Bon Rouge'. The stress response in plants is a complicated network of pathways that intersect at various levels. Both differential display and mRNAseq identified a number of genes that could be involved in such stress responses. In addition, a number of these (for ELIP, PAPS, GST) are involved in retrograde signaling from the chloroplast to the nucleus (Kleine and Lester, 2016). Retrograde signaling in response to environmental stresses involves some metabolite-linked pathways in addition to the commonly described redox and ROS pathways. Elucidating the metabolite and other compound trafficking between chloroplasts, cytoplasm and the nucleus may clarify the generation and transmission of metabolites and other compounds in retrograde signals involved in stress responses. Kleine and Lester (2016) showed that some retrograde signaling pathways converge with other networks such as light, immune and developmental

signaling. Several genes participating in retrograde signaling have been identified in our differential gene expression analysis and the dissection of signaling networks will allow a better understanding of multiple and convergent retrograde signaling, and point to differences in the two phenotypes that may underlie the difference in pigment production.

LIST OF REFERENCES

Adamska I, Kloppstech K, Ohad I. 1992. UV light stress induces the synthesis of the early light-inducible protein and prevents its degradation. J Biol Chem. 267(34):24732-7.

Ahrazem O, Trapero A, Gómez MD, Rubio-Moraga A, Gómez-Gómez L. 2010. Genomic analysis and gene structure of the plant carotenoid dioxygenase 4 family: A deeper study in Crocus sativus and its allies. Genomics 96:239-250.

Akazawa T, Takabe T and Kobayashi H. 1984. Molecular evolution of ribulose-1,5-bisphosphate carboxylase/oxygenase. Trends Biochem Sci. 9:380-388.

Alonso JM, Stepanova AN. 2004. The ethylene signaling pathway. Science. 306:1513-1515.

Albrecht V, Weinl S, Blazevic D, D'Angelo C, Batistic O, Kolukisaoglu U, Bock R, Schulz B, Harter K, Kudla J. 2003. The calcium sensor CBL1 integrates plant responses to abiotic stresses. Plant J. 36(4):457-70.

Alfenito MR, Souer E, Goodman CD, Buell R, Mol J, Koes R, Walbot V. 1998. Functional complementation of anthocyanin sequestration in the vacuole by widely divergent glutathione S-transferases. Plant Cell. 10(7):1135-49.

Almeida J, Carpenter R, Robbins TP, Martin C, Coen ES. 1989. Genetic interactions underlying flower colour patterns in *Antirrhinum majus*. Genes Dev. 3:1758-1767.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215(3):403-10.

Anderson I, Taylor TC. 2003. Structural framework for catalysis and regulation in ribulose-1,5-bisphosphate carboxylase/oxygenase. Arch Biochem Biophys. 414:130-140.

Appelhagen I, Lu GH, Huep G, Schmelzer E, Weisshaar B, Sagasser M. 2011. TRANSPARENT TESTA1 interacts with R2R3-MYB factors and affects early and late steps of flavonoid biosynthesis in the endothelium of *Arabidopsis thaliana* seeds. Plant J. 67(3):406-19.

Arata Y, Nishi T, Kawasaki-Nishi S, Shao E, Wilkens S, Forgac M. 2002. Structure, subunit function and regulation of the coated vesicle and yeast vacuolar H^+-ATPases. Biochim Biophys Acta. 1555(1-3):71-4.

Asahi T. 1960. Reduction of sulfite to sulfide in mung bean leaf. J of Biochemistry 48, 772-773.

Aviezer-Hagai K, Padler-Karavani V, Nelson N. 2003. Biochemical support for the V-

ATPase rotary mechanism: antibody against HA-tagged Vma7p or Vma16p but not Vma10p inhibits activity. J Exp Biol. 206(Pt 18):3227-37.

Avila J, Nieto C, Canas L, Benito MJ, Paz-Ares J. 1993. Petunia hybrida genes related to the maize regulatory C1 gene and to animal myb proto-oncogenes. Plant J. 3(4):553-62.

Baek GH, Kim I, Rao H. 2011. The Cdc48 ATPase modulates the interaction between two proteolytic factors Ufd2 and Rad23. Proc Natl Acad Sci USA. 108(33):13558-63.

Baranowski N, Frohberg C, Prat S, Willmitzer L. 1994. A novel DNA binding protein with homoly to Myb oncoproteins containing only one repeat can function as a transcriptional activator. EMBO J. 13:5383-5392.

Bariola PA, MacIntosh GC, Green PJ. 1999. Regulation of S-like ribonuclease levels in *Arabidopsis*. Antisense inhibition of RNS1 or RNS2 elevates anthocyanin accumulation. Plant Physiol. 119(1):331-42.

Bartels D, Hanke C, Schneider K, Michel D, Salamini F. 1992. A desiccation-related Elip-like gene from the resurrection plant *Craterostigma plantagineum* is regulated by light and ABA. EMBO J. 11(8):2771-8.

Beetham JK, Grant D, Arand M, Garbarino J, Kiyosue T, Pinot F, Oesch F, Belknap WR,

Shinozaki K, Hammock BD. 1995. Gene evolution of epoxide hydrolases and recommended nomenclature. DNA Cell Biol. 14(1):61-71.

Bègue H, Jeandroz S, Blanchard C, Wendehenne D, Rosnoblet C. 2017. Structure and functions of the chaperone-like p97/CDC48 in plants. Biochimica *et* Biophysica Acta - General Subjects 1861(1) Part A:3053-3060.

Beld M, Martin C, Huits H, Stuitjie AR, Gerats AGM. 1989. Flavonoid synthesis in *Petunia hybrida*: partial characterisation of dihydroflavonol-4-reductase gene. Plant Mol Biol. 13:491-502.

Bhuiyan NH, Selvaraj G, Wei Y, King J. 2009. Gene expression profiling and silencing reveal that monolignol biosynthesis plays a critical role in penetration defence in wheat against powdery mildew invasion. J Exp Bot. 60:509-521.

Bieza K, Lois R. 2001. An *Arabidopsis* mutant tolerant to lethal ultraviolet-B levels shows constitutively elevated accumulation of flavonoids and other phenolics. Plant Physiol. 126(3):1105-1.

Bilang J, Macdonald H, King PJ and Sturm A. 1993. A soluble aux-in-binding protein from *Hyoscyamus muticus* is a glutathione S-transferase. Plant Physiol. 102: 29-34.

Booij-James IS, Dube SK, Jansen MA, Edelman M, Mattoo AK. 2000. Ultraviolet-B radiation impacts light-mediated turnover of the photosystem II reaction center heterodimer in *Arabidopsis* mutants altered in phenolic metabolism. Plant Physiol. 124(3):1275-84.

Boss PK, Davies C, Robinson SP. 1996a. Expression of anthocyanin biosynthetic pathway genes in red and white grapes. Plant Mol Biol. 32:565-569.

Boss PK, Davies C, Robinson SP. 1996b. Analysis of the expression of anthocyanin pathway genes in developing *Vitis vinifera* L.cv Shiraz grape berries and the implications for pathway regulation. Plant Mol Biol. 111:1059-1066.

Boutet S, Vazquez F, Liu J, Béclin C, Fagard M, Gratias A, Morel JB, Crété P, Chen X, Vaucheret H. 2003. *Arabidopsis* HEN1: a genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. Curr Biol. 13(10):843-8.

Britsch L, Ruhnau-Brich B, Forkmann G. 1992. Molecular cloning, sequence analysis, and *in vitro* expression of flavonone β-hydroxylase from *Petunia hybrida*. J Biol Chem. 267:5380-5387.

Broekaert WF, Terras FR, Cammue BP, Osborn RW. 1995. Plant defensins: novel antimicrobial peptides as components of the host defense system. Plant Physiol.

108(4):1353-8.

Brown BA, Cloix C, Jiang GH, Kaiserli E, Herzyk P, Kliebenstein DJ, Jenkins GI. 2005. A UV-B-specific signaling component orchestrates plant UV protection.

Brugliera F, Holton TA, Stevenson TW, Farcy E, Lu C-Y, Cornish EC. 1994. Isolation and characterisation of a cDNA clone corresponding to the *Rt* locus of *Petunia hybrida*. Plant J. 5:85-92.

Burbulis IE, Iacobucci M, Shirley BW. 1996. A null mutation in the first enzyme of flavonoid biosynthesis does not affect male fertility in *Arabidopsis*. Plant Cell. 8(6):1013-25.

Callis J, Fromm M, Walbot V. 1987. Introns increase gene expression in cultured maize cells. Genes Dev. (10):1183-200.

Campos N, Bako L, Feldwisch J, Schell J and Palme K. 1992. A protein from maize labeled with azido-IAA has novel β-glucosidase activity. Plant J. 2:675-684.

Cao H, Glazebrook J, Clarke JD, Volko S, Dong X. 1997. The *Arabidopsis* NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. Cell. 88(1):57-63.

Carmell MA, Xuan Z, Zhang MQ, Hannon GJ. 2002. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis.

Genes Dev. 16(21):2733-42.

Catalanotto C, Azzalin G, Macino G, Cogoni C. 2000. Gene silencing in worms and fungi. Nature 404:24.

Cayli S, Klug J, Chapiro J, Fröhlich S, Krasteva G, Orel L, Meinhardt A. 2009. COP9 signalosome interacts ATP-dependently with p97/valosin-containing protein (VCP) and controls the ubiquitination status of proteins bound to p97/VCP. J Biol Chem. 284(50):34944-53.

Cervone F, Hahn MG, De Lorenzo G, Darvill A, Albersheim P. 1989. Host-Pathogen Interactions : XXXIII. A Plant Protein Converts a Fungal Pathogenesis Factor into an Elicitor of Plant Defense Responses. Plant Physiol. 90(2):542-548.

Chabregas SM, Luche DD, Van Sluys MA, Menck CF, Silva-Filho MC. 2003. Differential usage of two in-frame translational start codons regulates subcellular localization of *Arabidopsis thaliana* THI1. J Cell Sci. 116(Pt 2):285-91.

Chagné D, Carlisle CM, Blond C, Volz RK, Whitworth CJ, Oraguzie NC, Crowhurst RN, Allan AC, Espley RV, Hellens RP, Gardiner SE. 2007. Mapping a candidate gene (MdMYB10) for red flesh and foliage colour in apple. BMC Genomics. 3(8):212-222.

Chagné D, Crowhurst RN, Pindo M, Thrimawithana A, Deng C, Ireland H, Fiers M, Dzierzon H, Cestaro A, Fontana P, Bianco L, Lu A, Storey R, Knäbel M, Saeed M, Montanari S, Kim YK, Nicolini D, Larger S, Stefani E, Allan AC, Bowen J, Harvey I, Johnston J, Malnoy M, Troggio M, Perchepied L, Sawyer G, Wiedow C, Won K, Viola R, Hellens RP, Brewer L, Bus VG, Schaffer RJ, Gardiner SE, Velasco R. 2014. The Draft Genome Sequence of European Pear (Pyrus communis L. 'Bartlett'). PloS One 9:4 e92644.

Chalker-Scott L, Scott JD. 2004. Elevated ultraviolet-B radiation induces cross-protection to cold in leaves of Rhododendron under field conditions. Photochem Photobiol. 79(2):199-204.

Chandler VL, Radicella JP, Robbins TP, Chen J, Turks D. 1989. Two regulatory genes of the maize anthocyanin pathway are homologous: isolation of B utilizing R genomic sequences. Plant Cell. 1(12):1175-83.

Chang CS, Li YH, Chen LT, Chen WC, Hsieh WP, Shin J, Jane WN, Chou SJ, Choi G, Hu JM, Somerville S and Wu SH. 2008. LZF1, a HY5-regulated transcriptional factor, functions in *Arabidopsis* de-etiolation. Plant J. 54:205–219.

Chao Q, Rothenberg M, Solano R, Roman G, Terzaghi W, Ecker JR. 1997. Activation of the ethylene gas response pathway in *Arabidopsis* by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. Cell 89(7):1133-44.

Charron JB, Ouellet F, Pelletier M, Danyluk J, Chauve C, Sarhan F. 2005. Identification, expression, and evolutionary analyses of plant lipocalins. Plant Physiol. 139(4):2017-28.

Cheng FS, Weedon, NF, Brown SK. 1996. Identification of co-dominant AFLP markers tightly linked to fruit skin colour in apple. Theor Appl Genet. 93:222-227.

Cheng VWT, Johnson A, Rothery RA and Weiner JH. 2008. Alternative Sites for Proton Entry from the Cytoplasm to the Quinone Binding Site in *Escherichia coli* Succinate Dehydrogenase. Biochemistry 47:9107–16.

Chinnusamy V, Schumaker K, Zhu JK. 2004. Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants. J Exp Bot. 55(395):225-36.

Chopra S, Brendel V, Zhang J, Axtell JD, Petersen T. 1999. Molecular characterisation of a mutable pigmentation phenotype and isolation of the first active transposable element from *Sorghum bicolor*. Proc Nat Acad Sci. 96:15330-15335.

Chrispeels MJ, Holuigue L, Latorre R, Luan S, Orellana A, Pena-Cortes H, Raikhel NV, Ronald PC, Trewavas A. 1999. Signal transduction networks and the biology of plant cells. Biol Res. 32(1):35-60.

Christensen A, Svensson K, Thelin L, Zhang W, Tintor N, Prins D, Funke N, Michalak M, Schulze-Lefert P, Saijo Y, Sommarin M, Widell S, Persson S. 2010. Higher plant calreticulins have acquired specialized functions in *Arabidopsis*. PLoS One 5(6):e11342.

Christian M, Steffens B, Schenck D, Burmester S, Böttger M, Lüthen H. 2006. How does auxin enhance cell elongation? Roles of auxin-binding proteins and potassium channels in growth control. Plant Biol. 8:346-352.

Coe EH, Nueffer MG, Hosington DA. 1988. The genetics of corn: In Corn and Corn improvement, ed. G. F. Sprague, J. W. Dudley, pp 81-258. Madison, WI: USA.

Cogoni C, Macino G. 1999. Posttranscriptional gene silencing in Neurospora by a RecQ DNA helicase. Science. 286(5448):2342-4.

Cone KC, Cocciolone SM, Burr FA, Burr B. 1993. Maize anthocyanin regulatory gene pl is a duplicate of c1 that functions in the plant. Plant Cell. 5(12):1795-805.

Cosgrove DJ. 2000. Loosening of plant cell walls by expansins. Nature. 407(6802):321-6.

Craigon DJ, James N, Okyere J, Higgins J, Jotham J, May S. 2004. NASCArrays: A repository for microarray data generated by NASC's transcriptomics service. Nucleic Acids Res. 32(Database issue):D575-D577.

Croteau R, Kutchan TM, Lewis NG. 2000. Natural Products (Secondary Metabolites) in B. Buchanan, W. Gruisem, R. Jones Eds. Biochemistry and Molecular Biology of Plants, Chapter 24. American Society of Plant Physiologists. Curr Opin Plant Biol. 4(2):123-9.

Cummins I, Dixon DP, Freitag-Pohl S, Skipsey M, Edwards R. 2011. Multiple roles for plant glutathione transferases in xenobiotic detoxification. Drug Metab Rev. 43: 266-280.

Cushman JC, Bohnert HJ. 1997. Molecular Genetics of Crassulacean Acid Metabolism. Plant Physiol. 113:667-676.

Dalmay T, Hamilton A, Mueller E, Baulcombe DC. 2000. Potato virus X amplicons in *Arabidopsis* mediate genetic and epigenetic gene silencing. Plant Cell. 12(3):369-79.

Dalmay T, Horsefield R, Braunstein TH, Baulcombe DC. 2001. SDE3 encodes an RNA helicase required for post-transcriptional gene silencing in *Arabidopsis*.
EMBO J. 20(8):2069-78.

Datta S, Johansson H, Hettiarachchi C, Irigoyen ML, Desai M, Rubio V, Holm M. 2008. LZF1/SALT TOLERANCE HOMOLOG3, an *Arabidopsis* B-box protein involved in light-dependent development and gene expression, undergoes COP1-mediated ubiquitination. Plant Cell. 20(9):2324-38.

Davies KM. 1993a. A cDNA clone for flavonone 3-hydroxylase from *Malus*. Plant Physiol. 103:291.

Davies KM. 1993b. A Malus cDNA with Homology to the Antirrhinum Candica and Zea A2 Genes. Plant Physiol. 103:1015.

de Vetten N, Quattrocchio F, Mol J, Koes R. 1997. The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. Genes Dev. 11(11):1422-34.

de Vetten TE, ter Horst J, van Schalk H-P, de Boer A, Mol J, Koes R. 1999. A cytochrome b5 is required for full activity of flavonoid 3',5'-hydrolxylase, a cytochrome P450 involved in the formation of blue flower colors. Proc Natl Acad Sci. 96:778-783.

Dedonder A, Rethy R, Fredericq H, Van Montagu M, Krebbers A. 1993. *Arabidopsis rbcS* Genes Are Differentially Regulated by Light. Plant Physiol. 101:801-808.

Deng CH, Tian Y, Zhang C, Chagné D, Gleave A, Gardiner SE, Bus V, Sullivan S, Liachko I, Cong P, Yao, J-L. 2018. Fast construction of a reference genome: challenges and opportunities using 'Royal Gala' apple as a case study. Acta Hort. 1203:35-40.

Derrien B, Genschik P. 2014. When RNA and protein degradation pathways meet. Front Plant Sci. 5:161.

Dixon DP, Steel PG, Edwards R. 2011. Roles for glutathione transferases in antioxidant recycling. Plant Signal Behav. 6(8):1223-7.

Dixon RA, Paiva NL. 1995. Stress-induced phenylpropanoid metabolism. Plant Cell. 7:1085-1097.

Domgall I, Venzke D, Lüttge U, Ratajczak R, Böttcher B. 2002. Three dimensional map of a plant V-ATPase based on electron microscopy. J Biol Chem. 277:13115-13121.

D'Ovidio R, Roberti S, Di Giovanni M, Capodicasa C, Melaragni M, Sella L, Tosi P, Favaron F. 2006. The characterization of the soybean polygalacturonase-inhibiting proteins (Pgip) gene family reveals that a single member is responsible for the activity detected in soybean tissues. Planta. 224(3):633-45.

Dussi MC, Sugar D, Wrolstad RE. 1995. Characterizing and quantifying anthocyanins in red pears and the effect of light quality on fruit color. J Am Soc Hort Science 120:785-789.

Espley RV, Hellens RP, Putterill J, Stevenson DE, Kutty-Amma S, Allan AC. 2007. Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. Plant J. 49(3):414-27.

Falcone Ferreyra ML, Rius, SP and Casati, P. 2012. Flavonoids: biosynthesis, biological functions, and biotechnological applications. Front Plant Sci. 3(222):1-15.

Farmer LM, Book AJ, Lee KH, Lin YL, Fu H, Vierstra RD. 2010. The RAD23 family provides an essential connection between the 26S proteasome and ubiquitylated proteins in *Arabidopsis*. Plant Cell. 22(1):124-42.

Feild TS, Lee DW, Holbrook NM. 2001. Why leaves turn red in autumn. The role of anthocyanins in senescing leaves of red-osier dogwood. Plant Physiol. 127(2):566-74. Comment in: Plant Physiol. 2002 Mar;128(3):783.

Feldwisch J, Zettl R, Campos N and Palme, K. 1995. Identification of a 23 kDa protein from maize photoamnity-labelled with 5-azido-[7- 3H]indol-3-ylacetic acid. Biochem. J. 305:853-857.

Feng JX, Liu D, Pan Y, Gong W, Ma LG, Luo JC, Deng XW, Zhu YX. 2005. An annotation update via cDNA sequence analysis and comprehensive profiling of developmental, hormonal or environmental responsiveness of the *Arabidopsis* AP2/EREBP transcription factor gene family. Plant Mol Biol. 59:853–868.

Ferrari S, Vairo D, Ausubel FM, Cervone F, De Lorenzo G. 2003. Tandemly duplicated *Arabidopsis* genes that encode polygalacturonase-inhibiting proteins are regulated coordinately by different signal transduction pathways in response to fungal infection.

Plant Cell. 15(1):93-106.

Figueiredo P, Elhabiri M, Toki K, Saito N, Dangles O, Brouillard R. 1996. New aspects of anthocyanin complexation. Intramolecular copigmentation as a means for colour loss. Phytochemistry 41:301–308.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double stranded RNA in *Caenorhabdatisis elegans*. Nature 391:806-811.

Fischer TC, Gosch C, Pfeiffer J, Halbwirth H, Halle C, Stich K, Forkmann G. 2007. Flavonoid genes of pear (*Pyrus communis*). Trees 21:521-529.

Fischer-Schliebs E, Ball E, Berndt E, Besemfelder-Butz E, Binzel ML, Drobny M, Mühlenhoff D, Müller ML, Rakowski K, Ratajczak R. 1997. Differential immunological cross-reactions with antisera against the V-ATPase of *Kalanchoë daigremontiana* reveal structural differences of V-ATPase subunits of different plant species. Biol Chem. 378(10):1131-9.

Fishbain S, Prakash S, Herrig A, Elsasser S, Matouschek A. 2011. Rad23 escapes degradation because it lacks a proteasome initiation region. Nat Commun. 2:192.

Ford CM, Boss PK, Hoj PB. 1998. Cloning and characterisation of *Vitis vinifera* UDP-glucose: flavonoid 3-*O*-glucosyltransferase, a homologue of the enzyme encoded by the maize *bronze-1* locus that may primarily serve to glucosylate anthocyanidins *in vivo*. J Biol Chem. 273:9224- 9233.

Forgac M. 1998. Structure, function and regulation of the vacuolar H^+-ATPases. FEBS Lett. 440(3):258-63.

Franco-Zorrilla JM, Martin AC, Solano R, Rubio V, Leyva A, Paz-Ares J. 2002. Mutations at CRE1 impair cytokinin-induced repression of phosphate starvation responses in *Arabidopsis*. Plant J. 32(3):353-60.

Franks TK, Hayasaka Y, Choimes S, van Heeswijck R. 2005. Cyanogenic glucosides in grapevine: polymorphism, identification and developmental patterns. Phytochemistry. 66(2):165-73.

Furbank RT, Taylor WC. 1995. Regulation of photosynthesis in C3 and C4 plants: a molecular approach. Plant Cell. 7:797-807.

Furtek D, Schiefelbein JW, Johnston F, Nelson OE. (Jr.). 1988. Sequence analysis of three wild-type *Bronze-1* alleles from *Zea mays*. Plant Mol Biol 11. 473-481.

Gaff DF. 1971. Desiccation-Tolerant Flowering Plants in Southern Africa. Science.

174(4013):1033-1034.

Gallois JL, Drouaud J, Lécureuil A, Guyon-Debast A, Bonhomme S, Guerche P. 2013. Functional characterization of the plant ubiquitin regulatory X (UBX) domain-containing protein AtPUX7 in *Arabidopsis thaliana*. Gene **526**:299-308.

Garcia-Martinez JF, Moyano E, Alocer MJC, Martin C. 1998. Two bZIP proteins from *Antirrhinum* flowers preferentially bind a C-box/G-box motif and help to define a new sub-family of bZIP transcription factors. Plant J. 13:489-505.

Gehring CA, Irving HR. 2003. Natriuretic peptides - a class of heterologous molecules in plants. Int J Biochem Cell Biol. 2003 Sep;35(9):1318-22.

Gerbrandy SJ. 1974. Glycogen phosphorylase of potatoes. Purification and thermodynamic properties of the adsorption on glycogen. Biochim Biophys Acta. 370(2):410-8.

Gibson SI, Laby RJ, Kim D. 2001. The sugar-insensitive1 (sis1) mutant of *Arabidopsis* is allelic to ctr1. Biochem Biophys Res Commun 280:196-203.

Giri J, Vij S, Dansana PK, Tyagi AK. 2011. Rice A20/AN1 zinc-finger containing stress-associated proteins (SAP1/11) and a receptor-like cytoplasmic kinase (OsRLCK253) interact via A20 zinc-finger and confer abiotic stress tolerance in

transgenic *Arabidopsis* plants. New Phytol. 191(3):721-32.

Godoi PH, Galhardo RS, Luche DD, Van Sluys MA, Menck CF, Oliva G. 2006. Structure of the thiazole biosynthetic enzyme THI1 from *Arabidopsis thaliana*. J Biol Chem. 281(41):30957-66.

Goff SA, Cone KC, Chandler VL. 1992. Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. Genes Dev. 6(5):864-75.

Gong Z, Yamazaki M, Sugiyama M, Tanaka Y, Saito K. 1997. Cloning and molecular analysis of structural genes involved in anthocyanin biosynthesis are expressed in a forma-specific manner in *Perilla frutescens*. Plant Mol Biol. 35:915-927.

Gong Z-Z, Yamagishi E, Yamazaki M, Saito K. 1999a. A constitutively expressed *Myc*-like gene involved in anthocyanin biosynthesis form *Perilla frutescens*: molecular characterisation, heterologous expression in transgenic plants and transactivation in yeast cells. Plant Mol Biol. 41:33-44.

Gong Z-Z, Yamazaki M, Saito K. 1999b. A light-inducible *Myb*-like gene that is specifically expressed in red *Perilla frutescens* and presumably acts as a determining factor of the anthocyanin forma. Mol Gen Genet. 262:65-72.

Goodrich J, Carpenter R, Coen ES. 1992. A common gene regulates pigmentation pattern in diverse plant species. Cell. 68(5):955-64.

Gowri G, Paiva NL, Dixon RA. 1991. Stress responses in alfalfa (*Medicago sativa* L.). Sequence analysis of phenylalanine ammonium lyase (PAL) cDNA clones and appearance of PAL transcripts in elicitor-treated cell cultures and developing plants. Plant Mol Biol. 17:415-429.

Grant-Downton RT, Dickinson HG. 2006. Epigenetics and its implications for plant biology 2. The 'epigenetic epiphany': epigenetics, evolution and beyond. Ann Bot (Lond). 97(1):11-27.

Grimm B, Kruse E, Kloppstech K. 1989. Transiently expressed early light-inducible thylakoid proteins share transmembrane domains with light-harvesting chlorophyll binding proteins. Plant Mol Biol. 13(5):583-93.

Gu Y, Li S, Lord EM, Yang Z. 2006. Members of a novel class of *Arabidopsis* Rho guanine nucleotide exchange factors control Rho GTPase-dependent polar growth. Plant Cell. 18(2):366-81.

Güler S, Seeliger A, Härtel H, Renger G, Benning C. 1996. A null mutant of *Synechococcus sp*. PCC7942 deficient in the sulfolipid sulfoquinovosyl diacylglycerol. J. Biol. Chem. 271:7501-7507.

Guo, S. and Kemphues, K.J. 1995. *par-1*, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser Thr kinase that is asymmetrically distributed. Cell. 81:611-620.

Hafke JB, Hafke Y, Smith JA, Lüttge U, Thiel G. 2003. Vacuolar malate uptake is mediated by an anion-selective inward rectifier. Plant J. 35(1):116-28.

Hamberger B, Ellis M, Friedmann M, de Azevedo Sousa C, Barbazuk B, Douglas C. 2007. Genome-wide analyses of phenylpropanoid-related genes in *Populus trichocarpa*, *Arabidopsis thaliana*, and *Oryza sativa*: the Populus lignin toolbox and conservation and diversification of angiosperm gene families. Can J Bot. 85:1182-1201.

Harari-Steinberg O, Ohad I, Chamovitz DA. 2001. Dissection of the light signal transduction pathways regulating the two early light-induced protein genes in *Arabidopsis*. Plant Physiol. 127(3):986-97.

Harborne JB. 1976. Functions of flavonoids in plants. In TW Goodwin, Ed. Chemistry and Biochemistry of Plant Pigments, Ed 2, Vol 1. Academic Press, London.

Harborne JB. 1988. The Flavonoids: Advances in Research since 1980. (New York: Chapman and Hall).

Harborne JB, Williams CA. 2000. Advances in flavonoid research since 1992. Phytochemistry. 55:481-504.

Hardtke CS, Gohda K, Osterlund MT, Oyama T, Okada K, Deng XW. 2000. HY5 stability and activity in *Arabidopsis* is regulated by phosphorylation in its COP1 binding domain. Embo J. 19(18):4997-5006.

Haruko U, Haruko O. 1999. Glycobiology of the plant glycoprotein epitope: structure, immunogenicity and allergenicity of plant glycotopes. Trends Glycosci. Glycotech. 11:413-428.

Hasenfratz MP, Tsou CL, Wilkins TA. 1995. Expression of two related vacuolar H^+-ATPase 16-kilodalton proteolipid genes is differentially regulated in a tissue-specific manner. Plant Physiol. 108(4):1395-404.

Heddad M, Noren H, Reiser V, Dunaeva M, Andersson B, Adamska I. 2006. Differential expression and localization of early light-induced proteins in *Arabidopsis*. Plant Physiol. 142(1):75-87.

Heller W, Forkmann G. 1988. Biosynthesis. In *The Flavonoids*, ed J. B. Harborne, pp399-425. London, Chappel and Hall.

Hieber AD, Bugos RC, Yamamoto HY. 200. Plant lipocalins: violaxanthin de-epoxidase and zeaxanthin epoxidase. Biochim Biophys Acta. 1482(1-2):84-91.

Holm M, Hardtke CS, Gaudet R and Deng XW. 2001. Identification of a structural motif that confers specific interaction with the WD40 repeat domain of *Arabidopsis* COP1. EMBO J. 20:118–127.

Hinz G, Flügge UI. 1988. Phosphorylation of a 51-kDa envelope membrane polypeptide involved in protein translocation into chloroplasts. Eur J Biochem. 175(3):649-59.

Hiwasa K, Rose JK, Nakano R, Inaba A, Kubo Y. 2003. Differential expression of seven alpha-expansin genes during growth and ripening of pear fruit. Physiol Plant. 117(4):564-572.

Hoch WA, Singsaas EL, McCown BH. Resorption protection. 2003. Anthocyanins facilitate nutrient recovery in autumn by shielding leaves from potentially damaging light levels. Plant Physiol. 133(3):1296-1305.

Hocquigny S, Pelsy F, Dumas V, Kindt S, Heloir M-C, Merdinoglu D. 2004. Diversification within grapevine cultivars goes through chimeric states. Genome. 47(3):579-89.

Holton TA, Cornish EC. 1995. Genetics and Biochemistry of anthocyanin biosynthesis. Plant Cell. 7:1071-1083.

Hsu SW, Cheng CL, Tzen TC, Wang CS. 2010. Rop GTPase and its target Cdc42/Rac-interactive-binding motif-containing protein genes respond to desiccation during pollen maturation. Plant Cell Physiol. 51(7):1197-209.

Huang F-C, Horváth G, Molnár P, Turcsi E, Deli J, Schrader J, Sandmann G, Schmidt H, Schwab W. 2009. Substrate promiscuity of RdCCD1, a carotenoid cleavage oxygenase from *Rosa damascena*. Phytochemistry 70:457-464.

Hutin C, Nussaume L, Moise N, Moya I, Kloppstech K, Havaux M. 2003. Early light-induced proteins protect *Arabidopsis* from photooxidative stress. Proc Natl Acad Sci. 100(8):4921-6.

Isshiki A, Akimitsu K, Yamamoto M, Yamamoto H. 2001. Endopolygalacturonase is essential for citrus black rot caused by *Alternaria citri* but not brown spot caused by *Alternaria alternata*. Mol Plant Microbe Interact. 14(6):749-57.

Jackson D, Roberts K, Martin C. 1992. Temporal and spatial control of expression of anthocyanin biosynthetic genes in developing flowers of *Antirrhinum majus*. Plant J. 2:425-434.

Jacob-Wilk D, Goldschmidt EE, Riov J, Sadka A, Holland D. 1997. Induction of a Citrus gene highly homologous to plant and yeast thi genes involved in thiamine biosynthesis during natural and ethylene-induced fruit maturation. Plant Mol Biol.

35(5):661-6.

Jeong S-W, Das PK, Jeoung SC, Song J-Y, Lee HK, Kim Y-K, Kim WJ, Park Y I, Yoo S-D, Choi S-B, Choi G, Park Y-I. 2010. Ethylene Suppression of Sugar-Induced Anthocyanin Pigmentation in *Arabidopsis*. Plant Physiol. 154(3): 1514-1531.

Jiang H-W, Liu M-J, Chen IC, Huang, CH, Chao L-Y, Hsieh H-L. 2010. A Glutathione *S*-Transferase Regulated by Light and Hormones Participates in the Modulation of *Arabidopsis* Seedling Development. Plant Physiol. 154(4):1646-1658.

Jolly, M. 1993. INFRUITEC INFO, Number 627.

Jones L, McQueen-Mason S. 2005. A role for expansins in dehydration and rehydration of the resurrection plant *Craterostigma plantagineum*. FEBS Lett. 559(1-3):61-5.

Jones L, Ratcliff F, Baulcombe DC. 2001. RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance. Curr Biol. 11(10):747-57.

Jonsson LMV, Aaraman MEG, van Diepen J, de Vlaming P, Smit N, Schram AW. 1984). Properties and genetic control of anthocyanin 5-*O*-glucosyltransferase in flowers of *Petunia hybrida*. Planta. 160:341-347.

Jonsson LMV, de Vlaming P, Wiering H, Aaraman MEG, Schram AW. 1983. Genetic control of anthocyanin 5-*O*-methyl transferase activity in flowers of *Petunia hybrida*. Theor Appl Genet. 66:349-355.

Joubert DA, Slaughter AR, Kemp G, Becker JV, Krooshof GH, Bergmann C, Benen J, Pretorius IS, Vivier MA. 2006. The grapevine polygalacturonase-inhibiting protein (VvPGIP1) reduces *Botrytis cinerea* susceptibility in transgenic tobacco and differentially inhibits fungal polygalacturonases. Transgenic Res. 15(6):687-702.

Kaiserli E, Jenkins GI. 2007. UV-B Promotes Rapid Nuclear Translocation of the *Arabidopsis* UV-B Specific Signaling Component UVR8 and Activates Its Function in the Nucleus. Plant Cell. 19(8):2662-73.

Kajava AV. 1998. Structural diversity of leucine-rich repeat proteins. J Mol Biol. 277(3):519-27.

Kamsteeg J, van Brederode J, van Nigtevecht G. 1978. Identification, properties, and genetic control of UDP-glucose: cyanidin-3-rhamnosyl-(1 leads to 6)-glucoside-5-O-glucosyltransferase isolated from petals of the red campion (*Silene dioica*). Biochem Genet. 16(11-12):1059-71.

Kanevski I, Maliga P, Rhaoades DF, Gutteridge S. 1999. Plastome engineering of ribulose-1,5-bisphophate carboxylase/oxygenase in tobacco to form a sunflower large subunit and tobacco small subunit hybrid. Plant Physiol. 119:133-141.

Karpinski S, Gabrys H, Mateo A, Karpinska B, Mullineaux PM. 2003. Light perception in plant disease defence signalling. Curr Opin Plant Biol. 6(4):390-6.

Kendrick MD, Chang C. 2008. Ethylene signaling: new levels of complexity and regulation. Curr Opin Plant Biol 11:479–485.

Kitamura S. 2006. Transport of flavonoids: from cytosolic synthesis to vacuolar accumulation. In E Grotewold, Ed, The Science of Flavonoids. Springer, New York, 123-146.

Kleine T, Leister D. 2016. Retrograde signaling: Organelles go networking. Biochim Biophys Acta 1857:1313-1325.

Kliebenstein DJ, Lim JE, Landry LG, Last RL. 2002. *Arabidopsis* UVR8 regulates ultraviolet-B signal transduction and tolerance and contains sequence similarity to human regulator of chromatin condensation 1. Plant Physiol. 130(1):234-43

Kluge C, Lamkemeyer P, Tavakoli N, Golldack D, Kandlbinder A, Dietz KJ. 2003. cDNA cloning of 12 subunits of the V-type ATPase from *Mesembryanthemum*

crystallinum and their expression under stress. Molecular Memb Biol. 20:171-183.

Kodaira KS, Qin F, Tran LS, Maruyama K, Kidokoro S, Fujita Y, Shinozaki K, Yamaguchi-Shinozaki K. 2011. *Arabidopsis* Cys2/His2 Zinc-Finger Proteins AZF1 and AZF2 Negatively Regulate Abscisic Acid-Repressive and Auxin-Inducible Genes under Abiotic Stress Conditions. Plant Physiol. 157(2):742-56.

Koes RE, Quattrocchio F, Mol JNM. 1994. The flavonoid biosynthetic pathway in plants: function and evolution. Bioessays 16:123-132.

Koes RE, Spelt CE, van den Elzen PJ, Mol JN. 1989. Cloning and molecular characterization of the chalcone synthase multigene family of Petunia hybrida. Gene. 81(2):245-57.

Komjanc M, Festi S, Rizzotti L, Cattivelli L, Cervone F, De Lorenzo G. 1999. A leucine-rich repeat receptor-like protein kinase (LRPKm1) gene is induced in *Malus x domestica* by *Venturia inaequalis* infection and salicylic acid treatment. Plant Mol Biol. 40(6):945-57.

Kong J, Chia L, Goh N, Chia t, Brouillard R. 2003. Analysis of biological activities of anthocyanin. Phytochemistry 64:923-933.

Koprivova A, Mugford ST, Kopriva S. 2010. *Arabidopsis* root growth dependence on glutathione is linked to auxin transport. Plant Cell Rep. 29:1157-1167.

Kopriva S, Fritzemeier K, Wiedemann G, Reski R. 2007. The putative moss 3'-phosphoadenosine-5'-phosphosulfate reductase is a novel form of adenosine-5'-phosphosulfate reductase without an iron-sulfur cluster. J Biol Chem. 282(31):22930-8.

Kopriva S. 2006. Regulation of sulfate assimilation in *Arabidopsis* and beyond. Ann Bot (Lond). 97(4):479-95.

Kore-eda S, Cushman MA, Akselrod I, Bufford D, Fredrickson M, Clark E, Cushman JC. 2004. Transcript profiling of salinity stress responses by large-scale expressed sequence tag analysis in *Mesembryanthemum crystallinum*. Gene. 341:83-92.

Kortekamp A. 2006. Expression analysis of defence-related genes in grapevine leaves after inoculation with a host and a non-host pathogen. Plant Physiol Biochem. 44(1):58-67.

Kossmann J, Lloyd J. 2000. Understanding and influencing starch biochemistry. Crit Rev Biochem Mol Biol. 35(3):141-96.

Kranz HD, Denekamp M, Greco R, Jin H, Leyva A, Meissner RC, Petroni K, Urzainqui A, Bevan M, Martin C, Smeekens S, Tonelli C, Paz-Ares J. 1998. Towards functional

characterisation of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. Plant J. 16:263-276.

Krisa S, Teguo PW, Decendit A, Deffieux GM, Vercauteren J, Merillon J-M. 19??////. Production of ^{13}C-labelled anthocyanins by *Vitis vinfera* cell suspension cultures. Phytochemistry. 51:651-656.

Kristiansen KN, Rohde W. 1991. Structure of the *Hordeum vulgare* gene encoding dihydroflavonol-4-reductase and molecular analysis of *ant*18 mutants blocked in flavonoid synthesis. Mol Gen Genet. 2430:49-59.

Krol M, Ivanov AG, Jansson S, Kloppstech K, Huner NP. 1999. Greening under high light or cold temperature affects the level of xanthophyll-cycle pigments, early light-inducible proteins, and light-harvesting polypeptides in wild-type barley and the chlorina f2 mutant. Plant Physiol. 120(1):193-204.

Kroon J, Souer E, de Graaf A, Xue Y, Mol J, Koes R. 1994. Cloning and structural analysis of the anthocyanin pigmentation locus *Rt* of *Petunia hybrida*: characterisation of insertion sequences in two mutant alleles. Plant J. 5:69-80.

Kubasek WL, Shirley BW, McKillop A, Goodman HM, Briggs W, Ausubel FM. 1992. Regulation of Flavonoid Biosynthetic Genes in Germinating *Arabidopsis* Seedlings. Plant Cell. 4(10):1229-1236.

Kubo H, Peeters AJM, Aarts MGM, Pereira A, Koornneef M. 1999. ANTHOCYANINLESS2, a homeobox gene affecting anthocyanin distribution and root development in *Arabidopsis*. Plant Cell. 11:1217-1226.

Kumar LS. 1999. DNA markers in plant improvement: An overview. Biotechnology Advances. 17:143-182.

Laloi M. 1999. Plant mitochondrial carriers: an overview. Cell Mol Life Sci. 56(11-12):918-44.

Lancaster JE. 1992. Regulation of skin colour in apples. Crit Rev Plant Sci 10:487-502.

Landry LG, Chapple CC, Last RL. 1995. *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. Plant Physiol. 109(4):1159-66.

Lay FT, Brugliera F, Anderson MA. 2003. Isolation and properties of floral defensins from ornamental tobacco and petunia. Plant Physiol. 131(3):1283-93.

Lee J, He K, Stolc V, Lee H, Figueroa P, Gao Y, Tongprasit W, Zhao H, Lee I and Deng XW. 2007. Analysis of transcription factor HY5 genomic binding sites revealed its hierarchical role in light regulation of development. Plant Cell. 19:731–749.

Lelièvre JM, Latché A, Jones B, Bouzayen M, Pech JC. 1997. Ethylene and fruit ripening. Physiologia Plantarum 101:727–739.

Levi A, Galau GA, Wetzstein HY. 1992. A rapid procedure for the isolation of RNA from high-phenolic-containing tissue of pecan. HortSci. 27:1316-1318.

Leyva A, Jarillo JA, Salinas J, Martinez-Zapater JM. 1995. Low Temperature Induces the Accumulation of Phenylalanine Ammonia-Lyase and Chalcone Synthase mRNAs of *Arabidopsis thaliana* in a Light-Dependent Manner. Plant Physiol. 108(1):39-46.

Li J, Ou-Lee TM, Raba R, Amundson RG, Last RL. 1993. *Arabidopsis* Flavonoid Mutants Are Hypersensitive to UV-B Irradiation. Plant Cell. 5(2):171-179.

Li L, Ban Z-J, Li X-H, Wu M-Y, Wang A-L, Y-Q Jiang and Y-H Jiang. 2012. Differential Expression of Anthocyanin Biosynthetic Genes and Transcription Factor PcMYB10 in Pears (*Pyrus communis* L.). PLoS ONE 7(9): e46070.

Li L, Deng CH, Knäbel M, Chagné D, Kumar S, Sun J, Zhang S, Wu J. 2017. Integrated high-density consensus genetic map of Pyrus and anchoring of the 'Bartlett' v1.0 (*Pyrus communis*) genome. DNA Res. 24(3):289-301.

Li R, Rimmer R, Yu M, Sharpe AG, Séguin-Swartz G, Lydiate D, Hegedus DD. 2003.

Two *Brassica napus* polygalacturonase inhibitory protein genes are expressed at different levels in response to biotic and abiotic stresses. Planta. 217(2):299-308.

Li Y, Jones L, McQueen-Mason S. 2003. Expansins and cell growth. Curr Opin Plant Biol. 6(6):603-10.

Liang P, Pardee AB. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science. 257(5072):967-71.

Licciardello C, D'Agostino N, Traini A, Recupero GR, Frusciante L, Chiusano ML. 2014. Characterization of the glutathione S-transferase gene family through ESTs and expression analyses within common and pigmented cultivars of *Citrus sinensis* (L.) Osbeck. BMC Plant Biol. 14:39.

Lievens S, Goormachtig S, Holsters M. 2001. A critical evaluation of differential display as a tool to identify genes involved in legume nodulation: looking back and looking forward. Nucleic Acids Res. 29(17):3459-3468.

Lillo C, Lea US, Ruoff P. 2008. Nutrient depletion as a key factor for manipulating gene expression and product formation in different branches of the flavonoid pathway. Plant Cell & Environment 31:587-601.

Lloyd AM, Walbot V, Davies RW. 1992. *Arabidopsis* and *Nicotiniana* anthocyanin production activated by maize regulators *R* and *C1*. Science. 258:1773-1775.

Logemann E, Hahlbrock K. 2002. Crosstalk among stress responses in plants: pathogen defense overrides UV protection through an inversely regulated ACE/ACE type of light-responsive gene promoter unit. Proc Natl Acad Sci. 99(4):2428-32.

Logemann E, Tavernaro A, Schulz W, Somssich IE, Hahlbrock K. 2000. UV light selectively coinduces supply pathways from primary metabolism and flavonoid secondary product formation in parsley. Proc Natl Acad Sci. 97(4):1903-7.

Loyall L, Uchida K, Braun S, Furuya M, Frohnmeyer H. 2000. Glutathione and a UV light-induced glutathione S-transferase are involved in signaling to chalcone synthase in cell cultures. Plant Cell. 12(10):1939-50.

Lu Y, Hall DA, Last RL. 2011. A small zinc finger thylakoid protein plays a role in maintenance of photosystem II in *Arabidopsis thaliana*. Plant Cell. 23(5):1861-75.

Ludwig SR, Habera LF, Dellaporta SL, Wessler SR. 1989. Lc, a member of the maize R gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the myc-homology region.
Proc Natl Acad Sci USA. 86(18):7092-6.

Lüttge U, Ratajczak R. 1997. The physiology, biochemistry, and molecular biology of the plant vacuolar ATPase. Adv Bot Res. 25:253-296.

Lüttge U. 2002. CO_2-concentrating: consequences in crassulacean acid metabolism. J Exp Bot. 53(378):2131-42.

Macdonald H, Jones AM and King PJ. 1991. Photoaffinity labeling of soluble auxin-binding proteins. J Biol Chem. 266:7393-7399.

Machado CR, de Oliveira RL, Boiteux S, Praekelt UM, Meacock PA, Menck CF. 1996. Thi1, a thiamine biosynthetic gene in *Arabidopsis thaliana*, complements bacterial defects in DNA repair. Plant Mol Biol. 31(3):585-93.

Mahalingam R, Gomez-Buitrago A, Eckardt N, Shah N, Guevara-Garcia A, Day P, Raina R, Fedoroff NV. 2003. Characterizing the stress/defense transcriptome of *Arabidopsis*. Genome Biol. 4(3):R20.

Maine EA. 2000. A conserved mechanism for post-transcriptional gene silencing? Genome Biol. 2000; 1(3): reviews1018.1

Marcotrigiano M, Bernatzky. 1995. Arrangement of cell layers in the shoot apical meristem of periclinal chimeras influences cell fate. Plant J. 7:193-202.

Marcotrigiano M. 1991. Understanding foliar variegation as it relates to propagation. Comb Proc Intl Plant Prop Soc. 41:410-415.

Marcotrigiano M. 2000. Herbivory could unlock mutations sequestered in stratified shoot apices of genetic mosaics. Am J Bot. 87(3): 355-361.

Maric M, Maculins T, De Piccoli G, Karim Labib K. 2014. Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication. Science 346(6208):1253596-1-1243596-12.

Markham KR, Gould KS, Winefield CS, Mitchell KA, Bloor SJ, Boase MR. 2000. Anthocyanic vacuolar inclusions - their nature and significance in flower colouration. Phytochemistry.55(4):327-36.

Markwell J, Bruce BD, Keegstra K. 1992. Isolation of a carotenoid containing sub-membrane particle from the chloroplastic envelope outer membrane of pea (Pisum sativum). Biol Chem. 267:13933–37.

Marrs KA, Alfenito MR, Lloyd AM, Walbot V. 1995. A glutathione *S*-transferase involved in vacuolar transfer encoded by the maize gene *Bronze-2*. Nature 375:397-400.

Martienssen RA, Colot V. 2001. DNA methylation and epigenetic inheritance in plants and filamentous fungi. Science 293(5532):1070-1074.

Martin C, Gerats T. 1993. Control of Pigment Biosynthesis Genes during Petal Development. Plant Cell. 5(10):1253-1264.

Martin T, Oswald O, Graham IA. 2002. *Arabidopsis* seedling growth, storage lipid mobilization, and photosynthetic gene expression are regulated by carbon:nitrogen availability. Plant Physiol. 128(2):472-81.

Martinez-Garcia JF, Moyano E, Alocer MJC, Martin C. 1998. Two bZIP proteins from *Antirrhinum* flowers preferentially bind a C-box/G-box motif and help to define a new sub-family of bZIP transcription factors. Plant J. 13:489-505.

Maruyama-Nakashita A, Nakamura Y, Watanabe-Takahashi A, Inoue E, Yamaya T, Takahashi H. 2005. Identification of a novel cis-acting element conferring sulfur deficiency response in *Arabidopsis* roots. Plant J. 42(3):305-14.

Matzke MA, Birchler JA. 2005. RNAi-mediated pathways in the nucleus.
Nat Rev Genet. 6(1):24-35.

McCarty DR, Carson CB, Stinard PS, Robertson DS. 1989. Molecular Analysis of viviparous-1: An Abscisic Acid-Insensitive Mutant of Maize. Plant Cell. 1(5):523-532.

McQueen-Mason SJ, Cosgrove DJ. 1995. Expansin mode of action on cell walls. Analysis of wall hydrolysis, stress relaxation, and binding. Plant Physiol. 107(1):87-100.

Mendoza-Cozatl D, Loza-Tavera H, Hernandez-Navarro A, Moreno-Sanchez R. 2005. Sulfur assimilation and glutathione metabolism under cadmium stress in yeast, protists and plants. FEMS Microbiol Rev. 29(4):653-71.

Menzlaff E, Flügge UI. 1993. Purification and functional reconstitution of the 2-oxoglutarate/malate translocator from spinach chloroplasts. Biochim Biophys Acta. 1147(1):13-8.

Meyer B, Houlné G, Pozueta-Romero J, Schantz ML, Schantz R. 1996. Fruit-specific expression of a defensin-type gene family in bell pepper. Upregulation during ripening and upon wounding. Plant Physiol. 112(2):615-22.

Michalak M, Corbett EF, Mesaeli N, Nakamura K, Opas M. 1999. Calreticulin: one protein, one gene, many functions. Biochem J. 344:281-292.

Michalak M, Groenendyk J, Szabo E, Gold L, Opas M. 2009. Calreticulin, a multi-process calcium-buffering chaperone of the endoplasmic reticulum. Biochem J. 417(3):651-66.

Milla MA, Butler E, Huete AR, Wilson CF, Anderson O, Gustafson JP. 2002. Expressed sequence tag-based gene expression analysis under aluminum stress in rye. Plant Physiol. 130(4):1706-16.

Miller AN, Walsh CS, Cohen JD. 1987. Measurement of indole-3-acetic acid in peach fruits (Prunus persica L. Batsch cv. Redhaven) during development. Plant Physiology 84:491-494.

Minami E, Ozeki Y, Matsuoka N, Tanaka Y. 1989. Structure and some characterisation of the gene for phenylalanine ammonia-lyase from rice plants. Eur J Biochem. 185:19-25.

Miura K, Rus A, Sharkuu A, Yokoi S, Karthikeyan AS, Raghothama KG, Baek D, Koo YD, Jin JB, Bressan, RA, Hasegawa PM. 2005. The *Arabidopsis* SUMO E3 ligase SIZ1 controls phosphate deficiency responses. Proc Natl Acad Sci. 102(21):7760-5.

Montané M-H, Kloppstech K. 2000. The family of light-harvesting-related proteins (LHCs, ELIPs, HLIPs): was the harvesting of light their primary function? Gene 258, 1-8.

Monticolo F, Colantuono C, Maria Luisa Chiusano ML. 2017. Shaping the evolutionary tree of green plants: evidence from the GST family. Sci Reports 7:Article number 14363.

Moons A. 2005. Regulatory and functional interactions of plant growth regulators and plant glutathione S-transferases (GSTs). Vitam Horm. 72:155-202.

Mori H, Tanizawa K, Fukui T. 1991. Potato tuber type H phosphorylase isozyme.

Molecular cloning, nucleotide sequence, and expression of a full-length cDNA in *Escherichia coli*. J Biol Chem. 266(28):18446-53.

Mori H, Tanizawa K, Fukui T. 1993. A chimeric alpha-glucan phosphorylase of plant type L and H isozymes. Functional role of 78-residue insertion in type L isozyme. J Biol Chem. 268(8):5574-81.

Morisseau C, Beetham JK, Pinot F, Debernard S, Newman JW, Hammock BD. 2000. MUCILAGE-MODIFIED4 encodes a putative pectin biosynthetic enzyme developmentally regulated by APETALA2, TRANSPARENT TESTA GLABRA1, and GLABRA2 in the *Arabidopsis* seed coat. Plant Physiol. 134(1):296-306.

Mourrain P, Beclin C, Elmayan T, Feuerbach F, Godon C, Morel JB, Jouette D, Lacombe AM, Nikic S, Picault N, Remoue K, Sanial M, Vo TA, Vaucheret H. 2000. *Arabidopsis* SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell. 101(5):533-42.

Moyano E, Martinez-Garcia J, Martin C. 1996. Apparent redundancy in *myb* gene function provides gearing for the control of anthocyanin biosynthesis in *Antirrhinum* flowers. Plant Cell. 8:1519-1532.

Mueller LA, Goodman CD, Silady RA, Walbot V. 2000. AN9, a petunia glutathione S-transferase required for anthocyanin sequestration, is a flavonoid-binding protein. Plant Physiol. 123(4):1561-70.

Mulichak AM, Theisen MJ, Essigmann B, Benning C, Garavito RM. 1999. Crystal structure of SQD1, an enzyme involved in the biosynthesis of the plant sulfolipid headgroup donor UDP-sulfoquinovose. Proc Natl Acad Sci. 96:13097-13102.

Mundree SG, Whittaker A, Thomson JA, Farrant JM. 2000. An aldose reductase homolog from the resurrection plant Xerophyta viscosa Baker. Planta. 211(5):693-700.

Nakano T, Suzuki K, Fujimura T, Shinshi H. 2006. Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. Plant Physiol. 140:411–432.

Napoli CA, Fahy D, Wang H-Y, Taylor LP. 1999. white anther: A Petunia Mutant That Abolishes Pollen Flavonol Accumulation, Induces Male Sterility, and Is Complemented by a Chalcone Synthase Transgene. Plant Physiol. 120:615-622.

Nei M, Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci. 76(10):5269-5273.

Newgard CB, Hwang PK, Fletterick RJ. 1989. The family of glycogen phosphorylases: structure and function. Crit Rev Biochem Mol Biol. 24(1):69-99.

Niehl A, Amari K, Gereige D, Brandner K, Mély Y, Manfred Heinlein M. 2012. Control of *Tobacco mosaic virus* Movement Protein Fate by CELL-DIVISION-CYCLE Protein

48. Plant Physiol. 160(4): 2093-2108.

Nielson KJ, Hill JM, Anderson MA, Craik DJ. 1996. Synthesis and structure determination by NMR of a putative vacuolar targeting peptide and model of a proteinase inhibitor from *Nicotiana alata*. Biochemistry. 35:369–378.

Nikiforova V, Freitag J, Kempa S, Adamik M, Hesse H, Hoefgen R. 2003. Transcriptome analysis of sulfur depletion in *Arabidopsis thaliana*: interlacing of biosynthetic pathways provides response specificity. Plant J. 33(4):633-50.

Nueffer MG, Jones L, Zuber MS. 1968. The Mutants of Maize. (Madison, WI. Crop Science Society of America).

Nybom H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol Ecol. 13(5):1143-55.

Oeser B, Heidrich PM, Müller U, Tudzynski P, Tenberge KB. 2002. Polygalacturonase is a pathogenicity factor in the *Claviceps purpurea*/rye interaction. Fungal Genet Biol. 36(3):176-86.

Owen-Hughes T. 2003. Colworth memorial lecture. Pathways for remodelling chromatin. Biochem Soc Trans. 31(Pt 5):893-905.

Palmieri F, Bisaccia F, Capobianco L, Dolce V, Fiermonte G, Iacobazzi V, Indiveri C, Palmieri L. 1996. Mitochondrial metabolite transporters. Biochim Biophys Acta. 1275(1-2):127-32.

Passeri V, Koes R, Quattrocchio FM. 2016. New Challenges for the Design of High Value Plant Products: Stabilization of Anthocyanins in Plant Vacuoles. Front Plant Sci. 7:153, doi: 10.3389/fpls.2016.00153.

Pastori GM, Foyer CH. 2003. Common components, networks, and pathways of cross-tolerance to stress. The central role of "redox" and abscisic acid-mediated controls. Plant Physiol. 129(2):460-8.

Pellerin P, Vidal S,Williams P, Brillouet JM. 1995. Characterization of five type II arabinogalactan-protein fractions from red wine of increasing uronic acid content. Carbohydr Res. 277:135–143.

Pelletier ML, Murrel JR, Shirley BW. 1997. Characterisation of flavonol synthase and leucoanthocyanidin dioxygenase genes in *Arabidopsis*. Plant Physiol. 113:1437-1445.

Penninckx IA, Eggermont K, Terras FR, Thomma BP, De Samblanx GW, Buchala A, Metraux JP, Manners JM, Broekaert WF. 1996. Pathogen-induced systemic activation of a plant defensin gene in *Arabidopsis* follows a salicylic acid-independent pathway. Plant Cell. 8(12):2309-23.

Perera IY, Li X, Sze H. 1995. Several distinct genes encode nearly identical to 16 kDa proteolipids of the vacuolar H^+-ATPase from *Arabidopsis thaliana*. Plant Mol Biol. 29(2):227-44.

Pérez-Hernández J, Zaldívar-Machorro VJ, Villanueva-Porras D, Vega-Ávila E, and Chavarría A. 2016. A Potential Alternative against Neurodegenerative Diseases: Phytodrugs. Oxidative Medicine and Cellular Longevity. Article ID 8378613, http://dx.doi.org/10.1155/2016/8378613

Perrot GH, Cone KC. 1989. Nucleotide sequence of the maize R-S gene. Nucleic Acids Res. 17(19):8003.

Pfeiffer J, Kuhnel C, Brandt J, Duy D, Punyasiri PA, Forkmann G, Fischer TC. 2006. Biosynthesis of flavan 3-ols by leucoanthocyanidin 4-reductases and anthocyanidin reductases in leaves of grape (*Vitis vinifera* L.), apple (*Malus x domestica* Borkh.) and other crops. Plant Physiol Biochem. 44(5-6):323-34.

Pierantoni L, Dondini L, De Franceschi P, Musacchi S, Winkel BS, Sansavini S. 2010. Mapping of an anthocyanin-regulating MYB transcription factor and its expression in red and green pear, *Pyrus communis*. Plant Physiol Biochem. 48(12):1020-6.

Poethig RS. 1988. Heterochronic mutations affecting shoot development in maize. Genetics 119:959-73.

Pomeranz M, Zhang L, Finer J, Jang JC. 2011. Can AtTZF1 act as a transcriptional activator or repressor in plants? Plant Signal Behav. 6(5):719-22.

Pötter E, Kloppstech K. 1993. Effects of light stress on the expression of early light-inducible proteins in barley. Eur J Biochem. 214(3):779-86.

Quan Z, Fei H, Reeves JM, Pan Q, Changqing D, Jun W. Expression of Structural Genes Related to Anthocyanin Biosynthesis of *Vitis amurensis*, Mol Plant Breeding 4(31):254-264.

Quattrocchio FM, Wing JF, Leppen HTC, Mol JNM, Koes R. 1993. Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. Plant Cell. 5(11):1497-1512.

Quattrocchio FM, Wing JF, van der Woude K, Mol JNM, Koes R. 1998. Analysis of bHLH and MYB domain proteins: species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. Plant J. 13:475-448.

Quattrocchio FM, Wing JF, van der Woude K, Souer E, de Vetten N, Mol JN M, 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.

Quattrocchio FM, Verweij W, Kroon A, Spelt C, Mol JNM, Koes R. 2006. PH4 of Petunia Is an R2R3 MYB Protein That Activates Vacuolar Acidification through Interactions with Basic-Helix-Loop-Helix Transcription Factors of the Anthocyanin Pathway. Plant Cell. 18(5):1274-1291.

Radchuk R, Radchuk V, Weschke W, Borisjuk L, Weber H. 2006. Repressing the Expression of the SUCROSE NONFERMENTING-1-RELATED PROTEIN KINASE Gene in Pea Embryo Causes Pleiotropic Defects of Maturation Similar to an Abscisic Acid-Insensitive Phenotype. Plant Physiol. 140(1): 263–278.

Raes J, Rohde A, Christensen JH, Van de Peer Y, Boerjan W. 2003. Genome-wide characterization of the lignification toolbox in *Arabidopsis*. Plant Physiol. 133:1051-1071.

Ratajczak R. 2000. Structure, function and regulation of the plant vacuolar H^+-translocating ATPase. Biochim Biophys Acta. 1465:17-36.

Reisinger V, Plöscher M, Eichacker LA. 2008. Lil3 assembles as chlorophyll-binding protein complex during deetiolation. FEBS Lett 582:1547-1551.

Reymond P, Weber H, Damond M, Farmer EE. 2000. Differential gene expression in response to mechanical wounding and insect feeding in *Arabidopsis*. Plant Cell. 12(5):707-20.

Ridley BL, O'Neill MA, Mohnen D. 2001. Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. Phytochemistry 57:929–967.

Rohde A, Morreel K, Ralph J, Goeminne G, Hostyn V, De Rycke R, Kushnir S, Van Doorsselaere J, Joseleau JP, Vuylsteke M, Van Driessche G, Van Beeumen J, Messens E, Boerjan W. 2004. Molecular phenotyping of the *pal1* and *pal2* mutants of *Arabidopsis thaliana* reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. Plant Cell. 16(10):2749-71.

Romano N, Macino G. 1992. Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. Mol Microbiol. 6(22):3343-53

Romero I, Fuertes A, Benito MJ, Malpica JM, Leyva A, Paz-Ares J. 1998. More than 80 R2R3-MYB regulatory genes in the genome of *Arabidopsis thaliana*. Plant J. 14(3):273-84.

Römisch K. 2016. A Case for Sec61 Channel Involvement in ERAD. Trends Biochem

Sci. DOI: http://dx.doi.org/10.1016/j.tibs.2016.10.005.

Rosli HG, Civello PM, Martínez GA. 2009. alpha-l-Arabinofuranosidase from strawberry fruit: cloning of three cDNAs, characterization of their expression and analysis of enzymatic activity in cultivars with contrasting firmness. Plant Physiol Biochem. 47(4):272-81.

Rosnoblet C, Bègue H, Blanchard C, Pichereaux C, Besson-Bard A, Aimé S, Wendehenne D. 2016. Functional characterization of the chaperon-like protein CDC48 in cryptogein-induced immune response in tobacco. Plant Cell and Environment. doi: 10.1111/pce.12686.

Rossini S, Casazza AP, Engelmann EC, Havaux M, Jennings RC, Soave C. 2006. Suppression of both ELIP1 and ELIP2 in *Arabidopsis* does not affect tolerance to photoinhibition and photooxidative stress. Plant Physiol. 141(4):1264-73.

Rubio A, Rambla JL, Santaella M, Gómez MD, Orzaez D, Granell A, Gómez-Gómez L. 2008. Cytosolic and plastoglobuletargeted carotenoid dioxygenases from Crocus sativus are both involved in b-ionone-release. Biol Chem. 283:24816-25.

Ruggiano A, Foresti O, Carvalho P. 2014. ER-associated degradation: Protein quality control and beyond. J Cell Biol. 204(6):869-879.

Ruiz MT, Voinnet O, Baulcombe DC. 1998. Initiation and maintenance of virus-

induced gene silencing. Plant Cell. 10(6):937-46.

Ryan KG, Swinny EE, Markham KR, Winefield C. 2002. Flavonoid gene expression and UV photoprotection in transgenic and mutant Petunia leaves. Phytochemistry. 59(1):23-32.

Ryan KG, Swinny EE, Winefield C, Markham KR. 2001. Flavonoids and UV photoprotection in *Arabidopsis* mutants. Z Naturforsch [C]. 56(9-10):745-54.

Sablowski RWM, Moyano E, Culianez-Macia FA, Schuch W, Martin C, Bevan M. 1994. A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. EMBO J. 13:128-137.

Saijo Y, Tintor N, Lu X, Rauf P, Pajerowska-Mukhtar K, Häweker H, Dong X, Robatzek S, Schulze-Lefert P. 2009. Receptor quality control in the endoplasmic reticulum for plant innate immunity. EMBO J. 200928(21):3439-49.

Sainz MB, Grotewold E, Chandler VL. 1997. Evidence for direct activation of an anthocyanin promoter by the maize C1 protein and comparison of DNA binding by related Myb domain proteins. Plant Cell. 9(4):611-25.

Saito K, Kobayashi M, Gong Z, Tanaka Y, Yamazaki M. 1999. Direct evidence for anthocyanin synthase as a 2-oxoglutarate-dependent oxygenase: molecular cloning and

functional expression of cDNA from a red forma of *Perilla frutescens*. Plant J. 17:181-189.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, NY, Vol. 1, 2, 3.

Sampedro J, Cosgrove DJ. 2005. The expansin superfamily. Genome Biol. 6(12):242.

Saraste M, Walker JE. 1982. Internal sequence repeats and the path of polypeptide in mitochondrial ADP/ATP translocase. FEBS Lett. 144(2):250-4.

Satina S, Blakeslee AF, Avery AG. 1940. Demonstration of the three germ layers in the shoot apex of *Datura* by means of induced polyploidy in periclinal chimeras. Am J Bot. 27:895-905.

Sävenstrand H, Olofsson M, Samuelsson M, Strid A. 2004. Induction of early light-inducible protein gene expression in *Pisum sativum* after exposure to low levels of UV-B irradiation and other environmental stresses. Plant Cell Rep. 22(7):532-6.

Schaberg PG, Van Den Berg AK, Murakami PF, Shane JB, Donnelly JR. 2003. Factors influencing red expression in autumn foliage of sugar maple trees. Tree Physiol. 23(5):325-33.

Schauer SE, Jacobsen SE, Meinke DW, Ray A. 2002. *DICER-LIKE1*: blind men and

elephants in *Arabidopsis* development. Trends Plant Sci. 7(11):487-91.

Schmidt A, Jäger K. 1992. Open questions about sulfur metabolism in plants. Annual Review of Plant Physiol and Plant Mol Biol. 43:325–349

Schmitz G, Theres K. 1992. Structural and functional analysis of the *Bz2* locus of *Zea mays*: Characterisation of overlapping transcripts. Mol Gen Genet. 233:269-277.

Schneidereit J, Häusler RE, Fiene G, Kaiser WM, Weber AP. 2006. Antisense repression reveals a crucial role of the plastidic 2-oxoglutarate/malate translocator DiT1 at the interface between carbon and nitrogen metabolism. Plant J. 45(2):206-24.

Schulz C, Rehling P. 2014. Powering the cell. Science 346:1059-1060.

Sehitoglu MH, Farooqi AA, Qureshi MZ, Butt G, Aras A. 2014. Anthocyanins: targeting of signaling networks in cancer cells. Asian Pac J Cancer Prev. 15(5):2379-81.

Seidel T, Kluge C, Hanitzsch M, Ross J, Sauer M, Dietz KJ, Golldack D. 2004. Colocalization and FRET-analysis of subunits c and a of the vacuolar H^+-ATPase in living plant cells. J Biotechnol. 112(1-2):165-75.

Settembre E, Begley TP, Ealick SE. 2003. Structural biology of enzymes of the thiamin biosynthesis pathway. Curr Opin Struct Biol. 13(6):739-47.

Setya A, Murillo M, Leustek T. 1996. Sulphate reduction in higher plants—molecular evidence for a novel 5'-adenylylsulfate reductase. Proc Natl Acad Sci USA. 93:13383-13388.

Shaked-Sachray L, Weiss D, Reuveni M, Nissim-Levi A, Oren-Shamir M. 2002. Increased anthocyanin accumulation in aster flowers at elevated temperatures due to magnesium treatment. Physiol Plant. 114(4):559-565.

Sharma YK, Leon J, Raskin I, Davis KR. 1996. Ozone-induced responses in *Arabidopsis thaliana*: the role of salicylic acid in the accumulation of defense-related transcripts and induced resistance. Proc Natl Acad Sci USA. 93(10):5099-104.

Shi L, Katavic V, Yu Y, Kunst L, Haughn G. 2012. *Arabidopsis glabra2* mutant seeds deficient in mucilage biosynthesis produce more oil. Plant J. 69(1):37-46.

Shirley BW, Kubasek WL, Storz G, Bruggemann E, Koornneef M, Ausubel FM, Goodman HM. 1995. Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. Plant J. 8(5):659-71.

Smardon A, Spoerke JM, Stacey SC, Klein ME, Mackin N, Maine EM. 2000. EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr Biol. 10(4):169-78.

Smith AM, Zeeman SC, Smith SM. 2005. Starch degradation. Annu Rev Plant Biol. 56:73-98.

Snyder BA, Nicholson RL. 1990. Synthesis of phytoalexins in sorghum as a site-specific response to fungal ingress. Science 248(4963):1637-9.

Solano R, Nieto C, Avila J, Canas L, Dias I, Paz-Ares J. 1995. Dual DNA binding specificity of a petal epidermis-specific MYB transcription factor (MYB.Ph3) from *Petunia hybrida*. EMBO J. 14(8):1773-84.

Solano R, Nieto C, Paz-Ares J. 1995. MYB.Ph3 transcription factor from *Petunia hybrida* induces similar DNA-bending/distortions on its two types of binding sites. Plant J. 8:673-682.

Solano R, Stepanova A, Chao Q, Ecker JR. 1998. Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. Genes Dev. 12:3703-3714.

Sparvoli F, Martin C, Scienza A, Gavazzi G, Tonelli C. 1994. Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). Plant Mol Biol. 24:743-755.

Spelt C, Quattrocchio F, Mol J, Koes R. 2002. ANTHOCYANIN1 of petunia controls pigment synthesis, vacuolar pH, and seed coat development by genetically distinct mechanisms. Plant Cell. 14(9):2121-35.

Spelt C, Quattrocchio F, Mol JN, Koes R. 2000. *anthocyanin1* of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. Plant Cell. 12(9):1619-32.

Spreitzer RJ. 2003. Role of the small subunit in ribulose-1,5-bisphophate carboxylase/oxygenase. Arch. Biochem. Biophys. 414:141-149.

Stapleton AE. 1992. Ultraviolet Radiation and Plants: Burning Questions. Plant Cell. 4(11):1353-1358.

Steyn W, Wand SJE, Holcroft, DM, Jacobs G. 2004. Regulation of pear colour development in relation to activity of flavonoid enzymes. J Am Hort Science. 129 (1):6-12.

Stockinger EJ, Gilmour SJ, Thomashow MF. 1997. *Arabidopsis thaliana CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. Proc Natl Acad Sci USA 94:1035-1040.

Stone SL. 2014. The role of ubiquitin and the 26S proteasome in plant abiotic stress signaling. Front Plant Sci. 5:135.

Sugaya S. and Sakai S. 1996. Identification of a soluble auxin-binding protein as a glutathione-dependent formaldehyde dehydrogenase. Plant Sci. 114:1-9.

Sun Y, Li H, Huang J-R. 2012. *Arabidopsis* TT19 Functions as a Carrier to Transport Anthocyanin from the Cytosol to Tonoplasts. Molecular Plant. 5(2):387-400.

Suter M, von Ballmoos P, Kopriva S, Op den Camp R, Schaller J, Kuhlemeier C, Schürmann P, Brunold C. 2000. Adenosine 5'-phosphosulfate sulfotransferase and adenosine 5'-phosphosulfate reductase are identical enzymes. Journal of Biol Chem. 275:930-93.

Sze H, Schumacher K, Müller ML, Padmanaban S, Taiz L. 2002. A simple nomenclature for a complex proton pump: VHA genes encode the vacuolar H^+-ATPase. Trends Plant Sci. 7(4):157-61.

Sze H, Ward JM, Lai S. 1992. Vacuolar H^+-translocating ATPases from plants: structure, function, and isoforms. J Bioenerg Biomembr. 24(4):371-81.

Sze H. 1985. H^+-translocating ATPases: advances using membrane vesicles. Annu. Rev. Plant Physiol. 36:175-208.

Tacken EJ, Ireland H, Wang Y-Y, Putterill J, Schaffer RJ. 2012. Apple EIN3 BINDING F-box 1 (EBF1) inhibits the activity of three apple EIN3-like transcription factors. Annals of Botany Plants Advance Access published October 9.

Takahashi K, Takabayashi A, Tanaka A, Tanaka R. 2014. Functional analysis of light-harvesting-like protein 3 (LIL3) and its light-harvesting chlorophyll-binding motif in *Arabidopsis*. J Biol Chem. 289(2):987-99.

Takos AM, Jaffe FW, Jacob SR, Bogs J, Robinson SP, Walker AR. 2006. Light-induced expression of a MYB gene regulates anthocyanin biosynthesis in red apples. Plant Physiol. 142(3):1216-1232.

Tan B-C, Schwartz SH, Zeevaart JA, McCarty DR. 1997. Genetic control of abscisic acid biosynthesis in maize. Proc Natl Acad Sci USA. 94:12235-40.

Tanaka R, Oster U, Kruse E, Rudiger W, Grimm B. 1999. Reduced activity of geranylgeranyl reductase leads to loss of chlorophyll and tocopherol and to partially geranylgeranylated chlorophyll in transgenic tobacco plants expressing antisense RNA for geranylgeranyl reductase. Plant Physiol. 120:695-704.

Tanaka R, Rothbart M, Oka S, Takabayashi A, Takahashi K, Shibata M, Myouga F, Motohashi R, Shinozaki K, Grimm B, Tanaka A. 2010. LIL3, a light-harvesting-like

protein, plays an essential role in chlorophyll and tocopherol biosynthesis. Proc Natl Acad Sci USA. 107(38):16721-5.

Taniguchi M, Sugiyama T. 1997. The Expression of 2-Oxoglutarate/Malate Translocator in the Bundle-Sheath Mitochondria of *Panicum miliaceum*, a NAD-Malic Enzyme-Type C4 Plant, Is Regulated by Light and Development. Plant Physiol. 114(1):285-293.

Taran N, Okanenko A, Musienko N. 2000. Sulpholipid reflects plant resistance to stress-factor action. Biochem Soc Trans. 28(6):922-4.

Tateishi A, Mori H, Watari J, Nagashima K, Yamaki S, Inoue H. 2005. Isolation, characterization, and cloning of {alpha}-L-Arabinofuranosidase expressed during fruit ripening of Japanese pear. Plant Physiol. 138(3):1653-64.

Taylor LP, Briggs WR. 1990. Genetic regulation and photocontrol of anthocyanin accumulation in maize seedlings. Plant Cell. 2(2):115-27.

ten Have A, Mulder W, Visser J, van Kan JA. 1998. The endopolygalacturonase gene Bcpg1 is required for full virulence of *Botrytis cinerea*. Mol Plant Microbe Interact. 11(10):1009-16.

Thilo C, Fischer TC, Gosch C, Pfeiffer J, Halbwirth H, Halle C, Stich K, Forkmann G.

2007. Flavonoid genes of pear (*Pyrus communis*). Trees 21:521-529.

Thomma BP, Cammue BP, Thevissen K. 2002. Plant defensins. Planta. 216(2):193-202.

Tonelli C, Consonni G, Dolfini SF, Dellaporta SL, Viotti A, Gavazzi G. 1991. Genetic and molecular analysis of Sn, a light-inducible, tissue specific regulatory gene in maize. Mol Gen Genet. 225(3):401-10.

Toubart P, Desiderio A, Salvi G, Cervone F, Daroda L, De Lorenzo G. 1992. Cloning and characterization of the gene encoding the endopolygalacturonase-inhibiting protein (PGIP) of *Phaseolus vulgaris* L. Plant J. 2(3):367-73.

Treutter D. 2005. Significance of flavonoids in plant resistance and enhancement of their biosynthesis. Plant Biol. 7:581-591

Tucker ML, Burke A, Murphy CA, Thai VK, Ehrenfried ML. 2007. Gene expression profiles for cell wall-modifying proteins associated with soybean cyst nematode infection, petiole abscission, root tips, flowers, apical buds, and leaves. J Exp Bot. 58(12):3395-406.

Tyagi W, Rajagopal D, Singla-Pareek SL, Reddy MK, Sopory SK. 2005. Cloning and regulation of a stress-regulated *Pennisetum glaucum* vacuolar ATPase c gene and

characterization of its promoter that is expressed in shoot hairs and floral organs. Plant Cell Physiol. 46(8):1411-22.

Tyagi W, Singla-Pareek S, Nair S, Reddy MK, Sopory SK. 2006. A novel isoform of ATPase c subunit from pearl millet that is differentially regulated in response to salinity and calcium. Plant Cell Rep. 25(2):156-63.

Uimari A, Strommer J. 1997. *Myb26*: a MYB-like protein of pea flowers with affinity for promoters of phenylpropanoid genes. Plant J. 12:1273-1284.

Ulm R, Nagy F. 2005. Signalling and gene regulation in response to ultraviolet light. Curr Opin Plant Biol. 8(5):477-8.

Uptmoor R, Wenzel W, Friedt W, Donaldson G, Ayisi K, Ordon F. 2003. Comparative analysis on the genetic relatedness of *Sorghum bicolor* accessions from Southern Africa by RAPDs, AFLPs and SSRs. Theor Appl Genet. 106(7):1316-25.

Valledor L, Canal MJ, Pascual J, Rodrıguez R and Meijo M. 2012. Early induced protein 1 (*PrELIP1*) and other photosynthetic, stress and epigenetic regulation genes are involved in *Pinus radiata* D. don UV-B radiation response. Physiologia Plantarum 146:308-320.

Vandenbussche F, Habricot Y, Condiff AS, Maldiney R, Van der Straeten D, Ahmad M.

2007. HY5 is a point of convergence between cryptochrome and cytokinin signalling pathways in *Arabidopsis thaliana*. Plant J. 49(3):428-41.

Vos P, Hogers R, Bleeker M, Reijans M van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, et al. 1995. AFLP: a new technique for DNA fingerprinting. Nucl Acid Res. 23(21):4407-14.

Vosman B, Visser D, van der Voort JR, Smulders MJ, van Eeuwijk F. 2004. The establishment of 'essential derivation' among rose varieties, using AFLP. Theor Appl Genet. 109(8):1718-25.

Wade HK, Bibikova TN, Valentine WJ, Jenkins GI. 2001. Interactions within a network of phytochrome, cryptochrome and UV-B phototransduction pathways regulate chalcone synthase gene expression in *Arabidopsis* leaf tissue. Plant J. 25(6):675-85.

Wade SL, Auble DT. 2010. The Rad23 ubiquitin receptor, the proteasome and functional specificity in transcriptional control. Transcr. 1(1):22-26.

Walbot V, Mueller LA, Silady RA, Goodman CD. 2000. Do glutathione S-transferases act as enzymes or carrier proteins for their substrates. Sulfur Nutr. Sulfur Assimilation in Higher Plants, 155-165.

Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks MD, Gray JC. 1999. The *TRANSPARENT TESTA GLABRA1* locus, which regulated trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encode a WD40 repeat protein. Plant Cell. 11:1337-1349.

Wang H, Nair MG, Strasburg GM, Chang YC, Booren AM, Gray JI, DeWitt DL. 1999. Antioxidant and anti-inflammatory activities of anthocyanins and their aglycon, cyanidin, from tart cherries. J Nat Prod. 62(2):294-6. Erratum in: J Nat Prod. 62(5):802.

Wang J, Ji Q, Jiang L, Shen S, Fan Y, Zhang C. 2009. Overexpression of a cytosol-localized rhamnose biosynthesis protein encoded by *Arabidopsis* RHM1 gene increases rhamnose content in cell wall. Plant Physiol Biochem. 47(2):86-93.

Wang J, Wang Y, Yang J, Ma C, Zhang Y, Ge T, Qi Z, Kang Y. 2015. *Arabidopsis ROOT HAIR DEFECTIVE3* is involved in nitrogen starvation-induced anthocyanin accumulation. J Integr Plant Biol. 57(8):708-21.

Wang S, Liu S, Yu G, Gong P, Zhao D, Xuanjun Fang X. 2010. Studies on Rice Epoxide Hydrolase: Gene Cloning and Expressing in *E. coli*, Yeast and *Arabidopsis*. Rice Genomics and Genetics. Vol. 1 No. 2.

Wasternack C, Hause B. 2002. Jasmonates and octadecanoids: signals in plant stress responses and development. Prog Nucleic Acid Res Mol Biol. 72:165-221.

Wei N, Serino G, Deng XW. 2008. The COP9 signalosome: more than a protease. Trends Biochem Sci. 33(12):592-600.

Western TL, Young DS, Dean GH, Tan WL, Samuels AL, Haughn GW. 2004. MUCILAGE-MODIFIED4 encodes a putative pectin biosynthetic enzyme developmentally regulated by APETALA2, TRANSPARENT TESTA GLABRA1, and GLABRA2 in the *Arabidopsis* seed coat. Plant Physiol. 134(1):296-306.

Wiering H, De Vlaming P. 1984. Genetics of flower and pollen colour. In Monographs on Theoretical and Applied Genetics: Petunia ed. K. C. Sink 9:49-67. Springer-Verlag.

Wiering H. 1984. Genetics of flower color in *Petunia hybrida*. Hort Genen Phaenen. 17:117-134.

Wilson LG. 1962. Metabolism of sulfate: sulfate reduction. Annual Review of Plant Physiology 13:201–224.

Winkel-Shirley B, Kubasek WL, Storz G, Bruggemann E, Koorneef M, Ausubel M, Goodman H. 1995. Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. Plant J. 8(5):659-671.

Winkel-Shirley B. 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol. 126(2):485-93.

Winkel-Shirley B. 2002. Biosynthesis of flavonoids and effects of stress. Curr Opin Plant Biol. 5(3):218-23.

Wong LH, Čopič A, Levine TP. 2017. Advances on the Transfer of Lipids by Lipid Transfer Proteins. Trends Biochem Sci. 42(7):516-530.

Woo KC, Flügge UI, Heldt HW. 1987. A two-translocator model for the transport of 2-oxoglutarate and glutamate in chloroplasts during ammonia assimilation in the light. Plant Physiol. 84:624-632.

Wu J, Zhao G, Yang Y-N, Le W-Q, Khan MA, Zhang S-L, Gu C and Huang W-J. 2012. Identification of differentially expressed genes related to coloration in red/green mutant pear (*Pyrus communis* L.). Tree Genet Genomes: 1–9.

Xiong L, Zhu JK. 2001. Abiotic stress signal transduction in plants: Molecular and genetic perspectives. Physiol Plant. 112(2):152-166.

Yamazaki M, Gong Z, Fukuchi M, Fukui Y, Tanaka Y, Kusumi T, Saito K. 1999. Molecular cloning and biochemical characterisation of a novel anthocyanin 5-*O*-

glucosyltransferase by mRNA differential display for plant forms regarding anthocyanin. J Biol Chem. 274:7405-7411.

Yan A, Xu G, Yang ZB. 2009. Calcium participates in feedback regulation of the oscillating ROP1 Rho GTPase in pollen tubes. Proc Natl Acad Sci USA. 106(51):22002-7.

Yang Y, Yao G, Yue W, Zhang S, Wu J. 2015. Transcriptome profiling reveals differential gene expression in proanthocyanidin biosynthesis associated with red/green skin color mutant of pear (*Pyrus communis* L.). Front Plant Sci. 6:795.

Yoshida K, Kawachi M, Mori M, Maeshima M, Kondo M, Nishimura M, Kondo T. 2005. The involvement of tonoplast proton pumps and $Na^+(K^+)/H^+$ exchangers in the change of petal color during flower opening of Morning Glory, *Ipomoea tricolor* cv. Heavenly Blue. Plant Cell Physiol. 46(3):407-15.

Yoshimoto M, Okuno S, Yoshinaga M, Yamakawa O, Yamaguchi M, Yamada J. 1999. Antimutagenicity of sweetpotato (*Ipomoea batatas*) roots. Biosci Biotechnol Biochem. 63(3):537-41.

Yu B, Xu C, Benning C. 2002. *Arabidopsis* disrupted in SQD2 encoding sulfolipid synthase is impaired in phosphate-limited growth. Proc Natl Acad Sci. 99(8):5732-7.

Yu Y and Huang R. 2017. Integration of Ethylene and Light Signaling Affects Hypocotyl Growth in *Arabidopsis*. Front. Plant Sci. 8:57.

Zakhleniuk OV, Raines CA, Lloyd JC. 2001. pho3: a phosphorus-deficient mutant of *Arabidopsis thaliana* (L.) Heynh. Planta. 212(4):529-34.

Zeeman SC, Thorneycroft D, Schupp N, Chapple A, Weck M, Dunstan H, Haldimann P, Bechtold N, Smith AM, Smith SM. 2004. Plastidial α-Glucan Phosphorylase Is Not Required for Starch Degradation in *Arabidopsis* Leaves But Has a Role in the Tolerance of Abiotic Stress. Plant Physiol. 135(2): 849–858.

Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18(5):821-9.

Zettl R, Schell J, Palme K. 1994. Photoaffinity labeling of *Arabidopsis thaliana* plasma membrane vesicles by 5-azido-[7–3H]indole-3-acetic acid: identification of a glutathione *S*-transferase. Proc Natl Acad Sci USA. 91:689–693.

Zhao J, Pang Y, Dixon RA. 2010. The Mysteries of Proanthocyanidin Transport and Polymerization. Plant Physiol. 153(2):437-443.

Zheng DC, Han X, An Y, Guo HW, Xia XL, Yin WL. 2013. The nitrate transporter NRT2.1 functions in the ethylene response to nitrate deficiency in *Arabidopsis*. Plant Cell Environ 36: 1328-1337.

Zhu Z, An F, Feng Y, Li P, Xue L, A M, Jiang Z, Kim JM, To TK, Li W, Zhang X, Yu Q, Dong Z, Chen WQ, Seki M, Zhou JM, Guo H. 2011. Derepression of ethylene-stabilized transcription factors (EIN3/EIL1) mediates jasmonate and ethylene signaling synergy in *Arabidopsis*. Proc Natl Acad Sci USA. 108(30):12539-44.

Zimmermann IM, Heim MA, Weisshaar B, Uhrig JF. 2004. Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. Plant J. 40(1):22-34.

APPENDIX 1

NUCELOTIDE SEQUENCES OF cDNAS OBTAINED FROM DIFFERENTIAL DISPLAY ANALYSIS.

M1R
ATACAGAAGCACAACACAGGTAGGCATTAACCAAAATATCAAAGGTATATN
AGTTACATTACAAAAAGGAGTTCACATATTTAGTCCACCAAATATGTGTAATT
CTNGAAGGGCGAGCGCTTTACAAATGGTTTTCACATGGAGTCCTCCTTACCCT
GTCCCTCTCCCTGAAACCTGGCCTTAAATTCTCACGAGAGAGCTTCTCGTACC
CGTCCATNCGCTTCCGTCAAATAGGAGGGCCCGCTATGATGTTNCCTTCTGCT
GCTGCTGCTCTGGACCCAAAAGCTTTAATNCACTTAGTG
M1F
ACTAGTGATTAAGCTTTGGTCAGAGCAGCAGCAGCAGAAGGAACATCATAGC
GGCCCTCTATTTGACGAAGCGATGGTCGGGTACAGAGNAAGCTCTCTCGTGA
GNAATTTAAGGCCAGGTTTCAGGGAGAGGGACAGGGTAAGGAGGACTCCAT
GTGAAAACCATTTGTAAAGCGCTCGCCCTTCAGAATTACACATATTTGGTGGA
CTAAATATGTGAACTCCTTTTTGTAATGTAACGATATACCTTTGATATTTTNG
GTTAATGCCACCTGTGTTGTGCTTCTGTAT

M2F
GATTAAGCTTCTCAACGGCATATTTGGATTAGGCTGGTTGATTAGGACATTAT
TTTCGAGTACAAATCGCCTACCTGTAAAGAGTAGTTTAGAATATTGCATCCCA
TAAAATAAAAATGCTTCTCTACT
M2R
AGTAGAGAAGCATTTTTATTTTATGGGATGCAATATTCTAAACTACTCTTTAC
AGGTAGGCGATTTGTACTCGAAAATAATGTCCTAATCAACCAGCCTAATCCA
AATATGCCGTTGAGAAGCTTAATC

M3F
ACTAGTGATTAAGCTTAGTAGGCATGTAGTGTAAATATGGTGGTAAGATCTC
AGCTGCATGTTTTTCTGAATAAAAAAGGGCCTTCATCTTTTGT
M3R
ACAAAAGATGAAGGCCCTTTTTTATTCAGAAAAACATGCAGCTGAGATCTTA
CCACCATATTTACACTACATGCCTACTAAGCTTAATCACTAGT

M4R
ACAAAAGATGAAGGCCCTTTTTTATTCAGAAAAACATGCAGCTGAGATCTTA
CCACCATATTTACACTACATGCCTACTAAGCTTAATCGAATTCCCGCGGCCGC

CATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTC
GTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTG
GCGTTACCCAACTTAATGCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCG
TAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCA

M4F

GCTTAATCACTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAG
CTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCTAAATAG
CTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCA
CAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGC
CTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCC
AGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAA

M5F

GATTAAGCTTAGTAGGCAGGTCAAATTTATGTTTTCAGAAAAATGAAACAGT
AAGAATTTAAGCAAAATATATCATCCTTATTGTTGTTATGTTCTTGTTGATTTG
TACGCTAGTTGTATTATATTATAGAGCTTGATATACATATGTATCTTTGATGG
TT

M5R

AACCATCAAAGATACATATGTATATCAAGCTCTATAATATAATACAACTAGC
GTACAAATCAACAAGAACATAACAACAATAAGGATGATATATTTTGCTTAAA
TTCTTACTGTTTCATTTTTCTGAAAACATAAATTTGACCTGCCTACTAAGCTTA
ATC

M6F

GATTAAGCTTAACGAGGATAAAGAATGCATTGTCCATATTTGAGAATCCACA
AAGATGGCTGATTGTGGAGTTTATAGATTATGAATGTCGCATAGGTATGTATC
CAACTTGTAATGTTGGTTTCTTTTTTAGTTTTAAAAGTAGGTAAACAAAGGGA
TCGATTAGCGGCCCTAGTTTCTTGAGTATGGCCTCTACAATGGGGGATGAGA
GTGACTGTGTCAACTTCTCTTGCTTCAATTGAGTAATTGGAAACTAATCTGCT
AGTCGTGGTTCTATTGCTTTTCC

M6R

GGAAAAGCAATAGAACCACGACTAGCAGATTAGTTTCCAATTACTCAATTGA
AGCAAGAGAAGTTGACACAGTCACTCTCATCCCCCATTGTAGAGGCCATACT
CAAGAAACTAGGGCCGCTAATCGATCCCTTTGTTTACCTACTTTTAAAACTAA
AAAAGAAACCAACATTACAAGTTGGATACATACCTATGCGACATTCATAATC
TATAAACTCCACAATCAGCCATCTTTGTGGATTCTCAAATATGGACAATGCAT
TCTTTATCCTCGTTAAGCTTAATC

M7F

GATTAAGCTTTCCTGGAGGCCATTGTCGTGGCTTTCGCCGCAGATGCTTCTGC
ACTAAACATTGTTAATTAGCTATTAATTAATGACGAGATGATCATCATTAATC
ATACATGTATATGTGTATATGTGTGACATGTGGGGGTATTAAGATTAAATAAT
CGCTTAATTATCATTCCGTGCATGGATACCTACGTATGTGCATGCTTGTGTGC
TACTAGAATAAATTAATAACCCAATCTTTCACAGTTGGGTTATCATTAATTGT
TCTTTTGTTCTTGTTTATTAAGTAAAACTATCTCCGCGTGAGTTACTTTGT

M7R
ACAAAGTAACTCACGCGGAGATAGTTTTACTTAATAAACAAGAACAAAAGAA
CAATTAATGATAACCCAACTGTGAAAGATTGGGTTATTAATTTATTCTAGTAG
CACACAAGCATGCACATACGTAGGTATCCATGCACGGAATGATAATTAAGCG
ATTATTTAATCTTAATACCCCCACATGTCACACATATACACATATACATGTAT
GATTAATGATGATCATCTCGTCATTAATTAATAGCTAATTAACAATGTTTAGT
GCAGAAGCATCTGCGGCGAAAGCCACGACAATGGCCTCCAGGAAAGCTTAA
TCC

M8F
ACTAGTGATTAAGCTTTCCTGGAGAAATGATATTCACCTCATTCAGGTTGTAA
CCAATTTCTCAGTTGTACTTGTAACCTTAATGATATATATATTTAT
M8F
No sequence information

M9F
ACTAGTAGATTAAGCTTTCCTGGAGAAATGATATTCACCTCATTCAGGTTGTA
ACCAATTTCTCAGTTGTACTTGTAACCTTAATGATATATATATTTAT
M9R
ATAAATATATATATCATTAAGGTTACAAGTACAACTGAGAAATTGGTTACAA
CCTGAATGAGGTGAATATCATTTCTCCAGGAAAGCTTAATCACTAGT

M10R
CCAAACGGGGNAAATCACATTGAGAAAGTACATCCAAACAATTCATGGTCTA
AGTTCTTCAAAGGAAGCAGGAAACACCTCTCAAATTGAGCAACTATTTAGCC
CAGNATTATTACATACCAAATTTGAACAAAGCGAGAGACGCATTTAGTTATA
CCATCTAGCATCGCCTTCTATGCAGTATGCCCCTGACTCTTGCAGAGAAAATG
GTACTGACCGAGCTAGTCTTATTTGGTGGATTTAGGATCGTTCTTGAGAGCGT
AACCAGGAGGGGAATAGATGACAGCTGCGGCCATGGCTCCAGGAAAGCTTA
ATCACTAAGT
M10F
ACTAGTGATTAAGCTTTCCTGGAGCCATGGCCGCAGCTGTCATCTATTCCCCT
CCTGGTTACGCTCTCAAGAACGATCCTAAATCCACCAAATAAGACTAGCTCG
GTCAGTACCATTTTCTCTGCAAGAGTCAGGGGCATACTGCATAGAAGGCGAT
GCTAGATGGTATAACTAAATGCGTCTCTCGCTTTGTTCAAATTTGGTATGTAA
TAATCTGGGCTAAATAGTTGCTCAATTTGAGAGGTGTTTCCTGCTTCCTTTGA
AGAACTTAGACCATGAATTGTTTGGATGTACTTTCTCAATGTGATTTCCCCGT
TTAGG

M11F
ACTAGTGATTAAGCTTGCCATGGTAGGGTTTGTTTCAGCTCTGGCAGTGGAAC
TGTCCAAGGGCCAGGATGTGTTTACTCAGATCTCCGATGGTGGAGTTTCGTTG
TTCCTTGCCACCAGTATTTTGCTCTCAGTGGCATCTGCGATTCCTCTATTCAAA
GGAGTCAGCGTGGAGTCCAAATCAAATGGGCTCATGACGTCTGATGCCGAGC
TCTGGAACGGAAGGTTGGCCATGTTAGGTTTGGTTGCTTTGGCGTTCACTGAG
TATGTGAAGGGAGGGACTCTTGTTTGATCAGGTCCTAATTAATANGGGTGT

M11R
GGAGTGAAAAATTTCTTATAAAAACTGGTAAACCCTAATTACATGTGGAAGC
TTTTCGAAGGAATAATTAAGTTACATAACACCCGTATTAATTAGGACCTGATC
AAACAAGAGTCCCTCCCTTCACATACTCAGTGAACGCCAAAGCAACCAAACC
TAACATGGCCAACCTTCCGTTCCAGAGCTCGGCATCAGACGTCATGAGCCCA
TTTGATTTGGACTCCACGCTGACTCCTTTGAATAGAGGAATCGCAGATGCCAC
TGAGAGCAAAATACTGGTGGCAAGGAACAACGAAACTCCACCATCGGAGAT
CTGAGTAAACACATCCTGGCCCTTGGACAGTTCCACTGCCAGAGCTGAAACA
AACCCTACCATGGCAAAGCTTAATCACTAGT

M12F
GATTAAGCTTGCCATGGACTCGTAACCTCTGAACACAGATGGAAGCTACAGA
GTGCTAACCCCATTATATCGATCTCCCTAGTTTAAGTTTAAGATTCTTTTGGCT
TATTATTCCGGACTTTTTTTTCTTAGTTTCTTATTTTGGTTAAGCATATTTTTCA
GAGGCCTTAGAAAAGTTTAGGTTTGAGTGCGGCAAAATTCATGGTTCGGAAT
ATGTTCCAGGATTATTACGAGGATAAACTAACCGCTGTTCTTTCCCCTATTTTT
TACTTCTTACCCTTTCAAACCATATGGGATGTAACAGTTACTTTAATGGAAGG
AATGATACAATGATGGTGGTGCTTTATTTCC
M12R
GGAAATAAAAGCACCACCATCATTGTATCATTCCTTCCATTAAAGTAACTGTT
ACATCCCATATGGTTTGAAAGGGTAAGAAGTAAAAAATAGGGGAAAGAACA
GCGGTTAGTTTATCCTCGTAATAATCCTGGAACATATTCCGAACCATGAATTT
TGCCGCACTCAAACCTAAACTTTTCTAAGGCCTCTGAAAAATATGCTTAACCA
AAATAAGAAACTAAGAAAAAAAAGTCCGGAATAATAAGCCAAAAGAATCTT
AAACTTAACTAGGGAGATCGATATAATGGGGTTAGCACTCTGTAGCTTCCAT
CTGTGTTCAGAGGTTACGAGTCCATGGCAAGC

M13F
GATTAAGCTTGCCATGGNAACCCAAAGACTCATAGAAGGAACTTAGTGCTTG
CATGTGGAATGTTGAGTTTCTTGTTCTATTTGTGAGTTTTAGAGTTTTCCTAGT
TCTTTGTGTTTGTTTGACGCTATTTAATTTTTTGAAGGAGCCTCATCCCCCTTG
TTCCTTGTTCATGTGTAGTTTCTAAATTATACTTTTGTAAAAGGTAGATAAAT
AAATTGGTAAAATGTTTGATGTTGCTGTACTTGGGTGATGCAAATGTTTAAAA
CGGTTGAGCAGCACGCATGCTTGTAATGTATTCAAATTGGCCCC
M13R
GGGGCCAATTTGAATACATTACAAGCATGCGTGCTGCTCAACCGTTTTAAAC
ATTTGCATCACCCAAGTACAGCAACATCAAACATTTTACCAATTTATTTATCT
ACCTTTTACAAAAGTATAATTTAGAAACTACACATGAACAAGGAACAAGGGG
GATGAGGCTCCTTCAAAAAATTAAATAGCGTCAAACAAACACAAAGAACTAG
GAAAACTCTAAAACTCACAAATAGAACAAGAAACTCAACATTCCACATGCAA
GACAAGTTCCTTCATGAGTCTTTGGGTTCCATGGCAAGCTTAATC

M14F
GATTAAGCTTCTGCTGGAGTTGAGTTTAAACTTCAAACAATAAGTTTTCTAGT
TGTATTGCAGTCCTGTTTCTGCAAAATCACGCCTACAGAGTCTAACTGATTAT
TGAGAATTTTAATTTCGGAGTTGGAGTAGGATTTACTTGAATTTATGATATTT

GTACTAGTAATTTTTCGGGTGTGACTGAAATGGTAGGAAGGATTACACCCCTT
GTGATTGATTCTGAAGTTGAGTACTCGAGAGGGATCTTCTCGTTAGGAACAAT
TGTTCAAGGAATGAATGTCGAGGCGAAATAATAAATTTTTTACACCAC

M14R

GTGGTGTAAAAAATTTATTATTTCGCCTCGACATTCATTCCTTGAACAATTGT
TCCTAACGAGAAGATCCCTCTCGAGTACTCAACTTCAGAATCAATCACAAGG
GGTGTAATCCTTCCTACCATTTCAGTCACACCCGAAAAATTACTAGTACAAAT
ATCATAAATTCAAGTAAATCCTACTCCAACTCCGAAATTAAAATTCTCAATAA
TCAGTTAGACTCTGTAGGCGTGATTTTGCAGAAACAGGACTGCAATACAACT
AGAAAACTTATTGTTTGAAGTTTAAACTCAACTCCAGCAGAAGCTTAATC

M15F

GATTAAGCTTCTGCTGGAGAACAAGGACTTAAGGGAAACAATTGGCAAGGCA
GCACGTGAGGAAATGGAGAAATACGATTGGAAAGCAGCCACGCGAGTAATT
CGAAATGAGCAATACAATGCTGCCATTTGGTTCTGGAGGAAGAAGAGAGCGC
AATTTCTTAGACCTCTGCAGTGGTTCATGAAGCGCATTTTTCCTCCAACCACA
TCAGCAATCAAGTACAGGTGATTTCAGGTCCTACTGCTTTCAGCACAGGTAAC
TAGTTATTTGTAAACTTGCTTTGCGCTTCAAAAGAAGTTATTGTTGGATTATA
CTTGGGATAATTGTAGATTATTTGTTTGTAAATATGGAACGGTAGACAATGTC
AGTTACGTCGTTCTCTTGCT

M15R

AGCAAGAGAACGACGTAACTGACATTGTCTACCGTTCCATATTTACAAACAA
ATAATCTACAATTATCCCAAGTATAATCCAACAATAACTTCTTTTGAAGCGCA
AAGCAAGTTTACAAATAACTAGTTACCTGTGCTGAAAGCAGTAGGACCTGAA
ATCACCTGTACTTGATTGCTGATGTGGTTGGAGGAAAAATGCGCTTCATGAAC
CACTGCAGAGGTCTAAGAAATTGCGCTCTCTTCTTCCTCCAGAACCAAATGGC
AGCATTGTATTGCTCATTTCGAATTACTCGCGTGGCTGCTTTCCAATCGTATTT
CTCCATTTCCTCACGTGCTGCCTTGCCAATTGTTTCCCTTAAGTCCTTGTTCTC
CAGCAGAAGCTTAATC

M16F

GATTAAGCTTACGTTAGTTAGATTAGGGCAATTTGTCCCCTTGCACCCAAGTT
CGAATTTTCCTCCCCGTAGTATGGAATAGTTTAGAATAAGGAAATTATTTCTA
TACACCTTTTTCTCCATTTTTAT

M16R

ATAAAAATGGAGAAAAAGGTGTATAGAAATAATTTCCTTATTCTAAACTATT
CCATACTACGGGGAGGAAAATTCGAACTTGGGTGCAAGGGGACAAATTGCCC
TAATCAACTAACGTAAGCTTAATC

M18F

GATTAAGCTTCGAAATGGGCAGCACAAACGTGAGAATTGGATCGACAATATT
TGGGCCGAGAGAATATCCAAAGAAACAATGAAATTAGCTGGACTAATTTGAT
TCATGTTCTTGTAATTTTGTGCTATGATTTTGGAAAATTGTAGTGAGAGTTCA
ATGCCATTTGAAGGCTGGACATGCAAAACTCTGTTTCTTTAGAACTTATTGTT
TATGATAAGAACCATTCATATTATAAATTCTGTTGTATACCG

M18R

CGGTATACAACAGAATTTATAATATGAATGGTTCTTATCATAAACAATAAGTT
CTAAAGAAACAGAGTTTTGCATGTCCAGCCTTCAAATGGCATTGAACTCTCAC
TACAATTTTCCAAAATCATAGCACAAAATTACAAGAACATGAATCAAATTAG
TCCAGCTAATTTCATTGTTTCTTTGGATATTCTCTCGGCCCAAATATTGNCGAT
CCAATTCTTACGTTTGGGCTGCCCATTTCCAAGCTTAATC

M20F
ACTAGTGATTAAGCTTTTGAGGTGTTACTCATGCACATAATCAACTCGGTTGA
AAGCCCTTGTTCGTGTAAATGACTAATAAGTTCACCGGGAATTCCCAACTTCA
AGTTTATTTGAAGGGAAAATTTGTGAAATGTCATATAAACTTGTATCCGTTTT
TCAATTTGTCATATAAACTAAAATTTTAAACAATGTCATACCAANTTTCCG
M20R
CGGAAAGTTGGTATAGACATTGTTTAAAATTTTAGTTTATATGACAAATTGAA
AAACGGATACAAGTTTATATGACATTTCACAAATTTTCCCTTCAAATAAACTT
GAAGTTGGGAATTCCCGGTGAACTTATTAGTCATTTACACGAACAAGGGCTTT
CAACCGAGTTGATTATGTGCATGAGTAACACCTCAAAAGCTTAATCACTAGT

M21F
GATTAAGCTTACGTTAGAGGGATTTGGATGTTGGCTTAAATTATACAAGACTT
CGACTAAAACAACACTATCTGGAGTTTAAGCGCAGACGCGTTGGTTAGAGCA
AGTGTATTCTTTCTTGTGCATCCGAATTTAAATACTTTTTCACAATAATTCTGA
TTAGTTCCANGTAAATT
M21R
ATACAAATAATATTCTACTTGAAACTAATCAGAATTATTGTGAAAAAGTATTT
AAATTCGGATGCACAAGAAAGAATACACTTGCTCTAACCAACGCGTCTGCGC
TTAAACTCCAGATAGTGTTGTTTTAGTCGAAGTCTTGTATAATTTAAGCCAAC
ATCCAAATCCCTCTAACGTAAGCTTAATC

M22F
GATTAAGCTTACGTTAGATGTTAATAAACAAGGGAAAACCCAGAATCAAGCC
GAAAATAATAGAGGTTTAGGAGTGAATTTAGGAGTGAGTTTTTATGGTGGCT
GAGTGATGGTTATAGCTAGTGGATATTTGATGCGGTTTGGGACTTATGTGAGT
TTNGTTAATAAAGCGGATGGAGGGATGTTTGTCATCTTCTCTCTGCGTGTGCT
GG
M22R
CCAGCACACGCAGAGAGAAGATGACAAACATCCCTCCATCCGCTTTATTAAC
NAAACTCACATAAGTCCCAAACCGGATCAAATATCCACTAGCTATAACCATC
ACTCAGCCACCATAAAAACTCACTCCTAAATTCACTCCTAAACCTCTATTATT
TTCGGCTTGATTCTGGGTTTTCCCTTGTTTATTAACATCTAACGTAAGCTTAAT
C

M23F
ACTAGTGATTAAGCTTACGTTAGAGAATAAAAGAAGCCAACAGAAGTAGCA
ACCACCATACAGTACCCTGTTTGTATTATTGTGCCTGAGATCTAATTGAAATA
CCGCACAACGATCCAAAGACAAAAATAAACAAGTGTATGTAAGAAGAGAAA
TAATGCGTTTGG

M23R
CCAAACGCATTATTTCTCTTCTTACATACACTTAGTTTATTTTTGTCTTTGGAT
CGTTGTGCGTATTTCAATTAGATCTCAGGCACAATAANACAAACAGGGTACG
TATGGTGGNTGCTACTTCTGTTGGCTTCTTTTATTCTC

M24R
GGGGTAATGATAGGTAATTAAGCATAACAAGATAATACATTCAAACTCAGCA
ACCGCCTAAAACCCTGCACCCTGACGAAACATNATATTGATACATACATGGA
CTAAAGCTTAATC
M24F
GATTAAGCTTTATTCCATGTATNGTATNAATATGATAGTTACNNCAGGGTGCA
GGGTTTTAGGCGGCTGCTGNGTTNGAATGTATTATCTTGTTATGCTTAATTAC
CATCATTACCCCA

M25F
ACTAGTGATTAAGCTTCGAAATGATGATGCGTTGCTTAGTAAAACTTGGCATC
ACATAACGACAATGTTACGTGTAATATGTTCTTTTCACCATTAGGAATGTTAA
AGCTGCTGTGGAAATATTATTTTTCCGTATTGTATCAATCTTAGGTTGTATTTA
ATCCGATGGGAATGAGTTGTGCGAGAGATGTATAAGGATGAGTTAAGCATTT
CCTCAATCATGGTGGTAGACAATATTATCTTAGTTTTGAATGAATAACAGAAT
CATTTGTTC
M25R
GAACAAATGATTCTGTTATTCATTCAAAACTAAGATAATATTGTCTACCACCA
TGATTGAGGAAATGCTTAACTCATCCTTATACATCTCTCGCACAACTCATTCC
CATCGGATTAAATACAACCTAAGATTGATACAATACGGAAAAATAATATTTC
CACAGCAGCTTTAACATTCCTAATGGTGAAAAGAACATATTACACGTAACAT
TGTCGTTATGTGATGCCAAGTTTTACTAAGCAACGCATCATCATTTCGAAGCT
TAATCACTAGT

M27F
GATTAAGCTTACGTTAGATGTTAATAAACAAGGGAAAACCCAGAATCAAGCC
GAAAATAATAGAGGTTTAGGAGTGAATTTAGGAGTGAGTTTTTATGGTGGCT
GAGTGATGGTTATAGCTAGTGGATATTTGATGCGGTTTGGGACTTATGTGAGT
TTGGTTAATAAAGCGGATGGAGGGATGTTTGTCATCTTCTCTCTGCGTGTGCT
GGATATTTGGCTCAGACTCTCTTGGAAAAGAAATGGTAGTGTCACTTGCACTT
GCTTCCTTAATTATTATCTTCTATGATCTG
M27R
GCAGATCATAGAAGATAATAATTAAGGAAGCAAGTGCAAGTGACACTACCAT
TTCTTTTCCAAGAGAGTCTGAGCCAAATATCCAGCACACGCAGAGAGAAGAT
GACAAACATCCCTCCATCCGCTTTATTAACCAAACTCACATAAGTCCCAAACC
GCATCAAATATCCACTAGCTATAACCATCACTCAGCCACCATAAAAACTCAC
TCCTAAATTCACTCCTAAACCTCTATTATTTTCGGCTTGATTCTGGGTTTTCCC
TTGTTTATTAACATCTAACGTAAGCTTAATC

M28F
GATTAAGCTTACGTTAGGGCAAGGTTTGAGAGGGTGGTTGGATCAAAGGACT

CTGAAGCTTTGTACATGATAAATCCAGATGGTGCTGCAGGGCCAGAATTAAG
TATTTTCTTTGTCAGAGCTCATTAGTTTAATTAGCAAAAAATTATGTACTTTTT
TGATTAATTAATCTAGGTTATATATAATATAACTTTGTGTCTATGTAAACTAA
TTAAGGTTCTCAAATGCATGAATTGATTGTTTTGAGCTTCAATGATTGGAGTA
GTTGAGTCCC

M28R

GGAACTCAACTACTCCAATCATTGAAGCTCAAAACAATCAATTCATGCATTTG
GAGAACCTTAATTAGTTTACATAGACACAAAGTTATATNATATATAACCTAG
ATTAATTAATCAAAAAAGTACATAATTTTTTGCTAATTAAACTAATGAGCTCT
GACAAAGAAAATACTTAATTCTGGCCCTGCAGCACCATCTGGATTTATCATGT
ACAAAGCTTCAGAGTCCTTTGATCCAACCACCCTCTCAAACCTTGCCCTAACG
TAAGCTTAATC

M29F

GATTAAGCTTACGTTAGGGCAAGGTTTGAGAGGGTGGTTGGATCAAAGGACT
CTAGAAGCTTTGTACATGATTAATCCAGATGGTGCTGCAGGGCCAGAATTAA
GTATTTTCTTTGTCAGAGCTCATTAGTTTAATTAGCAAAAAATTATGTACTTTT
TTGATTAATTAATCTAGGTTATATATAATATAACTTTGTGTCTATGTAAACTA
ATTAAGGTTCTCAAATGCATGAATTGATTGTTTTGAGCTTC

M29R

AGAAGCTCAAAACAATCAATTCATAGCATTTGAGAACCTTAATTAGTTTACAT
AGACACAAAGTTATATTATATATAACCTAGATTAATTAATCAAAAAAGTACA
TAATTTTTTGCTAATTAAACTAATGAGCTCTGACAAAGAAAATACTTAATTCT
GGCCCTGCAGCACCATCTGGATTAATCATGTACAAAGCTTCAGAGTCCTTTGA
TCCAACCACCCTCTCAAACCTTGCCCTAACGTAAGCTTAATCC

M30F

GGGAATCGATTAAGCTTACGTTAGGAGGAGGCATCCGTGGGCTGATCCAAAA
ATTGTAAAGCAATTTATTGATGCAAAATTGCGTGAATTACTTGGTGAAAGGA
CAGCAGCAGATGATGAGAAGGTTCC

M30R

GGAACCTTCTCATCATCTGCTGCTGTCCTTTCACCAAGTAATTCACGCAATTTT
GCATCAATAAATTGCTTTACAATTTTTGGATCAGCCCACGGATGCCTCCTCCT
AACGTAAGCTTAATC

M31F

GATTAAGCTTACGTTAGGAAGAGGCATCCGTGGGCTGATCCAAAAATTGTGA
AGCAATTTATTGATGCAAAATTGCGTGAATTACTTGGTGAAAGGACAGCAGC
AGATGATGAGAAGGTTCC

M31R

GGAACCTTCTCATCATCTGCTGCTGTCCTTTCACCAAGTAATTCACGCAATTTT
GCATCAATAAATTGCTTCACAATTTTTGGATCAGCCCACGGATGCCTCTTCCT
AACGTAAGCTTAATC

APPENDIX 2

LIST OF PRIMERS USED FOR Q-RT-PCR OF cDNAs OBTAINED FROM DIFFERENTIAL DISPLAY.

M1F 5' GGCGAGCGCTTTACAAAT 3'
M1F 5' GGTTTCAGGGAGAGGGACA 3'

M4F 5' CATACGAGCCGGAAGCA 3'
M4R 5' TAATGTGAGTTAGCTCACTCATTAGG 3'

M5F 5' CCTTATTGTTGTTATGTTCTTGTTGAT 3'
M5R 5' CAAGCTCTATAATATAATACAACTAGCGT 3'

M6F 5' CCTAGTTTCTTGAGTATGGCCT 3'
M6R 5' AGCAAGAGAAGTTGACACAG 3'

M7F 5' CTTTCACAGTTGGGTTATCATTAATTG 3'
M7R 5' AAGTAACTCACGCGGAGA 3'

M10F 5' GAACGATCCTAAATCCACCAAATAAG 3'
M10R 5' ACCATCTAGCATCGCCT 3'

M11F 5' TTGCTCTCAGTGGCATCT 3'
M11R 5' TGATTTGGACTCCACGCT 3'

M12F 5' GTTTAGGTTTGAGTGCGGC 3'
M12R 5' GGTTAGTTTATCCTCGTAATAATCCTGG 3'

M13F 5' CATTTGCATCACCCAAGTACAG 3'
M13R 5' TGTTCCTTGTTCATGTGTAGTTTCTAA 3'

M14 5' CGCCTACAGAGTCTAACTGATTAT 3'
M14R 5' AGTAAATCCTACTCCAACTCCG 3'

M15F 5' AAGAAGTTATTGTTGGATTATACTTGGG 3'
M15R 5' TAACTGACATTGTCTACCGTTCCATATTTA 3'

M18F 5' TTCAATGCCATTTGAAGGCT 3'

M18R 5' CATAAACAATAAGTTCTAAAGAAACAGAGT 3'

M20F 5' CTCATGCACATAATCAACTCGG 3'
M20R 5' CGGTGAACTTATTAGTCATTTACACG 3'

M21F 5' TAAATTCGGATGCACAAGAAAGAATAC 3'
M21R 5' GGAGTTTAAGCGCAGACG 3'

M22F 5' ACCCAGAATCAAGCCGAA 3'
M22R 5' TATAACCATCACTCAGCCACCATA 3'

M23F 5' AAATACCGCACAACGATCC 3'
M23R 5' CCAAACGCATTATTTCTCTTCTTACAT 3'

M25F 5' AATGATGATGCGTTGCT 3'
M25R 5' CTACCACCATGATTGAGGAA 3'

M27F 5' ACCCAGAATCAAGCCGAA 3'
M27R 5' CCAAATATCCAGCACACGC 3'

M28F 5' GTTAGGGCAAGGTTTGAGAGG 3'
M28R 5' TGCAGCACCATCTGGATTTAT 3'

18SF 5' CCGACTTCTGGAAGGG 3'
18SR 5' TGTGGTAGCCGTTTCT 3'

APPENDIX 3

TABLE OF RPKM RATIOS FOR THE TOP 105 CONTIGS OVER-EXPRESSED IN THE RED PHENOTYPE cDNA LIBRARY GENERATED FROM mRNAseq.

Contig number	RPKM Green	RPKM Red	RPKM ratio	logbase10	logbase2
Consensus from Contig 1121	0	14,142	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 1974	0	7,115	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 2559	0	1,739	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 2572	0	6,294	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 2597	0	5,224	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 2598	0	5,123	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 3102	0	8,235	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 3148	0	0,000	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 3233	0	5,163	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 3273	0	6,063	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 3338	0	12,833	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 3655	0	2,754	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 4276	0	8,546	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 4736	0	2,789	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 4959	0	2,899	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 5046	0	0,732	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 5107	0	0,652	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 5209	0	57,469	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 5799	0	0,000	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 6069	0	7,203	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 699	0	2,287	#DIV/0!	#NUM!	#DIV/0!
Consensus from Contig 71	3,429	444,419	129,622	3,183	7,018
Consensus from Contig 7522	1,519	87,441	57,560	2,123	5,847
Consensus from Contig 4825	9,744	427,188	43,839	3,619	5,454
Consensus from Contig 4700	2,835	115,763	40,828	2,516	5,351
Consensus from Contig 482	0,365	13,189	36,142	0,682	5,176
Consensus from Contig 1847	2,175	75,697	34,804	2,217	5,121
Consensus from Contig 261	1,001	33,067	33,019	1,520	5,045
Consensus from Contig 3298	4,408	135,719	30,788	2,777	4,944
Consensus from Contig 3457	2,438	75,064	30,788	2,262	4,944

Consensus from Contig 4077	0,778	21,856	28,111	1,230	4,813
Consensus from Contig 5330	36,397	936,598	25,733	4,533	4,686
Consensus from Contig 1684	0,707	17,990	25,434	1,105	4,669
Consensus from Contig 7681	1,187	30,185	25,434	1,554	4,669
Consensus from Contig 1938	2,023	50,109	24,764	2,006	4,630
Consensus from Contig 6036	8,127	193,100	23,760	3,196	4,570
Consensus from Contig 3551	8,440	197,703	23,426	3,222	4,550
Consensus from Contig 1786	9,525	216,756	22,756	3,315	4,508
Consensus from Contig 1799	1,482	33,714	22,756	1,699	4,508
Consensus from Contig 654	1,282	29,182	22,756	1,573	4,508
Consensus from Contig 6748	1,062	24,162	22,756	1,409	4,508
Consensus from Contig 587	0,313	6,712	21,418	0,323	4,421
Consensus from Contig 882	7,314	156,656	21,418	3,059	4,421
Consensus from Contig 3664	1,492	29,950	20,079	1,650	4,328
Consensus from Contig 3536	0,362	7,279	20,079	0,421	4,328
Consensus from Contig 4853	5,960	114,360	19,187	2,834	4,262
Consensus from Contig 4319	11,385	217,526	19,106	3,394	4,256
Consensus from Contig 4539	3,325	62,863	18,908	2,320	4,241
Consensus from Contig 7251	1,050	19,686	18,741	1,316	4,228
Consensus from Contig 915	0,928	17,392	18,741	1,208	4,228
Consensus from Contig 7941	1,942	35,358	18,205	1,837	4,186
Consensus from Contig 1139	1,754	31,695	18,071	1,745	4,176
Consensus from Contig 3155	4,256	76,914	18,071	2,515	4,176
Consensus from Contig 4542	11,679	207,506	17,767	3,384	4,151
Consensus from Contig 413	9,434	166,275	17,625	3,196	4,140
Consensus from Contig 630	7,009	118,835	16,956	2,921	4,084
Consensus from Contig 2898	9,681	159,824	16,509	3,190	4,045
Consensus from Contig 2954	9,404	151,061	16,063	3,152	4,006
Consensus from Contig 7012	0,725	11,646	16,063	0,926	4,006
Consensus from Contig 3683	12,741	200,399	15,729	3,407	3,975
Consensus from Contig 3774	8,977	141,191	15,729	3,103	3,975
Consensus from Contig 7915	13,907	214,092	15,394	3,474	3,944
Consensus from Contig 3048	1,421	21,871	15,394	1,492	3,944
Consensus from Contig 1661	4,428	68,164	15,394	2,480	3,944
Consensus from Contig 7264	4,937	74,901	15,171	2,568	3,923
Consensus from Contig 7600	4,377	65,620	14,992	2,458	3,906
Consensus from Contig 4939	14,011	208,092	14,852	3,465	3,893
Consensus from Contig 3046	1,892	27,854	14,725	1,722	3,880
Consensus from Contig 2239	12,522	181,985	14,533	3,358	3,861
Consensus from Contig 7791	7,931	111,476	14,055	2,947	3,813
Consensus from Contig 4159	1,572	21,048	13,386	1,520	3,743
Consensus from Contig 3681	26,225	348,762	13,299	3,961	3,733
Consensus from Contig 1593	7,087	85,378	12,047	2,782	3,591
Consensus from Contig 7877	1,664	19,678	11,824	1,515	3,564
Consensus from Contig 701	3,622	42,669	11,780	2,189	3,558
Consensus from Contig 6124	7,729	90,533	11,713	2,845	3,550
Consensus from Contig 6699	6,672	78,146	11,713	2,717	3,550

Consensus from Contig 5722	2,025	23,313	11,512	1,674	3,525
Consensus from Contig 2224	2,133	24,266	11,378	1,714	3,508
Consensus from Contig 2756	16,121	183,431	11,378	3,471	3,508
Consensus from Contig 4543	0,591	6,728	11,378	0,600	3,508
Consensus from Contig 4718	27,759	310,540	11,187	3,936	3,484
Consensus from Contig 3957	5,906	64,824	10,977	2,583	3,456
Consensus from Contig 7298	11,186	122,513	10,952	3,137	3,453
Consensus from Contig 4522	4,149	44,430	10,709	2,266	3,421
Consensus from Contig 6888	7,443	79,706	10,709	2,773	3,421
Consensus from Contig 7588	4,926	52,754	10,709	2,415	3,421
Consensus from Contig 2703	1,130	12,099	10,709	1,136	3,421
Consensus from Contig 6731	11,977	125,587	10,486	3,177	3,390
Consensus from Contig 6805	7,177	74,532	10,384	2,728	3,376
Consensus from Contig 2306	15,292	154,988	10,135	3,375	3,341
Consensus from Contig 7363	6,711	68,014	10,135	2,659	3,341
Consensus from Contig 74	2,009	20,365	10,135	1,612	3,341
Consensus from Contig 1710	4,200	42,477	10,114	2,251	3,338
Consensus from Contig 4009	15,209	153,477	10,091	3,368	3,335
Consensus from Contig 7893	72,727	726,585	9,991	4,723	3,321
Consensus from Contig 7450	19,216	189,957	9,885	3,562	3,305
Consensus from Contig 6734	19,575	192,154	9,816	3,575	3,295
Consensus from Contig 80	20,325	199,515	9,816	3,608	3,295
Consensus from Contig 5449	9,101	88,322	9,705	2,905	3,279
Consensus from Contig 1345	11,482	110,661	9,638	3,104	3,269
Consensus from Contig 6678	9,785	94,073	9,614	2,964	3,265
Consensus from Contig 596	2,349	22,464	9,561	1,722	3,257
Consensus from Contig 7193	10,742	102,384	9,531	3,041	3,253
Consensus from Contig 888	15,537	147,892	9,519	3,361	3,251

APPENDIX 5

BLAST2GO ANNOTATION FOR THE TOP 105 CONTIGS OVER-EXPRESSED IN THE RED PHENOTYPE cDNA LIBRARY GENERATED FROM mRNAseq.

Contig number	Seq. Description	Seq bp	#Hits	min. eValue	mean Similarity
ConsensusfromContig1121	cell division cycle protein 48 homolog	697	20	3.73733E-131	97.95%
ConsensusfromContig1974	lil3 protein	833	20	1.57357E-116	76.95%
ConsensusfromContig2559	cobalt ion binding protein	380	20	3.43177E-31	91.95%
ConsensusfromContig2572	probable rhamnose biosynthetic enzyme 1-like	210	20	2.05548E-41	96.15%
ConsensusfromContig2597	epoxide hydrolase 2-like	253	20	7.24803E-46	86.0%
ConsensusfromContig2598	rac-like gtp-binding protein arac11	258	20	1.65488E-55	100.0%
ConsensusfromContig3102	uv excision repair protein	641	20	1.88488E-75	71.55%
ConsensusfromContig3148	alpha-l-arabinofuranosidase	254	20	7.18534E-48	93.4%
ConsensusfromContig3233	ethylene-responsive transcription factor rap2-4-like	768	20	2.61763E-90	66.95%
ConsensusfromContig3273	probable mitochondrial-processing peptidase subunit beta-like	327	20	4.67246E-63	93.75%
ConsensusfromContig3338	temperature-induced lipocalin	618	20	1.05599E-111	89.9%
ConsensusfromContig3655	26s protease regulatory subunit 8 homolog a-like	480	20	1.05826E-74	98.55%
ConsensusfromContig4276	mitochondrial 2-oxoglutarate malate carrier	232	20	2.29948E-24	90.25%
ConsensusfromContig4736	nadh dehydrogenase	237	20	4.99309E-34	87.1%
ConsensusfromContig4959	zinc transporter 11-like	228	20	1.0107E-27	71.6%
ConsensusfromContig5046	methyl-cpg-binding domain isoform 1	903	20	1.08508E-66	77.8%
ConsensusfromContig5107	protein tic chloroplastic-like	1013	20	0.0	81.75%
ConsensusfromContig5209	---NA---	207	0		
ConsensusfromContig5799	50s ribosomal protein chloroplastic-like	594	20	1.54678E-57	89.45%
ConsensusfromContig6069	mitochondrial import receptor subunit tom20-like	367	20	6.49967E-52	83.85%
ConsensusfromContig699	atp phosphoribosyltransferase-like	578	20	2.555E-53	82.95%
ConsensusfromContig71	glutathione s-transferase	864	20	1.1562E-127	87.25%
ConsensusfromContig7522	lecithin-cholesterol acyltransferase-like 1-like	325	20	6.56027E-48	80.75%

ConsensusfromContig4825	auxin-binding protein abp19a-like	608	20	5.09289E-104	89.3%
ConsensusfromContig4700	gdsl esterase lipase at5g33370-like	1392	20	6.13314E-173	88.5%
ConsensusfromContig482	succinate dehydrogenase	1353	20	0.0	97.55%
ConsensusfromContig1847	alpha-xylosidase 1-like	227	20	1.04182E-38	88.5%
ConsensusfromContig261	3-ketoacyl- thiolase peroxisomal-like	1479	20	0.0	90.3%
ConsensusfromContig3298	clathrin heavy chain 1-like	224	20	2.87136E-43	99.6%
ConsensusfromContig3457	cell number regulator 8-like	405	20	6.10587E-32	92.7%
ConsensusfromContig4077	clathrin heavy chain 1-like	635	20	4.47509E-122	97.45%
ConsensusfromContig5330	bifunctional inhibitor lipid-transfer protein seed storage 2s albumin superfamily protein	1153	20	1.97222E-45	94.65%
ConsensusfromContig1684	rhodanese-like domain-containing protein chloroplastic-like	698	20	2.90773E-103	76.8%
ConsensusfromContig7681	PREDICTED: uncharacterized protein LOC101301324	416	1	1.26629E-8	77.0%
ConsensusfromContig1938	Protein	488	20	1.67294E-67	84.0%
ConsensusfromContig6036	60s ribosomal protein l44-like	243	20	2.9062E-32	98.7%
ConsensusfromContig3551	cell wall	234	6	1.05247E-9	69.5%
ConsensusfromContig1786	10-formyltetrahydrofolate synthetase	311	20	2.39115E-60	99.45%
ConsensusfromContig1799	protein notum homolog	1333	20	0.0	82.2%
ConsensusfromContig654	cation transport regulator-like protein 2-like	770	20	2.65789E-138	87.1%
ConsensusfromContig6748	sec61 transport protein	465	20	1.0357E-103	99.8%
ConsensusfromContig587	protein root hair defective 3 homolog 2-like	3151	20	0.0	70.0%
ConsensusfromContig882	beta-xylosidase alpha-l-arabinofuranosidase 2-like	270	20	4.40879E-50	82.2%
ConsensusfromContig3664	PREDICTED: uncharacterized protein LOC100253416	331	1	4.41976E-4	50.0%
ConsensusfromContig3536	actin 1	1362	20	0.0	99.8%
ConsensusfromContig4853	40s ribosomal protein s17-4	497	20	9.90853E-84	92.0%
ConsensusfromContig4319	alpha-xylosidase 1-like	477	20	7.37391E-54	77.35%
ConsensusfromContig4539	protein root hair defective 3 homolog 2-like	2376	20	0.0	70.0%
ConsensusfromContig7251	atpase plasma membrane-type-like	470	20	7.21803E-68	95.9%
ConsensusfromContig915	areb-like protein	532	20	1.59557E-60	79.75%
ConsensusfromContig7941	cytochrome p450 77a3-like	1271	20	0.0	89.8%
ConsensusfromContig1139	30s ribosomal protein chloroplastic-like	563	20	5.29394E-120	83.65%
ConsensusfromContig3155	udp-galactose udp-glucose transporter 3-like	232	20	2.35005E-42	89.45%
ConsensusfromContig4542	---NA---	465	0		
ConsensusfromContig413	---NA---	314	0		
ConsensusfromContig630	non-symbiotic hemoglobin	634	20	1.0714E-73	95.7%
ConsensusfromContig2898	transaldolase-like protein	306	20	5.23066E-62	96.35%
ConsensusfromContig2954	squalene partial	210	20	2.22339E-40	99.8%
ConsensusfromContig7012	high mobility group family	681	20	1.01014E-25	84.75%

ConsensusfromContig3683	---NA---	310	0		
ConsensusfromContig3774	alpha-xylosidase 1-like	220	20	1.09788E-40	96.45%
ConsensusfromContig7915	---NA---	213	0		
ConsensusfromContig3048	sorbitol transporter	695	20	9.96527E-156	89.35%
ConsensusfromContig1661	clathrin interactor epsin 2-like	223	20	8.69075E-17	90.75%
ConsensusfromContig7264	phosphoenolpyruvate phosphate translocator chloroplastic-like	300	20	7.88303E-56	89.9%
ConsensusfromContig7600	eukaryotic translation initiation factor 1a-like	564	20	1.81288E-79	95.2%
ConsensusfromContig4939	auxin-binding protein abp19a-like	740	20	5.07922E-119	88.3%
ConsensusfromContig3046	tripeptidyl-peptidase 2-like	261	20	9.27076E-49	94.35%
ConsensusfromContig2239	probable pectate lyase 15-like	276	20	8.65231E-50	86.65%
ConsensusfromContig7791	---NA---	249	0		
ConsensusfromContig4159	probable xaa-pro aminopeptidase p-like	314	20	8.68207E-54	94.85%
ConsensusfromContig3681	glutathione s-transferase u17	866	20	2.26796E-140	84.85%
ConsensusfromContig1593	heat shock 70 kda protein 17-like	209	20	3.90558E-34	90.05%
ConsensusfromContig7877	probable nitrite transporter at1g68570-like	1780	20	0.0	85.8%
ConsensusfromContig701	1-aminocyclopropane-1-carboxylate oxidase	1363	20	0.0	94.45%
ConsensusfromContig6124	beta-galactosidase 3-like	511	20	2.33181E-107	91.45%
ConsensusfromContig6699	hypothetical protein PRUPE_ppa002580mg	296	1	4.32489E-4	81.0%
ConsensusfromContig5722	elongation factor 1-gamma-like	1219	20	5.91561E-92	96.3%
ConsensusfromContig2224	60s ribosomal protein l34-like	463	20	4.76251E-60	98.8%
ConsensusfromContig2756	abc transporter b family member 1-like	245	20	1.07737E-41	91.65%
ConsensusfromContig4543	uncharacterized partial	1670	20	8.98511E-49	72.6%
ConsensusfromContig4718	proline-rich protein 4-like	249	20	2.9626E-33	75.3%
ConsensusfromContig3957	omega-3 fatty acid desaturase	418	20	1.16672E-89	92.5%
ConsensusfromContig7298	transmembrane 9 superfamily member 4-like	971	20	0.0	97.5%
ConsensusfromContig4522	40s ribosomal protein s3-3-like	238	20	8.473E-48	96.55%
ConsensusfromContig6888	---NA---	199	0		
ConsensusfromContig7588	peroxisomal biogenesis factor 11 family protein	902	20	1.46909E-155	92.2%
ConsensusfromContig2703	protein argonaute 1	437	20	2.35228E-64	88.65%
ConsensusfromContig6731	psbb mrna maturation factor chloroplastic-like	742	20	6.4916E-170	93.15%
ConsensusfromContig6805	phenylalanine ammonia lyase	2270	20	0.0	95.1%
ConsensusfromContig2306	uncharacterized loc101217823	226	20	7.05597E-40	91.45%
ConsensusfromContig7363	heavy metal-associated isoprenylated plant protein 26-like	515	20	5.56534E-74	87.95%
ConsensusfromContig74	auxin efflux carrier component	1720	20	0.0	85.05%
ConsensusfromContig1710	translation initiation factor if- chloroplastic-like	1058	20	1.77696E-142	91.9%
ConsensusfromContig4009	60s ribosomal protein l17-2-like	422	20	2.01282E-	96.15%

				52	
ConsensusfromContig7893	auxin-binding protein abp19a-like	835	20	1.64766E-118	88.3%
ConsensusfromContig7450	alpha-xylosidase 1-like	334	20	4.28836E-60	86.7%
ConsensusfromContig6734	atp-citrate synthase alpha chain protein 1-like	227	20	3.06261E-42	95.0%
ConsensusfromContig80	subtilisin-like protease-like	583	20	8.90394E-97	78.4%
ConsensusfromContig5449	---NA---	217	0		
ConsensusfromContig1345	pleiotropic drug resistance protein 2-like	215	20	4.42457E-23	91.5%
ConsensusfromContig6678	plasma membrane h+-atpase	555	20	1.76726E-100	96.9%
ConsensusfromContig596	subtilisin-like protease-like	1471	20	0.0	81.8%
ConsensusfromContig7193	serpin-like protein	1149	20	0.0	81.3%
ConsensusfromContig888	delta 9 desaturase	286	20	2.22566E-39	85.1%