Identification of transcriptional regulators determining nutritional quality of tomato

Eleanor Fearnley

A thesis submitted for the degree of Doctor of Philosophy to the University of East Anglia

John Innes Centre

September 2018

Abstract

Consumption of dietary antioxidants and vitamins is advantageous for the prevention of diet-related diseases. Increasing the consumption of Vitamin E (VTE) in the human diet should have a plethora of health benefits; increased anti-inflammatory responses, reduced cardio-vascular risk in patients with diabetes and improved plasma membrane repair. Therefore, there are predicted benefits to increasing VTE in our diets. Biofortification attempts to increase VTE have been limited to manipulating enzymes of the VTE pathway, which often lead to the greater production of less bioactive vitamers. Transcriptional regulation of VTE synthesis in plants remains poorly understood, but identification of transcriptional regulators could overcome some of the limitations on biofortification of VTE in crops used for food.

In this thesis, I have used expression quantitative trait loci (eQTLs) to identify candidate transcriptional regulators of VTE biosynthesis in tomato fruit. Using the *S.pennellii* x *S.lycopersicum* cv. M82 tomato introgression lines (ILs), I have identified several trans-eQTLs, which contain candidate transcriptional regulators of VTE biosynthesis. The trans-eQTL IL9-3-2 revealed a MYB transcription factor – a candidate transcriptional repressor of the methyl erythritol phosphate (MEP) pathway, which supplies precursors for VTE biosynthesis. The trans-eQTL IL6-2-2 also revealed two MYB 1R transcriptional repressors of VTE biosynthesis in tomato fruit. Additionally, analysis of fruits of the IL parents revealed that *S.pennellii* fruits contain normally absent forms of VTE; tocotrienols. Overall, my work to identify transcriptional regulators of VTE biosynthesis has resulted in the production of VTE-enriched tomatoes.

Contents

Chapter 1: General introduction 2

1.1 Vitamin E 2

1.1.1 Vitamin E as an antioxidant 2

1.1.2 Vitamin E in humans 3

1.1.3 Vitamin E pathway 11

1.1.4 Vitamin E in plants 15

1.2 Plant metabolic engineering of VTE 18

1.3 Transcription factors 19

1.3.1 Transcriptional regulation of the VTE biosynthesis 20

1.3.2 Transcriptional regulation of the MEP pathway 22

1.3.3 Transcriptional regulation of the shikimate pathway 24

1.3.4 Post transcriptional regulation of the VTE pathway 26

1.3.5 Regulation of the VTE pathway is limited and needs clarification 26

1.4 Experimental approach to identify transcriptional factors that regulate VTE biosynthesis in tomato 27

1.4.1 *Solanum pennellii* introgression lines provide a unique genetic resource 27

1.4.1.1 Using ILs to discover expression QTLs – combining expression with phenotypes 28

1.4.1.2 Expression QTLs or metabolic QTLs, and their limitations 30

1.4.2 Viral Induced gene silencing (VIGS) 31

1.4.3 Transient over-expression using tobacco rattle virus 35

1.4.4 CRISPR/Cas9 35

Chapter 2: Materials and Methods 40

2.1 Materials 40

2.1.1 Chemicals 40

2.1.2 Antibiotics 40
2.1.3 Plant materials 40
2.1.4 Bacterial strains 42
2.1.5 Plasmids 42
2.1.6 Media recipes 42
2.2 Methods 42
2.2.1 Primer design 42
2.2.2 Polymerase Chain Reaction (PCR) 42
2.2.3 Purification of DNA from PCR products or agarose 44
2.2.4 Preparation of competent *Escherichia coli* cells 44
2.2.5 *E.coli* transformation 45
2.2.6 *Agrobacterium tumefaciens* competent cell preparation 45
2.2.7 *Agrobacterium* transformation 45
2.2.8 Plasmid DNA isolation from *E.coli* 46
2.2.9 Quantification of RNA and DNA 47
2.2.10 DNA isolation from plants 47
2.2.11 RNA isolation from plants 47
2.2.12 cDNA synthesis 49
2.2.12.1 First strand cDNA synthesis 49
2.2.12.2 Applied Biosystems High Capacity cDNA synthesis 49
2.2.13 Real time quantitative PCR (qRT-PCR) 49
2.2.14 Cloning 50
2.2.14.1 Restriction enzyme cloning 50
2.2.14.2 Gateway® cloning 51
2.4.14.3 Golden gate cloning 51
2.2.15 Viral Induced Gene Silencing (VIGS) 52
2.2.16 Transient over expression of genes in tomato fruit 53
2.2.17 Tomato transformation 53

2.2.18 Tocochromanol extraction and analysis 54
2.2.19 Carotenoid extraction and analysis 55
2.2.20 Statistical tests 56

Chapter 3: Identification of candidate genes using co-expression QTL analysis of *Solanum pennellii* x *S.lycopersicum* cv. M82 introgression lines 58

3.1 Introduction 58
3.1.1 Expression Quantitative Trait Loci (eQTL) analysis 58
3.1.2 *S.lycopersicoides* x *S.lycopersicum* cv. VF36 introgression lines 58
3.2 Materials and Methods 60
3.2.1 *S.pennellii* x *S.lycopersicum* cv. M82 IL RNA sequencing data and analysis 60
3.2.2 *Solanum pennellii* x *S.lycopersicum* cv. M82 IL parents RNA sequencing data analysis 61
3.2.3 *Solanum pennellii* x *S.lycopersicum* cv. M82 leaf RNA sequencing data analysis 61
3.2.4 *S.lycopersicoides* x *S.lycopersicum* cv. VF36 IL RNA sequencing and metabolite data analysis 61
3.3 Results 62
3.3.1 Screening of *Solanum pennellii* introgression lines RNA sequencing data for co-expression of the VTE pathway 62
3.3.2 Four trans-eQTL loci are identified for the VTE pathway 67
3.3.3 Co-expression analysis of *Solanum pennellii* x *S.lycopersicum* cv. M82 introgression lines reveals a trans-eQTL on chromosome 9 71
3.3.4 Co-expression analysis of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines reveals a trans-eQTL on chromosome 6 82
3.4 Discussion 90
3.4.1 Co-expression analysis of eQTLs is a powerful tool to find candidate genes 90

3.4.2 Cis-eQTLs are often linked to structural genes and mQTLs 94
3.4.3 The trans-eQTL IL9-3-2 contains putative trans-acting regulators of the VTE pathway and corresponds to *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs 95
3.4.4 The trans-eQTLs IL6-2-2 contains a trans-acting regulator putatively regulating the VTE and SK pathway 97
3.4.5 Candidate gene mining of trans-eQTLs has enabled the identification of several putative TFs regulating the VTE pathway 99
Chapter 4: Analysis of the Vitamin E pathway for the parents of the *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression lines 101
4.1 Introduction 101
4.1.1 *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression lines 101
4.1.2 The Vitamin E pathway 101
4.1.3 Regulation of the VTE pathway is limited and needs clarification 103
4.2 Materials and Methods 103
4.2.1 Plant Materials 103
4.3 Results 107
4.3.1 Spatio- and temporal control of VTE pathway gene expression exists within *S.lycopersicum* cv. M82 107
4.3.2 The MEP and SK pathway showed differential expression profiles compared to the VTE pathway in *S.lycopersicum* cv. M82 109
4.3.3 Tocopherol levels were altered during development of *S.lycopersicum* cv. M82 fruit in the epidermis of fruit, but, not in the pericarp 111
4.3.4 *S.pennellii* fruit demonstrated similar expression patterns for the VTE pathway genes as *S.lycopersicum* cv. M82 fruit 114
4.3.5 Differential expression of the MEP and SK pathway genes during *S. pennellii* fruit development 116

4.3.6 *S.pennellii* fruit contained more tocochromanols than *S.lycopersicum* cv. M82 119

4.4 Discussion 119

4.4.1 Transcriptional regulation of the VTE pathway is essential to maintain VTE levels in *S.lycopersicum* cv. M82 tomato fruit 119

4.4.2 MEP pathway gene expression could modulate VTE levels in *S.lycopersicum* cv. M82 fruit 124

4.4.3 Expression of SK pathway genes was induced early in *S.lycopersicum* cv. M82 fruit 125

4.4.4 Expression of VTE, MEP and SK pathway genes in *S.pennellii* fruit showed smaller changes during fruit development than the same genes in *S.lycopersicum* cv. M82 125

4.4.5 *S.pennellii* fruit contains a type of tocochromanol that is not normally found in tomato 128

4.4.6 *S.lycopersicum* cv. M82 showed differential expression of VTE biosynthetic genes and different metabolite profiles to *S.pennellii* fruit..... 130

Chapter 5: Characterisation of SlMYB79 from the trans-eQTL IL9-3-2..... 133

5.1 Introduction 133

5.1.1 Transcriptional regulation of the VTE, MEP and SK pathways 133

5.1.2 The trans-eQTL analyses revealed two trans-eQTL for candidate gene screening 133

5.2 Materials and Methods 134

5.2.1 Candidate gene mining of the trans-eQTL IL9-3-2 134

5.2.2 Plant Materials 134

5.2.3 Viral Induced Gene Silencing (VIGS)..... 135

5.2.4 Plasmid construction for stable transformation for over expression.... 135

5.2.5 Plasmid construction for CRISPR/Cas9 gene editing..... 136

5.2.6 Trolox equivalent antioxidant capacity assay x 136

5.2.7 Analysis of candidate R2R3 MYB transcription factors 137

5.3 Results 137

5.3.1 Screening TFs in the trans-eQTL IL9-3-2 interval provided one candidate for further analysis 137

5.3.2 SlMYB79 CRISPR lines suggested that SlMYB79 was a repressor of the genes encoding enzymes of the MEP pathway 147

5.3.3 Overexpression of SlMYB79 results in higher levels of tocopherols 159

5.3.4 Tomatoes overexpressing SlMYB79 had a higher antioxidant capacity than controls 168

5.3.5 Overexpression of SlMYB71 in Moneymaker tomatoes 168

5.3.6 Expression of SlMYB79 and SlMYB71 during tomato fruit development and ripening 177

5.3.7 Analysis of protein alignments between SlMYB79 and SlMYB71 179

5.3.8 Analysis of promoters of orthologs of SlMYB79 and SlMYB71 182

5.4 Discussion 188

5.4.1 Gene mining of the trans-eQTL IL9-3-2 and transient assays revealed one gene encoding a MYB TF that might regulate VTE biosynthesis transcriptionally 188

5.4.2 The SlMYB79 CRISPR knock-out lines suggested that SlMYB79 is a transcriptional repressor 189

5.4.3 Tomato fruit overexpressing SlMYB79 have higher tocopherol contents 190

5.4.4 The transcription factors SlMYB79 and SlMYB71 might regulate VTE biosynthesis indirectly 193

5.4.5 SlMYB79 may be a repressor in non-photosynthetic tissues 195

5.4.6 MYB79 and MYB71 might be differentially expressed in *S.pennellii* fruits, compared to *S.lycopersicum* cv. M82 196

5.4.7 Metabolic engineering of SMYB79 leads to nutritionally enhanced tomatoes 199

Chapter 6: Transient screening of candidate transcription factors from the trans-eQTL IL6-2-2 202
6.1 Introduction 202
6.1.1 The trans-eQTL analyses elucidated two regions for candidate gene screening 202
6.1.2 Transcriptional regulation of the VTE pathway 202
6.2 Materials and Methods 202
6.2.1 Candidate gene mining of the trans-eQTL IL6-2-2 202
6.2.2 Viral Induced Gene Silencing (VIGS) 203
6.2.3 Transient over expression of genes in tomato fruit 203
6.2.4 Phylogenetic analysis 203
6.3 Results 204
6.3.1 Screening TFs in the trans-eQTL IL6-2-2 interval provided three candidate TFs for further analysis 204
6.3.2 Transient silencing of the candidate TFs in *S.lycopersicum* cv. Moneymaker *Del/Ros* fruit 209
6.3.3 The candidate transcriptional regulators are differentially expressed during *S.lycopersicum* cv. M82 tomato development and ripening 216
6.3.4 Transient over expression of candidate TFs in *S.lycopersicum* cv Moneymaker tomato fruit suggested that they are repressors of VTE biosynthesis 216
6.3.5 Transient assays that silenced and overexpressed the candidate TFs suggested that they regulate the VTE biosynthesis 221
6.4 Discussion 223
6.4.1 Screening of the trans-eQTL IL6-2 TF candidates resulted in three putative TFs of the VTE pathway. 223
6.4.2 Transient assays of *SlTF2* showed that it may have an indirect role in transcriptional regulation of the VTE pathway 223

6.4.3 *SlTF4* is a possible negative regulator of the VTE pathway 225
6.4.4 *SlTF7* encodes a novel repressor of the VTE pathway, the expression of which may be modulated by light..... 226
6.4.5 The three candidate TFs may interact to control VTE biosynthesis and genes encoding enzymes of VTE biosynthesis 227
Chapter 7: Tocotrienol biosynthesis in *Solanum pennellii* tomato fruit 230
7.1 Introduction 230
7.1.1 *Solanum pennellii* fruits contain tocotrienols 230
7.1.2 Tocotrienol biosynthesis 230
7.1.3 The Vitamin E salvage pathway 230
7.2 Materials and Methods 231
7.2.1 Plant Materials 231
7.2.2 RNA sequencing data of *S.pennellii, S.lycopersicum* and *S.pennellii* x *S.lycopersicum* cv. M82 ILs 232
7.2.3 SNP analysis and genotyping of the BILs..... 232
7.2.4 Candidate gene mining of IL6-4 233
7.3 Results 233
7.3.1 Screening of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs and BILs revealed that tocotrienols are produced in IL6-4 233
7.3.2 SNP genotyping of the *S.pennellii* backcrossed ILs 237
7.3.3 Analysis of the backcrossed inbred lines revealed different metabolite and expression profiles 244
7.3.4 The gene(s) responsible for tocotrienol biosynthesis in *S.pennellii* fruit was not 'gained' or 'lost' from *S.pennellii* or *S.lycopersicum*..... 249
7.3.5 Gene mining of IL6-4 revealed that a gene encoding a transcription factor might affect flux of available substrates into the VTE pathway 252
7.3.6 Gene mining of IL6-4 revealed several candidate transcription factors 257
7.4 Discussion 257

7.4.1 Tocotrienols are present in *S.pennellii* and were identified in one IL, IL6-4 257
7.4.2 Tocotrienol biosynthesis correlated with beta carotene contents in the BIL tomato fruit 260
7.4.3 *S.pennellii* fruits are still photosynthetically active and might synthesise tocotrienols to maintain ROS homeostasis. 261
Chapter 8: General discussion and outlook 265
8.1.1 General summary 265
8.1.2 Expression Quantitative Trait Loci analysis 266
8.1.3 Transcriptional regulation of VTE biosynthesis in tomato fruits 268
8.1.4 VTE biosynthesis might be regulated by several TFs 269
8.1.4.1 The trans-eQTL IL9-3-2 revealed a putative transcriptional repressor that regulates genes encoding enzymes of the MEP pathway 270
8.1.4.2 Transient screening of candidate TFs of the trans-eQTL IL6-2-2 revealed two negative regulators of VTE biosynthesis 271
8.1.5 Tocotrienols 272
8.1.6 Future outlook 273
Appendix 275
List of plasmids 275
List of primers 280
Media recipes 301
Chapter 1 appendix 304
List of *S.pennellii* x *S.lycopersicum* introgression lines 304
Chapter 4 appendix 305
Chapter 5 appendix 317
Chapter 7 appendix 319
List of *S.pennellii* x *S.lycopersicum* introgression lines that were not analysed in Chapter 7 319

List of figures

Figure number	Title	Page
1-1	Chemical structures of tocopherols and tocotrienols	4
1-2	Outline of tocochromanol pathway	12-13
1-3	Schematic diagram of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines	29
1-4	Plasmid map of the viral induced gene silencing vectors	33
1-5	Viral induced gene silencing of Moneymaker *Del//Ros* tomatoes	34
1-6	Plasmid map of pTRV2 overexpression vectors	36
1-7	Schematic diagram of CRISPR/Cas9 gene editing	38
3-1	Schematic diagram of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines	59
3-2	Outline of tocochromanol pathway	63-64
3-3	Heatmap of gene expression of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines for the VTE pathway	65
3-4	Heatmap of gene expression of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines for the VTE pathway, including cis-eQTLs	66
3-5	Heatmap of gene expression of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines for the VTE pathway, including trans-eQTLs	68
3-6	Schematic diagram of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines on chromosome 9	72
3-7	Heatmap of gene expression of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines for the MEP pathway, including trans-eQTLs	73
3-8	Heatmap of gene expression of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines for the SK pathway, including trans-eQTLs	76
3-9	Heatmap of gene expression of leaf RNA sequencing data *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines for the VTE pathway, including trans-eQTLs	77
3-10	Schematic diagram of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines and *S.lycopersicoides* x *S.lycopersicum* cv. VF36 introgression lines on chromosome 9	80

3-11	Relative fold change of alpha tocopherol in *S.lycopersicoides* x *S.lycopersicum* cv. VF36 introgression lines	81
3-12	Schematic diagram of candidate gene mining of the trans-eQTL IL9-3-2	83
3-13	Schematic diagram of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines on chromosome 6	87
3-14	Schematic diagram of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines and *S.lycopersicoides* x *S.lycopersicum* cv. VF36 introgression lines on chromosome 6	89
3-15	Schematic diagram of candidate gene mining of the trans-eQTL IL6-2-2	91
4-1	Tomato fruit stages of *S.lycopersicum* and *S.pennellii*	102
4-2	Outline of the tocochromanol pathway	104-105
4-3	Relative fold change of gene expression of the VTE pathway during nine stages of *S.lycopersicum* fruit development	108
4-4	Heatmap of $\log_2$ relative fold changes of MEP pathway gene expression for nine stages of *S.lycopersicum* fruit development	110
4-5	Heatmap of $\log_2$ relative fold changes of SK pathway gene expression for nine stages of *S.lycopersicum* fruit development	112
4-6	Tocopherol contents of nine stages of *S.lycopersicum* fruit development	113
4-7	Relative fold change of gene expression of the VTE pathway during *S.pennellii* fruit development	115
4-8	Heatmap of $\log_2$ relative fold changes of MEP pathway gene expression during *S.pennellii* fruit development	117
4-9	Heatmap of $\log_2$ relative fold changes of SK pathway gene expression during *S.pennellii* fruit development	118
4-10	Tocochromanol contents during *S.pennellii* fruit development	120-121
5-1	Schematic diagram of candidate gene mining of the trans-eQTL IL9-3-2	138
5-2	Schematic diagram of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines and *S.lycopersicoides* x *S.lycopersicum* cv. VF36 introgression lines on chromosome 9	141
5-3	Relative fold change of tocopherols in silenced sectors of Moneymaker *Del/Ros* fruit	144
5-4	Relative fold change in expression of genes encoding enzymes of the VTE pathway and *SlMYB79* in silenced sectors of Moneymaker *Del/Ros* fruit	145-146
5-5	Schematic diagram of CRISPR guide RNAs and genotypes identified in the T1 generation	148

5-6	Truncated protein alignments of *SlMYB79* CRISPR lines	149
5-7	Tocopherol contents of *SlMYB79* CRISPR knock-out lines	150
5-8	Relative fold change in expression of genes encoding transcription factors in the *SlMYB79* CRISPR lines	152
5-9	Relative fold change in expression of genes encoding enzymes of the VTE pathway in the *SlMYB79* CRISPR lines	153
5-10	Heatmap of log_2 relative expression of genes encoding enzymes of the MEP pathway in the *SlMYB79* CRISPR lines	154
5-11	Relative fold change in expression of genes encoding enzymes of the MEP pathway in the *SlMYB79* CRISPR lines	155
5-12	Relative fold change in expression of genes encoding enzymes of the MEP pathway in the *SlMYB79* CRISPR lines	156
5-13	Heatmap of log_2 relative expression of genes encoding enzymes of the SK pathway in the *SlMYB79* CRISPR lines	157
5-14	Relative fold change in expression of genes encoding enzymes of the SK pathway in the *SlMYB79* CRISPR lines	158
5-15	Relative fold change in expression of genes encoding transcription factors in tomatoes overexpressing *SlMYB79*	160
5-16	Tocopherol contents in tomatoes overexpressing *SlMYB79*	161
5-17	Relative fold change in expression of genes encoding enzymes of the VTE pathway in tomatoes overexpressing *SlMYB79*	162
5-18	Heatmap of log_2 relative expression of genes encoding enzymes of the MEP pathway in tomatoes overexpressing *SlMYB79*	163
5-19	Relative fold change in expression of genes encoding enzymes of the MEP pathway in tomatoes overexpressing *SlMYB79*	165
5-20	Relative fold change in expression of genes encoding enzymes of the MEP pathway in tomatoes overexpressing *SlMYB79*	166
5-21	Heatmap of log_2 relative expression of genes encoding enzymes of the SK pathway in tomatoes overexpressing *SlMYB79*	167
5-22	Trolox-equivalent antioxidant capacity assay of tomatoes overexpressing *SlMYB79*	169
5-23	Phylogenetic tree of SlMYB79 and SlMYB71 with *Arabidopsis* R2R3 MYB transcription factors	170-171
5-24	Protein alignments of SlMYB79 and SlMYB71	172
5-25	Protein alignments of SlMYB79, SlMYB71, AtMYB79 and AtMYB71	174
5-26	Heatmap of log2 relative expression of genes encoding transcription factors and genes encoding enzymes of the	175

	VTE, MEP and SK pathway in tomatoes overexpressing *SlMYB71*	
5-27	Tocopherol contents of tomatoes overexpressing *SlMYB71*	176
5-28	Relative expression of SlMYB79 and SlMYB71 during *S.lycopersicum* cv. M82 tomato development and ripening	78
5-29	Protein alignments of SlMYB79 and SpMYB79	181
5-30	Protein alignments of SlMYB71 and SpMYB71	183
5-31	Promoter alignments of SlMYB79 and SpMYB79	185
5-32	Promoter alignments of SlMYB71 and SpMYB71	186
5-33	Model of the possible SlMYB79 regulatory mechanism	187
6-1	Schematic diagram of candidate gene mining of the trans-eQTL IL6-2-2	205
6-2	Relative fold change of tocopherol in silenced sectors of Moneymaker *Del/Ros* fruit that were transiently silenced with candidate transcription factors	211
6-3	Relative fold change in expression of candidate transcription factors in silenced sectors of Moneymaker *Del/Ros* fruit that were transiently silenced with candidate transcription factors	212
6-4	Relative fold change in expression of genes encoding enzymes of the VTE pathway in silenced sectors of Moneymaker *Del/Ros* fruit that were transiently silenced with candidate transcription factors	214-215
6-5	Relative fold change in expression of candidate transcription factors during *S.lycopersicum* cv. M82 fruit development and ripening	217
6-6	Tocopherol contents of fruits overexpressing candidate transcription factors transiently	218
6-7	Relative fold change in expression of candidate transcription factors in tomatoes overexpressing candidate transcription factors	219
6-8	Relative fold change in expression of genes encoding enzymes of VTE pathway in tomatoes overexpressing candidate transcription factors	220
6-9	Phylogenetic trees of candidate transcription factors with *Arabidopsis* transcription factors	222
7-1	Tocochromanol contents during *S.lycopersicum* cv. M82 fruit development and ripening	234
7-2	Screen of the *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines	235-236
7-3	Schematic diagram of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines on chromosome 6	238

7-4	Schematic diagram of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines and backcrossed inbred lines on chromosome 6	239
7-5	Photographs of fruit of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs and BILs	241
7-6	Outline of tocochromanol pathway	242-243
7-7	Tocochromanol and carotenoid contents of the *S.pennellii* introgression lines and backcrossed inbred lines	245-246
7-8	Heatmap of log2 relative expression of genes encoding enzymes of the VTE, MEP and SK pathways in the *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines and backcrossed inbred lines	248
7-9	Heatmap of relative expression of genes encoding enzymes of the MEP pathway from the RNA sequencing data of the *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines (IL6-3 and IL6-4)	256

List of tables

Table number	Title	Page
1-1	Relative affinities of α-tocopherol transfer protein of tocochromanol vitamers	5
1-2	Vitamin E composition of vegetables, fruits, nuts and oils	7-9
2-1	List of antibiotics, the concentrations used and use	41
3-1	Relative fold changes of genes encoding enzymes of the VTE pathway in the RNA sequencing data of *S.pennellii* x *S. lycopersicum* cv. M82 introgression lines	69
3-2	Expression ratios of genes encoding enzymes of the VTE pathway in parents of the *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines	70
3-3	Positions of genes encoding enzymes of the MEP and SK pathways in the *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines	75
3-4	Expression of genes encoding enzymes of the VTE pathway in the RNA sequencing data of *S.lycopersicoides* x *S.lycopersicum* cv. VF36 introgression lines	79
3-5	Candidate genes encoding transcription factors identified in the trans-eQTL IL9-3-2	84-85
3-6	Candidate genes encoding transcription factors identified in the trans-eQTL IL9-3-2	92-93
5-1	Relative expression values of genes encoding enzymes of the VTE pathway in the introgression lines of the trans-eQTL IL9-3-2	139
5-2	Relative expression values of genes encoding enzymes of the VTE pathway in the introgression lines of the *S.lycopersicoides* x *S.lycopersicum* cv. M82	142
5-3	Summary of viral induced gene silencing in Microtom *Del/Ros* fruit	143
5-4	Relative expression of candidate transcription factors in the RNA sequencing data of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines	180
6-1	Relative expression of genes encoding enzymes of the VTE pathway and genes encoding candidate transcription factors of the RNA sequencing data of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines in the trans-eQTL IL6-2-2	206
6-2	Expression ratios of genes encoding transcription factors in parents of the *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines	207

6-3	Relative expression values of candidate transcription factors identified in the trans-eQTL IL6-2-2 in the leaf RNA sequencing data of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines	208
6-4	Summary of viral induced gene silencing of candidate transcription factors identified in the trans-eQTL IL6-2-2 in Microtom *Del/Ros* fruit	210
7-1	SNP genotyping of the *S.pennellii* x *S.lycopersicum* cv. M82 backcrossed inbred lines	240
7-2	Genes that lie in the IL6-4 region in *S.pennellii,* but not in the *S.lycopersicum* region	250
7-3	Genes that lie in the IL6-4 region in *S.lycopersicum,* but not in the *S.pennellii* region	251
7-4	Relative expression of candidate genes from the RNA sequencing data of the *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines	253
7-5	Relative expression of genes encoding enzymes of the VTE pathway from the RNA sequencing data of the *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines (IL6-3 and IL6-4)	255
7-6	Relative expression of candidate genes identified in IL6-4 in the RNA sequencing data of the *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines and the parents	258

Abbreviations

AAA(s)	Aromatic amino acid(s)
AA	Amino acid
ABA	Abscisic acid
ACP	Acyl-acyl carrier protein
AP2	Apetala 2
AsA	Ascorbic Acid
At	*Arabidopsis thaliana*
αTTP	α tocopherol transfer protein
ATAF	*Arabidopsis* transcription activation factor
B	Breaker
bHLH	Basic Helix Loop Helix
BILs	Backcrossed inbred lines
BP	Base pairs
bZIP	Basic leucine zipper domain
CDS	Coding sequence
ChIP-seq	Chromatin immunoprecipitation sequencing
CMS	2C-Methyl-D-erthritol-4-phosphate cytidyltransferase
CM	Chorismate mutase
CP	Coat protein
CS	Chorismate synthase
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNAs	CRISPR RNAs
CRTR-B	β-ring hydroxylase
CRTR-E	ε-ring hydroxylase
CUC	Cup-shaped cotyledon
CVD	Cardiovascular disease
CYCβ	Chromoplast-specific lycopene cyclase
DAHPS	3-Deoxy-D-arabino-hepulosonate
Del	Delila
DET1	DE-ETIOLATED 1
DHQS	3-Dehydroquianate synthase
DNA	Deoxyribonucleic acid
DPA	Days post anthesis
dsRNA	double-stranded RNA
DXS	1-Deoxy-D-xylulose-5-P-synthase
DXR	2-C-Methyl-D-erythritol-4-phosphate synthase
DW	Dry weight
EIN3	Ethylene insensitive 3
EOBI	EMISSION OF BENZENOIDS I
EOBII	EMISSION OF BENZENOIDS II
EPSPS	5-Enolpyruvylshikimate-3-P-synthase
ERF	Ethylene responsive factor

FW	Fresh weight
eQTLs	Expression quantitative trait locus
GGDP	Geranylgeranyl diphosphate
GGDR	Geranylgeranyl reductase
GGPS	Geranylgeranyl pyrophosphate synthase
GOI	Gene of interest
GPPS	Geranyl pyrophosphate synthase
gRNA	guide RNA
GTFs	General transcription factors
HDR	4-Hydroxy-3-methylbut-2-enyl-diphosphate reductase
HDS	4-Hydroxy-3-methylbut-2-enyl-diphosphate synthase
HD-ZIP	Homeodomain and leucine zipper
HGGT	Homogentisate geranylgeranyl transferase
HPPD	4-Hydroxyphenylpyruvate dioxygenase
HR	Homologous recombination
IG	Immature green
IPI	Isopentyl diphosphate δ isomerase
ISPE	4-2-C-Methyl-D-erythritol kinase
ISPF	2-C-Methyl-D-erythritol-2-3-cyclodiphosphate synthase
IL(s)	Introgression line(s)
LB	Left border
LDL	Low density lipoprotein
LYCβ	Lycopene β cyclase
LYCε	Lycopene ε cyclase
MCS	Multiple cloning site
MGGBQ	2-Methyl-6-geranylgeranylquinol
MG	Mature green
MEP	Methyl erythritol phosphate
MP	Movement protein
MPBQ	2-Methyl-6-phytyllquinol
mQTLs	Metabolic quantitative trait locus
MT	Methyltransferase
MYB	Myeloblastosis
NAC	NAM, ATAF, and CUC
NAM	No apical meristem
NHEJ	Non-homologous end joining
Nos T	Nopaline synthase terminator
O_2^-	Singlet oxygen
ODO1	ODORANT1
OE	Over expression
OH	Hydroxyl group
ORF	Open reading frame
PAM	Protospacer adjacent motif
PAT	Prephenate aminotransferase

PC-8	Plastochromanol-8
PDP	Phytyl diphosphate
PEBV pro	Pea early browning virus promoter
PHY	Phytochromes
PIF	Phytochrome interacting factor
POI	Promoter of interest
PSII	Photosystem II
PSY	Phytoene synthase
PTGS	Post-transcriptional gene silencing
PUFA	Poly-unsaturated fatty acid
QTLs	Quantitative Trait Locus
qRT-PCR	Real time quantitative reverse transcription PCR
RB	Right border
RDA	Recommended daily allowance
RdRp	RNA dependent RNA polymerase
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
Ros	Rosea 1
ROS	Reactive oxygen species
Rz	Self-cleaving ribozyme
RPKM	Reads per kilobase million
SAM	S-adenosyl methionine
SDH-DHQS1	3-Dehydroquianate dehydratase
SDH-DHQS2	Shikimate 5-dehydrogenase
SK	Shikimate
SK	Shikimate kinase
ssRNA	single-stranded RNA
TAD	Transcriptional activation domain
TAT	Tyrosine aminotransferase
TBPs	TATA binding proteins
TCP	Teosinte branched 1, Cycloidea and PCF
TEAC	Trolox equivalent antioxidant capacity
TF(s)	Transcription factor(s)
Ti	Tumour inducing
trancrRNA	transactivating CRISPR RNA
TRV	Tobacco rattle virus
TyrA	Arogenate dehydrogenase
UTR	Untranslated region
VIGS	Viral induced gene silencing
VTE	Vitamin E
VTE2 (HPT)	Homogentisate phytyl transferase
VTE3 (MPBQMT)	Dimethyl-phytylquinol methyl transferase
VTE1 (TC)	Tocopherol cyclase
VTE4 (γ-TMT)	γ-Tocopherol C-methyl transferase

VTE5 (PK)	Phytyl kinase
VTE6 (PPK)	Phytyl-phosphate kinase
WRI1	WRINKLED 1
WT	Wild type
3'UTR	3' untranslated region
16K	16kDa cysteine rich protein
35S P	Cauliflower mosaic virus 35S constitutive promoter

Acknowledgements

First, I would like to thank my supervisor, Cathie Martin. Cathie has always encouraged me to be independent in my research and always provides advice when needed. I would also like to thank my supervisory team members, Cristobal Uauy and Janneke Balk for their scientific contributions in meetings.

I have been very lucky during my PhD as I have met so many wonderful and enthusiastic people at the John Innes Centre. There are many members of Cathie's lab, both past and present, who I am grateful to for their advice over the years. But, most importantly, have provided much amusement and laughter. I would like to thank Ingo, Dario and Eugenio in particular, as I would not be writing this thesis without their continued support and patience with my never-ending questions. I would also like to thank my fellow PhD students (Javi and Jie), who have been the moral support throughout my PhD.

I am grateful for all the support from Metabolite services and for their guidance. Baldeep has been a good teacher and I have really enjoyed learning about metabolite analysis.

I would like to thank my University friends, who have been asking since day one 'when are you going to start writing?'. I would also like to thank Stef, who has the most impeccable timing, and knows exactly when a call is needed!

My family have always supported me, and my University years have been some of the toughest that we have been through as a family, and often it was not appreciated. So, I would like to thank you for always being on the end of the phone, no matter the time or day.

And of course, there is Charlie. Where would I be without your support. Somehow you always seem to be there, no matter what happens. I guess I could thank you more, but we are not like that and there is only one thing that needs to be said. TTA.

Chapter 1: General Introduction

Chapter 1: General introduction

1.1 Vitamin E

Plant secondary metabolites offer a unique perspective to improve health. As global population increases, health related diseases are on the rise, and there is a significant advantage in consuming more secondary metabolites for prevention of chronic disease. It is well known that an increase in fruit and vegetable consumption, rich in health beneficial compounds such as polyphenols and isoprenoids, aids in prevention of cardiovascular diseases (CVD), cancer and diabetes (Fitzpatrick et al., 2012, Kris-Etherton et al., 2002, Wang et al., 2014). The health beneficial compounds found in fruits and vegetables are known as vitamins, minerals and phytonutrients. Vitamins are considered essential micronutrients that are needed for human metabolism to function. Vitamins are not synthesised by humans, therefore all vitamins must be sourced from our diets. Most vitamins are obtained from plant sources, which often are products of specialised metabolism in plants. These vitamins often have beneficial effects in plants, such as an antioxidant ability, as well as having health benefits in humans. Some of the vitamins produced by plants include; vitamin A (carotenoids), vitamin C (ascorbic acid), vitamin E (VTE - tocochromanols), vitamin K (quinones), vitamin D (calciferols) and B vitamins (B_1- thiamine, B_2-riboflavin, B_3- niacin, B_5 – pantothenic acid, B_6 – pyridoxine, pyridoxal and pyridoxamine, B_8 – biotin and B_9 -folate).

VTE is an essential vitamin that was discovered over 90 years ago, as an essential reproductive nutrient (Evans and Bishop, 1922). Since its discovery, VTE has been shown to be a potent antioxidant in human diets (Levy et al., 2004, Kirsh et al., 2006, Jiang, 2014). The recommended daily allowance (RDA) for VTE is often met (15mg per day for adults over the age of 14), therefore its health benefits are normally associated with increasing consumption over RDA (NIH, 2016).

1.1.1 Vitamin E as an antioxidant

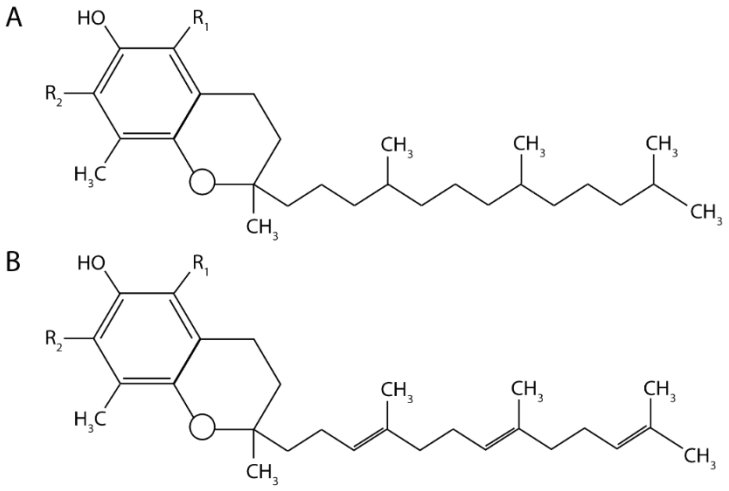
The term VTE describes a group of lipophilic compounds, which reside within the diverse prenylquinone group, including; plastoquinones and vitamin K_1

(Collakova and DellaPenna, 2001). Also known as tocochromanols, the group can be further classified into tocopherols and tocotrienols. Each has 4 forms (alpha - α, beta - β, gamma - γ and delta - δ); containing a polar chromanol head group and a prenyl chain. The isoforms differ only by their methyl group substitution on the chromanol head (figure 1-1). Variation between the tocopherols and tocotrienols exists due to the saturated prenyl tail of tocopherols, which contrasts with the three double bonds in the tail of tocotrienols (figure 1-1) (DellaPenna, 2005, Schneider, 2005).

The antioxidant ability of VTE vitamers stems from their ability to protect polyunsaturated fatty acids (PUFA) from oxidation and their roles in maintenance of membrane integrity. Donation of a hydrogen to a lipid peroxyl radical during non-enzymatic lipid peroxidation prevents formation of potentially damaging radicals and protects PUFAs. VTE can also scavenge reactive oxygen species (ROS) and quench singlet oxygen (O_2^-); preventing damage to cells. Alpha forms of tocochromanols are the most active form as an antioxidant of VTE because of the substitutions in their chromanol ring (figure 1-1). The chromanol ring facilitates transfer of hydrogen from the hydroxyl (OH) group to a peroxyl radical, during ROS scavenging (Schneider, 2005).

1.1.2 Vitamin E in humans

Within humans, α-tocopherol is the most bioactive form of VTE. This form of VTE does not have the highest antioxidant ability in the human body, but, is preferentially retained and bound by the hepatic α-tocopherol transfer protein (αTTP), resulting in higher levels of α-tocopherol in comparison with other tocochromanols in blood plasma (Hosomi et al., 1997). The αTTP enhances the transfer of VTE through membranes (Hosomi et al., 1997). Table 1-1 shows the



	Chromanol Head Substitution	
Tocochromanol form	R_1	R_2
Alpha	CH_3	CH_3
Beta	CH_3	H
Gamma	H	CH_3
Delta	H	H

Figure 1-1 Chemical structure of (A) Tocopherols and (B) Tocotrienols. There are eight forms of tocochromanols; alpha (α)-tocopherol, beta (β)- tocopherol, gamma (γ)-tocopherol, delta (δ)-tocopherol, α-tocotrienol, β-tocotrienol, γ-tocotrienol and δ-tocotrienol. The tocopherols and tocotrienols differ due to the saturation of the prenyl chain (A) and (B). Whereas, the tocopherol and tocotrienol vitamers differ based on the chromanol head substitution (R1 and R2). The substitutions are listed in the key.

Table 1-1 Relative alpha tocopherol transfer protein (αTTP) affinities of tocochromanol vitamers, taking α-tocopherol affinity as 100%. Adapted from Hosomi et al. (1997).

Competitors	Relative affinity (%)
α-tocopherol	100
β-tocopherol	38.1 ± 9.3
γ-tocopherol	8.9 ± 0.6
δ-tocopherol	1.6 ± 0.3
α-tocotrienol	12.4 ± 2.3

relative affinity of the αTTP protein for other VTE vitamers, which have a reduced affinity of binding (Schneider, 2005, Hosomi et al., 1997, Lim and Traber, 2007). It has been suggested that γ-tocopherols and tocotrienols have a greater antioxidant capacity than the other VTE vitamers (Jiang et al., 2001, Jiang, 2014). However, less active forms of VTE and synthetic vitamers are excreted in the short term, via the urine, in contrast with natural forms (Traber et al., 1998, Burton et al., 1998).

As humans lack the ability to synthesise VTE - all VTE is obtained from the diet. Deficiencies tend to be limited to malnutrition or to individuals who cannot retain VTE within the body (Fitzpatrick et al., 2012). As VTE synthesis is exclusive to photosynthetic organisms; VTE-rich foods consist mainly of plant oilseeds (Ahsan et al., 2015). Notable examples include: wheat germ, sunflower seeds, almonds, spinach and olive oil (table 1-2). Most monocotyledonous plants such as wheat, accumulate more tocotrienols rather than tocopherols, which is not ideal nutritionally, since they are excreted (Burton et al., 1998). Oilseeds tend to have a higher tocotrienol content and in the US, dietary VTE mainly consists of oilseed products (oils and margarines). Thus, US diets might have suboptimal levels of VTE (Jiang et al., 2001, Maras et al., 2004). In general, the public are not gaining the benefits of a high VTE diet, let alone acquiring the most bioactive form of VTE, which highlights the potential importance of improving VTE levels and the nutritional quality of crops.

VTE's ability to aid in the prevention of chronic diseases stems from increasing consumption above RDA. Its role in improving CVD has been shown *in vitro* – it prevents atherosclerosis by oxidising low density lipoprotein (LDL) (Bowry et al., 1992, Loffredo et al., 2015). This has had a knock on effect on reducing CVD risk for those affected by Type 2 diabetes or those suffering from high oxidative stress (Blum et al., 2010, Boaz et al., 2000, Milman et al., 2008). VTE also has an anti-inflammatory effect and a possible role in preventing prostate cancer (Dietrich et al., 2006, Venkateswaran et al., 2004). However, there is controversy surrounding VTE supplementation and the associated risk of prostate cancer (Ju et al., 2010, Kirsh et al., 2006)

Table 1-2 Vitamin E composition of vegetables, fruits, nuts and oils. The tables shows alpha tocopherol, alpha tocotrienol and gamma tocopherol contents (mg per 100mg). This table was generated using the USDA Food composition database (U.S. Department of Agriculture, 2018). The dashes (--) denote when tocochromanols were not determined.

Description	alpha-tocopherol (mg per 100g)	alpha-tocotrienol (mg per 100g)	gamma-tocopherol (mg per 100g)
Oil, wheat germ	149.4	--	--
Oil, hazelnut	47.2	--	--
Peanut butter, smooth, vitamin and mineral fortified	43.2	--	--
Oil, sunflower, linoleic (less than 60%)	41.08	--	--
Oil, almond	39.2	--	--
Margarine-like spread, benecol light spread	37.22	0	13.8
Seeds, sunflower seed kernels, oil roasted, without salt	36.33	0	0.46
Nuts, almonds	25.63	--	0.64
Nuts, hazelnuts or filberts, blanched	17.5	--	2.15

Table 1-2 continued.

Oil, canola	17.46	0.03	27.34
Margarine-like spread, smart balance omega plus spread (with plant sterols & fish oil)	11.78	2.66	15.9
Peanuts, all types, raw	8.33	--	--
Peppers, jalapeno, raw	3.58	0.02	0.05
Spinach, frozen, chopped or leaf, unprepared	2.9	0	0.14
Turnip greens, raw	2.86	--	0.16
Egg, yolk, raw, fresh	2.58	0.03	1.33
Broccoli, cooked	2.53	--	0.32
Blueberries, dried, sweetened	2.35	--	--
Cranberries, dried, sweetened (Includes foods for USDA's Food Distribution Program)	2.1	--	0.17
Spinach, raw	2.03	0	0.18
Tomato products, canned, puree, without salt added	1.97	--	0.24
Kale, cooked, boiled, drained, without salt	1.61	0	0.12

Table 1-2 continued.

Asparagus, cooked, boiled, drained	1.5	--	0.21
Kiwifruit, green, raw	1.46	0	0.03
Squash, winter, butternut, raw	1.44	--	--
Blackberries, raw	1.17	0	1.34
Asparagus, raw	1.13	--	0.09
Parsnips, cooked, boiled, drained, without salt	1	--	--
Beans, white, mature seeds, cooked, boiled, without salt	0.94	--	--
Mangos, raw	0.9	0.03	0.01
Brussels sprouts, raw	0.88	--	--
Collards, cooked, boiled, drained, without salt	0.88	--	--
Raspberries, raw	0.87	0	1.42
Tomatoes, red, ripe, cooked	0.56	0	0.21
Oil, flaxseed, cold pressed	0.47	0.87	28.76

and there is a suggestion that only δ-tocopherol can inhibit prostate cancer cell growth *in vitro* (Wang et al., 2015).

Controversy surrounds VTE and its putative role in improving health, because high doses of VTE have been associated with negative side effects (maximum RDA 1000/800mg for men and women, respectively). However, moderate consumption of VTE has been linked to enhanced immune responses of individuals with high oxidative stresses. α-tocopherol supplementation reduced pulmonary inflammation of aged mice that were infected with bacterial pneumonia (Bou Ghanem et al., 2015). These authors showed that tocopherol could modulate innate immune response in aged individuals. Additionally, moderate increases in VTE supplementation (200mg per day) of patients over 65 years resulted in enhanced immune responses to vaccinations (diptheria, hepatitis B and tetanus) (Meydani et al., 1997). The moderate supplementation of VTE in this group resulted in elevated antibody levels compared to the placebo, control group, and those supplemented with 60mg per day or 800mg per day of α-tocopherol (Meydani et al., 1997). It has been suggested VTE's role in immune response stems from its antioxidant capacity. ROS in human contribute to age-related T-cell decline, and many enzymes involved in the immune and inflammatory response are susceptible to oxidative stress (Mocchegiani et al., 2014). Therefore, VTE's role in humans might be to limit ROS-mediated damage of the immune response that is exacerbated by age (Mocchegiani et al., 2014).

Recently, tocotrienols have been associated with having beneficial effects on health (Aggarwal et al., 2010). Human subjects were supplemented with a tocotrienol rich fraction (160mg per day consisting of 74% tocotrienols and 26% tocopherols) and it resulted in lower plasma cholesterol (Chin et al., 2008). Less DNA damage was observed in these subjects, which is a contributing factor to age-related DNA instability associated with carcinogenesis (Chin et al., 2008, Chin et al., 2011). There is also great interest in VTE's role and its neuroprotective properties. Tocotrienol supplementation in mice prevented ischemic stroke-induced injury of the brain compared with control (Khanna et al., 2005, Park et al.,

2011) . Tocotrienols may have other beneficial effects on health that have not yet been studied, and research in this area is increasing (Aggarwal et al., 2010).

Studies which have described negative side effects of VTE involve consumption of VTE as a supplement, rather than acquiring VTE from food. Synthetic forms of VTE (supplementation) are preferentially excreted by the body, in contrast to natural forms (Burton et al., 1998, Traber et al., 1998), and doses up to 800 mg have been shown to be harmless (Meydani et al., 1998). Therefore, the idea that changes in diet can improve health, coupled with the ability to improve nutritional quality in crops, as a natural product, makes VTE an interesting candidate to study.

1.1.3 Vitamin E pathway

VTE is synthesised in photosynthetic organisms only. The genes encoding enzymes of the biosynthetic VTE pathway were characterised initially in cyanobacteria (*Synechocystis* sp. PCC 6803) – the equivalent genes have been identified within *Arabidopsis thaliana* (Schledz et al., 2001, Collakova and DellaPenna, 2001, Savidge et al., 2002, Cheng et al., 2003, Valentin et al., 2006, vom Dorp et al., 2015, Bergmuller et al., 2003).

VTE precursors derive from the plastidic non-mevalonate/methyl erythitol phosphate (MEP) and the shikimate (SK) pathways (figure 1-2). The MEP pathway produces many precursors for synthesis of other metabolites, including carotenoids, chlorophyll, gibberellins and abscisic acid (ABA). Similarly, the SK pathway produces precursor metabolites for biosynthesis of phenylpropanoids, as well as metabolites derived from tyrosine and tryptophan. The VTE pathway bifurcates to form tocopherols and tocotrienols (figure 1-2), but these pathway branches are dependent on precursors derived from the afore mentioned MEP and SK pathways.

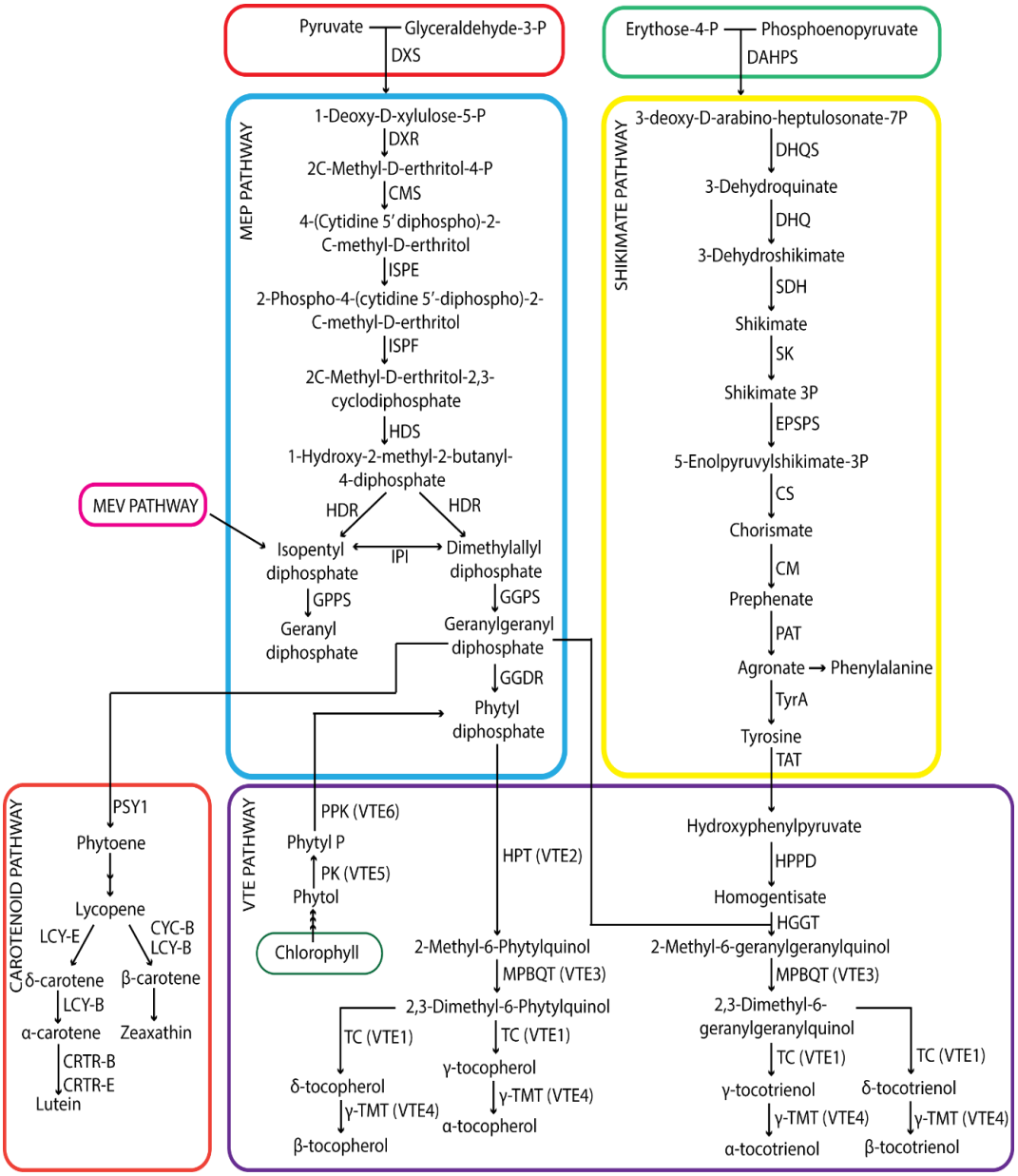

Pyruvate
Glyceraldehyde-3-P
DXS
Erythose-4-P
Phosphoenopyruvate
DAHPS
MEP PATHWAY
1-Deoxy-D-xylulose-5-P
DXR
2C-Methyl-D-erthritol-4-P
CMS
4-(Cytidine 5′diphospho)-2-C-methyl-D-erthritol
ISPE
2-Phospho-4-(cytidine 5′-diphospho)-2-C-methyl-D-erthritol
ISPF
2C-Methyl-D-erthritol-2,3-cyclodiphosphate
HDS
1-Hydroxy-2-methyl-2-butanyl-4-diphosphate
HDR
HDR
MEV PATHWAY
Isopentyl diphosphate
IPI
Dimethylallyl diphosphate
GPPS
Geranyl diphosphate
GGPS
Geranylgeranyl diphosphate
GGDR
Phytyl diphosphate
SHIKIMATE PATHWAY
3-deoxy-D-arabino-heptulosonate-7P
DHQS
3-Dehydroquinate
DHQ
3-Dehydroshikimate
SDH
Shikimate
SK
Shikimate 3P
EPSPS
5-Enolpyruvylshikimate-3P
CS
Chorismate
CM
Prephenate
PAT
Agronate → Phenylalanine
TyrA
Tyrosine
TAT
CAROTENOID PATHWAY
PSY1
Phytoene
Lycopene
LCY-E
CYC-B
LCY-B
δ-carotene
β-carotene
LCY-B
α-carotene
Zeaxathin
CRTR-B
CRTR-E
Lutein
VTE PATHWAY
PPK (VTE6)
Phytyl P
PK (VTE5)
Phytol
Chlorophyll
HPT (VTE2)
Hydroxyphenylpyruvate
HPPD
Homogentisate
HGGT
2-Methyl-6-Phytylquinol
MPBQT (VTE3)
2-Methyl-6-geranylgeranylquinol
MPBQT (VTE3)
2,3-Dimethyl-6-Phytylquinol
TC (VTE1)
TC (VTE1)
δ-tocopherol
γ-tocopherol
γ-TMT (VTE4)
γ-TMT (VTE4)
α-tocopherol
β-tocopherol
2,3-Dimethyl-6-geranylgeranylquinol
TC (VTE1)
TC (VTE1)
γ-tocotrienol
δ-tocotrienol
γ-TMT (VTE4)
γ-TMT (VTE4)
α-tocotrienol
β-tocotrienol

Figure 1-2 Outline of the tocochromanol pathway. MEP, SK, carotenoid and VTE pathway in blue, yellow, orange, and purple, respectively. Enzyme names are as follows: **DXS**; 1-Deoxy-D-xylulose-5-P synthase, **DXR**; 2C-Methyl-D-erythritol-4-phosphate synthase, **CMS**; 2C-Methyl-D-erythritol-4-phosphate cytidyltransferase, **ISPE**; 4-2-CMethyl-D-erythritol kinase, **ISPF**; 2C-Methyl-D-erythritol-2-3-cyclodiphosphate synthase, **HDS;** 4-Hydroxy-3-methylbut-2-enyl-diphospahte synthase, **HDR**; 4-Hydroxy-3-methylbut-2-enyl-diphosphate reductase, **IPI**; Isopentyl diphosphate δ isomerase, **GPPS;** Geranyl pyrophosphate synthase, **GGPS**; Geranylgeranyl pyrophosphate synthase **GDDR**; Geranylgeranyl reductase, **PSY**; Phytoene synthase, **DAHPS**; 3-Deoxy-D-arabino-hepulosonate, **DHQS**; 3-Dehydroquianate synthase**, SDH-DHQ1**; 3-Dehydroquinate dehydratase**, SDH-DHQ2**; Shikimate 5-dehydrogenase**, SK**; Shikimate kinase**, EPSPS**; 5-Enolpyruvylshikimate-3-P-synthase, **CS**; Chorismate synthase, **CM**; Chorismate mutase, **PAT**; Prephenate aminotransferase, **TyrA**; Arogenate dehydrogenase, **TAT**; Tyrosine aminotransferase, **HPPD**; 4-Hydroxyphenylpyruvate dioxygenase, **HPT (VTE2)**; Homogentisate phytyl transferase, **MPBQMT (VTE3)** Dimethyl-phytylquinol methyl transferase, **TC (VTE1)**; Tocopherol cyclase, **γ-TMT (VTE4)**; γ-Tocopherol C-methyl transferase, **PK (VTE5)**; Phytol kinase, **PPK (VTE6)**; Phytyl-phosphate kinase, **PSY1**; Phytoene synthase 1, **CYC-B**; chromoplast-specific lycopene cyclase, **LYC-B**; lycopene β cyclase , **LYC-E** lycopene ε cyclase , **CRTR-B**; β-ring hydroxylase , **CRTR-E**; ε-ring hydroxylase.

Homogentisate is a product of the SK pathway, which provides the chromanol head group of tocochromanols (figure 1-1 and figure 1-2). Phytyl diphosphate (PDP) and geranylgeranyl phosphate (GGDP) from the MEP pathway, are used to provide the isoprenyl tail for tocopherols and tocotrienols, respectively. Through modulation of homogentisate phytyl transferase (HPT-VTE2) and homogentisate geranylgeranyl transferase (HGGT), homogentisate and PDP/GGDP undergo condensation reactions to form 2-methyl-6-phytylquinol (MPBQ) and 2-methyl-6-geranylgeranylquinol (MGGBQ), respectively (figure 1-2). HPT (VTE2) is the first committed step of the *Arabidopsis* VTE biosynthetic pathway (Savidge et al., 2002, Collakova and DellaPenna, 2003), because its precursors undergo an irreversible reaction to form MPBQ/MGGBQ, which are devoted solely to the VTE biosynthesis. Interestingly, dicotyledonous plants tend to accumulate more tocopherols than tocotrienols as they lack HGGT to synthesise tocotrienols (Lu et al., 2013). In contrast, monocotyledonous plants tend to accumulate more tocotrienols, particularly in their seeds, compared to dicotyledonous plants as they carry a gene that encodes HGGT (Yang et al., 2011, Lu et al., 2013).

The methyltransferases (VTE3(1) and VTE3(2)) use MPBQ and MGGBQ to form 2, 3-dimethyl-5-phytyl-1, 4-benzoquinone (DMPBQ) and 2, 3-dimethyl-5geranylgeranyl-1, 4-benzoquinone (DMGGBQ), respectively (figure 1-2). Both VTE3(1) and VTE3(2) proteins have chloroplast signal peptides, but *VTE3(1)* is more highly expressed in tomato fruit than *VTE3(2)* (Almeida et al., 2011, Quadrana et al., 2013). The VTE1 enzyme cyclises DMPBQ and DMGGBQ to generate δ-tocopherol/tocotrienol respectively (figure 1-2) (Kanwischer et al., 2005). This requires an additional substrate, S-adenosyl methionine (SAM) in the reaction (not shown in figure 1-2). After cyclisation, γ- and δ-tocochromanols undergo a final methylation step with SAM, catalysed by γ-tocopherol methyltransferase (γ-TMT/VTE4), to form α- and β-tocochromanols respectively (figure 1-2) (Bergmuller et al., 2003).

The regeneration of PDP from free phytol during the chlorophyll degradation pathway has established two new VTE enzymes. VTE5 phosphorylates free phytol derived from chlorophyll degradation to form phytyl phosphate (Valentin et al., 2006) (figure 1-2). This precursor is further phosphorylated by VTE6 to form PDP for the biosynthesis of VTE (vom Dorp et al., 2015) (figure 1-2). The importance of these two enzymes was established because total tocopherols were nearly abolished in *Arabidopsis* knock-out mutant lines of *vte5* and *vte6* (vom Dorp et al., 2015, Valentin et al., 2006). The *vte5* mutant showed that in seeds tocopherol levels were reduced to 20% of wild type (WT) levels (Valentin et al., 2006). Similarly, in leaves of the *vte6* mutant tocopherol levels were reduced by 98%, compared to the WT (vom Dorp et al., 2015). This highlights the importance of the chlorophyll degradation pathway in salvaging free phytol for VTE biosynthesis.

VTE does not undergo enzyme-mediated degradation, but as it is an antioxidant, VTE reacts with lipid peroxyl radicals formed from lipid peroxidation to form a tocopherol radical (Mene-Saffrane and DellaPenna, 2010). The tocopherol radical is converted back to the tocopherol vitamer by other antioxidants, such as ascorbate (Mene-Saffrane and DellaPenna, 2010).

1.1.4 Vitamin E in plants

The function of VTE in plants is centred on its antioxidant ability. The different tocochromanol vitamers have been shown to have differing antioxidant capacities. α-tocopherol and α-tocotrienol have a greater scavenging potential, compared to the other forms of VTE (Serbinova et al., 1991). Whereas, α-tocotrienols have significantly greater antioxidant capacities than α-tocopherol in phosphatidylcholine liposomes (Suzuki et al., 1993). The significant antioxidant ability of α-tocotrienol has been attributed to its distribution in membranes, and it is distributed more uniformly than α-tocopherol. The prenyl structure of α-tocotrienol can affect the structure of cell membranes, which enables more

efficient interactions between the chromanol head groups and lipid radicals (Serbinova et al., 1991).

In photosynthetic organisms, the light harvesting complexes located in thylakoids of chloroplasts are the site for the light-dependent photosynthetic reactions. In the light harvesting complex (PSII), light (photons) are absorbed in the PSII reaction centre to excite chlorophyll (Smith et al., 2009). This process initiates a chain of redox reactions which drives the photochemical reactions of photosynthesis. However, PSII is sensitive to too much light and too many excited electrons result in the production of O_2^- and cause photo-oxidative damage. Many cofactors reside within PSII to quench the excess photoexcitation energy, such as β-carotene and plastoquinone (Chrost et al., 1999, Trebst, 2003). Initially, it was thought that VTE would also play a role during photosynthesis, because VTE is synthesised in plastoglobuli that are associated with thylakoid membranes of chloroplasts (Austin et al., 2006). Porfirova et al. (2002) suggested that VTE may also be a photoprotective antioxidant of photosynthesis. The *Arabidopsis* vte1 mutant of the VTE pathway showed that inhibition of the synthesis of γ-/δ-tocopherol, results in the accumulation of DMPBQ. The *vte1* mutant showed increased levels of zeaxanthin, which is a photo-inhibitory antioxidant (Porfirova et al., 2002), suggesting that VTE may function during photoinhibition, like zeaxanthin. This is supported by the *vte4* mutant that accumulates γ-tocopherol but cannot produce α-tocopherol. Under high light stress, the γ-tocopherol in this mutant was able to reduce photo-oxidative stress (Bergmuller et al., 2003). Excess photoexcitation energy of PSII results in the production of O_2^-, which can degrade the reaction centre protein (D1). Using inhibitors of HPPD in *Chlamydomonas reinhardtii,* tocopherol levels were depleted and the D1 protein was degraded (Kruk et al., 2005). This phenotype was partially recovered using O_2^- quenchers, this suggests that tocopherols are O_2^-scavengers produced from PSII (Trebst et al., 2002, Kruk et al., 2005, Krieger-Liszkay and Trebst, 2006). This confirmed the role of VTE as a photo-inhibitory antioxidant during photosynthesis.

Tocopherols can also protect PUFAs from lipid peroxidation because VTE is lipophilic and therefore resides within membranes. The knock-out *vte2 Arabidopsis* mutant showed that lack of tocopherols and their intermediates negatively impact early seedling development and germination (Sattler et al., 2004). These authors showed that tocopherols have a role in maintaining lipid peroxidation of lipid storage molecules during seed storage and suggested that seeds are significant sources of VTE in the human diet (table 1-2). This also provided evidence that VTE can have additional antioxidant roles that are unrelated to photosynthesis.

VTE has been suggested to play a role in photo-assimilate transport in potato and maize plants. The maize *sucrose export defective1 (sxd1)* mutant is a mutation of VTE1 which results in callose deposition in plasmodesmata between the bundle sheath and vascular parenchyma (Botha et al., 2000). This mutation affects the accumulation of anthocyanins and starch in source leaves in maize (Botha et al., 2000). RNA interference (RNAi) of the potato *SXD1* gene demonstrated a similar phenotype to the maize *sxd1* mutant (Hofius et al., 2004). These authors showed that this was likely to be due to callose deposition, which was not observed in the *Arabidopsis vte1* mutant (Sattler, 2003), rather than a lack of tocopherol accumulation (Hofius et al., 2004).

The antioxidant roles of VTE in plants have been elucidated from *Arabidopsis* mutant lines and inhibitor studies (Kruk et al., 2005, Porfirova et al., 2002, Bergmuller et al., 2003). Similarly, in humans, the roles of VTE as potent antioxidants have been shown in individuals with high oxidative stresses and in immune responses (Blum et al., 2010, Boaz et al., 2000, Milman et al., 2008, Bou Ghanem et al., 2015, Meydani et al., 1998, Mocchegiani et al., 2014). Therefore, understanding the regulatory nature of VTE biosynthesis in plants is important and may provide useful for metabolic engineering.

1.2 Plant metabolic engineering of VTE

Improving the VTE content in tomatoes would be advantageous. Tomatoes are an important nutritional crop; their consumption has increased to over 1 billion tonnes in 2010, approximately a 40% rise since 2000 (FAOSTAT, 2010). Tomatoes already contain many beneficial compounds, such as carotenoids therefore they are considered a functional food (Canene-Adams et al., 2005). This, coupled with the fact that tomatoes already produce VTE, demonstrates that they are a good model crop for engineering VTE levels using biotechnology or breeding (Chun et al., 2006). Metabolic engineering designed to accumulate high levels of specialised metabolites in tomato has been successful for anthocyanins, flavonols, isoflavones, stilbenoids and betalains (Butelli et al., 2008, Zhang et al., 2015, Polturak et al., 2017, Luo et al., 2008). Therefore, tomato is a suitable candidate crop for biofortification of VTE.

There have been some efforts to increase VTE levels in tomato. Lu et al. (2013) produced transgenic tomatoes overexpressing *VTE2* which accumulated 10-fold more VTE in unripe, green tomato fruit, compared to ripe, red fruit and WT. Interestingly, the tomato fruits also contained tocotrienols, which are not normally synthesised in tomatoes. This implied that during fruit ripening and the transition from chloroplast to chromoplast, there is negative regulation of plastid gene and transgene expression. Therefore, within red tomato fruit, *VTE2* would not normally be highly expressed, resulting in low levels of VTE and no tocotrienols. Similarly, overexpression of apple fruit *MdVTE2* in microtom tomato fruit, increased α-tocopherol by 1.7-fold and γ-tocopherol by 3.1-fold, compared to WT fruits (Seo et al., 2011). However, the increases in VTE in these transgenic fruits are still well below the RDA of VTE.

Most attempts of VTE biofortification have been focused on seed crops. A soybean mutant of homogentisate dioxygenase (HGO) showed that seeds accumulate up to two-fold increases in VTE compared to WT (Stacey et al., 2016). HGO is a homogentisate dioxygenase, which is the first committed step for homogentisate catabolism. The *hgo* mutant also showed a 27-fold increase in

tocotrienols, which suggested that homogentisate is limiting tocotrienol production (Stacey et al., 2016). However, enzymatic assays of VTE2 and HGGT have shown that VTE2 preferentially uses PDP over GGDP when PDP pools are high (Yang et al., 2011). However, VTE2 uses GGDP when PDP pools are low to synthesise tocotrienols (Yang et al., 2011), which suggests that GGDP is a limiting factor of tocotrienol biosynthesis as well.

The majority of biofortification strategies have included the overexpression of genes encoding enzymes of the VTE pathway (Karunanandaa et al., 2005, Li et al., 2010, Kanwischer et al., 2005, Seo et al., 2011, Van Eenennaam et al., 2003). However, there should be advantages to understanding the regulation of VTE biosynthesis in order to overcome the limitations on flux along biosynthetic pathways. My thesis focuses on understanding the transcriptional regulation of VTE biosynthesis and the identification of candidate transcriptional regulators of this pathway.

1.3 Transcription factors

Transcription is the primary point for the regulation of gene expression. This process involves a number of transcription-related proteins for its initiation. Transcription factors (TFs) are one of many DNA-binding proteins which modulate gene expression (Yanagisawa, 1998). Initially, transcription begins when a TATAA (TATA box) is recognised by TATA binding proteins (TBPs) in a promoter, which is upstream of the start codon (ATG) of a gene (Smith et al., 2009). However, not all gene promoters that are recognised by RNA polymerase II carry TATA boxes. TBPs form a complex with many general TFs (GTFs) and recruit RNA polymerase II to the TATA box. The GTFs can have multiple functions, such as unwinding DNA which allows RNA polymerase II to change its conformation and start transcription. However, other TFs can aid the transcription process. Generally, TFs bind to DNA often referred to as cis-elements, upstream of the TATA box or downstream of genes. These are also known as enhancer regions or enhancer elements if they bind transcriptional activators. If transcriptional repressors bind to an enhancer

region, they are called repressor elements, The TFs can act alone or in large TF complexes to allow DNA to fold to form complexes with the RNA polymerase II. These TFs can act as activators of gene expression, or repressors that inhibit work by multiple mechanisms to bind DNA by other TFs (Smith et al., 2009). Either repressors can bind directly via repression elements, such as the EAR domain or by competing for binding sites.

TFs can be separated into different TF families based either on the sequence of their DNA binding motifs, or on their protein domains. There are several common plant TF families that are abundant in higher plants, including; **My**elo**b**lastosis (MYB), **N**o **A**pical **M**eristem, ***A****rabidopsis* transcription activation factor and **C**up-shaped cotyledon (NAM/NAC), WRKY, **B**asic leucine **zip**per (bZIP) and bHLH (**B**asic **H**elix **L**oop **H**elix) TFs. Different members of plant TF families often share similar functions, or a TF family can be divided into subgroups which share common roles. MYB TFs are an example of a large plant TF family which can be divided structurally and functionally into distinct sub-groups (Stracke et al., 2001).

MYB TFs and bZIP TFs have been suggested to regulate the VTE pathway (Quadrana et al., 2013). Promoter analysis of the isoprenoid pathway genes encoding biosynthetic enzymes (VTE and MEP pathways) showed that they contain common TF binding motifs, which implies that they could be co-regulated (Quadrana et al., 2013). The most abundant motifs identified in the promoter analysis were bZIP and MYB TF binding motifs.

1.3.1 Transcriptional regulation of the VTE biosynthesis

Metabolic engineering to improve VTE content has been successful for some VTE biosynthetic genes (Hunter and Cahoon, 2007, Shintani and DellaPenna, 1998, Li et al., 2010), yet it is evident that tocochromanol composition can be skewed towards less bioactive isoforms (Kanwischer et al., 2005, Van Eenennaam et al., 2003, Falk et al., 2003, Cahoon et al., 2003, Tzin et al., 2009). The VTE pathway is fully characterised, but transcriptional control over the MEP, SK and VTE pathways requires further elucidation. Understanding of the transcriptional

regulation of VTE production is very limited, but, it is evident that transcriptional control does exist (Quadrana et al., 2013). Transcription factors regulating this pathway have not yet been identified. Similarly, the MEP and SK pathways also have few characterised transcriptional regulators.

Evidence for transcriptional control of VTE production has been demonstrated through differential transcript levels of MEP genes in *Arabidopsis*. The precise modulation of transcript levels demonstrated it was likely that common transcriptionally-activated expression of structural genes occurs within the MEP pathway (Guevara-Garcia et al., 2005). This, coupled with spatial and temporal control for α-tocopherol production in tomato (Quadrana et al., 2013), suggested that transcriptional control occurs during tomato fruit ripening.

Analyses of the promoter regions of genes encoding enzymes of the MEP and the VTE pathway have suggested transcriptional co-regulation exists (Quadrana et al., 2013). Although, Quadrana et al (2013) provided evidence that transcriptional regulation must exist, they did not identify any transcription factors, and rather, described the common TF binding motifs within promoters of the VTE pathway genes. QTL analysis of soybean varieties that accumulate VTE showed that a polymorphism in the *VTE4* promoter resulted in increased expression of the gene encoding VTE4 (Dwiyanti et al., 2011). This provided evidence that VTE is transcriptionally regulated, at least in soybean, and probably in other species as well.

Recently, a TF has been identified by analysis of an *Arabidopsis wrinkled 1 (wri1)* mutant that lacks triacylglycerols in seeds, resulting in a wrinkled seed phenotype (Baud et al., 2007, Cernac and Benning, 2004). Triacylglycerols are one of the storage lipids in *Arabidopsis* embryos. The *wri1* mutant encodes for a seed-specific AP2/ERF TF and accumulates 1.2-fold more α-tocopherol than WT seeds (Pellaud et al., 2018). Interestingly this mutant also accumulates γ-tocomonenol, which is another type of tocochromanol that is found in just a few plant species (Pellaud et al., 2018, Puah et al., 2007). γ-tocomonenol is very similar in its structure to other γ-tocochromanols, but it contains one double bond between C-

11’ and C-12’ of the prenyl chain. Expression analyses were not carried out on the *wri1,* as the focus of the study was the synthesis of γ-tocomonenol (Pellaud et al., 2018). Therefore, it is not clear how this TF regulates VTE biosynthesis transcriptionally, if at all, or whether the effects of its loss of function are indirect. It was clear from this study that because this mutant was identified based on the composition of lipid storage molecules, there had been little analysis of its effects on VTE biosynthesis (Baud et al., 2007, Cernac and Benning, 2004).

There are no TFs, known to date to regulate the VTE pathway directly, and those identified, regulate upstream pathways, which effect substrate availability for tocopherol biosynthesis (Cordoba et al., 2009, Dal Cin et al., 2011, Pepper et al., 1994). However, there is evidence that epigenetic factors play a prominent role in modulating VTE content (Quadrana et al., 2014). Consequently, transcriptional control over VTE is likely to be complex.

1.3.2 Transcriptional regulation of the MEP pathway

Some studies propose that the MEP pathway and putatively the VTE pathway can be regulated by environmental factors (Liu et al., 2004, Bino et al., 2005, Azari et al., 2010). Transcripts of genes encoding enzymes of the MEP accumulate in response to light and during early seedling development. This may be due to accompanying increases in ROS, resulting in an increase in demand for MEP pathway products for their antioxidant capacity. In addition, down regulation of *DE-ETIOLATED1* (*DET1)* in tomato plants (involved in the light signal transduction pathway) showed increased levels of antioxidants like tocopherols, although this occurred with a fitness cost to the plant, contributing to low yields (Davuluri et al., 2005, Enfissi et al., 2010). Tocopherol biosynthetic gene transcript levels were increased in tomato fruit, as were many other genes involved in the synthesis of antioxidants, such as carotenoids and flavonoids. Therefore, it is unclear whether light affects tocopherol synthesis directly, but it does show that transcriptome changes take place, and the light signal pathway affects transcriptional regulation of the VTE pathway.

Transcriptional regulators of the VTE pathway have not yet been identified. However, there is evidence that environmental factors can regulate the MEP pathway, to provide substrates for VTE biosynthesis. Light and plant hormones have been shown to interact with TFs (either directly or indirectly) to induce gene expression and regulate biosynthetic pathways, such as the MEP and phenylpropanoid pathways, in many species (Gonzalez et al., 2008, Zuluaga et al., 2008, Albert et al., 2009). However, light is the only known signal of regulating transcription of the genes encoding enzymes of the MEP pathway (Davuluri et al., 2005, Enfissi et al., 2005, Fraser et al., 2007, Enfissi et al., 2010).

Responses to light are controlled by phytochromes (PHY), which are photoreceptors that can perceive red and far-red light during photomorphogenesis. There are five tomato PHYs; *PHYA, PHYB1, PHYB2, PHYE* and *PHYF.* The mutant, *phyA,* showed that *PHYA* modulates accumulation of lycopene temporally, during tomato fruit ripening (Gupta et al., 2014) - lycopene is a carotenoid (figure 1-2). Additionally, transgenic tomatoes lines that over expressed *PHYB2,* accumulated anthocyanins in young leaves compared to WT, under continuous red light (Husaineid et al., 2007).

PHYs are transcriptionally regulated by phytochrome interacting factors (PIFs), which are bHLH TFs. In tomato, *PIF1* negatively regulates the first enzyme for carotenoid biosynthesis – *PSY1* (Toledo-Ortiz et al., 2010). It is likely that modulation of *PSY1* activity will have a downstream effect on VTE content since carotenoids and VTE use the same precursors for their synthesis (figure 1-2). PIFs can have other roles in plants, such as *PIF4 in Arabidopsis,* which binds to the same targets as another light responsive TF, *HY5.* This allows for transcriptional regulation of thermosensory plant responses (Delker et al., 2014, Gangappa and Kumar, 2017). The interaction between these two TFs highlights the myriad of roles PIFs play and their possible interactions with other TFs. *HY5* is regulated directly by light (Hardtke et al., 2000, Nawkar et al., 2017) and *HY5* RNAi lines in tomato showed reduced carotenoid and chlorophyll content of tomato fruit and leaves, respectively (Liu et al., 2004). The 'pull' of carotenoid and chlorophylls is

linked intimately to VTE biosynthesis (Lira et al., 2017). *HY5* has been shown to bind to another bZIP TF (*ABI5),* which is important for ABA signalling (Chen et al., 2008). ABA biosynthesis also takes precursors from the MEP pathway. ABA is an important hormone during stress responses, and the precursors for its synthesis are derived from the MEP pathway (figure 1-2). Accumulation of VTE in seeds has been shown to be important for seed longevity (Sattler et al., 2004) and ABA is also important for seed dormancy (Linkies and Leubner-Metzger, 2012). Therefore, it is likely that there will be cross talk between the pathways, but there could also be feedback loops for inter-regulation between the pathways, as both originate from the MEP pathway.

1.3.3 Transcriptional regulation of the shikimate pathway

Regulation of the SK pathway, essential for the generation of homogentisate for VTE production (figure 1-2), is not well understood. Flux within the pathway is mediated through DAHPS, which is thought to be rate limiting (Tzin et al., 2012). Phenylalanine is a precursor for flavonol synthesis, which is produced by the SK pathway. AtMYB12 has been identified as a transcriptional activator of the flavonoid pathway (Mehrtens et al., 2005), and tomato plants overexpressing *AtMYB12*, under the control of the fruit specific ripening promoter, E8, showed that *AtMYB12* can induce expression of genes encoding enzymes of primary metabolism as well as those of secondary flavanol metabolism (Luo et al., 2008, Zhang et al., 2015). Chromatin immunoprecipitation sequencing of *AtMYB12* showed that this TF could bind to the DAHPS promoter and the plastidial enolase (ENO) promoter (Zhang et al., 2015). ENO encodes for an enzyme that synthesises pyruvate from D-glyceraldehyde 3-phosphate in primary metabolism (Prabhakar et al., 2009). DAHPS has been shown to be limit the rate of the shikimate pathway (Tzin et al., 2012). Tomatoes overexpressing *AtMYB12* showed that carbon flux was redistributed to tyrosine, in comparison to WT fruit. Tyrosine is a precursor for phenylalanine, of which the latter is a precursor for flavonols. Tyrosine is a precursor for VTE as well, to produce homogentisate (figure 1-2). Therefore,

AtMYB12 may have an indirect impact on VTE metabolism by increasing the supply of tyrosine.

Transcriptional control has been suggested following studies of floral fragrance in *Petunia hybrida.* Petunia RNAi lines of *ODORANT1 (ODO1)* showed that abundance of transcripts involved in the SK pathway was altered compared to WT (Verdonk et al., 2005). *ODO1* is a R2R3 MYB TF, which activates the EPSPS promoter early in the synthesis of shikimate (Verdonk et al., 2005). EMISSION OF BENZENOIDS I AND II (EOBI and EOBII) TFs modulate phenylpropanoid-related scent volatiles in petunia, through transcriptional regulation of ODO1. These TFs act upstream of ODO1, but Petunia RNAi lines of *EOBI* showed downregulation of *EPSPS, DAHPS, CS* and *CM1* expression (Spitzer-Rimon et al., 2012). There is a negative feedback loop between EOB1 and ODO1 (Spitzer-Rimon et al., 2012). Therefore, both ODO1 and EOB1 can alter transcript abundance of genes encoding enzymes of the SK pathway (Van Moerkercke et al., 2011). EOBII acts upstream of both ODO1 and EOBI, and thus can regulate SK pathway gene expression through transcriptional regulation of these TFs (Spitzer-Rimon et al., 2010, Van Moerkercke et al., 2011). Therefore, these TFs may impact VTE biosynthesis by altering substrate availability.

HPPD has been identified in many species as a senescence related gene (Singh et al., 2011, Quirino et al., 2000, Lee et al., 2001, Chrost et al., 1999), and is characterised as catalysing one of the key branch points of VTE synthesis. There is evidence for the important role of HPPD in many species, since increased *HPPD* expression is associated with increased tocopherol and carotenoids in fruits (Singh et al., 2011, Tsegaye et al., 2002). An increase of 37% was observed when *HPPD* was overexpressed in *A.thaliana* (Tsegaye et al., 2002). Addition of a HPPD inhibitor to *Chlamydomonas reinhardtii*, reduced tocopherol levels by 20%, showing the importance of this gene for VTE biosynthesis (Trebst et al., 2002). This provides evidence that *HPPD* is important for the synthesis of shikimate, and therefore could be an internally regulated control point of the SK pathway.

Studies of floral volatiles and senescence have elucidated several TFs that regulate the SK pathway transcriptionally, however, ambiguity still surrounds SK pathway regulation, and the effects on VTE biosynthesis.

1.3.4 Post transcriptional regulation of the VTE pathway

Knowledge of transcriptional control of the VTE, MEP and SK pathways is limited. However, internal regulation of VTE accumulation has been described. Changes in the MEP pathway, which precedes the VTE pathway (figure 1-2), can alter flux and availability of precursors for VTE biosynthesis. The 'pull' from carotenoids must be considered in attempting to understand the control of VTE biosynthesis. Post-transcriptional control over the MEP pathway has been proposed; accumulation of *DXS* can modulate transcript and protein levels of MEP pathway genes (Cordoba et al., 2009). However, *DXS* transcripts accumulate when the MEP pathway is inhibited, so DXS accumulation is mediated by the demand for pathway products. MEP inhibitor studies have suggested that a feedback mechanism for DXS regulation might exist, which alters DXS protein levels, and consequently affects flux through the MEP pathway (Guevara-Garcia et al., 2005, Fraser et al., 2007).

Similarly, in the SK pathway, homogentisate production is a result of prenylation by *HPPD* generating hydoxyphenylpyruvate (HPP). As shown in figure 1-2, arogenate dehydrogenase (TyrA) converts arogenate to tyrosine. The tyrosine, however internally inhibits *TyrA* and *CM* (Tzin and Galili, 2010a). Most of the carbon that is fixed by photosynthesis for the production of arogenate is converted to phenylalanine, limiting flux into the VTE pathway (figure 1-2). Therefore, sensitive inhibition occurs post chorismate biosynthesis (figure 1-2), to enable the generation of phenylalanine and possible phenylpropanoid products, such as anthocyanins and flavonols (Tzin et al., 2012).

1.3.5 Regulation of the VTE pathway is limited and needs clarification

Transcription and post-translational regulation of the VTE, MEP and SK pathways is limited. It is clear that VTE biosynthesis may be inhibited by pathway

flux of substrates, or metabolite products which can internally inhibit transcript levels and by transcriptional regulators. VTE biosynthesis is likely to be controlled by TFs, as precise modulation of transcript levels has been demonstrated during tomato development and ripening (Quadrana et al., 2013). Therefore, this thesis aimed to identify transcriptional regulators of the VTE pathway which may be used to increase the biosynthesis of VTE.

1.4 Experimental approach to identify transcriptional factors that regulate VTE biosynthesis in tomato

There is very little knowledge known about the transcriptional regulation of VTE biosynthesis and it is likely to be complex. Therefore, I have used a variety of methods to identify transcriptional regulators of VTE biosynthesis and characterise their function.

1.4.1 *Solanum pennellii* introgression lines provide a unique genetic resource

Introgression lines (ILs) have been developed for many species as a genetic resource to introduce new genes into an existing gene pool to harness new genetic variation. ILs typically involve a cross between two different species, and progeny from this cross are then backcrossed to a single recurrent parent, until a single introgressed donor segment is achieved in a number of lines. ILs are useful as they can be used to identify quantitative trait loci (QTLs), which leads to the elucidation of phenotypes and their molecular characterisation. Complex traits are often controlled by QTLs, which have pleiotropic and additive effects on phenotypes. However, many QTLs have not been characterised at the molecular level, and rather, are masked by epistatic interactions. ILs provide an alternative approach to this problem (Lippman et al., 2007), as they segregate transgressively without the masking effects from epistatic interactions, enabling identification of differential phenotypic expression (Fernie et al., 2006, Lippman et al., 2007).

1.4.1.1 Using ILs to discover expression QTLs – combining expression with phenotypes

Traditionally, ILs have been used to characterise QTLs molecularly, for a given specific genotype. However, with the advancement of high throughput gene expression profiling, more complex traits can be identified through the combination of ILs and transcriptomics. Expression QTLs (eQTLs) can enable the identification of causal genes that are responsible for phenotypes by combining gene expression with ILs. eQTLs can act either locally (cis) or distally (trans) to the locus of interest. Identification of trans-eQTLs are preferred as they are unlinked to the gene of interest and are likely due to variation in activity of a master regulator, like a transcription factor. Trans-eQTLs are associated with a broad spectrum of effects, for example controlling many genes, and are often few in number (Hansen et al., 2008). This is in contrast with cis-eQTLs which act locally on a gene of interest and can affect TF binding sites and/or chromatin structure (Alberts et al., 2007, Cubillos et al., 2012). The identification of trans-eQTLs are the basis of this thesis and the use of ILs between *Solanum* species to elucidate trans-acting regulators of the VTE pathway.

To find trans-eQTLs I used the well-established *Solanum pennellii* x *Solanum lycopersicum* cv. M82 IL population (Eshed and Zamir, 1995). These ILs are a cross between *S.lycopersicum* cv. M82 and *S.pennellii* (figure 1-3). The progeny from this cross were backcrossed to the *S.lycopersicum* cv. M82 parent recurrently, until only one part of one chromosome had a *S.pennellii* donor segment in an *S.lycopersicum* cv. M82 background. *S.lycopersicum* cv. M82 is a red-fruited commercial processing tomato, and *S.pennellii* is a wild, green fruited relative of tomato. The seventy -six *S.pennellii* x *S.lycopersicum* cv. M82 ILs were generated over 20 years ago for the identification of novel traits for the improvement of yield of tomato fruit (Eshed and Zamir, 1995) and they provide complete genome coverage (a full list of the ILs are available in the appendix). Therefore, the *S.pennellii* x *S.lycopersicum* cv. M82 ILs, combined with expression data, provide a unique resource to identify new genetic variation.

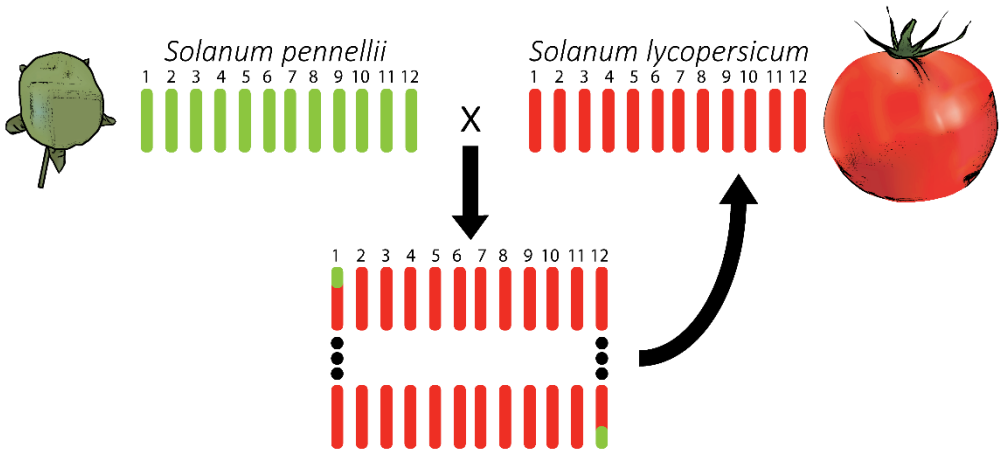


Figure 1-3 Schematic diagram of *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression lines (ILs). *S.pennellii* (green fruit) was crossed with *Solanum lycopersicum* cv.M82 (red fruit). Progeny from this cross were backcrossed to *S.lycopersicum* to form ILs, providing full genome coverage.

1.4.1.2 Expression QTLs or metabolic QTLs, and their limitations

The use of eQTLs has not been widely adopted in tomato-based studies for metabolism, which tend to focus on metabolic QTLs (mQTLs), rather than take an eQTL approach. Metabolic analysis undertaken on a whole plant phenotype scale can generate a vast mQTL array for the *S.pennellii* x *S.lycopersicum* cv. M82 ILs. mQTLs enable the identification of genomic regions influencing metabolite contents and link phenotypes with metabolite contents, which are much more specific than QTLs, and are less time consuming to map finely. However, they are often linked to structural genes. mQTLs that are linked to structural genes, often identify the cause of differential metabolite profiles. However, mQTLs rarely increase our understanding of regulatory networks underlying metabolite contents. Therefore, the use of eQTLs would provide a novel approach to understand vitamin E biosynthesis in tomato, as candidate genes could be identified, which are unrelated to the VTE structural mQTLs.

eQTL studies have been used in humans (Majewski and Pastinen, 2011, Gilad et al., 2008) and plant models (Potokina et al., 2008, Druka et al., 2008, Chapman et al., 2012) to elucidate possible regulatory components. One of the first eQTL studies in tomato used a microarray platform to establish the transcriptome which showed six ILs that harbour alterations in Brix (soluble solid content). The authors linked downstream effects of ILs and changes in gene expression (Baxter et al., 2005). A similar approach could be used to identify regulators of VTE synthesis.

Several trans- and cis-eQTLs have been identified for the *S.pennellii* x *S.lycopersicum* cv. M82 ILs, using leaf RNA sequencing data, for various traits including resistance traits to leaf development (Ranjan et al., 2016). But, this approach to find trans-eQTLs has not been adopted fully for tomato fruit, and not at all for the VTE pathway.

Quadrana et al (2014) showed that a cis-eQTL to which VTE3(1) in tomato was linked, was an mQTL on IL9-2-6, resulting in an altered α-tocopherol content.

VTE3(1) is a methyltransferase enzyme in the VTE pathway responsible for the addition of a methyl group to 2-methyl-6-phytylquinol (a precursor for VTE) (figure 1-2). This IL was mapped to VTE3(1) and VTE5 (Almeida et al., 2011). The fine mapping of IL9-2-6 showed that VTE3(1) mapped to a sub region classified as IL9-2-6-1, and, most remarkably, alters tocopherol levels, which are dependent on the methylation state of the promoter. The promoter region, proximal to VTE3(1), showed epigenetic modification; it is differentially expressed depending on its parental allele. The *S.lycopersicum* allele is highly methylated in comparison to *S.pennellii* IL parent, due to a SINE transposable element in the promoter of the gene encoding the methyl transferase. The hyper methylation of *VTE3(1)* in *S.lycopersicum* results in low tocopherol content (Schauer et al., 2008, Quadrana et al., 2014). Although, the discovery of epialleles involved in the VTE pathway is important, the *VTE3(1)* eQTL is a cis-eQTL controlling expression. No candidate gene has yet been identified as a regulator of VTE content, which more likely could be identified through trans-eQTL analyses. This highlights the potential of identifying trans-eQTL regulators that control expression of genes encoding enzymes of the VTE pathway.

1.4.2 Viral Induced gene silencing (VIGS)

Viral induced gene silencing (VIGS) is a tool to test the function of genes by transiently silencing genes of interest (GOI). VIGS has been adapted for many plant species. The VIGS used in this thesis is based on gene silencing in the host plant by the tobacco rattle virus (TRV).

TRV is part of the tobravirus cluster, which are bipartite positive, single-strand viruses that normally require both single strand RNAs (ssRNAs) for successful infection of hosts (Hernandez et al., 1995). The tobravirus genome consists of RNA1 and RNA2. RNA1 encodes for a RNA dependent RNA polymerase (RdRp) – the replicase, a 16kDa cysteine rich protein, a movement protein and a 3’ untranslated region (3’UTR) (Hernandez et al., 1995). The function of the 16kDa protein is not fully understood, but it is needed for infection. RNA2 encodes for a

coat protein and a 3'UTR, of which, the latter is homologous to the 3'UTR in RNA1 (Hernandez et al., 1995).

VIGS used in this thesis is based on the system designed and created by Orzaez et al. (2009). This method uses the transgenic tomato lines, expressing two *Antirrhinum majus* transcription factors: *Delila (Del)* and *Rosea1 (Ros1),* under the control of two fruit specific E8 promoters. The authors use modified TRV RNA1/2 containing plasmids (Orzaez et al., 2009). The modified pTRV2, which includes RNA2 was created with an additional *Del* and *Ros1* fragment and a GOI fragment (figure 1-4). These fragments are approximately 200-300 nucleotides in length. Upon infection of pTRV2 and pTRV1, the plant's defence system is triggered, which activates post-transcriptional gene silencing (PTGS).

Double-stranded RNAs (dsRNAs) are synthesised from the viral ssRNAs using the viral RdRp (Shao et al., 2008). The dsRNAs then trigger PTGS and are recognised by RNAse III-type endonuclease Dicer enzymes, which bind the dsRNAs and cleaves them into short segments (Vaucheret et al., 2001, Waterhouse et al., 2001, Shao et al., 2008). The Dicer-bound dsRNAs bind to ARGONAUTE 1 (AGO1) to form small interfering RNAs (siRNAs), which are 21-24 nucleotides in length. One of the strands binds to AGO1 and the other strand is degraded. These AGO1 bound siRNAs form a RNA-induced silencing complex (RISC), which targets the GOI mRNAs and cleaves nucleotide sequences that base pair with the target sequence (Vaucheret et al., 2001, Waterhouse et al., 2001), resulting in the silencing of the GOI.

Combining VIGS with the *Del/Ros1* over-expressing tomatoes means it can be used as a scoring system to determine where your *GOI* is silenced. Figure 1-5 shows a *Del/Ros1* tomato which was injected with the pTRV2 *Del/Ros1* plasmid (at the mature green (MG) stage) and harvested fourteen days later. MG is defined when the tomato fruit has reached its full size and the seeds are mature before the onset of ripening. Figure 1-5 shows that the infection of TRV has spread throughout the tomato fruit, creating patches of red-silenced sectors and purple,

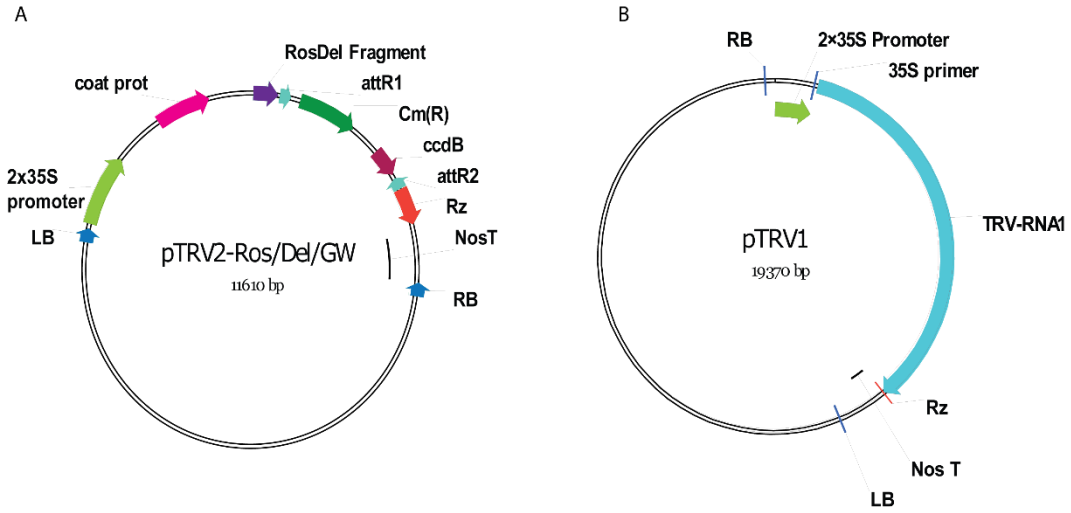


Figure 1-4 Plasmid map of the viral induced gene silencing (VIGS) vectors, (A) pTRV2 *Del/Ros* Gateway® and (B) pTRV1. Tobacco rattle virus (pTRV2) is a bipartite virus that required co-infection with pTRV1 for replication within a host. The parts of the plasmid are as follows; LB – left border, RB – right border, 2x35S promoter– 2x35S constitutive promoter, coat prot – Coat protein, *RosDel* fragment – Del/Ros VIGS fragments, attR1/2 – Gateway® sites, Cm(R) – Chloramphenicol resistance, ccdB – ccdB survival gene, Rz – self-cleaving ribozyme, NosT – Nopaline synthase terminator and TRV-RNA1 – RNA1 contains RdRp (RNA dependent RNA polymerase), a movement protein and 16kDa cysteine rich protein that are required for infection.

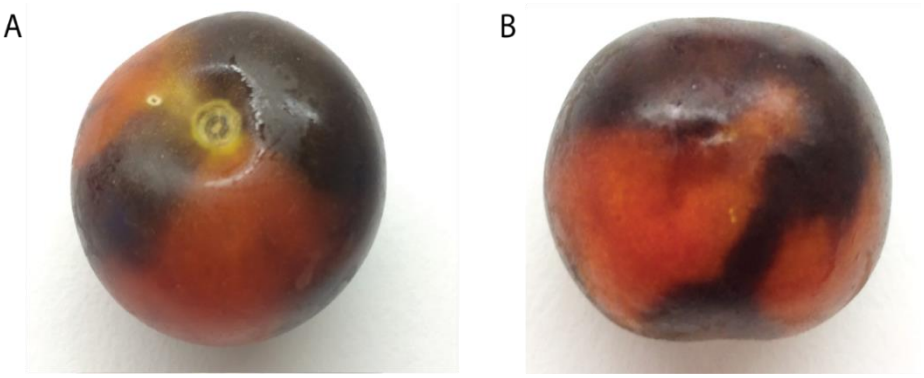


Figure 1-5 An example of a tomato injected with pTRV2 *Del/Ros* construct in Moneymaker *Del/Ros* tomatoes, (A) *Del/Ros* silenced tomato showing the injection point adjacent to the vascular attachment site and (B) Underside of the Del/Ros silenced tomato, showing how the virus has spread throughout the tomato fruit. The Moneymaker *Del/Ros* tomatoes are purple throughout. Therefore, when *Del/Ros* are silenced, the tomatoes return to WT colour (red). The silencing is cell specific and spreads from cell to cell, which creates patches of red-silenced sectors and purple-non-silenced sectors. Metabolite and expression analyses are conducted on the red silenced sectors of *Del/Ros* silenced fruit, which are compared to red-silenced sectors of *Del/Ros GOI* silenced fruit.

non-silenced sectors. The silenced sectors of the pTRV2 *Del/Ros1* fruit can be compared to pTRV2 *Del/Ros1 GOI* fruit, to determine the function of the GOI.

In this thesis, I used this VIGS system to transiently silence candidate genes encoding TFs to determine their function.

1.4.3 Transient over-expression using tobacco rattle virus

The TRV over expression (OE) system used in this thesis is a modified pTRV2 plasmid created by Dr Vera Thole (figure 1-6). This OE system is based on TRV, which is also the basis for VIGS. As previously described, RNA2 requires RNA1 for successful infection in a host. Therefore, a pBIN19 construct containing RNA1, under the control of the 35S promoter, and an intron was generated, which is called pTRV1 (figure 1-4). The intron prevents expression of the RdRp in *E.coli* and *Agrobacterium*. The pTRV2 over expression plasmid is based on RNA2 of TRV2 strain PPK20 (Hernandez et al., 1995), under the control of a 35S promoter, which is inserted into the pGREEN background (Hellens et al., 2000). This involves an additional promoter from another tobravirus called pea early browning virus (PEBV), which was inserted in the RNA2 of pTRV2. The PEBV promoter was chosen as it had previously been shown to be expressed in infected leaf tissues (Johnsen et al., 1991). PEBV promoter drives the expression of your *GOI* and enables the GOI to be replicated by the viral RdRp from pTRV1. This results in OE of your *GOI*.

The pGREEN vectors are *Agrobacterium* tumour inducing (Ti) binary vectors (Hellens et al., 2000) which require co-transformation of pSOUP. The pGREEN and pSOUP vectors were originally made to minimise the size of unnecessary plasmid DNA for plant transformations and to increase the number of selectable markers.

In this thesis, I used the transient OE system to OE candidate genes encoding TFs to determine their function.

1.4.4 CRISPR/Cas9

Clustered regularly interspaced, short palindromic repeats (CRISPR)/Cas9 is a gene editing technology that was developed from bacteria that use the

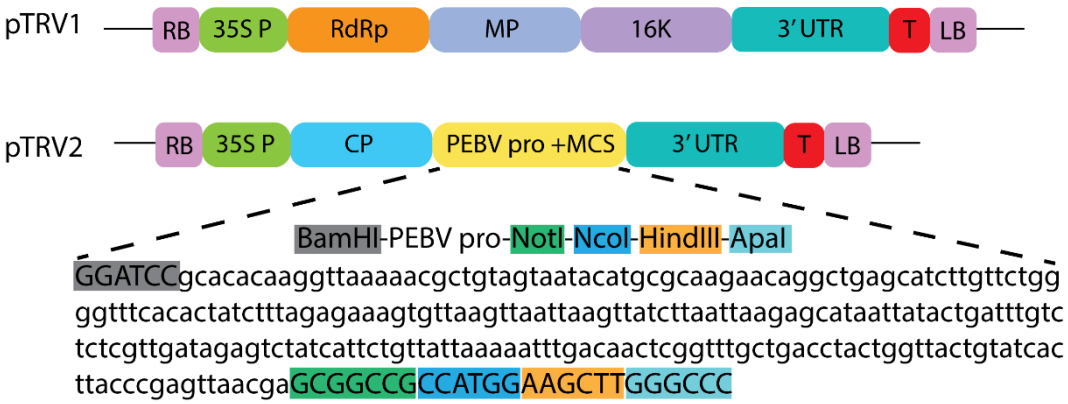


Figure 1-6 Plasmid map of the pTRV2 overexpression vectors. Tobacco rattle virus is a bipartite virus that required co-infection with pTRV1 for replication within a host. The letters represent the following; LB – left border, RB – right border, 35S P – 35S constitutive promoter, RdRp – RNA dependent RNA polymerase, MP – Movement protein, 16K - 16 kDa cysteine rich protein, 3'UTR – 3' untranslated region, CP – coat protein, PEBV pro - pea early-browning virus promoter, MCS – multiple cloning site and T – Nopaline synthase terminator. The MCS is shown in the pTRV2 plasmid map, this region was used to insert the gene of interest (GOI) into the plasmid. The coloured boxes in the MCS represent restriction enzymes.

CRISPR/Cas9 system for defence against viruses and foreign DNA (Mojica et al., 2005, Bolotin et al., 2005, Pourcel et al., 2005, Barrangou et al., 2007). In the native system, the Cas9 protein is guided by small interfering CRISPR RNAs (crRNAs) to the foreign DNA which contains a protospacer-adjacent motif (PAM) downstream of the target DNA sequence (5' NGG 3') (Brouns et al., 2008, Marraffini and Sontheimer, 2008). This forms a complex with transactivating CRISPR RNA (trancrRNA) (Deltcheva et al., 2011) to allow Cas9 to cleave the foreign DNA in the protospacer region, upstream of the PAM sequence (Garneau et al., 2010).

This system was modified to direct the Cas9 protein to a PAM sequence of target GOI DNA using a guide RNA (gRNA) for bacteria, mice and plants (Sapranauskas et al., 2011, Cong et al., 2013, Jiang et al., 2013, Jiang et al., 2015). The gRNAs consist of a PAM of the GOI target, fused to the 3' end of the crRNA with the 5' end of the trancrRNA (Cong et al., 2013) (figure 1-7). The reconstituted Cas9 introduces double strand breaks (DSBs) in DNA to induce DNA repair mechanisms by hosts. This system has been shown to be versatile in many plant species (Mao et al., 2013, Xie and Yang, 2013, Li et al., 2013, Svitashev et al., 2016).

There are two methods for DNA repair: non-homologous end joining (NHEJ) and homologous recombination (HR). NHEJ is the most common form of DNA repair in plants (Puchta, 2005). The CRISPR/Cas9 system used in this thesis relies on NHEJ DNA repair to create mutations in the DNA sequence which can result in knock-out mutations (Bortesi and Fischer, 2015). Whereas, there are some studies which use CRISPR/Cas9 with donor DNA with homology to the target sequence to insert genes or modify them through HR (Li et al., 2013, Sun et al., 2016). This method has been used in this thesis to understand the function of a candidate transcriptional regulator of the VTE pathway.

These methods (eQTL analysis, VIGS, transient OE and CRISPR/Cas9) were used to identify and characterise the function of candidate transcriptional regulators of the VTE pathway, in this thesis.

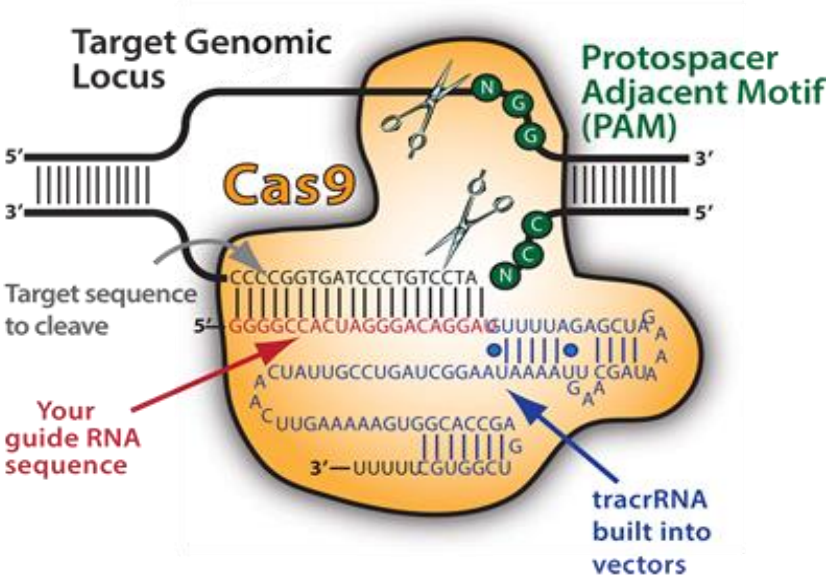


Figure 1-7 Schematic diagram of the modified CRISPR/Cas9 system to a protospacer adjacent motif (PAM) sequence of target gene of interest DNA. The guide RNA (gRNA) is introduced into a vector with the 5’ end of the trancrRNA. The Cas9 protein introduces double strand breaks of DNA to induce non-homologous end joining in hosts. This image was reprinted from a CRISPR/Cas9 information website (BioCat, 2018).

Chapter 2: Materials and Methods

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Chemicals

Chemicals were purchased from Invitrogen, Promega, Roche, Qiagen, Sigma, BioRad, Cambridge BioSciences and New England Biolabs. For LC-FDA, chemicals were purchased from Sigma, Merck Millipore, Cambridge Bioscience Ltd., Honeywell and Fisher Scientific.

2.1.2 Antibiotics

Antibiotics were used for selection of bacteria and transgenic plants. The following antibiotics were used: ampicillin, kanamycin, streptomycin, gentamycin, rifampicin and chloramphenicol. Table 2-1 shows the working concentrations for each purpose.

2.1.3 Plant materials

Solanum lycopersicum varieties: Moneymaker, M82 and Microtom tomatoes were used as wild type (WT) controls in different experiments, as specified.

Transgenic tomato lines were generated using pBINE8:SlMYB79 in a Microtom background and pBINE8:SlMYB71 in a Moneymaker background. The CRISPR tomato lines were generated using PICSL002203 gRNA1/2 in a Moneymaker background.

Table 2-1 List of antibiotics, the concentrations used and use.

Antibiotic	**Final concentration**	**Solvent**	**Purpose**
Ampicillin	100 µg ml^{-1}	ddH_20	*E.coli* /*A.tumefaciens* selection
Kanamycin	50µg ml^{-1}	ddH_20	50µg ml^{-1} for *E.coli* selection
Kanamycin	100 µg ml^{-1}	ddH_20	100 µg ml^{-1} for plant transformation
Streptomycin	100 µg ml^{-1}	ddH_20	*E.coli* selection
Gentamycin	50 µg ml^{-1}	ddH_20	*E.coli* selection
Rifampicin	50 µg ml^{-1}	DMSO	*A.tumefaciens* selection
Chloramphenicol	34 µg ml^{-1}	Ethanol	*E.coli* selection

2.1.4 Bacterial strains

The *Escherichia coli* strain, DH5α, was used for plasmid transformation. Gateway® plasmids were transformed into the ccdB survival strain, DB3.1 for propagation. *Agrobacterium tumefaciens*, strain Agl1 was used for stable transformations in tomato plants. For transient expression experiments, the *Agrobacterium tumefaciens* strain Agl1 was used.

2.1.5 Plasmids

All plasmid vectors used are listed in appendix.

2.1.6 Media recipes

For growth of bacteria and plants, Luria-Bertani broth (LB), Murashige and Skoog (MS), Super Optimal broth with Catabolite repression (SOC) and tryptone yeast (TY) were used. All recipes are listed in the appendix.

2.2 Methods

2.2.1 Primer design

Primers were designed using primer blast software (Ye et al., 2012). For molecular cloning of coding sequences and qRT-PCR, all primers were designed to be 20-26 nucleotides in length. Gateway® and restriction enzyme cloning primers were designed between 30-50 nucleotides in length. Primers to amplify 80-120 fragments were designed for qRT-PCR. The GC content for all primers was designed to be between 45-60% and a melting temperature of 60 degrees. A list of the primers used in thesis can be found in appendix.

2.2.2 Polymerase Chain Reaction (PCR)

Several polymerases were used for different PCR reactions; either Phusion, Go Taq or Taq polymerase. PCR reactions were undertaken by using G-Storm Thermal Cyclers (Kapa Biosystems) or ThermoFisher cyclers. The Phusion polymerase is a thermostable, high fidelity polymerase which is quicker and has a very low error rate, due to its proof reading ability. Phusion polymerase is useful for cloning with a low error rate, such as gene cloning. The taq based polymerases

are slower and have a much higher error rate, however, they are useful for colony PCRs of *E.coli* and *Agrobacterium* cultures and genotyping plants.

A standard Phusion PCR reaction (Invitrogen) consisted of a 20 µl reaction, with the following final concentrations of components: up to 250ng of DNA template, 200 µM dNTPS, 0.5 µM of F and R primers, 1 x HF Phusion buffer (1.5mM $MgCl_2$), 3% DMSO and 0.4 units of Phusion polymerase. A GC rich buffer (1.5mM $MgCl_2$) was used for GC rich templates, which was used for certain PCR reaction, such as for promoter cloning. The GC buffer was used at the same concentration as the HF buffer. The thermocycling conditions are: initial denaturation (98°C for 30 seconds), then 25-35 cycles of denaturation (98°C for 5-10 seconds), annealing (45-72°C for 10-30 seconds), extension (72°C for 15-30 seconds, per kb), and a final extension (72°C for 5 minutes).

PCR amplification using Go Taq polymerase (Promega) consisted of a final volume of 25 µl with the following final concentrations of components: up to 250ng of DNA template, 0.1-1.0 µM of F and R primers and 1 x Go Taq Green Master Mix (proprietary mix). The thermocycling conditions were: initial denaturation (95°C for 2 minutes), then 25-30 cycles of denaturation (95°C for 30 seconds), annealing (45-72°C for 30 seconds), extension (72°C for 1 minute, per kb), and a final extension (72°C for 5 minutes).

PCR using Taq polymerase (Qiagen) consisted of a final volume of 25 µl with the following final concentrations of components: up to 250ng of DNA template, 0.2 µM of F and R primers, 1 x Taq reaction buffer (Tris HCl, KCl, $(NH_4)_2SO_4$, 15 mM $MgCl_2$ (pH 8.7)) and 0.625 units of Taq DNA polymerase. The thermocycling conditions were: initial denaturation (95°C for 30 seconds), then 25-30 cycles of denaturation (95°C for 30 seconds), annealing (45-72°C for 30 seconds), extension (68°C for 1 minute, per kb), and a final extension (68°C for 5 minutes).

All PCR reactions were run on a 1% agarose gel at 130V for 30 minutes or longer for separation of DNA fragments.

2.2.3 Purification of DNA from PCR products or agarose

PCR products were purified from 1% agarose gels using the QIAquick Gel Purification Kit (Qiagen). Selected gel bands were cut from the agarose gel and weighed. Buffer QG (5.5 M guanidine thiocyanate, 20 mM Tris HCl (pH 6.6)) was added to the gel slice at a ratio of 3 x volume QG: 1 x gel slice volume (e.g. 200 µl = 200mg). The reaction was incubated for 10 minutes at 50°C, until the slice had dissolved. 1 volume of isopropanol was added to the solution and was pipetted onto a QIAquick column, and centrifuged for 1 minute at 17,900 x g The column was washed using 750µl of buffer PE (10 mM Tris-HCl (pH 7.5) in 80% ethanol) and centrifuged for 1 minute at 17,900 x g. The product was then eluted using water, or, buffer EB (10 mM Tris-Cl (pH 8.5)). PCR reactions or restriction digest products were purified directly from the reactions using the QIAquick PCR Purification Kit (Qiagen) following the manufacturer's instructions. Buffer PB (5M guanidine-HCl in 30% isopropanol) was added to the PCR sample at a ratio of 5 volumes PB : 1 volume of PCR sample. The solution was added to a QIAquick spin column and centrifuged for 1 minute at 17,900 x g, to bind DNA. The column was washed using 750µl of buffer PE and centrifuged for 1 minute at 17,900 x g. The product was then eluted using water of buffer EB.

2.2.4 Preparation of competent *Escherichia coli* cells

A single colony of *E.coli* (DH5α and other strains) was grown overnight in a 10ml LB culture at 37°C, at 220rpm. An aliquot (3ml) of this culture was used as inoculum in 400ml of LB and grown at 37°C, and 220rpm, until an OD_{550} of 0.5 was reached. The culture was centrifuged for 10 minutes at 3,000 x g. The remaining pellet was resuspended in 1/3rd of its original volume of TBD medium (recipe in appendix). The bacterial culture was incubated on ice for 90 minutes, before being centrifuged for 5 minutes at 1000 x g. The supernatant was removed and the pellet was resuspended in 1/25th of its original volume, and 0.2ml aliquots were placed into Eppendorf tubes. Cells were immediately frozen in liquid nitrogen and frozen at -80°C for long term storage.

2.2.5 *E.coli* transformation

Competent cells were transformed with plasmids using heat shock transformation. Cells were thawed on ice and approximately 100ng of plasmid were added to the cells and mixed. The cells were then kept on ice for 30 minutes and heat shocked at 42°C for 1 minute. The tube was then put on ice for 5 minutes and 500µl of SOC was added (recipe in appendix). The cells were incubated at 37°C for 1-2 hours in a shaker at 220rpm and then centrifuged at 9,000 x g for 3 minutes and 30 seconds. The pellet was resuspended in 30µl of SOC and plated onto LB agar plates with the appropriate antibiotic for the plasmid. Plates were put into the 37°C oven overnight and colonies were picked the next day.

2.2.6 *Agrobacterium tumefaciens* competent cell preparation

A single colony of *Agrobacterium* was grown for 1-2 days (strain dependent) in a 10ml TY culture at 28°C (recipe in appendix). This culture was used as an inoculum in 200ml of TY and grown at 28°C, at 220rpm, for 1-2 days. The culture was centrifuged for 10 minutes at 4,000 x g and the pellet was re-suspended in 50ml of ice cold water. This step was repeated for three successive washes with water and 10% glycerol. The solution was then centrifuged for 10 minutes at 1,400 x g and the pellet was re-suspended in 2ml of 10% glycerol. A 50µl sample was taken from the bacteria and electroporated at the following settings: 400 ohms Ω, 25 µFD and 2.5 kV to check if the time constant was between 4-5 seconds. The time constant is a measure of how efficient the cells acquire DNA. If the time constant was between 4-5 seconds, the cells were aliquoted into Eppendorf tubes and frozen at -80°C for long term storage. If the time constant was higher than 5 seconds, the cells were washed with 10% glycerol until the optimum time constant was achieved.

2.2.7 *Agrobacterium* transformation

Competent cells were transformed with plasmids using electroporation. Cells were thawed on ice and 250ng of plasmid DNA was added to cells and mixed. Cells were kept on ice of 30 minutes and then transferred to an electroporation

cuvette. A BioRad Pulser (BioRad Laboratories) was used to transform the electrocompetent Agrobacterium cells with the following settings: 400 Ω, 25 µFD and 2.5 kV. 500µl of SOC was added to the cells and incubated at 28° C for 2-3 hours in a shaker at 220rpm and then centrifuged at 1,400 x g for 3 minutes and 30 seconds. The pellet was resuspended in 50µl of SOC and plated onto TY agar plates with the appropriate antibiotic for the plasmid. Plates were put into the 28°C oven overnight and colonies were picked 2-3 days later, depending on the strain.

2.2.8 Plasmid DNA isolation from *E.coli*

Plasmid DNA was extracted from 10ml LB cultures grown overnight, with an appropriate antibiotic selection. All plasmid DNA was extracted and purified using the QIAprep Miniprep Kit (Qiagen), which uses alkaline lysis to release the plasmid DNA which is then adsorbed onto a silica membrane, in a high salt environment. The bacterial culture was centrifuged for 10 minutes at 1,400 x g, to form a pellet. The pellet was resuspended in 250µl of buffer P1 (50mM Tris-Cl (pH 8.0), 10mM EDTA, 100µg/mL RNase A), and then added to 250µl of buffer P2 (200mM NaOH, 1% SDS) to start the lysis. The reactions were incubated for 5 minutes, then 350µl of buffer N3 (4.2 M Gu-HCl, 0.9M potassium acetate, (pH 4.8)) was added to stop the lysis. All reactions were centrifuged for 10 minutes at 18,000 x g, which formed a supernatant and white pellet. The supernatant (approximately 800µl) was added to a QIAprep 2.0 spin column (with the silica membrane) and centrifuged for 1 minute at 18,000 x g, and the flow through was discarded. 750µl of buffer PE (10mM Tris-HCl, in 80% ethanol (pH 7.5)) was added to the column and centrifuged for 1 minute at 18,000 x g and the flow through was discarded. The latter centrifugation was repeated to dry the column. To elute the plasmid DNA from the column, 50µl of EB (10mM Tris-Cl (pH 8.5)) was added to the column and centrifuged for 1 minute at 18,000 x g.

2.2.9 Quantification of RNA and DNA

All RNA and DNA was quantified using a Thermo Nanodrop 2000C UV-Vis Spectrophotometer following the manufacturer's instructions.

2.2.10 DNA isolation from plants

Plant leaf tissue was ground in the genome grinder, using glass beads for 1 minute at 1,000 x g, to a fine powder. The powder was used to extract DNA using the DNeasy plant mini kit (Qiagen), which included proprietary buffer compositions. AP1 buffer (400µl) and 100mg ml^{-1} RNAse A (4µl) was added to approximately 100mg of powder. The samples were incubated for 10 minutes at 65 °C to lyse the cells. 130µl of buffer P3 was added to the lysate and incubated on ice for 5 minutes. This step precipitates proteins and polysaccharides. The samples were centrifuged for 5 minutes at 20,000 x g to remove the larger precipitates and glass beads. The supernatant was centrifuged in a QIAshredder mini spin column for 2 minutes at 20,000 x g. The flow through was moved to a new tube and 1.5x volume of buffer AW1 was added and mixed by pipetting. The mixture (650µl) was centrifuged for 1 minute at 6,000 x g in a DNeasy mini spin column. The flow through was discarded and this step was repeated until the all the mixture had been loaded onto the column. Buffer AW2 (500µl) was added to the column and centrifuged for 1 minute at 6,000 x g, and the flow through discarded. The latter step was repeated, however, the centrifugation step was increased to 2 minutes at 20,000 x g, to dry the membrane. The DNA was then eluted using buffer AE and incubated for 5 minutes at room temperature, before being centrifuged for 1 minute at 6,000 x g.

2.2.11 RNA isolation from plants

RNA from tomato fruits was extracted using TRIzol, which is a monophasic solution, containing; guanidine thiocyanate and phenol. The tomatoes were chopped, seeds were removed, and the remaining tissues were ground into a powder using liquid nitrogen and a pestle and mortar. Approximately 200mg of powder was added to 1.5ml of TRIzol reagent (TRI Reagent, Sigma) and vortexed.

The samples were left to incubate, whilst shaken, at room temperature for 5 minutes. 1-bromo 3-chloropropane (BCP) (150µl) was added to the sample which was then shaken vigorously. Once the phases started to separate, the sample was centrifuged for 10 minutes at 20,000 x g, at 4°C. The aqueous upper phase was transferred to a new tube and an additional 150µl of BCP was added and shaken. The sample was centrifuged again for 10 minutes at 20,000 x g, at 4°C. The aqueous upper phase was removed and transferred to a new tube. Isopropanol (750 µl) was added to the sample and mixed by inversion. The sample was then incubated at -20°C for 2 hours (at this stage the sample could also be left overnight). After incubation, the sample was centrifuged for 15 minutes at 20,000 x g, at 4°C. The supernatant was removed and the pellet was washed with 1ml of 75% ethanol and centrifuged for 5 minutes at 20,000 x g at 4°C. The latter step was repeated. The supernatant was completely removed and the remaining pellet was left to dry in the air for 5 minutes, before being resuspended in 40µl RNAse-free water.

RNA for qRT-PCR was re-precipitated to remove any phenol contamination from the TRIzol reagent. A solution of 4µl of 3M sodium acetate and 100µl of 100% ethanol was added to the RNA and incubated overnight at -20°C. The RNA was centrifuged for 1 hour at 20,000 x g, at 4°C. The remaining pellet was washed twice with 70% ethanol and centrifuged after each wash for 30 minutes at 20,000 x g at 4°C. The resulting pellet was resuspended in 40µl of RNAse free water and centrifuged for 5 minutes at 20,000 x g at 4°C, and the supernatant was moved to a new tube.

After extraction, all RNA was treated with DNAse. 8µl of RNA was added to 1µl of 10x reaction buffer and 1µl of DNAseI (1 unit/µl) (Sigma). The reactions were left for 15 minutes at room temperature. 1µl of stop solution (50mM EDTA) was added to the reactions (this binds calcium and magnesium ions) to inactivate the DNAse, and heat treated for 10 minutes at 70°C.

2.2.12 cDNA synthesis

cDNA was synthesised by one of two methods: first strand cDNA synthesis or applied Biosystems high capacity cDNA synthesis kit. For molecular cloning, the first strand cDNA synthesis method was applied, whereas for qRT-PCR, the applied Biosystems high capacity cDNA kit was used.

2.2.12.1 First strand cDNA synthesis

For first strand cDNA synthesis, Superscript™ III (Invitrogen) was used. Superscript™ III is a thermos-tolerant version of reverse transcriptase allowing first strand cDNA synthesis at elevated temperatures. Up to 3µg of RNA was used from the RNA extraction (above). There were two steps to the cDNA synthesis: primer annealing and reverse transcription. For the primer annealing, a primer mix of equal volumes of 10µM oligo dT and 10µM random primers were added. dNTPs (10mM) and water were also added, before being left at 65°C for 5 minutes, to the denature the RNA and allow primer annealing. After the primer annealing step, 5x first strand buffer is added, along with 0.1M DTT, RNAse out (34units/µl) and Superscript III. The reaction was incubated at 50°C for 60 minutes and then for 70°C for 15 minutes, to kill the enzyme. The cDNA was kept at -20°C for long term storage.

2.2.12.2 Applied Biosystems High Capacity cDNA synthesis

The applied biosystems high capacity cDNA synthesis requires 2µg of RNA. The reaction mixture consisted of: 10x RT buffer, 25x dNTPs (100mM), 10x RT Random Primers, MultiScribe™ Reverse Transcriptase, RNAse inhibitor and nuclease free H_2O, up to a total volume of 20µl. This was added to 10µl of RNA (RNA was extracted as previously described). The reaction was incubated at 25°C for 10 minutes, 37°C for 120 minutes and 85°C for 5 minutes. The cDNA was kept at -20°C for long term storage.

2.2.13 Real time quantitative PCR (qRT-PCR)

A LightCycler® 480 Instrument was used for qRT-PCR and required LightCycler® 480 SYBR Green I Master Mix for use in a LightCycler® 480 Multiwell

Plate 384. The reactions used cDNA synthesised using the applied Biosystems high capacity cDNA kit. The 2x SYBR Green I Master Mix (FastStart Taq DNA Polymerase, reaction buffer, dNTP mix (with dUTP instead of dTTP), SYBR Green I dye, and $MgCl_2$) consisted of half of the reaction volume, with a mix of 10µM forward and 10µM reverse primers and water. The qRT-PCR reactions consisted of four stages: pre-incubation, amplification, melting curve and cooling. The pre-incubation phase was at 95°C for 5 minutes. The amplification step consists of 45 cycles of: 95°C for 15 seconds, 60°C for 15 seconds and 72°C for 20 seconds. The melt curve consisted of: 95°C for 5 seconds, 60°C for 1 minute, the temperature was then continuous up to 97°C and 5 acquisitions were taken per °C. The plate was then cooled to 40°C for 30 seconds. The Ct/Cp values were calculated using the second derivative maximum method. This differs from the fit-point method, which determines Ct values by setting a threshold, during the amplification in the log-linear phase. The second derivative method determines the Cp value by using measuring the point of concavity of the log-linear amplification phase, and therefore, cannot be determined by the user. The Ct and Cp values describe the same given value that can be used to determine relative fold change of expression, but, are named differently depending on the method.

2.2.14 Cloning

2.2.14.1 Restriction enzyme cloning

Genes cloned for the pTRV2 over expression vector were cloned using traditional restriction enzyme cloning. PCR products were amplified with restriction enzyme sites in the primer oligos and were purified using QIAquick Gel Purification or PCR Purification Kits (Qiagen). PCR products and plasmids were digested using appropriate restriction enzymes (full list in appendix) and separated on a 1% agarose gel using electrophoresis. Selected bands were extracted from the gel using the QIAquick Gel Purification (Qiagen) kit. The restriction digest products were ligated into the appropriate plasmids, following digestion with selected restriction enzymes, with T4 DNA ligase, at a molar ratio of 3:1 (insert:

plasmid). This ligation reaction was transformed into *E.coli* (strain DH5α) for propagation of the plasmid, with appropriate antibiotics, and extracted using the QIAprep Miniprep Kit (Qiagen) kit.

2.2.14.2 Gateway® cloning

PCR products were amplified using oligonucleotides with additional Gateway® recombination sequences and purified using QIAquick Gel Purification or PCR Purification Kits (Qiagen). PCR products with attB1 and attB2 sites were recombined into pDONR207/pDONR221 using BP clonase™ (Invitrogen), following the manufacturer's instructions, at room temperature and overnight. The reaction was stopped using 1µl of proteinase K and incubated at 37°C for 10 minutes. The BP reaction mix containing the recombined plasmid was transformed into *E.coli* (strain DH5α) for propagation of the entry vector and plasmids were extracted using a QIAprep Miniprep Kit (Qiagen) kit.

The insert from the entry vector was recombined into the destination vector, using LR clonase™ (Invitrogen), following the manufacturer's instructions, at room temperature and overnight. The reaction was stopped using 1µl of proteinase K and incubated at 37°C for 10 minutes. The LR reaction, combining the insert in the destination vector, was transformed into *E.coli* (strain DH5α) for propagation of the destination vector and the plasmid was extracted using the QIAgen Miniprep kit (Qiagen).

2.4.14.3 Golden gate cloning

Guide RNAs (gRNAs) for CRISPR constructs were ordered as oligos for Golden gate cloning. This cloning technique is based on Type IIS restriction enzymes, which cut outside of their restriction site. This allows for multiple assembly of many parts at the same time. All plasmids used in this section were supplied by TSL SynBio platform.

The gRNAs were designed with flanking BsaI recognition sequences (known as Level 0 parts). These were ligated into a Level 1 acceptor plasmid, using conventional restriction enzyme cloning. This method combines the digestion and

ligation reactions in one step, which is called a Digestion-Ligation (Dig-Lig). The ligated Level 1 constructs containing the gRNAs, under the control of the U6 promoter were ligated into a Level 2 plasmid, using the Dig-Lig method. The level 2 plasmid contained the domesticated Cas9 under the control of a double 35S promoter. These plasmids were transformed into *E.coli* (strain DH5α) for propagation of the Level 2 plasmid and was extracted using the QIAgen Miniprep kit (Qiagen).

2.2.15 Viral Induced Gene Silencing (VIGS)

VIGS fragments were designed using the Solgenomics network (SGN) VIGS tool (Fernandez-Pozo et al., 2015b). All fragments were approximately 300 nucleotides in length, and the VIGS tool was set to check for possible homology of 21 nucleotides in length. The tool blasted VIGS fragments generated to the tomato genome, to check for off target sequences, and the fragment with the least matches was used. Primers were created for VIGS fragments and Gateway® primers were designed using Vector NTI (supplemental table 6-1).

Fragments were cloned using PCR and cloned into pDON207 using the Gateway® system (as described previously). The VIGS fragments were then transformed into the destination plasmid (pTRV2 *Del/Ros*) using the Gateway® system. These plasmids were transformed into Agl1 strain of *Agrobacterium tumefaciens.*

Single colonies of the transformed *Agrobacterium* were grown in a liquid 10ml culture with appropriate antibiotics for two days at 28°C. The cultures were used as inoculum for 100ml cultures with appropriate antibiotics and grown for 2 days at 28°C. After propagation, cultures were centrifuged at 1,400 x g for 15 minutes. The remaining pellet was washed twice with infiltration medium (10mM MES, 10mM $MgCl_2$, 200µM acetosyringone, pH5.6) and centrifuged at 4,000 x g for 15 minutes. The pellet remaining was resuspended to an OD_{600}=0.5 for Moneymaker fruit and OD_{600}=0.2 for Microtom fruit. Cultures of the pTRV2 *Del/Ros GOI* were mixed with cultures of pTRV1 at a ratio of 1:1.

All tomato fruit was injected at the mature green stage, through the top of the tomato fruit, adjacent to the peduncle. Microtom fruit were injected with 0.3ml of washed *Agrobacterium,* and Moneymaker fruit were injected with 0.5ml of *Agrobacterium.* Tomato fruit were harvested fourteen days post breaker, as described by Orzaez et al. (2009). All further RNA extractions, qRT-PCR and tocochromanol extraction was completed, as described in chapter 2. A full list of primers used in this experiment are included in the appendix.

2.2.16 Transient over expression of genes in tomato fruit

The transient viral over-expression system (pTRV2) was created, and kindly provided, by Dr Vera Thole. Coding sequences (CDS) of the candidate genes were cloned using Phusion PCR, with additional restriction enzymes added to the primers used (described previously). The sequences were ligated into the pTRV2 over expression vector and propagated in *E.coli* for plasmid extraction.

pTRV2 OE *GOI* vectors and pTRV1 were transformed into the *Agrobacterium* strain GV3101 containing pSOUP. Cultures were grown in the same way as the VIGS vectors. pTRV1 and pTRV2 OE *GOI* vectors were mixed at a ratio of 1:1 and injected into ten days post breaker Moneymaker tomato fruit. Tomato fruit was harvested five days post injection for analysis. All further RNA extractions, qRT-PCR and tocochromanol extraction was completed, as described in chapter 2. A full list of primers used in this experiment are included in the appendix.

2.2.17 Tomato transformation

Tomato seeds (the cultivar is specified for each transformation in section 2.1.3) were sterilised using 10% bleach for 30 minutes and were then washed liberally with ddH_20 water. Between 100-200 tomato seeds were plated onto germination media (recipe in appendix) into Sigma phytatray II tubs and placed at 4°C until needed. Seeds could be stored at this temperature for up to a month. The germination tubs were placed in the growth room (16-hour photoperiod of

incandescent light, at 23°C) for 7-10 days until the cotyledons had emerged, but before the true leaves had grown.

For tomato transformation, an *Agrobacterium* culture was propagated with the desired plasmid for transformation, and grown overnight at 28°C. During this time, suspension media plates were made and 1ml of *Nicotiana benthamiana* suspension culture cells were plated onto the media. The suspension plates were left covered in the growth room over night before use.

The next day, the plates with the *N.benthamiana* feeder layer were covered with sterilised Whatman No.1 filter paper, ensuring all air bubbles were excluded. Each cotyledon was cut into two pieces and the tip of the cotyledon was removed. Once completed, the *Agrobacterium* culture was spun down and washed with MS media (recipe in appendix), this step was repeated twice. The pellet was resuspended in the MS media and diluted to an OD_{600} of 0.4-0.5.

Cotyledon pieces were submerged in the *Agrobacterium* culture and then removed and blotted on sterile filter paper. The cotyledon pieces were then placed adaxial surface down onto the suspension plates and co-cultivated with the feeder layer in the growth room for 2 days.

After two days, cotyledons were transferred to regeneration plates, with the adaxial surface facing upwards, so the cotyledons curled into the media. The plates were placed in the growth room for two weeks. The cotyledons started to form callus and were moved to fresh regeneration plates every two weeks for approximately 4 months. Explants started to develop and were separated from the callus and placed into rooting medium. These explants were left in rooting medium for 6 weeks whilst they produced roots. Once a strong root had formed, plantlets were transferred to soil and placed in the greenhouse.

2.2.18 Tocochromanol extraction and analysis

Tocochromanols were extracted and either quantified by absolute or relative measurements using LC-FDA detection. The method has been adapted from (Almeida et al., 2011). To extract tocochromanols from plant tissues, samples

were ground into a powder using liquid nitrogen. 1.5ml of methanol containing 0.1% butylated hydroxytoluene (BHT) were added to 250-500mg of tomato powder and vortexed. 1ml of chloroform was added to the samples and they were sonicated at 4°C for 5 minutes. After the samples had been sonicated, 1ml of Tris buffer (50mM Tris (pH7.5, 1M NaCl) was added and the chloroform phase separated from the methanol phase. The chloroform phase was removed and placed in a new tube. The remaining methanol phase was re-extracted with 1ml of chloroform and the chloroform phase was placed into the new tube. The chloroform phases were dried under nitrogen gas until all liquid had evaporated. The samples were resuspended in 0.25ml of 99.5:0.5 heptane: isopropanol. The samples were filtered using a filter Eppendorf (CoStar SpinX 0.2µm filter) for 2 minutes at 20,000 x g. The samples were then diluted 1:4 or 1:10 for sample injection, as specified for each experiment.

Tocochromanol analysis was carried out using the Shimadzu single quad using LC with fluorescence detection and separation on a Merck Millipore LiChrospher ® Si 5µm LIChroCART ® 250-4 column. The mobile phase was an isocratic solvent solution of 99.5:0.5 heptane: isopropanol for 34 or 45 minutes, for tocopherols or tocochromanols, respectively. Eluting compounds were detected and quantified by fluorescence with excitation at 296 nm and emission at 340 nm. Identification and quantification of the compounds was obtained by comparison with tocochromanol standards. The tocopherol standards were obtained from Merck Millipore and the tocotrienols standards were from Cambridge Biosciences Ltd. The standard curves were set at four concentrations, from 1.78 to 95.60 pmoles μl^{-1} and 0.35 to 7.82 pmoles μl^{-1} for tocopherols and tocotrienols, respectively.

Carotenoid extraction and analysis

2.2.19 Carotenoid extraction and analysis

Tomato samples were freeze dried and 400mg was weighed out for each sample. An aliquot of 100µl of 5M NaCl and 50µl of hexane was added to the

sample and shaken for 30 seconds. The sample was centrifuged for 2 minutes at 20,000 x g. Dichloromethane (200µl) was added to the sample, shaken for 30 seconds and centrifuged for 2 minutes at 20,000 x g. Ethyl acetate (1ml) (including 0.1% butylated hydroxytoluene) was added to the sample. The sample was then agitated for 10 minutes and centrifuged (as described previously). The supernatant was moved to a new tube and the previous steps were repeated on the remaining pellet. The supernatant was centrifuged for 2 minutes at 20,000 x g in a CoStar SpinX 0.22m filter column. These samples were then used for injection.

All further metabolite analysis was completed by Baldeep Kular. Carotenoid analysis was carried out using the Agilent 12900 with UV detection and separation on a Luna 3µm C18(2) 2 X 100 mm column. The mobile phase was binary gradient comprising of acetonitrile (90% v/v) up to 20 minutes. A gradient of 40% acetonitrile (90% v/v) and 60% ethyl acetate (100% v/v) was set from 20-22 minutes, before 100% acetonitrile (90% v/v) for the remaining 3 minutes. Eluting compounds were detected using UV and measured at the following wavelengths; lycopene (454nm) and β-carotene (454nm). The carotenoid standards were obtained from Sigma. The standard curves were set at four concentrations for each carotenoid. Lycopene standards were from 0.03 to 0.61 µg ml^{-1} and β-carotene standards were set at 0.03 to 0.75 µg ml^{-1}.

2.2.20 Statistical tests

The statistical tests carried out were two tailed t-tests or Tukey tests. All p values are described as either being less than 0.05 or 0.01.

Chapter 3: Identification of candidate genes using co-expression QTL analysis of *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression lines

Chapter 3: Identification of candidate genes using co-expression QTL analysis of *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression lines

3.1 Introduction

3.1.1 Expression Quantitative Trait Loci (eQTL) analysis

Underlying transcriptional regulator networks of the VTE pathway have not yet been identified, even though transcriptional regulation of this pathway must exist (Quadrana et al., 2013). Therefore, I have used the well-established *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression lines (ILs) (figure 3-1) to identify candidate TFs of VTE biosynthesis.

ILs of many species have been used to characterise the molecular function of quantitative trait loci (QTL) (Lippman et al., 2007). Expression QTL (eQTL) is an advancement of this traditional method, which combines high-throughput genome studies with QTLs. This method can be used to identify regions of chromosomes that are specifically linked to expression of genes. eQTLs can act locally (cis-) or distally (trans-) to the loci of interest (Alberts et al., 2007, Cubillos et al., 2012). cis-eQTLs tend to locate to the enzymes controlling the trait of interest. For example, in VTE biosynthesis, a cis-eQTL was identified for VTE3(1) in tomato, that was linked to an mQTL on IL9-2-6 (Quadrana et al., 2014). Therefore, there has been great interest in identifying trans-eQTLs, to identify regions which may contain candidate regulators of a trait.

3.1.2 *S.lycopersicoides* x *S.lycopersicum* cv. VF36 introgression lines

Additionally, another *Solanum* species IL population was available for eQTL analysis (Canady et al., 2005). The *Solanum lycopersicoides* x *S.lycopersicum* cv. VF36 ILs are a cross between *S.lycopersicoides* and *S.lycopersicum* cv. VF36. *S.lycopersicoides* is a wild nightshade-like tomato, which is resistant to many abiotic and biotic stresses. Several of these ILs were heterozygous due to sterility issues. However, the lines that were homozygous were useful to compare trans-eQTL analysis between *S.pennellii* x *S.lycopersicum* cv. M82 ILs and *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs.

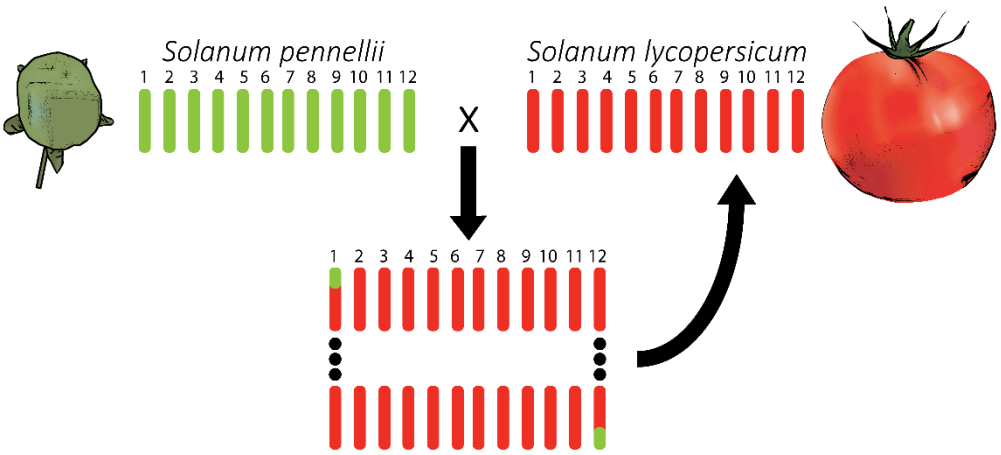


Figure 3-1 Schematic diagram of *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression lines (ILs). *S.pennellii* (green fruit) was crossed with *Solanum lycopersicum* cv.M82 (red fruit). Progeny from this cross were backcrossed to *S.lycopersicum* to form ILs, providing full genome coverage.

I have used this method using the RNA sequencing data of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs to identify candidate TFs that reside in trans-eQTLs for VTE biosynthesis.

3.2 Materials and Methods

3.2.1 *S.pennellii* x *S.lycopersicum* cv. M82 IL RNA sequencing data and analysis

RNA sequencing data for the 76 *S.pennellii* x *S.lycopersicum* cv. M82 ILs was obtained from the tomato functional genomics database (Lee and Giovannoni), which is publicly available. The transcriptome analysis was based on ripe tomato pericarp and was harvested when 80% of the tomatoes were ripe.

To screen the IL population RNA sequencing data, I looked for co-expression of the vitamin E (VTE) pathway genes and identified regions that showed at least 20% changes in expression (relative to *S.lycopersicum* cv. M82). Precursor pathways to the VTE pathway (methyl erythritol phosphate – (MEP), and shikimate – (SK) pathways) were not included in the trans-eQTL analysis, as using this method for many genes is difficult. Therefore, in the screening, only VTE structural genes were used for co-expression analysis. Once a trans-eQTL had been identified, precursor pathways were checked retrospectively.

Initially, VTE2 was used as the key enzyme in VTE biosynthesis and was the basis for trans-eQTL identification. VTE6 was identified as the missing phytol kinase and added to the analysis later. The gene shown in the results is the putative functional homolog of VTE6 in *Arabidopsis thaliana* (vom Dorp et al., 2015, Wang et al., 2017).

Chromosomal regions harbouring trans-eQTLs were mined for transcriptionally related genes. From this list of genes, TFs were selected. Finally, these candidate TFs were screened for co-expression with the pathway genes, and with precursor pathways (MEP and SK pathways). Candidate genes identified were cloned and screened in transient assays, as described in chapters 5 and 6.

3.2.2 *S.pennellii* x *S.lycopersicum* cv. M82 IL parents RNA sequencing data analysis

Published RNA sequencing data of the *S.pennellii* x *S.lycopersicum* cv. M82 IL parents; *S.lycopersicum* and *S.pennellii* were used to calculate expression ratios for VTE pathway genes and candidate genes, which was a ratio of expression between the two parents (Koenig et al., 2013). Statistical significance of the expression of a given gene was calculated between these two parents and used to support the trans-eQTL analysis.

3.2.3 *S.pennellii* x *S.lycopersicum* cv. M82 leaf RNA sequencing data analysis

RNA sequencing data for the leaf material from the 76 *S.pennellii* x *S.lycopersicum* cv. M82 ILs were published by Chitwood et al. (2013). M82 was not included in this analysis, therefore all fold changes in expression are expressed relative to the average expression of all *S.pennellii* x *S.lycopersicum* cv. M82 ILs. These data (Chitwood et al., 2013) were used for co-expression of VTE pathway genes to support trans-eQTLs identified from the fruit *S.pennellii* x *S.lycopersicum* cv. M82 RNA sequencing data.

3.2.4 *S.lycopersicoides* x *S.lycopersicum* cv. VF36 IL RNA sequencing and metabolite data analysis

The *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs were used as a resource to compare identified trans-eQTLs from the *S.pennellii* x *S.lycopersicum* cv. M82 fruit IL RNA sequencing data. Corresponding IL expression profiles were compared between the *S.pennellii* x *S.lycopersicum* cv. M82 and *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs and used to support the trans-eQTLs identified from the *S.pennellii* x *S.lycopersicum* cv. M82 fruit IL RNA sequencing data.

RNA sequencing data for 54 *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs (Canady et al., 2005) were completed by several collaborators. The draft genome and scaffold for transcriptome of *S.lycopersicoides* was carried out by Dr Björn Usadel (Aachen University, Germany). The RNA sequencing library construction was carried out on tomatoes grown in the US by Professor James Giovanni (Boyce

Thompson Institute and USDA-ARS, USA) and Professor Harry Klee (University of Florida, USA). Tomatoes were ripe when harvested. The libraries were drafted and annotated by Dr Björn Usadel and Professor Zhangjun Fei (Boyce Thompson Institute and USDA-ARS, USA).

S.lycopersicoides x *S.lycopersicum* cv. VF36 ILs GC-MS metabolite analysis was completed by Dr Saleh Alseekh (Max Planck Institute, Germany). The metabolite analysis was carried out on tomatoes grown in Israel by Professor Dani Zamir (Hebrew University, Israel) and harvested by Dr Dario Breitel (John Innes Centre, UK). All *S.lycopersicoides* x *S.lycopersicum* cv. VF36 IL work was a collaborative project called RegulaTomE and was funded by BBSRC under the ERA-CAPS II call.

3.3 Results

3.3.1 Screening of *S.pennellii* introgression lines RNA sequencing data for co-expression of the VTE pathway

The VTE pathway consists of seven genes which are needed for VTE biosynthesis (figure 3-2). The VTE pathway genes were analysed in the *S.pennellii* x *S.lycopersicum* cv. M82 IL RNA sequencing data using co-expression analysis. This co-expression analysis showed each IL had its own unique expression profile (figure 3-3). The ILs with the greatest changes in expression (relative to the *S.lycopersicum* parent - M82) are key for elucidating eQTLs.

Several ILs showed significant changes of VTE pathway gene expression, however, the majority tend to be cis-eQTLs (figure 3-4). The cis-eQTLs usually involve variation in expression of structural genes encoding enzymes in the pathway. ILs exhibiting cis-eQTLs include: IL7-2, IL7-3, IL7-4-1, IL8-2, IL8-2-1, IL9-1, IL9-2 and IL9-2-6. Cis-eQTLs that were identified are based on a single VTE pathway gene, as the rest of the pathway genes did not co-express with the cis-eQTL gene. *VTE3(1)* and *VTE3(2)* are two isoforms of *VTE3*, which is a methyltransferase

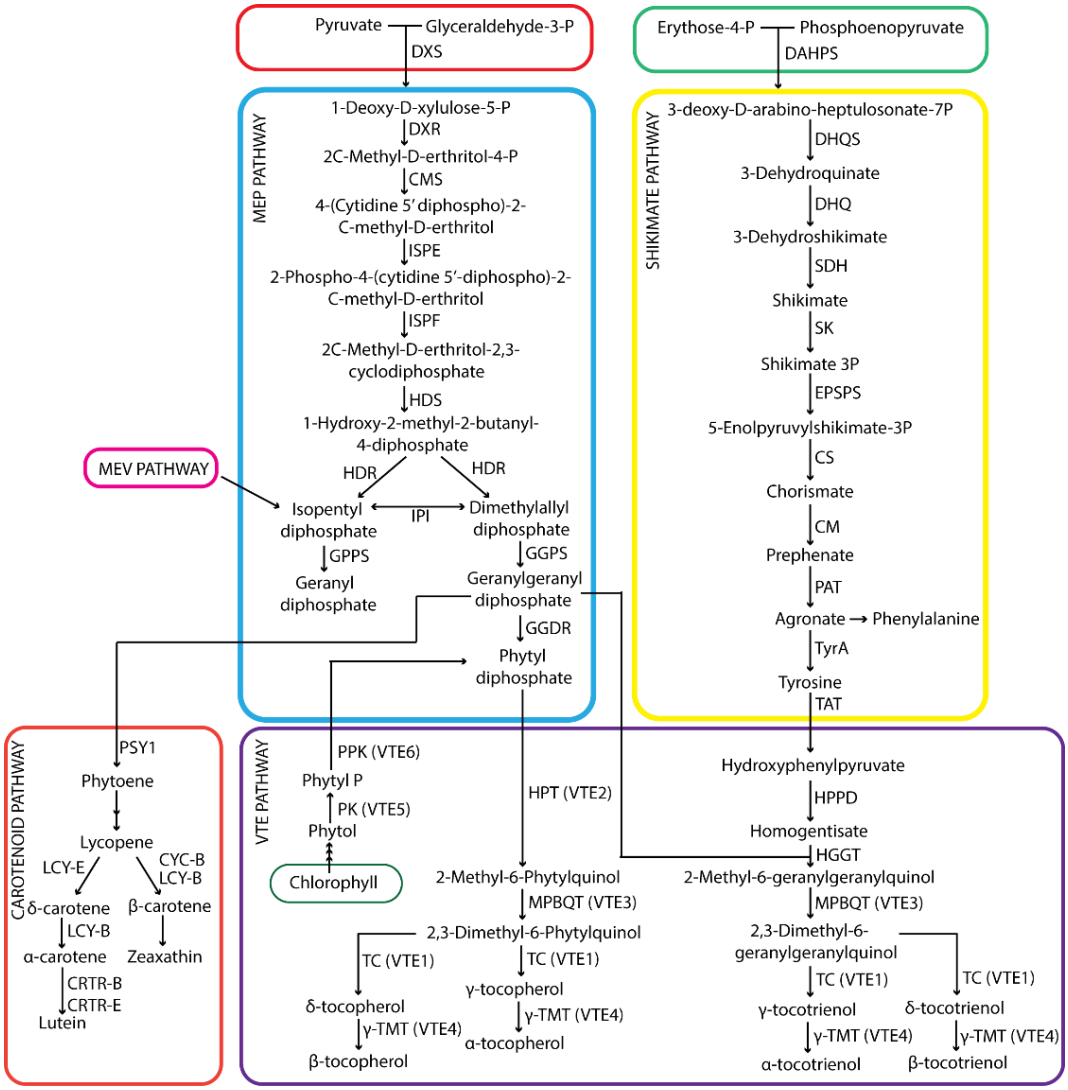

Pyruvate
Glyceraldehyde-3-P
DXS
Erythose-4-P
Phosphoenopyruvate
DAHPS
MEP PATHWAY
1-Deoxy-D-xylulose-5-P
DXR
2C-Methyl-D-erthritol-4-P
CMS
4-(Cytidine 5′diphospho)-2-C-methyl-D-erthritol
ISPE
2-Phospho-4-(cytidine 5′-diphospho)-2-C-methyl-D-erthritol
ISPF
2C-Methyl-D-erthritol-2,3-cyclodiphosphate
HDS
1-Hydroxy-2-methyl-2-butanyl-4-diphosphate
HDR
HDR
MEV PATHWAY
Isopentyl diphosphate
IPI
Dimethylallyl diphosphate
GPPS
GGPS
Geranyl diphosphate
Geranylgeranyl diphosphate
GGDR
Phytyl diphosphate
SHIKIMATE PATHWAY
3-deoxy-D-arabino-heptulosonate-7P
DHQS
3-Dehydroquinate
DHQ
3-Dehydroshikimate
SDH
Shikimate
SK
Shikimate 3P
EPSPS
5-Enolpyruvylshikimate-3P
CS
Chorismate
CM
Prephenate
PAT
Agronate → Phenylalanine
TyrA
Tyrosine
TAT
CAROTENOID PATHWAY
PSY1
Phytoene
Lycopene
LCY-E
CYC-B
LCY-B
δ-carotene
β-carotene
LCY-B
α-carotene
Zeaxathin
CRTR-B
CRTR-E
Lutein
VTE PATHWAY
PPK (VTE6)
Phytyl P
PK (VTE5)
Phytol
Chlorophyll
HPT (VTE2)
Hydroxyphenylpyruvate
HPPD
Homogentisate
HGGT
2-Methyl-6-Phytylquinol
2-Methyl-6-geranylgeranylquinol
MPBQT (VTE3)
MPBQT (VTE3)
2,3-Dimethyl-6-Phytylquinol
TC (VTE1)
TC (VTE1)
δ-tocopherol
γ-tocopherol
γ-TMT (VTE4)
γ-TMT (VTE4)
α-tocopherol
β-tocopherol
2,3-Dimethyl-6-geranylgeranylquinol
TC (VTE1)
TC (VTE1)
γ-tocotrienol
δ-tocotrienol
γ-TMT (VTE4)
γ-TMT (VTE4)
α-tocotrienol
β-tocotrienol

Figure 3-2 Outline of the tocochromanol pathway. MEP, SK, carotenoid and VTE pathway in blue, yellow, orange, and purple, respectively. Enzyme names are as follows: **DXS**; 1-Deoxy-D-xylulose-5-P synthase, **DXR**; 2C-Methyl-D-erythritol-4-phosphate synthase, **CMS**; 2C-Methyl-D-erythritol-4-phosphate cytidyltransferase, **ISPE**; 4-2-CMethyl-D-erythritol kinase, **ISPF**; 2C-Methyl-D-erythritol-2-3-cyclodiphosphate synthase, **HDS;** 4-Hydroxy-3-methylbut-2-enyl-diphospahte synthase, **HDR**; 4-Hydroxy-3-methylbut-2-enyl-diphosphate reductase, **IPI**; Isopentyl diphosphate δ isomerase, **GPPS;** Geranyl pyrophosphate synthase, **GGPS**; Geranylgeranyl pyrophosphate synthase **GDDR**; Geranylgeranyl reductase, **PSY**; Phytoene synthase, **DAHPS**; 3-Deoxy-D-arabino-hepulosonate, **DHQS**; 3-Dehydroquianate synthase**, SDH-DHQ1**; 3-Dehydroquianate dehydratase**, SDH-DHQ2**; Shikimate 5-dehydrogenase**, SK**; Shikimate kinase**, EPSPS**; 5-Enolpyruvylshikimate-3-P-synthase, **CS**; Chorismate synthase, **CM**; Chorismate mutase, **PAT**; Prephenate aminotransferase, **TyrA**; Arogenate dehydrogenase, **TAT**; Tyrosine aminotransferase, **HPPD**; 4-Hydroxyphenylpyruvate dioxygenase, **HPT (VTE2)**; Homogentisate phytyl transferase, **MPBQMT (VTE3)** Dimethyl-phytylquinol methyl transferase, **TC (VTE1)**; Tocopherol cyclase, **γ-TMT (VTE4)**; γ-Tocopherol C-methyl transferase, **PK (VTE5)**; Phytol kinase, **PPK (VTE6)**; Phytyl-phosphate kinase, **PSY1**; Phytoene synthase 1, **CYC-B**; chromoplast-specific lycopene cyclase, **LYC-B**; lycopene β cyclase , **LYC-E** lycopene ε cyclase , **CRTR-B**; β-ring hydroxylase , **CRTR-E**; ε-ring hydroxylase.

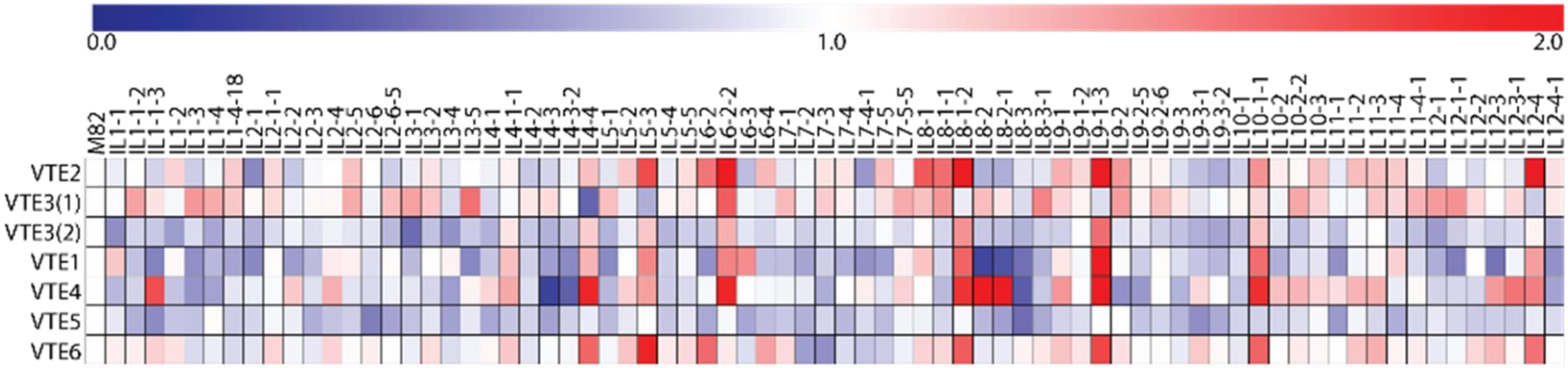


Figure 3-3 Heatmap of gene expression in fruit of *S.pennellii* x *S.lycopersicum* cv. M82 ILs showing the relative fold change of VTE pathway genes, relative to M82. These data was based on the *S.pennellii x S.lycopersicum* cv. M82 ILs RNA sequencing data (Lee and Giovannoni). This scale bar shows relative fold change of expression from 0.0 to 2.0.

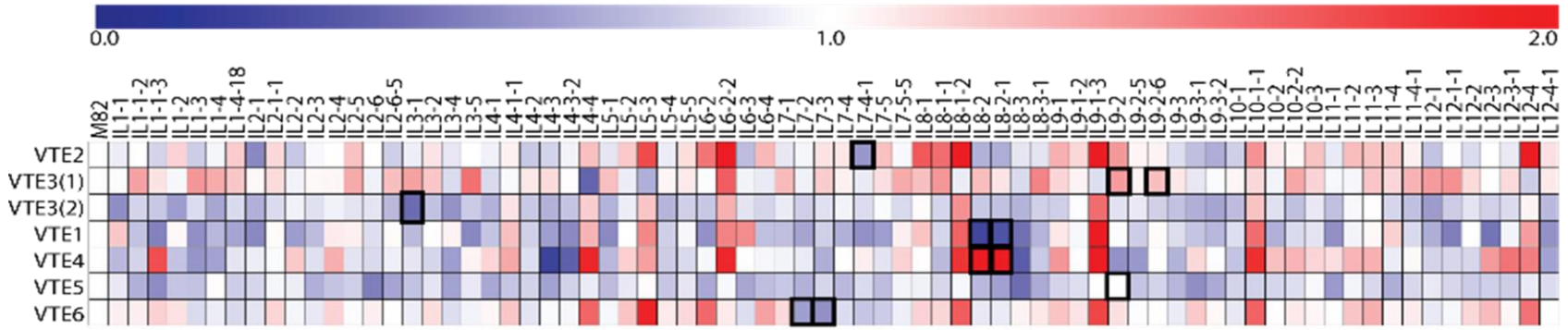


Figure 3-4 Heatmap of gene expression in fruit *S.pennellii x S.lycopersicum* cv. M82 ILs showing the relative fold change of VTE pathway genes, relative to M82. The boxes show where the structural genes lie and are cis-eQTLs. These data was based on the *S.pennellii x S.lycopersicum* cv. M82 ILs RNA sequencing data (Lee and Giovannoni). This scale bar shows relative fold change of expression from 0.0 to 2.0.

responsible for adding a methyl group to 2-methyl-6-phytylquinol (figure 3-2). According to the cis-eQTL data set (figure 3-4), *VTE3(1)* was highly expressed in IL9-2 and IL9-2-6. Whereas, *VTE3(2)* was lowly expressed in its cis-eQTL; IL3-1. *VTE3(1)* is the active isoform in fruit and this is reinforced in these data, which showed that it is highly expressed in tomato fruit.

3.3.2 Four trans-eQTL loci are identified for the VTE pathway

In contrast with the cis-eQTLs identified, four trans-eQTL loci were identified in the *Solanum pennellii* x *S.lycopersicum* cv. M82 ILs (figure 3-5). Identification of these trans-eQTLs was based on co-expression of the VTE pathway genes, rather than a single gene. Candidate trans-eQTLs, include: IL 6-2, IL6-2-2, IL 8-1, IL8-1-1, IL8-1-2, IL9-1-3, IL9-3, IL9-3-1 and IL9-3-2, which were candidates for trans regulation and do not contain VTE pathway genes in the region. The trans-eQTLs candidate ILs showed differing co-expression patterns. ILs: 6-2, 6-2-2, 8-1, 8-1-1, 8-1-2, 9-1-3, displayed higher expression of VTE pathway genes (figure 3-5), in comparison with M82 (table 3-1). Whereas, ILs 9-3, 9-3-1 and 9-3-2 (IL9-3 subgroup) showed reduced expression of VTE pathway genes (figure 3-5). These results suggested co-regulation exists for the VTE pathway.

RNA sequencing data were also available for the *S.pennellii* x *S.lycopersicum* cv. M82 introgression line parents; *S.lycopersicum* and *S.pennellii*. These data were used to find pathway genes and determine the expression ratio between the two parents. The significance of the VTE pathway gene expression ratios can be seen in table 3-2. *VTE3(1), VTE3(2)* and *VTE4* showed significant differences in expression between the two IL parents.

I decided to focus on two IL regions; IL6-2 and IL6-2-2 (IL6-2 subgroup) and the IL9-3 IL subgroup (IL9-3, IL9-3-1 and IL9-3-2), as these two regions showed the greatest changes of VTE structural gene expression.

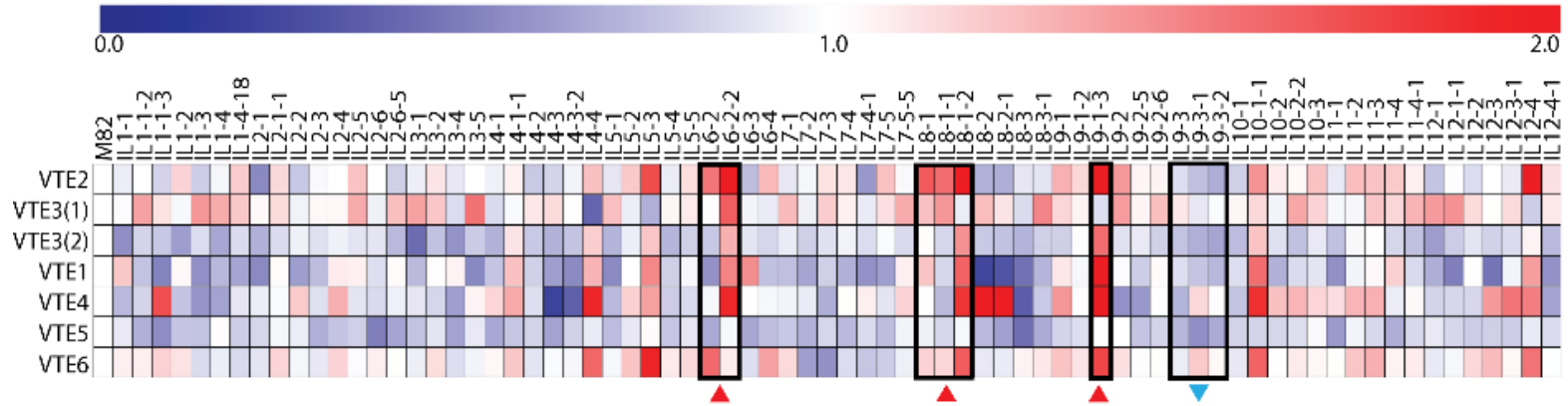


Figure 3-5 Heatmap of gene expression in fruit of *S.pennellii x S.lycopersicum* cv. M82 ILs showing the relative fold change of VTE pathway genes, relative to M82. The black boxes show where the trans-eQTLs lie and the arrow heads depict increases (red arrow head) or decreases (blue arrow head) in VTE pathway gene expression. These data was based on the *S.pennellii x S.lycopersicum* cv. M82 ILs fruit RNA sequencing data (Lee and Giovannoni). This scale bar shows relative fold change of expression from 0.0 to 2.0.

Table 3-1 Fold changes in expression of VTE pathway genes, relative to *S. lycopersicum* cv.M82. These data are based on RNA sequencing data from fruit of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs .

Gene name	Solgenomics identifier	IL position	Relative fold change of expression in *S.pennellii* x *S.lycopersicum* cv.M82 ILs, relative to M82 parent					
			M82	IL6-2	IL6-2-2	IL9-3	IL9-3-1	IL9-3-2
VTE2	Solyc07g017770	IL7-4-1	1	1.53	2.78	0.87	0.75	0.68
VTE3(1)	Solyc09g065730	IL9-2	1	0.98	1.63	1.10	0.90	0.98
VTE3(2)	Solyc03g005230	IL3-1	1	0.78	1.31	0.77	0.69	0.65
VTE1	Solyc08g068570	IL8-2 and IL8-2-1	1	0.57	1.49	0.87	0.78	0.72
VTE4	Solyc08g076360	IL8-2 and IL8-2-1	1	0.95	1.89	0.71	1.15	0.99
VTE5	Solyc09g018510	IL9-2	1	0.66	0.96	0.72	0.56	0.67
VTE6	Solyc07g062180	IL7-2 and IL7-3	1	1.59	1.09	0.92	1.21	1.04

Table 3-2 Expression ratios of genes encoding enzymes of the VTE pathway in parents of *S.pennellii* x *S.lycopersicum* cv. M82 IL population, from RNA sequencing data of fruit from both parents (Koenig et al., 2013). Statistical significance is indicated by the p value (p = <0.05).

Gene name	**Solgenomics identifier**	***S.lycopersicum* : *S.pennellii* ratio**	**P value**
VTE2	Solc07g017770	0.84	0.91
VTE3(1)	Solyc09g065730	2.09	$3.17x10^{-5}$
VTE3(2)	Solyc03g005230	1.31	0.05
VTE1	Solyc08g068570	1.36	0.14
VTE4	Solyc08g076360	20.90	$2.73x10^{-30}$
VTE5	Solyc09g018510	1.32	0.11
VTE6	Solyc07g062180	0.53	0.21

3.3.3 Co-expression analysis of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines reveals a trans-eQTL on chromosome 9

The trans-eQTL identified on chromosome 9 was found using co-expression analysis of the *S.pennellii* x *S.lycopersicum* cv. M82 fruit ILs. These results were then compared to *S.pennellii* x *S.lycopersicum* cv. M82 leaf RNA sequencing data and *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs RNA sequencing data. This is discussed in sections 3.3.3.1-4.

3.3.3.1 Analysis of the trans-eQTL IL9-3-2 showed the VTE and MEP pathway are co-expressed

First, I analysed the IL9-3 IL subgroup, as this region showed the greatest reduction of VTE pathway gene expression (figure 3-5 and table 3-1). *VTE2, VTE3(2)* and *VTE4* showed reduced expression in the IL9-3 IL subgroup. *VTE2* expression was the lowest in IL9-3-2, compared to all other ILs in the RNA sequencing data. Although, not all genes were down-regulated, this trans-eQTL showed significant decreases of the VTE pathway gene expression.

IL9-2 and IL9-2-6 overlap with the IL9-3 subgroup (figure 3-6), however, these lines do not show changes in VTE pathway gene expression (figure 3-5). These regions of the *S.pennellii* genome in these ILs were therefore removed from candidate gene screening, and the overlapping segment with IL9-3 was also removed from candidate gene mining. Within the IL9-3, there are two other ILs; IL9-3-1 and IL9-3-2 (figure 3-6). The IL9-3 encompasses both ILs, and IL9-3-2 also sits within the IL9-3-1 region. The VTE pathway gene expression profile in the RNA sequencing data (figure 3-5) is similar in all ILs, therefore, the data suggested that the trans-eQTL responsible, resides in IL9-3-2 region. This is the region that that is present in both IL9-3 and IL9-3-1. Thus, this was the interval that was mined for candidate transcription factors.

The trans-eQTL found in the IL9-3-2 was checked for co-expression with genes in the precursor pathways (MEP and SK pathway). There was a general reduction of gene expression for the MEP pathway genes in IL9-3-2 (figure 3-7).

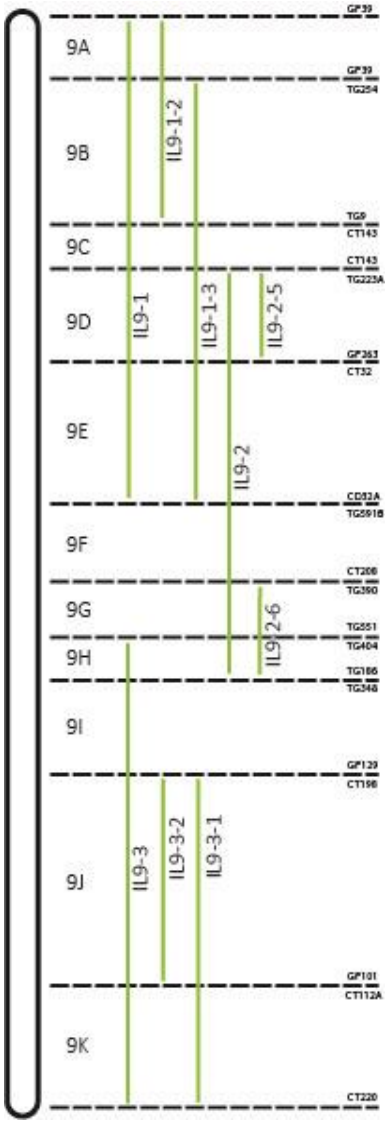


Figure 3-6 Schematic diagram showing *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines (ILs) on chromosome 9. 9A-K indicate bin mapping positions, which are indicated by the dashed lines. At the end of each dashed line, the markers that were used to determine the chromosome position are shown. The green lines with IL9-1-IL9-3-1 are physical introgression lines mapped and used for analysis.

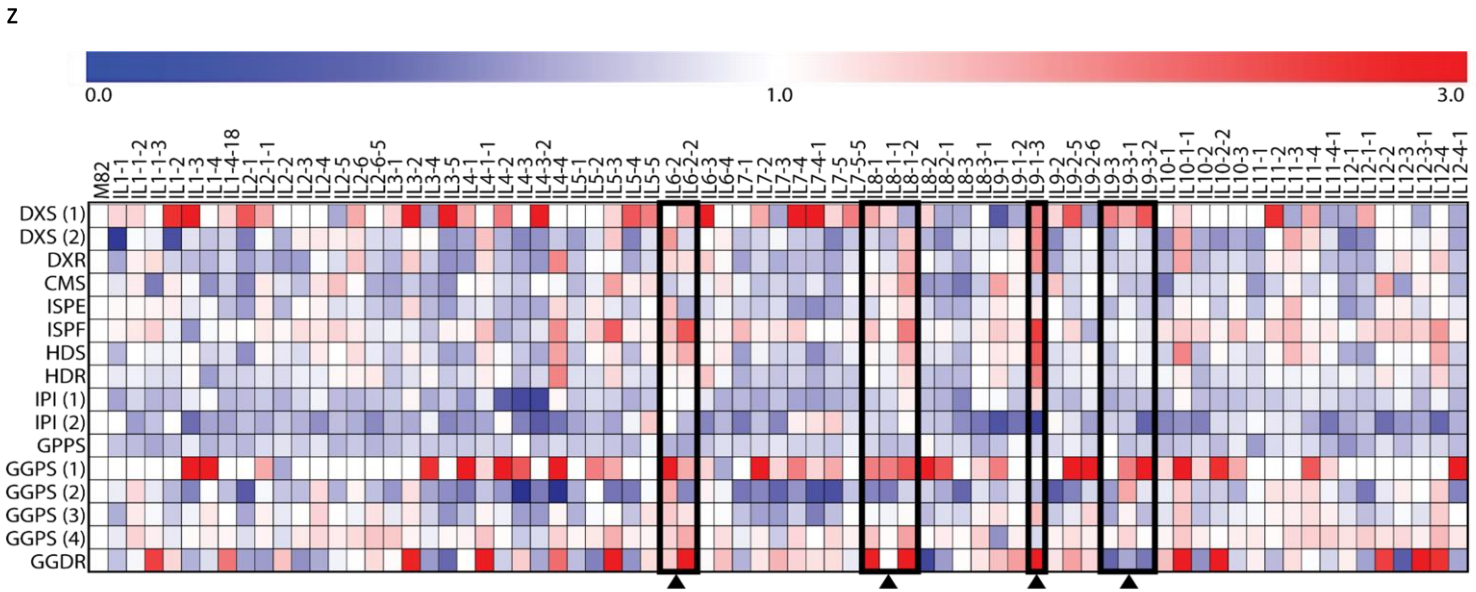


Figure 3-7 Heatmap of *S.pennellii* x *S.lycopersicum* cv. M82 introgression line (ILs) RNA sequencing data, showing relative fold change in comparison to M82, for the MEP pathway. These data was based on the *S.pennellii x S.lycopersicum* cv. M82 ILs fruit RNA sequencing data (Lee and Giovannoni). Scale bar shows relative fold change from 0.0 to 3.0. Black boxes, indicated by black arrow heads, show trans-eQTLs.

However, there were two genes which showed significantly higher expression profiles compared to M82; *DXS(1)* and *GGPS(1).* The trans-eQTL IL9-3-2 does not contain any MEP pathway genes (table 3-3), therefore it is not a cis-eQTL for this pathway, and rather could be a trans-eQTL for the MEP pathway as well as for the VTE pathway. This trans-eQTL was also identified as a trans-eQTL for lycopene biosynthesis in tomato, which showed genes encoding enzymes of the carotenoid pathway were also down-regulated (Li, 2018).

The expression profiles of the SK pathway were variable in the trans-eQTL IL9-3-2 (figure 3-8). Global transcript levels of the SK pathway were reduced compared to M82, which was consistent with the findings for the MEP and VTE pathways. There were many genes which showed no differences in expression. However, there were two genes; *EPSPS3* and *TAT2* which were highly expressed in this region. The trans-eQTL IL9-3-2 did not contain any structural genes for the SK pathway, therefore, it is not a cis-eQTL. But, this region did not present itself as a trans-eQTL for the whole SK pathway. It may be a trans-eQTL for individual genes such as *TAT2* or *EPSPS3*. There are several genes which are also down-regulated in this region, which include; *SDH-DHQS2, SDH-DHQS3, TYRA1, TAT1* and *TAT3,* which implied that this region is a trans-eQTL for these genes.

3.3.3.2 The trans-eQTL IL9-3-2 was compared to the *S.pennellii* x *S.lycopersicum* cv. M82 leaf RNA sequencing data

S.pennellii x *S.lycopersicum* cv. M82 ILs leaf RNA sequencing data was screened for changes in VTE pathway gene expression for the trans-eQTL IL9-3-2 (figure 3-9). This region did not show large changes in VTE pathway expression. IL9-3-2 did not show any great changes in expression for any of the genes, although, *VTE3(2)* was slightly up-regulated. This is the isoform that is most active in leaves, therefore, the trans-eQTL could affect the other isoform variant of *VTE3.* However, the lack of increase in the rest of the pathway suggested that this was unlikely. These data reinforced the idea that the candidate TFs in this trans-eQTL, which regulates the VTE pathway are probably fruit specific.

Table 3-3 Position of MEP and SK pathway genes in *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines (ILs). Gene names and solgenomics identifiers are displayed.

Solgenomics identifier	Gene name	IL region
Solyc11g010850	DXS(1)	IL11-1 and IL11-2
Solyc01g067890	DXS(2)	IL1-1 and IL1-2
Solyc03g114340	DXR	IL3-4
Solyc01g102820	CMS	IL1-3 and IL1-4
Solyc01g009010	ISPE	IL11-3
Solyc08g081570	ISPF	IL1-4-18
Solyc11g069380	HDS	IL11-3
Solyc01g109300	HDR	IL1-4-18
Solyc04g056390	IPI(1)	IL4-3
Solyc05g055760	IPI(2)	IL5-5
Solyc08g023470	GPPS	IL8-1-1
Solyc11g011240	GGPS(1)	IL11-2
Solyc04g079960	GGPS(2)	IL4-3
Solyc02g085700	GGPS(3)	IL2-5 and IL2-6
Solyc09g008920	GGPS(4)	IL9-1-3
Solyc03g115980	GGDR	IL3-4
Solyc04g074480	DAHPS1	IL4-3 and IL4-4
Solyc01g105420	DAHPS2	IL1-4
Solyc01g105390	DAHPS3	IL1-4
Solyc11g009080	DAHPS4	IL11-1 and IL11-2
Solyc02g083590	DHQS	IL6-4
Solyc01g067750	SDH-DHQS1	IL1-1 and IL1-2
Solyc06g084460	SDH-DHQS2	IL6-4
Solyc10g038080	DHQ	IL10-1
Solyc04g051860	SK	IL4-3-2
Solyc01g091190	EPSPS1	IL1-2
Solyc05g050980	EPSPS2	IL5-3
Solyc09g005460	EPSPS3	IL9-1-2
Solyc04g049350	CS1	IL4-3-2
Solyc04g009620	CS2	IL4-1
Solyc02g088460	CM1	IL2-5 and IL2-6
Solyc11g017240	CM2	IL11-2 and IL11-3
Solyc04g054710	PAT	IL4-3-2
Solyc07g007590	TyrA	IL7-5-5
Solyc07g053720	TAT1	IL7-3
Solyc10g007110	TAT2	IL10-1-1
Solyc07g053720	TAT3	IL7-3

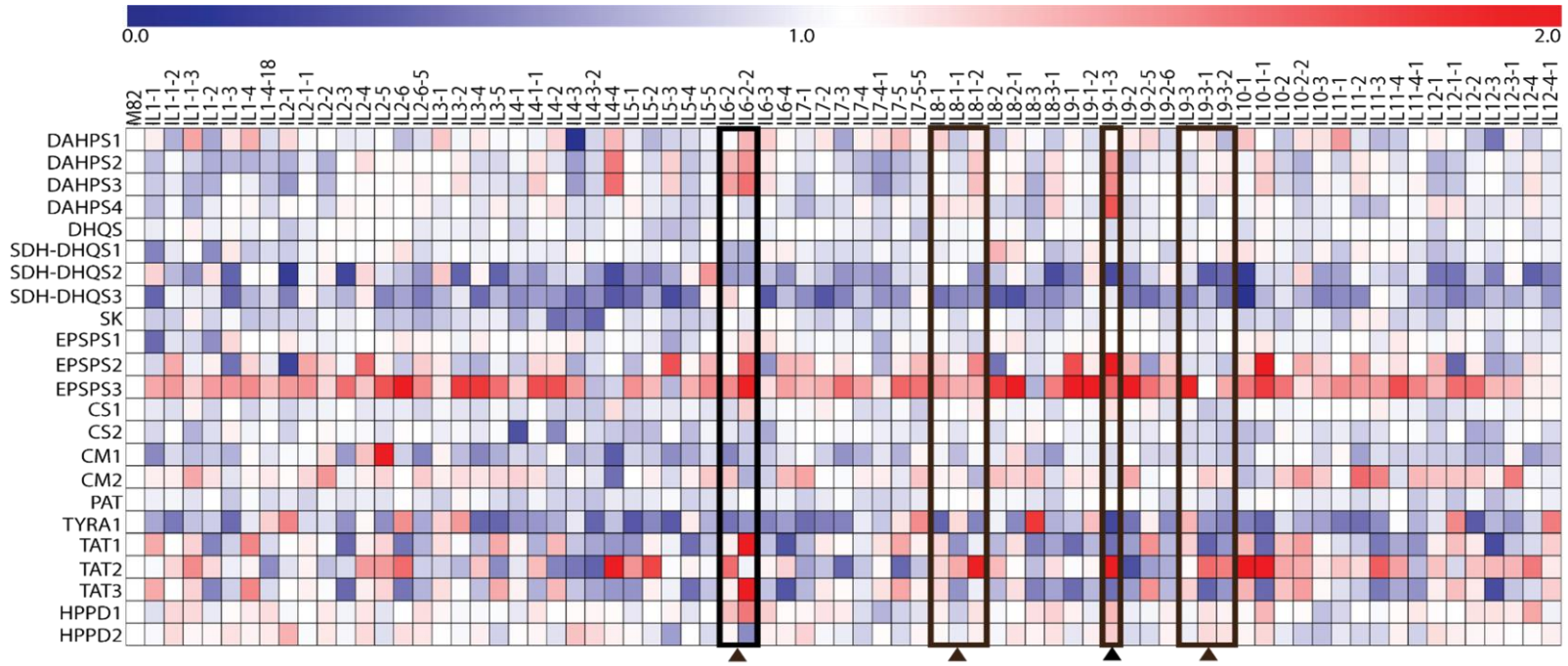


Figure 3-8 Heatmap of *S.pennellii* x *S.lycopersicum* cv. M82 introgression line (ILs) RNA sequencing data, showing relative fold change in comparison to M82, for the SK pathway. These data was based on the *S.pennellii x S.lycopersicum* cv. M82 ILs fruit RNA sequencing data (Lee and Giovannoni). Scale bar shows relative fold change from 0.0 to 2.0. Black boxes, indicated by black arrow heads, show trans-eQTLs.

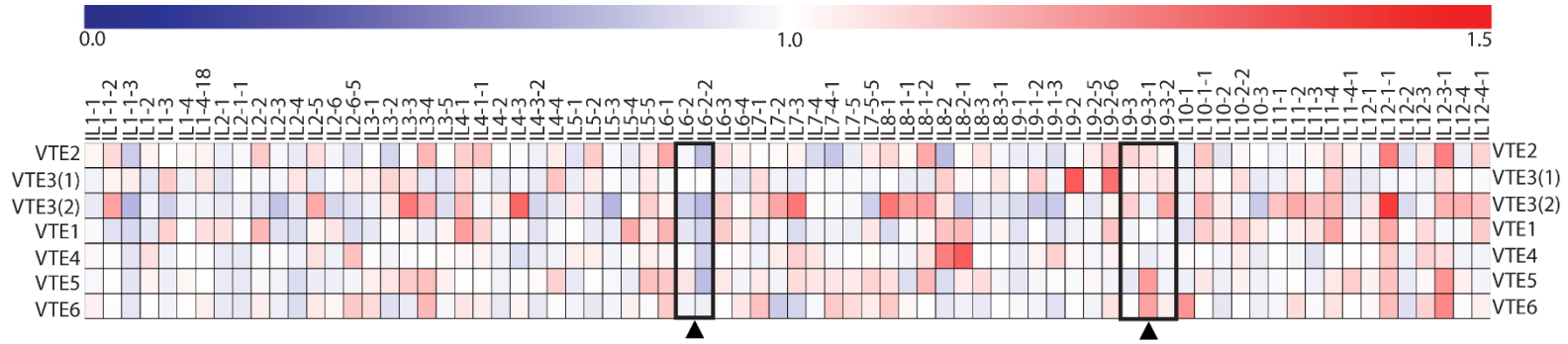


Figure 3-9 Heatmap of *S.pennellii* x *S.lycopersicum* cv. M82 introgression line (ILs) RNA sequencing data of leaves (Chitwood et al., 2013), showing relative fold change the of the average of each gene across all the ILs, for the VTE pathway. The black boxes, indicated by the black arrow heads show the trans-eQTLs identified in *S.pennellii* x *S.lycopersicum* cv. M82 ILs RNA sequencing data of fruit (IL6-2-2 and IL9-3-2). Scale bar shows relative fold change from 0.0 to 1.5.

3.3.3.3 The *S.pennellii* x *S.lycopersicum* cv. M82 IL9-3 trans-eQTL is correlated to *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs

The *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs are a cross between *S.lycopersicum* cv. VF36 and *S.lycopersicoides* (Canady et al., 2005). The population generated from this cross are still segregating, therefore there was still some heterozygosity within the ILs.

The IL trans-eQTL IL9-3-2 from *S.pennellii* x *S.lycopersicum* cv. M82 ILs was compared to its equivalent ILs in *S.lycopersicoides* x *S.lycopersicum* cv. VF36, as shown in figure 3-10. There were three *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs that overlapped with the IL9-3 subgroup of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs; 4270A, 4270B/4271 and 4272, which had different levels of zygosity (table 3-4). As 4270A was heterozygous, this line was not considered for further analysis. Line 4270B/4271 was homozygous, and, the introgressed segment was within IL9-3, but did not encompass the *S.pennellii* IL9-3-2 segment, which was previously identified as the interval of interest, as a trans-eQTL. Therefore, this line was also removed from the analysis. On the other hand, 4272, included the same region as the *S.pennellii* IL9-3-2 segment, but also IL9-3-1, and part of IL9-3 (figure 3-10). This line was homozygous and was used to compare to the trans-eQTL analysis identified in the *S.pennellii* x *S.lycopersicum* cv. M82 ILs. The line; 4272, showed reduced expression of *VTE2, VTE1* and *VTE5*. These data mirrored the trans-eQTL in the *S.pennellii* x *S.lycopersicum* cv. ILs, which showed reduced gene expression of the same VTE pathway genes. Therefore, this implied that this trans-eQTL contained a candidate transcriptional factor regulating this pathway, which is present, and differentially active, in more than one species.

Large scale metabolite analysis was completed for the *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs fruit tissue and alpha tocopherol was quantified. Alpha tocopherol contents in the ILs were variable and were measured relative to the *S.lycopersicum* parent; VF36. As figure 3-11 shows, the line 4270, showed higher

Table 3-4 *S.lycopersicoides* x *S.lycopersicum* cv. VF36 introgression line (IL) RNA sequencing data of VTE pathway genes. All values have been normalised to the *S.lycopersicoides* parent; VF36. N/A denoted not applicable.

Gene name	**Solgenomics identifier**	**Relative expression in *S.lycopersicoides* x *S.lycopersicum* cv. VF36 IL, relative to VF36**						
		VF36	**4253**	**4254**	**4255**	**4270A**	**4270B**	**4272**
Homozygous (L), Heterozygous (H) or N/A		N/A	H	L	L	H	L	L
VTE2	Solyc07g017770	1	0.88	0.46	1.44	1.68	1.46	0.77
VTE3(2)	Solyc09g065730	1	0.56	0.51	1.04	0.8	1.22	0.92
VTE3(1)	Solyc03g005230	1	0.75	0.71	1.00	1.11	0.99	0.84
VTE1	Solyc08g068570	1	0.43	0.84	0.74	1.05	1.00	0.59
VTE4	Solyc08g076360	1	0.81	1.05	0.91	1.61	1.34	1.02
VTE5	Solyc09g018510	1	0.83	1.06	0.75	0.83	0.93	0.70
VTE6	Solyc07g062180	1	0.86	0.97	1.08	1.31	1.37	0.96

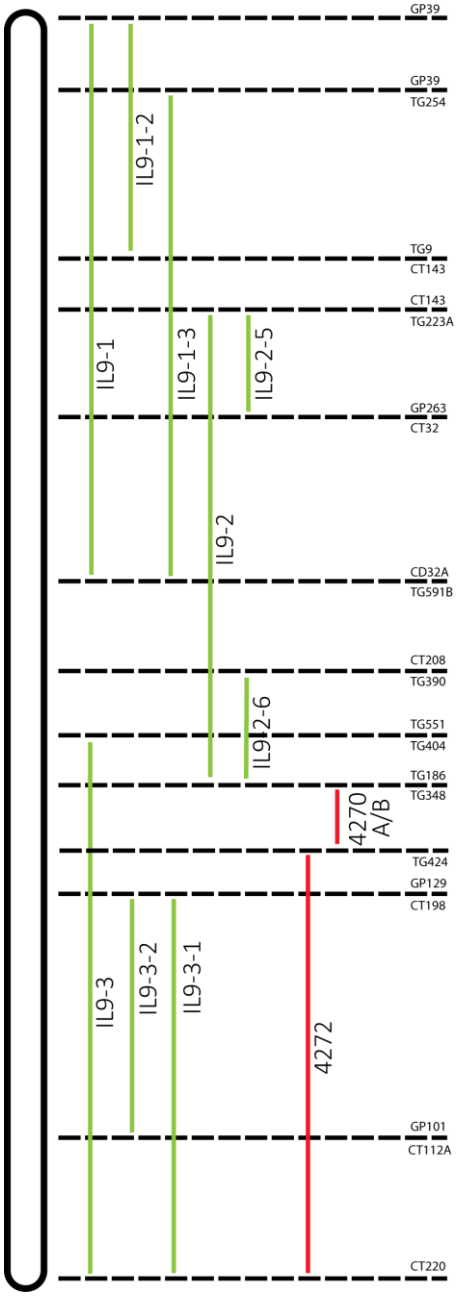


Figure 3-10 Schematic diagram showing *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines (ILs) and *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs on chromosome 9. The green lines indicate where the *S.pennellii* DNA lies and the red lines show where the *S.lycopersicoides* DNA lies. At the end of each dashed line, the markers that were used to determine the chromosome position are shown.

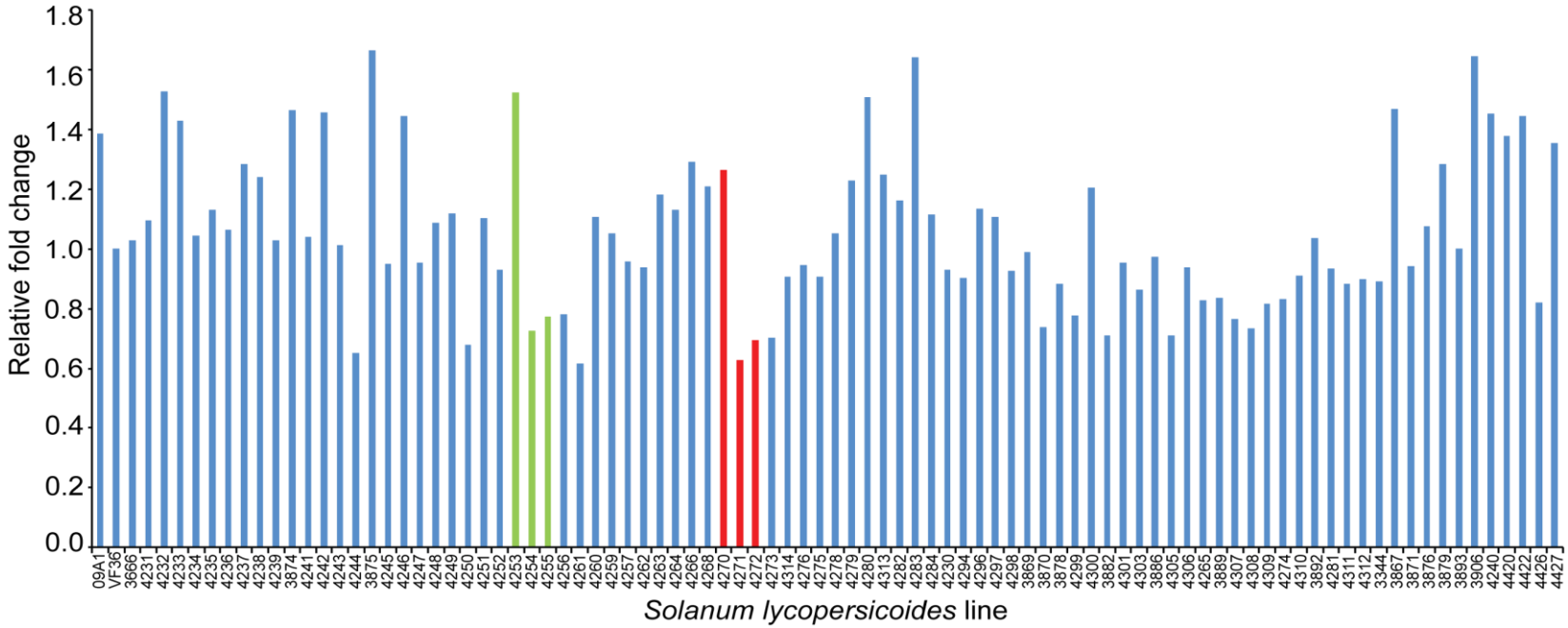


Figure 3-11 Relative fold change of alpha tocopherol in *S.lycopersicoides* x *S.lycopersicum* cv. VF36 introgression lines (ILs), in respect to parent line, VF36. Green bars indicate the *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs which are overlapping with the *S.pennellii* x *S.lycopersicum* cv.M82 IL6-2 subgroup. The red bars indicate the *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs which are overlapping with the *S.pennellii* x *S.lycopersicum* cv.M82 IL9-3 subgroup.

alpha tocopherol concentrations, compared to VF36, however, this was a heterozygous line. Lines 4271 and 4272 both had lower alpha tocopherol contents, which was consistent with the reduced expression of VTE pathway genes in these lines. These data strengthened the argument that the trans-eQTL interval controls VTE synthesis, not only *S.pennellii,* but also in *S.lycopersicoides.*

3.3.3.4 Candidate gene mining the trans-eQTL IL9-3

The trans-eQTL identified in the IL9-3 subgroup implied that a master transcriptional regulator may reside within the locus. I mined trans-eQTL IL9-3-2 region for candidate transcriptional regulators, that might regulate expression of genes in the VTE pathway. Within IL9-3-2, 1380 genes were identified in the region (figure 3-12). I screened for TFs within the trans-eQTL region and found 54 transcriptionally related genes. These transcriptionally related genes were reduced to 11 TFs based on the co-expression with the VTE pathway genes in IL9-3-2. The 11 candidate TFs are shown in table 3-5, which shows their statistical significance based on the expression ratio between the parents of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs; *S.lycopersicum* and *S.pennellii.*

Additionally, the gene list was reduced to TFs with certain motifs. (table 3-5). It is documented that certain TFs motifs exist within the VTE pathway promoters (Quadrana et al., 2013). Based on this study, I chose TFs that were MYB TFs and bZIP TFs. These candidates were then screened for activity using viral induced gene silencing (VIGS), as described in chapter 5.

3.3.4 Co-expression analysis of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines reveals a trans-eQTL on chromosome 6

The trans-eQTL identified on chromosome 6 was found using co-expression analysis of the *S.pennellii* x *S.lycopersicum* cv. M82 fruit ILs. These results were compared to *S.pennellii* x *S.lycopersicum* cv. M82 leaf RNA sequencing data and *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs RNA sequencing data. This is discussed in sections 3.3.4.1-4.

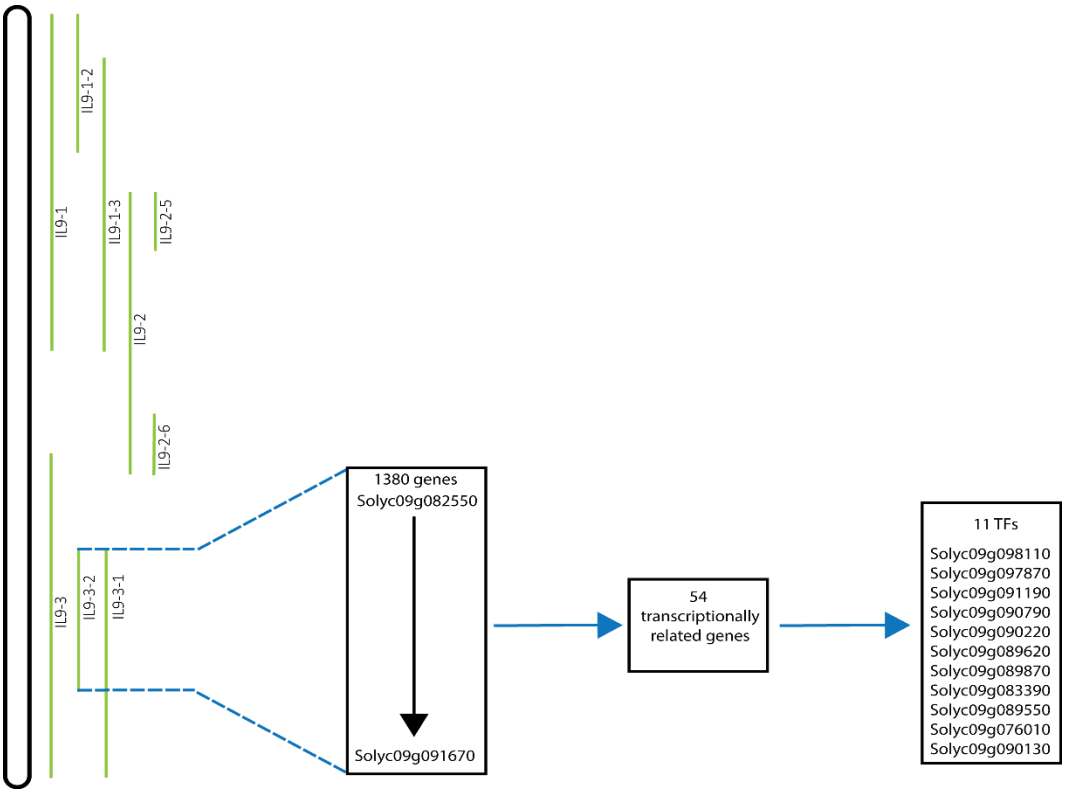


Figure 3-12 Schematic diagram of candidate gene mining for *S.pennellii* x *S.lycopersicum* cv. M82 trans-eQTL IL9-3-2. Eleven candidate transcription factors were identified, which showed co-expression patterns with genes encoding enzymes of the VTE pathway in the *S.pennellii* x *S.lycopersicum* cv. M82 fruit RNA sequencing data (Lee and Giovannoni).

Table 3-5 Candidate gene list for *S.pennellii* x *S.lycopersicum* cv. M82 IL9-3 subgroup. Genes that have statistically different expression between *S.pennellii* and *S.lycopersicum* (Lee and Giovannoni) are shown as p values <0.05. No statistical significant is shown by the dashed line (-).

Solgenomics identifier	**Gene description**	**Statistical significance between *S.pennellii* and *S.lycopersicum* RNA seq data? (P value < 0.05 or No)**
Solyc09g098110	basic helix-loop-helix (bHLH) family protein	$2.80x10^{-5}$
Solyc09g097870	basic helix-loop-helix (bHLH) family protein	$9.05x10^{-17}$
Solyc09g091190	PHD finger family	$5.93x10^{-3}$
Solyc09g090790	AtMYB79, MYB305 | MYB305 (myb domain protein 305); DNA binding / transcription factor	$4.04x10^{-6}$
Solyc09g090220	pentatricopeptide (PPR) repeat-containing protein	-
Solyc09g089620	DNA-binding protein-related	$4.48x10^{-16}$
Solyc09g089870	basic helix-loop-helix (bHLH) family protein	-
Solyc09g083390	zinc knuckle (CCHC-type) family protein	No

Table 3-5 continued.

Solyc09g089550	ATHB33 | ATHB33 (*Arabidopsis thaliana* HOMEOBOX PROTEIN 33); DNA binding / transcription factor	$2.01x10^{-5}$
Solyc09g076010	PHD finger transcription factor, putative	-
Solyc09g090130	MYB15, AtY19, AtMYB15 | AtMYB15/AtY19/MYB15 (myb domain protein 15); DNA binding	-

3.3.4.1 Analysis of the trans-eQTL IL 6-2-2 showed co-expression of the VTE pathway

IL6-2 was also considered to be a trans-eQTL for the VTE pathway. As figure 3-5 and table 3-1 showed, the VTE pathway genes were highly up-regulated in this region. Key genes up-regulated in this region included; *VTE2, VTE3(1), VTE1* and *VTE4.* These genes were all highly expressed in IL6-2-2, although, they did not show significant changes in IL6-2. Expression of *VTE5* and *VTE6* was not altered in IL6-2-2 relative to M82 (figure 3-5 and table 3-1).

IL6-2 encompasses the whole IL6-2-2 region (figure 3-13), this implied that there is a masking effect of IL6-2 on the IL6-2-2 region. Therefore, the IL6-2-2 region was taken further for trans-eQTL analysis. IL6-2-2 also overlaps with the IL6-1, although, RNA sequencing data was not carried out on this line. Therefore, I could not exclude that this overlapping segment with IL6-1, could contain a putative regulator for the VTE pathway. Therefore, the trans-eQTL mined for candidate TFs included the complete IL6-2-2 region.

The precursor pathways, preceding the VTE pathway (MEP and SK pathway), were analysed in the trans-eQTL IL6-2-2. The MEP pathway expression profile was variable for each gene (figure 3-7) and cannot be generalised as either up-regulated or down-regulated. Genes such as *GGDR* and *ISPF* showed increased expression relative to M82, these data implied that this also could be a trans-eQTL for these two genes.

In contrast with the MEP pathway, the SK pathway showed greater variability of expression of pathway genes (figure 3-8), in the trans-eQTL IL6-2-2. *EPSPS2, EPSPS3, TAT1* and *TAT3* displayed increased expression in the trans-eQTL IL6-2-2, which suggested that the trans-eQTL could also have a role regulating these genes. *DAHPS1, DAHPS2* and *DAHPS3* were up-regulated, but the levels are not as high as previously mentioned genes, which could indicate feedback or could also indicate a trans-eQTL.

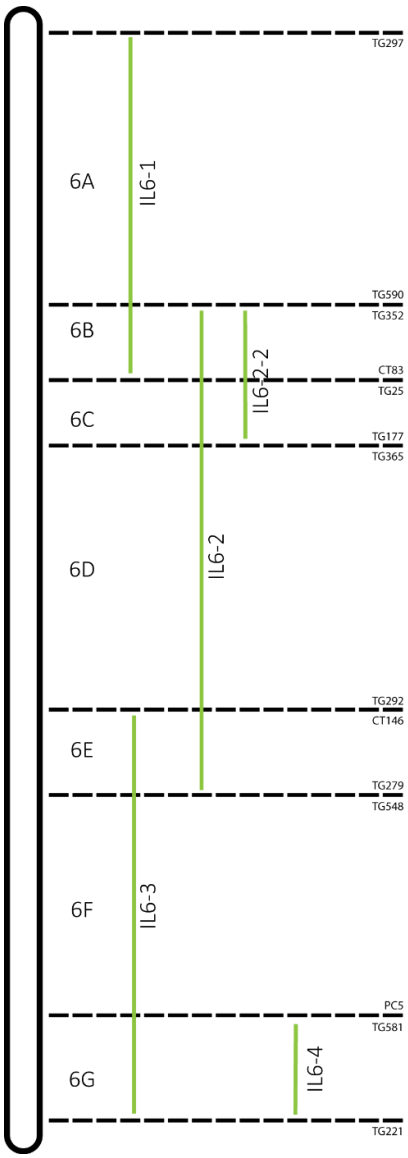


Figure 3-13 Schematic diagram showing *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines (ILs) on chromosome 6. 6A-G indicate bin mapping positions, which are indicated by the dashed lines. At the end of each dashed line, the markers that were used to determine the chromosome position are shown. The green lines with IL6-1-IL6-4 are physical ILs mapped and used for analysis.

3.3.4.2 *S.pennellii* x *S.lycopersicum* cv. M82 introgression line leaf RNA sequencing data showed an opposing expression profile compared to the *S.pennellii* IL6-2-2 trans-eQTL in fruit

Leaf RNA sequencing data for *S.pennellii* x *S.lycopersicum* cv. M82 ILs was screened for the trans-eQTL residing in the IL6-2 subgroup. This trans-eQTL showed an opposing expression profile in leaves, when compared to the gene expression profile of the fruit RNA sequencing data (figure 3-5 and 3-9). *VTE2, VTE3(2), VTE1, VTE4* and *VTE5* were down-regulated, in this region, in the leaf RNA sequencing data, whereas, in the fruit RNA sequencing data, the VTE pathway gene expression profile was up-regulated. Therefore, these data suggested that IL6-2-2 contained a trans-eQTL for leaves, but one which had the opposite effects on the expression profile in leaves to fruit, from the RNA sequencing data.

3.3.4.3 The *S.pennellii* IL trans-eQTLs IL6-2-2 was compared to the *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs

The trans-eQTL identified in *S.pennellii* IL6-2-2 was also analysed in the equivalent *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs. Figure 3-14 shows the overlapping *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs with the *S.pennellii* x *S.lycopersicum* cv. M82 ILs for the trans-eQTL IL6-2-2. The 4253 IL was heterozygous and was removed from further analysis (table 3-4). The other lines; 4254 and 4255, were homozygous, but 4255 included part of the *S.pennellii* IL6-2, which is thought to be masking IL6-2-2 (table 3-4). Thus, this line was also removed from the analysis. This left one line (4254) which covered the *S.pennellii* IL6-2-2, which was homozygous.

Co-expression analysis of the VTE pathway in the *S.lycopersicoides* x *S.lycopersicum* cv. VF36 IL RNA sequencing data showed that the genes were not co-expressed (table 3-4). There was a reduction of expression of *VTE2, VTE3(1)* and *VTE3(2),* however, this was the opposite expression profile to the *S.pennellii* IL6-2-2 trans-eQTL, which showed highly expressed VTE genes. Therefore, the

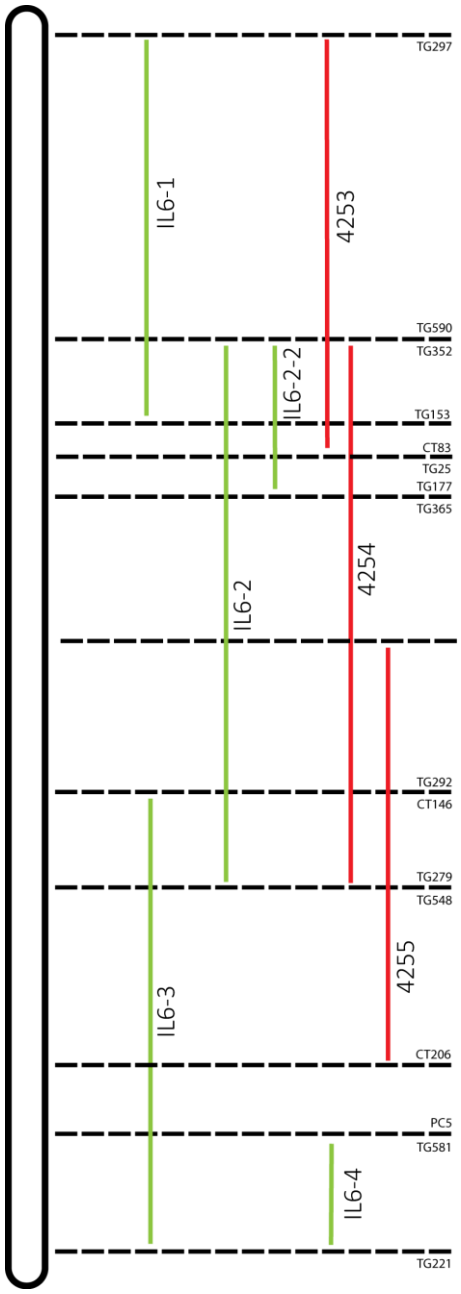


Figure 3-14 Schematic diagram showing *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines (ILs) and *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs on chromosome 6. The green lines indicate where the *S.pennellii* DNA lies and the red lines show where the *S.lycopersicoides* DNA lies in each IL. At the end of each dashed line, the markers that were used to determine the chromosome position are shown.

trans-eQTL identified for the *S.pennellii* trans-eQTL IL6-2-2 did not correlate well to the *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs.

3.3.4.4 Candidate gene mining of the *S.pennellii* trans-eQTL IL6-2-2

The trans-eQTL IL6-2-2 region was screened for candidate transcriptional regulators of the VTE pathway. The region contained 6374 genes (figure 3-15) and 73 transcriptionally related genes were found in IL6-2-2. Transcriptionally related genes were reduced to 26 TFs based on their co-expression with pathway genes. Candidates from the gene list were further reduced based on their statistically significant expression ratios between *S.lycopersicum* and *S.pennellii* (table 3-6).

Transcriptional motifs in the promoters of the VTE pathway were identified (as previously described) (Quadrana et al., 2013), and were used to reduce the candidates to MYB and bZIP TFs. Candidates that were selected for further analysis were transiently silenced using VIGS and as described in chapter 6.

3.4 Discussion

3.4.1 Co-expression analysis of eQTLs is a powerful tool to find candidate genes

Global eQTL analysis has been widely adopted in several plant species, including; *Arabidopsis* (West et al., 2007), barley (Potokina et al., 2008) and maize (Swanson-Wagner et al., 2009), melon (Galpaz et al., 2018), and used to identify many potential positive traits that are useful for breeders. However, these studies have not identified key genes for more elusive traits, such as regulatory networks of metabolite pathways. This suggested that eQTL mapping using co-expression analyses to identify transcriptional regulators might be rewarding as these might confer subtle traits.

Co-expression analysis of the VTE pathway, using the *S.pennellii* x *S.lycopersicum* cv. M82 ILs, to find candidate eQTLs is a complementation method to many trans-eQTL studies. This method does not incorporate SNPs and markers, which can reduce the size and region of a trans-eQTL. However, my approach, which used co-expression analysis, is a powerful tool to elucidate trans-eQTLs that are influencing the VTE pathway. This technique has been used in network analysis

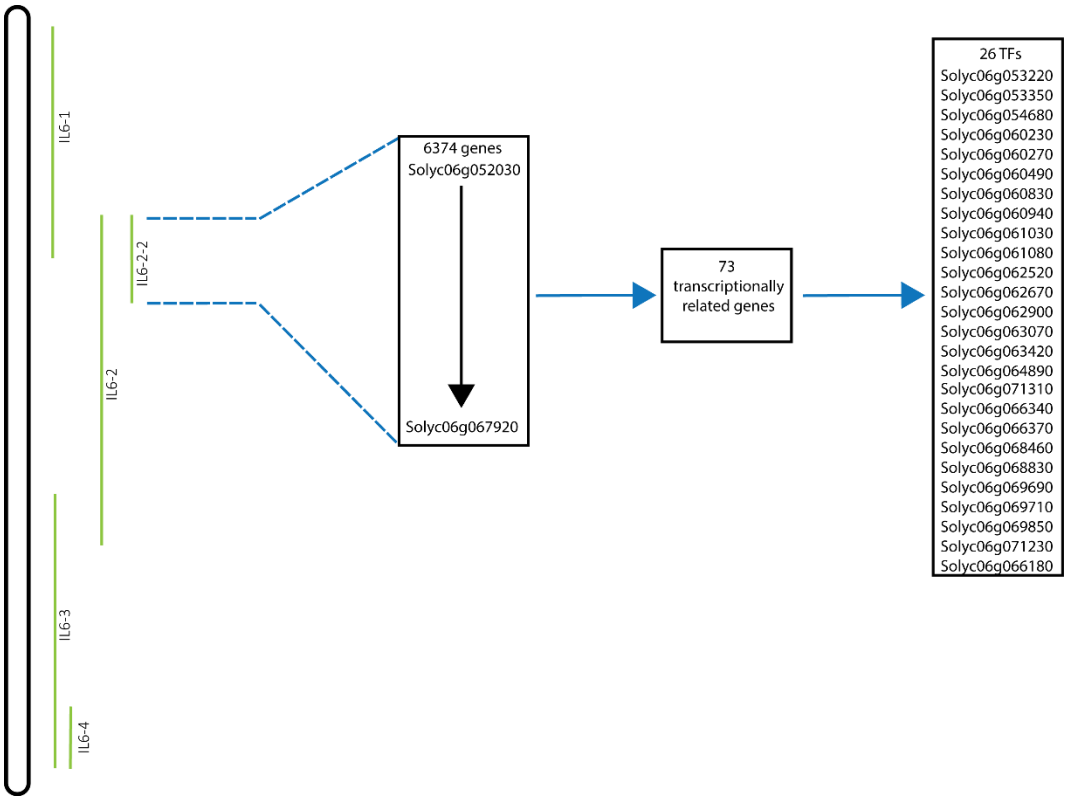


Figure 3-15 Schematic diagram of candidate gene mining for the trans-eQTL IL6-2-2. Twenty-six candidate transcription factors were identified, which showed co-expression patterns with genes encoding enzymes of the VTE pathway in the *S.pennellii* x *S.lycopersicum* cv. M82 fruit RNA sequencing data (Lee and Giovannoni).

Table 3-6 Candidate gene list for *S.pennellii* x *S.lycopersicum* cv. M82 IL6-2 subgroup. Genes that have statistically different expression between *S.pennellii* and *S.lycopersicum* (Koenig et al., 2013) are shown as p values <0.05. No statistical significant is shown by the dashed line (-).

Solgenomics identifier	**Gene description**	**Statistical significance between *S.pennellii* and *S.lycopersicum* RNA seq data? (P value < 0.05 or No)**
Solyc06g053220	ATHB7, ATHB-7 | ATHB-7 (*Arabidopsis thaliana* HOMEOBOX 7)	-
Solyc06g053350	bZIP transcription factor, putative (bZIP69)	-
Solyc06g054680	MADS box transcription factor family	-
Solyc06g060230	ATAF1 (*Arabidopsis* NAC domain containing protein 2)	-
Solyc06g060270	bZIP transcription factor family	-
Solyc06g060490	VIP1 (VIRE2-INTERACTING PROTEIN 1); transcription factor	-
Solyc06g060830	ATHB-2 (*Arabidopsis thaliana* HOMEOBOX PROTEIN 2); DNA binding / transcription factor	9.71×10^{-3}
Solyc06g060940	zinc finger (GATA type) family protein	-

Table 3-6 continued.

Solyc06g061030	APRR2 (PSEUDO-RESPONSE REGULATOR 2)	-
Solyc06g061080	ANAC071 (*Arabidopsis* NAC domain containing protein 71)	-
Solyc06g062520	(*Arabidopsis* dof zinc finger protein 1); DNA binding / transcription factor	$1.88x10^{-2}$
Solyc06g066370	WRKY33 (WRKY DNA-binding protein 33)	-
Solyc06g068460	WRKY40 (WRKY DNA-binding protein 40)	-
Solyc06g068830	AP2 domain-containing transcription factor, putative	-
Solyc06g069690	TOC1 (TIMING OF CAB EXPRESSION 1); transcription regulator	$1.51x10^{-7}$
Solyc06g069710	ANAC100/ATNAC5 (*Arabidopsis* NAC domain containing protein 100); transcription factor	-
Solyc06g069850	MYB30 (myb domain protein 30); DNA binding	$7.57x10^{-6}$
Solyc06g071230	myb family transcription factor	$3.28x10^{-10}$
Solyc06g066180	UNE16 (unfertilized embryo sac 16) G2-like transcription factor family, GARP	$1.26x10^{-6}$

studies using microarray databases to find causal genes for phenotypes in plants (Jimenez-Gomez et al., 2010, Wang et al., 2018). Therefore, this alternative co-expression analysis using RNA-seq data may be applied successfully to the VTE pathway to find causal TFs of expression phenotypes.

Co-expression analysis can present other challenges, as in human cell cultures (Yao et al., 2017), where co-expression analyses showed that it is hard to identify true trans-acting eQTLs due to neighbouring cis-acting eQTLs which may affect the trans-eQTL region. However, the *S.pennellii* x *S.lycopersicum* cv. M82 IL co-expression approach can provide detailed insights into regulatory networks, which other genome wide studies cannot elucidate, such as genome wide association studies (GWAS) often fail to identify TFs, to link phenotypes with candidate genes (Yao et al., 2017).

3.4.2 Cis-eQTLs are often linked to structural genes and mQTLs

Co-expression analysis in the *S.pennellii* x *S.lycopersicum* cv. M82 ILs, of the VTE pathway genes, revealed several cis-eQTLs. These cis-eQTLs can have large phenotypic effects, and they result in large changes in target metabolites, as I, and others, have observed. Cis-eQTLs are often linked to structural genes and it is likely they are caused by differences in promoter regions or polymorphisms in regulatory elements of structural genes from the two parents (Gilad et al., 2008, Hansen et al., 2008, Alberts et al., 2007).

There have not been many studies on VTE cis-eQTLs, except for IL9-2-6-1, a cis-eQTL found for *VTE3(2)*, which is epigenetically related (Quadrana et al., 2014). The majority of mQTLs studied have been linked to underlying cis-eQTLs, which contain structural genes (Wentzell et al., 2007), and the VTE pathway is no exception. Almeida et al., (2012) and Schauer et al., (2006) showed that mQTLs which showed altered tocopherol content within tomato fruit are linked to structural pathway genes. The core structural VTE pathway genes are localised on chromosomes 7, 8 and 9; except *VTE3(2),* which mapped to chromosome 3. Regulatory components have been hard to identify, and they are often missing

from mQTL studies, and this illustrates why there is a lack of identified TFs regulating the VTE pathway. Chromosome 9 harbours two QTLs (IL9-1 and IL9-2-6), altering total and compositional tocopherol. Out of these two QTLs, IL9-1 altered alpha and total tocopherols levels, yet co-localises with *GGPS(4)* and *TYRA2* (table 3-3). Therefore, genes in IL9-1 could be altering the precursor availability for the pathway, explaining the change in total tocopherols. QTLs on chromosome 6 affect total, delta (δ) and beta (β) tocopherol levels (Schauer et al., 2006, Almeida et al., 2011), but, IL6-1 and IL6-2 co-localises with genes upstream in the VTE pathway. Indirect effects of the co-localised chlorophyllase (*CHL*) enzyme, which degrades chlorophyll for regeneration of phytyl diphosphate by VTE5, may affect precursor availability (Almeida et al., 2011). Other VTE mQTLs on ILs 7-4, 7-4-1, 8-2, 8-2-1, have all been mapped to structural genes and alter total and compositional tocopherol levels. Thus, Almeida et al (2012) showed that tocopherol levels can be altered, but are linked to structural genes, and no regulatory elements were discovered. This suggested that an eQTL co-expression analysis on the *S.pennellii* x *S.lycopersicum* cv. M82 ILs might offer a powerful tool to identify regulatory components of the VTE pathway.

3.4.3 The trans-eQTL IL9-3-2 contains putative trans-acting regulators of the VTE pathway and corresponds to *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs

Expression in fruit of the *S.pennellii* IL9-3-2 was reduced for VTE pathway genes, which suggested that this is a trans-eQTL. Reduced expression of this pathway implied that an activator of the pathway would also be down-regulated, or, a repressor would be up-regulated. The candidate TFs identified in this region could also be trans-regulators for the precursor pathway; the MEP pathway, as this IL also showed reduced expression of the genes of this pathway. Not all VTE pathway genes were down-regulated in this region in IL9-3-2 (table 3-1), but the VTE pathway genes have differential expression profiles which are induced during fruit development and ripening, even though tocopherol levels are not significantly altered (Quadrana et al., 2013). Expression of *VTE2, VTE3(1)* and *VTE4*

was reduced in this trans-eQTL, and *VTE3(1)* and *VTE4* showed significantly different expression ratios between the *S.pennellii* x *S.lycopersicum* cv. M82 IL parents *(S.lycopersicum and S.pennellii).* The fact that expression of several VTE pathway genes was significantly different between the two *S.pennellii* x *S.lycopersicum* cv. M82 IL parents implied that candidate TFs are likely to regulate these genes. There are no known TFs that regulate the VTE pathway, but Quadrana et al. (2013) showed that the VTE, MEP and SK are differentially expressed during tomato development and ripening. Therefore, transcriptional regulation does exist. The co-expression of the VTE pathway genes and significant differences in transcript levels of my data, between the parents suggested that this analysis might identify TFs regulating the VTE pathway or the MEP pathway.

Up-regulated genes in the MEP pathway (figure 3-7), such as; *DXS(1)* and *GGPS(1)* could be more highly expressed due to specific isoforms. In tomato fruit isoforms of *DXS* and *GGPS* are differentially expressed in different organs (appendix figure 3-1), therefore in this trans-eQTL, they are also differentially expressed. Other genes in the MEP pathway showed reduced expression in the trans-eQTL IL9-3-2, which suggested co-regulation of the pathway. Very little is known about transcriptional regulators of the MEP pathway, and many studies have focused on post-translational mechanisms of regulation (Cordoba et al., 2009, Pepper et al., 1994) . Therefore, this trans-eQTL analysis could provide future insights into MEP transcriptional regulation.

These trans-eQTLs identified were compared to *S.pennellii* x *S.lycopersicum* cv. M82 IL RNA sequencing data of leaves. The trans-eQTL IL9-3-2 did not show the same co-expression results in leaves, as it did in fruit. An equivalent trans-eQTL was identified in the *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs. Reduced expression of VTE pathway genes was observed in the *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs, suggesting that this region of chromosome 9 is a strong candidate for carrying candidate transcriptional regulators of the VTE pathway. Using mQTLs has shown chlorogenic acid content in tomato fruit can be due to regulatory TFs, and not pathway genes (Alseekh et al., 2015). But, this is

the first instance that trans-eQTLs have been used to find regulatory TFs of a pathway in tomato fruit.

3.4.4 The trans-eQTLs IL6-2-2 contains a trans-acting regulator putatively regulating the VTE and SK pathway

The trans-eQTL IL6-2-2 showed large differences of VTE pathway gene expression compare to other ILs in fruit . This region showed increased expression of *VTE2, VTE3(1) VTE3(2), VTE1* and *VTE4* (table 3-1). *VTE5* and *VTE6* were not co-expressed in this region, but, *VTE5* and *VTE6* are not directly involved in VTE biosynthesis, and rather, synthesise the regeneration of free phytol from chlorophyll degradation (figure 3-2) (Almeida et al., 2016, Valentin et al., 2006, vom Dorp et al., 2015). Therefore, these two parts of the pathway can be considered separately.

Co-expression of the precursor pathways of the VTE pathway showed inconsistent results, compared to the expression of the VTE pathway. The MEP pathway did not show consistent co-expression for the trans-eQTL IL6-2-2. The SK pathway showed that some genes might be co-regulated using this co-expression analysis. *EPSPS2, EPSPS3, TAT1* and *TAT3* were highly expressed in the trans-eQTL IL6-2-2, and *DAHPS1, DAHPS2* and *DAHPS3* were up-regulated, but, not to such high expression levels.

Several SK pathway genes, including *DAHPS* and *EPSPS* are co-regulated by transcription factors in petunia (Spitzer-Rimon et al., 2010, Takatsuji et al., 1992, Verdonk et al., 2005), although none of these are present in this trans-eQTL. Therefore, this trans-eQTL could also contain a putative regulator responsible for the co-regulation of these SK pathway genes. On the other hand, *DAHPS* could be the gene that is transcriptionally regulated in this region and the other genes that are up-regulated could be due to downstream effects. *DAHPS* enzymes are not inhibited by the aromatic amino acids (AAAs) produced by the SK pathway, which include tyrosine and phenylalanine (Herrmann, 1995). *TAT* enzymes are regulated, in part, by AAAs. *TAT* enzymes are known to be important for VTE biosynthesis, as

shown by *Arabidopsis* knock-out mutants which have significantly altered levels of tocopherols and higher tyrosine levels (Riewe et al., 2012). Therefore, *DAHPS* could be a transcriptional target for the trans-eQTL in this region, resulting in subtle changes of expression. The genes that encode the *TAT* enzymes, which showed large changes in expression in IL6-2-2 compared to the other ILs, could be under the influence of feedback inhibition of tyrosine and phenylalanine, caused by an increase in flux in the pathway, initiated by the induction of *DAHPS* expression. Either possibility could be the cause for higher expression of VTE pathway genes in this trans-eQTL

The trans-eQTL IL6-2-2 in the *S.pennellii* x *S.lycopersicum* cv. M82 ILs from fruit was compared to *S.pennellii* x *S.lycopersicum* cv. M82 leaf RNA sequencing data. This IL showed a similar reduction of VTE pathway gene expression, indicating that this trans-eQTL could also contain a trans-acting regulator active in leaves, as well as fruit, controlling the VTE pathway. This region has previously been identified as a trans-eQTL for plant pathogen and disease defence genes (Ranjan et al., 2016), which contains master regulators for these responses. The SK pathway is known to be involved in defence responses (Ferrari et al., 2007), thus, the VTE pathway expression profile could be a downstream effect of trans-eQTLs affecting the SK pathway. Alternatively, this trans-eQTL could also contain a master regulator of the VTE pathway.

The trans-eQTL IL6-2-2 was also compared to the equivalent *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs for the fruit RNA sequencing data, but did not show changed co-expression of VTE pathway genes in the equivalent IL. Therefore, this trans-eQTL cannot be translated to other *Solanum* species, but, can be correlated to leaf RNA sequencing data for *S.pennellii* x *S.lycopersicum* cv. M82 ILs.

3.4.5 Candidate gene mining of trans-eQTLs has enabled the identification of several putative TFs regulating the VTE pathway

Candidate genes for the trans-eQTLs showed several possible candidates for transient screening assays (figure 3-12 and 3-15). The TFs selected were restricted based on co-expression analysis with the VTE pathway in the *S.pennellii* x *S.lycopersicum* cv. M82 ILs and promoter analysis from literature (Quadrana et al., 2013). Both MYB and bZIP TF binding motifs are highly conserved in the VTE pathway gene promoters. Using promoter enrichment analysis for trans-eQTLs in rice showed that several submergence response genes were co-regulated. The authors found the known regulator responsible for this trans-eQTL using this method (Kuroha et al., 2017). Therefore, it is likely that using promoter motifs to reduce TF candidates is a useful way to find putative regulators. TFs that were identified for the candidate trans-eQTLs were cloned and transiently screened in tomato fruit, as described in chapters 5 and 6.

Using the co-expression approach to identify trans-eQTLs is a powerful tool to identify candidate TFs regulating the VTE pathway. Using this method allows identification of previously unknown transcriptional regulators, that could be responsible for regulating the VTE pathway gene expression in tomato. Of course, this approach can be used for other metabolic pathways, in addition to the VTE pathway.

Chapter 4: Analysis of the Vitamin E pathway for the parents of the *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression lines

Chapter 4: Analysis of the Vitamin E pathway for the parents of the *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression lines

4.1 Introduction

4.1.1 *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression lines

The *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines (ILs) are a useful tool to identify putative transcriptional regulators of Vitamin E (VTE) synthesis in tomato fruit. In order to identify such regions, first, I needed to address the VTE content and transcriptional regulation of the parent ILs to determine the appropriate ILs for candidate gene mining. Therefore, I have completed a time course of VTE levels during fruit development and ripening within *S.lycopersicum* cv. M82 and *S.pennellii*. Previous studies had suggested that regulation of VTE synthesis occurs spatially and temporally within tomato fruit (*S.lycopersicum* cv. M82) (Quadrana et al., 2013), but, total tocopherol content remains unchanged over fruit development and ripening. That study failed to complete a fully comprehensive time course with several stages of development, and rather assessed just 4 stages: immature green (IG), mature green (MG), breaker (B) and red (R) tomato fruit (30, 45, 50 and 60 days post anthesis, respectively). I analysed VTE levels and expression profiles in both parental lines, over several stages of fruit development and ripening, which are summarised in figure 4-1. For M82, I analysed both tissue of the pericarp and the epidermis, as the available *S.pennellii* x *S.lycopersicum* cv. M82 IL RNA sequencing data are based solely on pericarp tissues.

4.1.2 The Vitamin E pathway

VTE is produced solely in photosynthetic organisms and the pathway was characterised in *Arabidopsis thaliana* (Shintani and DellaPenna, 1998, Schledz et al., 2001, Collakova and DellaPenna, 2001, Savidge et al., 2002, Shintani et al., 2002, Cheng et al., 2003). Identification of *Arabidopsis* mutants has elucidated the VTE biosynthetic enzymes as well as novel functions for VTE within plants

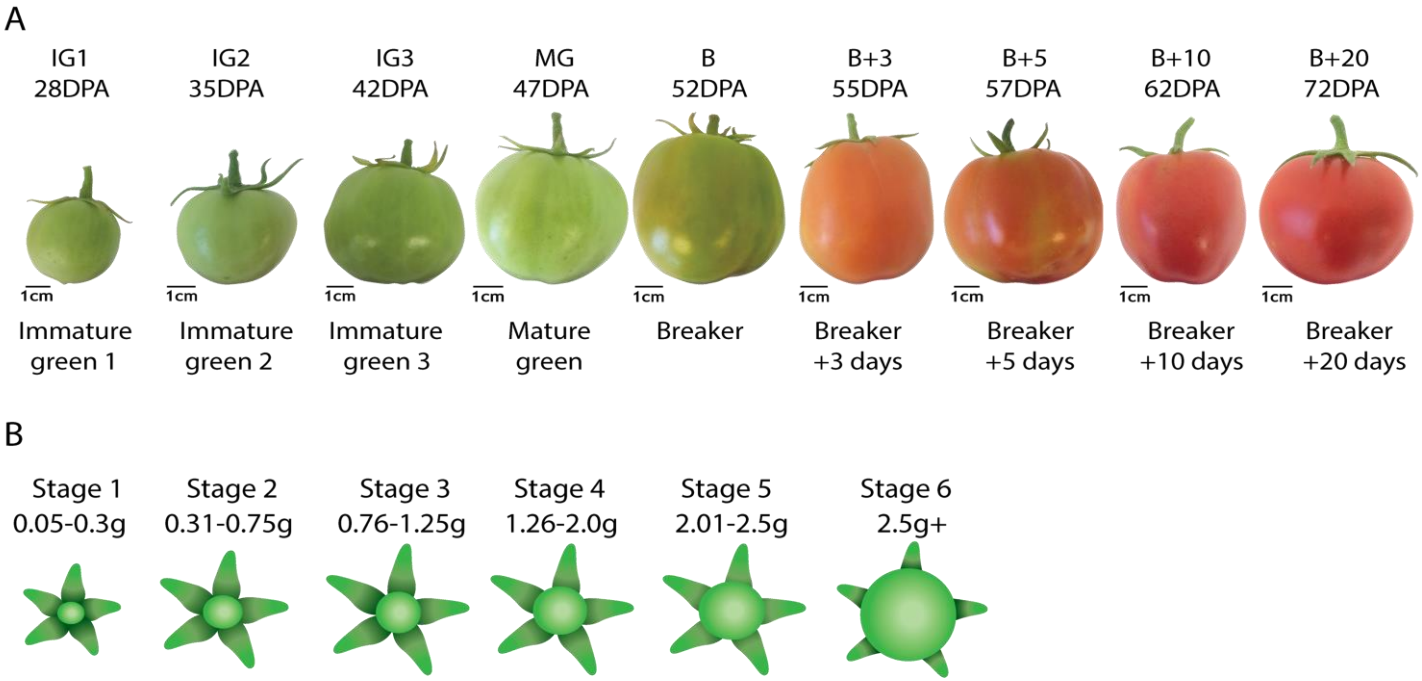


Figure 4-1 The tomato fruit stages of (A*) S.lycopersicum* cv. M82 and (B) *S.pennellii* used for the VTE time course. Photos were taken at the time of the experiment. The tomato stages and days post anthesis (DPA) are labelled above the *S.lycopersicum* cv. M82 tomatoes. The corresponding fruit stage is shown underneath the tomatoes of *S.lycopersicum* cv. M82 and the scale bar shows 1cm. The *S.pennellii* fruits were harvested based on their weights, which were grouped into six stages, as shown in the diagram.

(Maeda et al., 2006, Sattler et al., 2006, Maeda and DellaPenna, 2007, Maeda et al., 2014). Tocochromanol precursors derive from the plastidic methyl erythitol phosphate (MEP) and the shikimate (SK) pathways (figure 4-2). Both tocopherol and tocotrienols are synthesised from precursors derived from these pathways. Tocopherols are produced by VTE2, which catalyses a condensation reaction between phytyl diphosphate (PDP), from the MEP pathway, and homogentisate, from the SK pathway. Tocotrienols are produced by HGGT, which catalyses a condensation reaction between geranylgeranyl diphosphate (GGDP) from the MEP pathway and homogentisate. Tocopherols and tocotrienol biosynthesis differs only at this branch point, and all other enzymes in the VTE pathway can use either product for biosynthesis of VTE (figure 4-1).

4.1.3 Regulation of the VTE pathway is limited and needs clarification

Arabidopsis vte mutants have been shown to have altered tocopherol levels, although, not to substantial levels, and rather, tend to have altered composition of tocopherols (Porfirova et al., 2002, Collakova and DellaPenna, 2003, Kanwischer et al., 2005, Bergmuller et al., 2003). This highlights the requirement to identify transcriptional regulators of this pathway, to increase VTE levels, as transcriptional regulators are not limited by pathway flux in the same way as structural genes. This chapter aims to understand the transcriptional environment of the VTE, MEP and SK pathway, using analysis of expression and metabolite data.

4.2 Materials and Methods

4.2.1 Plant Materials

S.lycopersicum cv. M82 plants were grown under greenhouse conditions for fruit harvest. Nine stages of fruit development and ripening were harvested for analysis, which included; immature green 1 (IG1), immature green 2 (IG2), immature green 3 (IG3), mature green (MG), breaker (B), breaker + 3 days (B+3), breaker + 5 days (B+5), breaker + 10 days (B+10) and breaker + 20 days (B+20). These stages corresponded to days post anthesis (DPA); IG1 = 28 DPA, IG2 = 35

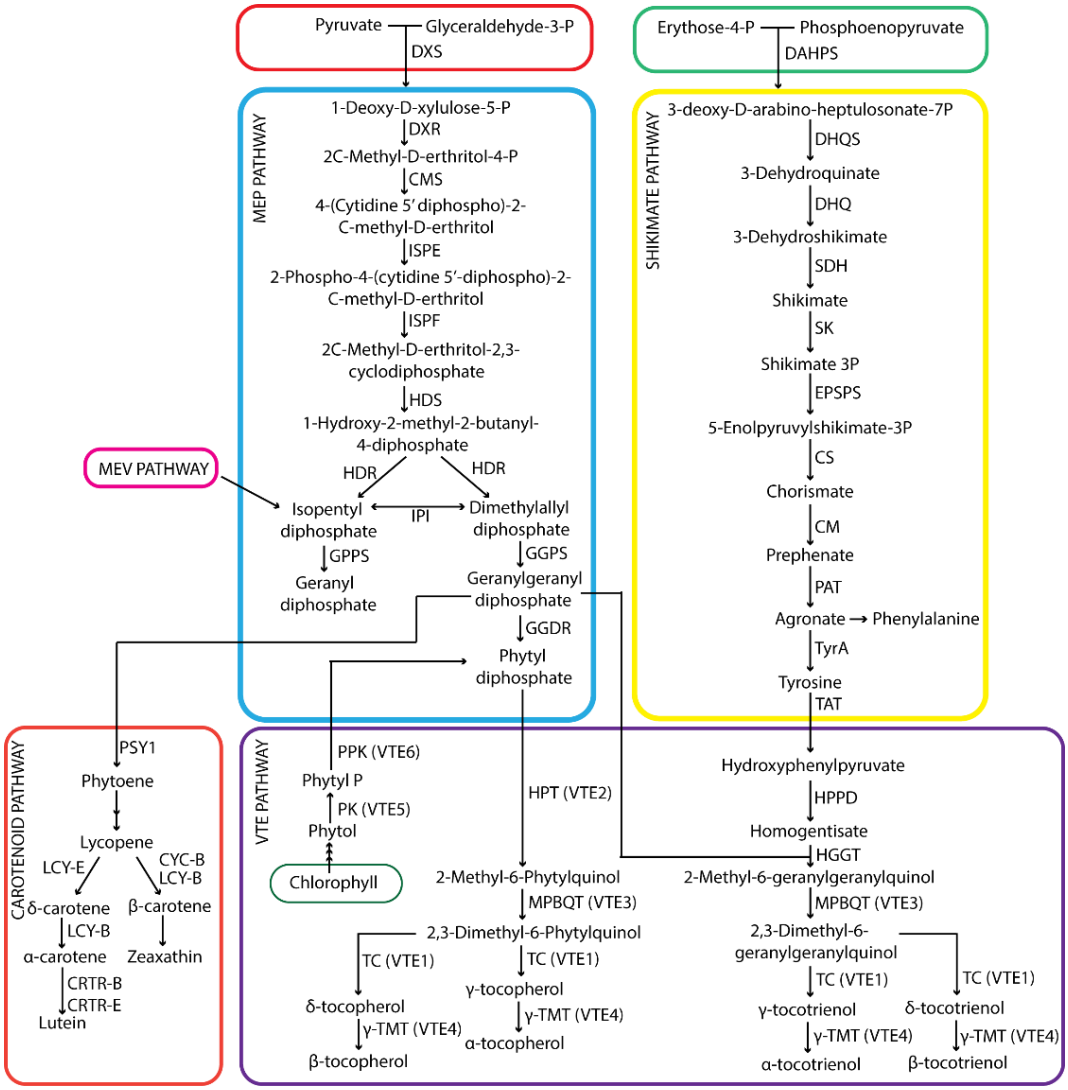

Pyruvate
Glyceraldehyde-3-P
DXS
Erythose-4-P
Phosphoenopyruvate
DAHPS
MEP PATHWAY
1-Deoxy-D-xylulose-5-P
DXR
2C-Methyl-D-erthritol-4-P
CMS
4-(Cytidine 5′diphospho)-2-C-methyl-D-erthritol
ISPE
2-Phospho-4-(cytidine 5′-diphospho)-2-C-methyl-D-erthritol
ISPF
2C-Methyl-D-erthritol-2,3-cyclodiphosphate
HDS
1-Hydroxy-2-methyl-2-butanyl-4-diphosphate
HDR
HDR
MEV PATHWAY
Isopentyl diphosphate
IPI
Dimethylallyl diphosphate
GPPS
GGPS
Geranyl diphosphate
Geranylgeranyl diphosphate
GGDR
Phytyl diphosphate
SHIKIMATE PATHWAY
3-deoxy-D-arabino-heptulosonate-7P
DHQS
3-Dehydroquinate
DHQ
3-Dehydroshikimate
SDH
Shikimate
SK
Shikimate 3P
EPSPS
5-Enolpyruvylshikimate-3P
CS
Chorismate
CM
Prephenate
PAT
Agronate → Phenylalanine
TyrA
Tyrosine
TAT
CAROTENOID PATHWAY
PSY1
Phytoene
Lycopene
LCY-E
CYC-B
LCY-B
δ-carotene
β-carotene
LCY-B
α-carotene
Zeaxathin
CRTR-B
CRTR-E
Lutein
VTE PATHWAY
PPK (VTE6)
Phytyl P
PK (VTE5)
Phytol
Chlorophyll
HPT (VTE2)
2-Methyl-6-Phytylquinol
MPBQT (VTE3)
2,3-Dimethyl-6-Phytylquinol
TC (VTE1)
TC (VTE1)
γ-tocopherol
δ-tocopherol
γ-TMT (VTE4)
γ-TMT (VTE4)
α-tocopherol
β-tocopherol
Hydroxyphenylpyruvate
HPPD
Homogentisate
HGGT
2-Methyl-6-geranylgeranylquinol
MPBQT (VTE3)
2,3-Dimethyl-6-geranylgeranylquinol
TC (VTE1)
TC (VTE1)
γ-tocotrienol
δ-tocotrienol
γ-TMT (VTE4)
γ-TMT (VTE4)
α-tocotrienol
β-tocotrienol

Figure 4-2 Outline of the tocochromanol pathway. MEP, SK, carotenoid and VTE pathway in blue, yellow, orange, and purple, respectively. Enzyme names are as follows: **DXS**; 1-Deoxy-D-xylulose-5-P synthase, **DXR**; 2C-Methyl-D-erythritol-4-phosphate synthase, **CMS**; 2C-Methyl-D-erythritol-4-phosphate cytidyltransferase, **ISPE**; 4-2-CMethyl-D-erythritol kinase, **ISPF**; 2C-Methyl-D-erythritol-2-3-cyclodiphosphate synthase, **HDS;** 4-Hydroxy-3-methylbut-2-enyl-diphospahte synthase, **HDR**; 4-Hydroxy-3-methylbut-2-enyl-diphosphate reductase, **IPI**; Isopentyl diphosphate δ isomerase, **GPPS;** Geranyl pyrophosphate synthase, **GGPS**; Geranylgeranyl pyrophosphate synthase **GDDR**; Geranylgeranyl reductase, **PSY**; Phytoene synthase, **DAHPS**; 3-Deoxy-D-arabino-hepulosonate, **DHQS**; 3-Dehydroquianate synthase**, SDH-DHQS1**; 3-Dehydroquianate dehydratase**, SDH-DHQS2**; Shikimate 5-dehydrogenase**, SK**; Shikimate kinase**, EPSPS**; 5-Enolpyruvylshikimate-3-P-synthase, **CS**; Chorismate synthase, **CM**; Chorismate mutase, **PAT**; Prephenate aminotransferase, **TyrA**; Arogenate dehydrogenase, **TAT**; Tyrosine aminotransferase, **HPPD**; 4-Hydroxyphenylpyruvate dioxygenase, **HPT (VTE2)**; Homogentisate phytyl transferase, **MPBQMT (VTE3)** Dimethyl-phytylquinol methyl transferase, **TC (VTE1)**; Tocopherol cyclase, **γ-TMT (VTE4)**; γ-Tocopherol C-methyl transferase, **PK (VTE5)**; Phytol kinase, **PPK (VTE6)**; Phytyl-phosphate kinase, **PSY1**; Phytoene synthase 1, **CYC-B**; chromoplast-specific lycopene cyclase, **LYC-B**; lycopene β cyclase , **LYC-E** lycopene ε cyclase , **CRTR-B**; β-ring hydroxylase , **CRTR-E**; ε-ring hydroxylase.

DPA, IG3 = 42 DPA, MG = 47DPA, B = 52DPA, B+3 = 55DPA, B+5 = 57 DPA, B+10 = 62 DPA and B+20 = 72 DPA. The DPA can be variable depending on the conditions in which the plants are grown, therefore the DPA was adjusted to visual B stages, for the relevant conditions. Breaker is defined as the stage at which the first change of colour is seen on the flesh of the tomato fruit. At least four fruits were harvested for each time point, all were from different M82 plants.

S.lycopersicum cv. M82 fruit were harvested and seeds were excised. The pericarp (flesh) was separated from the epidermis (peel). As much as possible of the pericarp was scraped off the epidermis, although not all of it could be removed, therefore this fraction has to be considered as enriched epidermis tissue. Both the tissues of the pericarp and the epidermis were ground to a powder using liquid nitrogen. This powder was stored at -80°C for further analyses. All RNA extractions, cDNA synthesis, qRT-PCR and metabolite analysis are described in chapter 2.

S.pennellii fruit growth is slow, therefore, the plant used for analysis was growing in the greenhouse at the time of the *S.lycopersicum* cv. M82 sowings. Tomatoes were not harvested based on their developmental stage as *S.pennellii* does not develop and mature in the same way as fruit of *S.lycopersicum*. *S.pennellii* fruit grow exponentially until they reach a mature size (Steinhauser et al., 2010). Based on the study by Steinhauser et al. (2010), *S.pennellii* fruit were harvested based on their weight, which correlates directly their fruit volume. Fruit of the *S.pennellii* were harvested into six groups based their weight, which include; stage 1 = 0.05-0.3g, stage 2 = 0.31-0.75g, stage 3 = 0.76-1.25g, stage 4 = 1.26-2.0g, stage 5 = 2.01-2.5g and stage 6 = 2.5g or more. The number of fruit in each sample was dependent on the stage and fruit weight (appendix table 4-1). However, four replicates of each stage were used for analyses. Tomatoes were harvested and ground to a fine powder, as previously described.

4.3 Results

4.3.1 Spatio- and temporal control of VTE pathway gene expression exists within *S.lycopersicum* cv. M82

Differential expression of the VTE pathway genes was observed in the *S.lycopersicum* cv. M82 time course. As figure 4-3 shows many of the VTE pathway genes had similar co-expression patterns in the tissue of the pericarp of M82 tomato fruit, throughout the development and ripening of the fruit. *VTE2, VTE1, VTE3(1), VTE3(2), VTE4* and *VTE5* did not show substantial changes in gene expression throughout tomato development and ripening. However, they showed a general decrease in expression until stage B+3. After this stage, the expression of these genes was increased slightly in B+5 and B+10, and then, transcript levels were reduced in the B+20 fruit stage. There were no significant changes in expression during tomato fruit development and ripening for these genes (table 4-2 in appendix). In contrast with these genes, *VTE6* displayed a differential expression pattern. The expression of *VTE6* in the tissue of the pericarp showed a general increase in transcript levels throughout development and ripening.

Expression of the VTE pathway genes in the epidermis, in comparison to the pericarp for *VTE2, VTE1, VTE3(2), VTE4* and *VTE5* genes (figure 4-3) was similar. These genes showed similar expression values (except *VTE1)*, however, at the B+3 stage, the expression level increased in the epidermis, which was unlike the reduction in expression seen in pericarp tissues. *VTE3(1)* showed a relatively stable expression profile during tomato development and ripening in the epidermis (figure 4-3). This is shown in appendix table 4-2, and the only significant changes in expression were between the B+5 stage and B+10 stage. However, there were significant differences in *VTE3(1)* expression between the tissue of the pericarp and epidermis. These tissues showed significant changes at MG, B+3, B+10 and B+20. *VTE6* showed a similar expression pattern in the epidermis, which matched its expression pattern in the pericarp. Expression of *VTE6* was reduced at the B+3 stage (figure 4-3), which was significantly different from the expression of *VTE6* in the pericarp. This contrasted with the other VTE pathway genes which

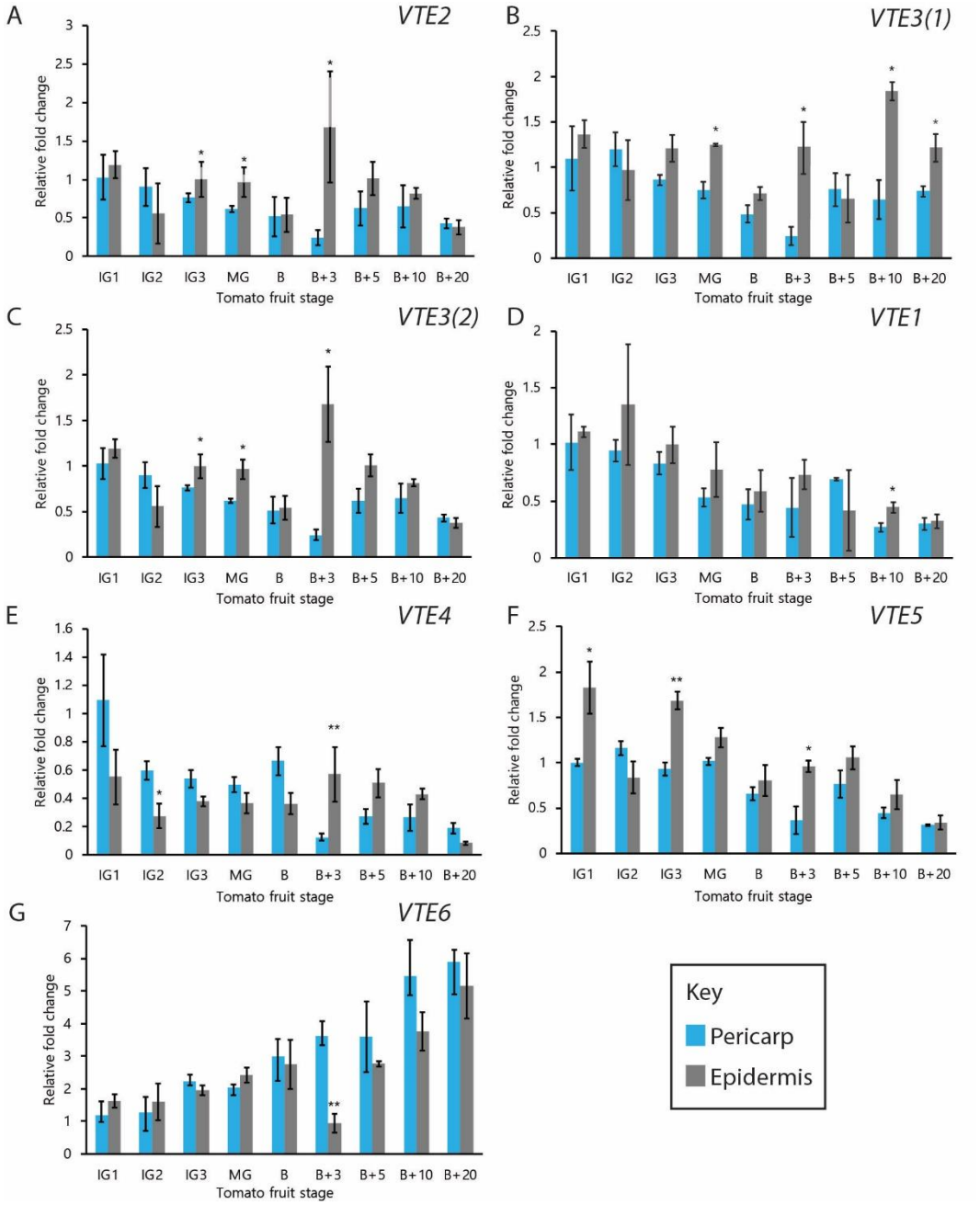


Figure 4-3 Relative fold changes of gene expression of the VTE pathway genes during nine stages of *S.lycopersicum* fruit development. Blue bars represent pericarp tissue and grey bars represent epidermis tissue (n=3). Figure legends represent the following genes: A = VTE2, B = VTE3(1), C = VTE3(2), D = VTE1, E = VTE4, F = VTE5, G= VTE6. The error bars show the standard error of the mean. The stars indicate significant differences (p value<0.05) = *, (p value <0.01) = **, between the pericarp and epidermis at a given tomato fruit stage, using student t-tests. Tukey tests for the statistical significance between the different stages of *S.lycopersicum* for tissues of the pericarp and epidermis are shown in appendix table 4-1. All relative fold changes are normalised to IG1 pericarp stage of tomato fruit development and the reference gene was SlCAC.

showed increased expression at this stage when compared to the tissue of the pericarp. This general increase in *VTE6* was statistically significant between the early (IG1, IG2, IG3 and MG) and late (B+10 and B+20) fruit stages (table 4-2 in appendix).

Overall, expression of the VTE pathway genes in the pericarp was constant during tomato fruit development and ripening. Expression of these genes in the tissue of the epidermis was generally for several genes. But, most interestingly, *VTE6* showed a general increase in expression during tomato fruit development and ripening, which has not been documented before.

4.3.2 The MEP and SK pathway showed differential expression profiles compared to the VTE pathway in *S.lycopersicum* cv. M82

Divergent expression profiles were observed for the MEP pathway genes during fruit development (figure 4-4). The expression of the genes within the MEP pathway are higher than the expression of the VTE pathway previously observed in figure 4-3. The expression profile, generally, showed an increase of MEP pathway gene expression as the tomato enters the ripening stages for tissues of both the pericarp and epidermis (figure 4-4). Therefore, there may be a transitional switch from the early tomato development stages and the later tomato ripening stages for this pathway. Both the pericarp and the epidermis showed similar expression profiles, with a few exceptions. Expression of *DXS(1), ISPE, ISPF, HDS, IPI(1), GGPS(2)* and *GGPS(3)* was higher in the epidermis and was also induced at B+3, which opposes the pericarp expression profile. The expression in the pericarp was reduced for these genes at B+3, and then induced later at B+5. *DXS(2), IPI(2)* and *GGPS(1)* expression was reduced during tomato fruit development and ripening stages. *CMS* expression was also reduced, but, it was low in the earlier stages of fruit development, and in the transition from MG to B fruit. Additionally, there are several genes; *GGPS(4), GGDR* and *HST,* which all displayed relatively stable expression profiles throughout the development and ripening processes.

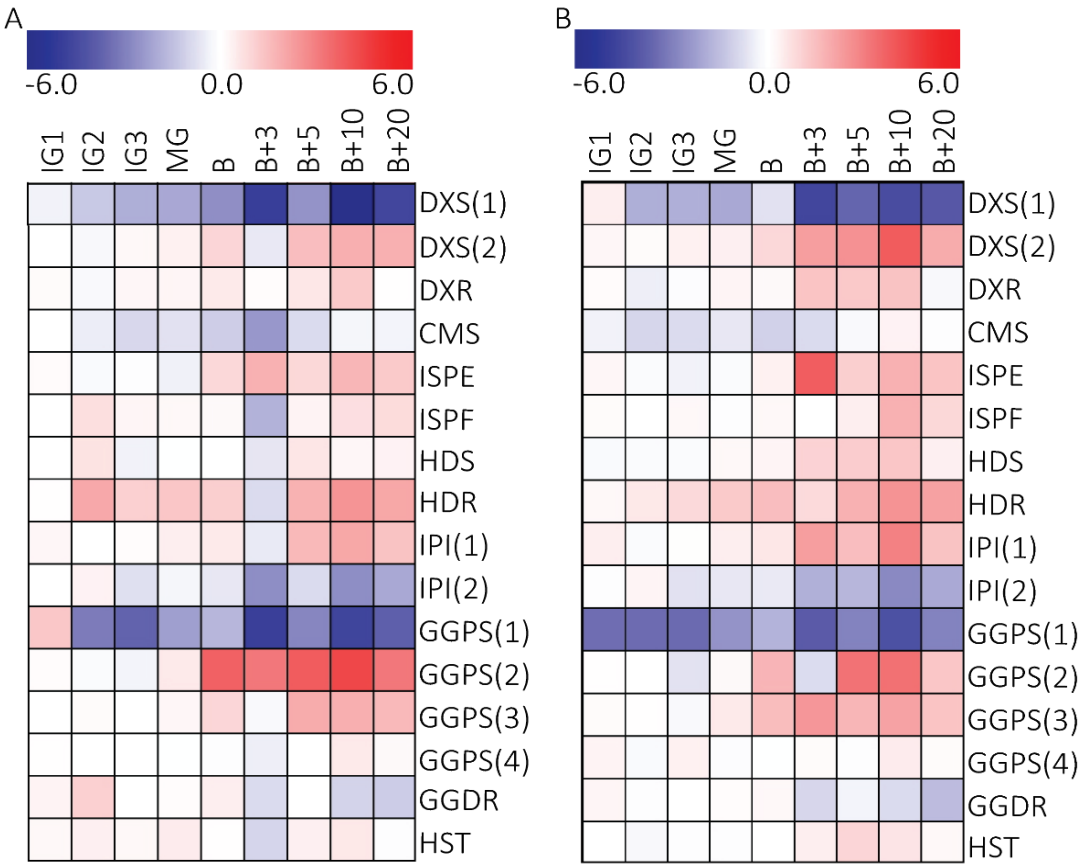


Figure 4-4 Heatmap showing log_2 of relative fold changes of the MEP pathway gene expression for nine stages of *S.lycopersicum* fruit ripening and development (n=3). Figure A shows data for the pericarp and B shows data for the epidermis. The scale bar shows log_2 changes from -5.0 to 5.0. The relative expression values for each gene are normalised to IG1 pericarp of tomato fruit development and SlCAC was used as a reference gene.

The SK pathway showed substantially greater changes in pathway gene expression compared to the VTE pathway (figure 4-5). Several genes of the SK pathway in the tissues of the pericarp and epidermis were highly expressed throughout tomato development and ripening stages. There was much more variation in expression of SK pathway genes between the epidermis and the pericarp, although, there were some similarities. Genes such as; *DHQS, SDH-DHQS1, SDH-DHQS2, CM2, TAT1, TAT2* and *HPPD2* displayed similar expression profiles in the tissues of the pericarp and the epidermis. However, *SK, EPSPS, CS1, CS2, CM1* and *HPPD)* were more highly expressed in the epidermis than in the pericarp, and the induction of expression of these genes occurred earlier in fruit development – as early as IG1 stage, which was not observed in the tissue of the pericarp. Interestingly, for several SK genes, there was a decrease in expression at the B+3 stage, which was also observed in the MEP gene expression profile of the tissue of the pericarp. The *DAHPS1* and *DAHPS2* genes showed opposing expression profiles during the ripening stages of tomato. Within the pericarp tissue, the expression of these two genes was reduced during ripening, whereas, in the epidermal tissue, expression of both was increased, or, at least there was not a large decrease in expression (figure 4-5).

These data demonstrated that there were spatial and temporal expression changes between the pericarp and epidermis. Therefore, differential transcriptional control could exist for the different tissues. For the MEP and SK pathway, there was a general increase in expression of many genes, which was unlike the expression of the VTE pathway genes.

4.3.3 Tocopherol levels were altered during development of *S.lycopersicum* cv. M82 fruit in the epidermis of fruit, but, not in the pericarp

The metabolite profile for VTE levels changed concomitantly with the expression profiles of VTE pathway genes during fruit development (figure 4-3). As figure 4-6 shows, there was no large change in total or compositional tocopherol

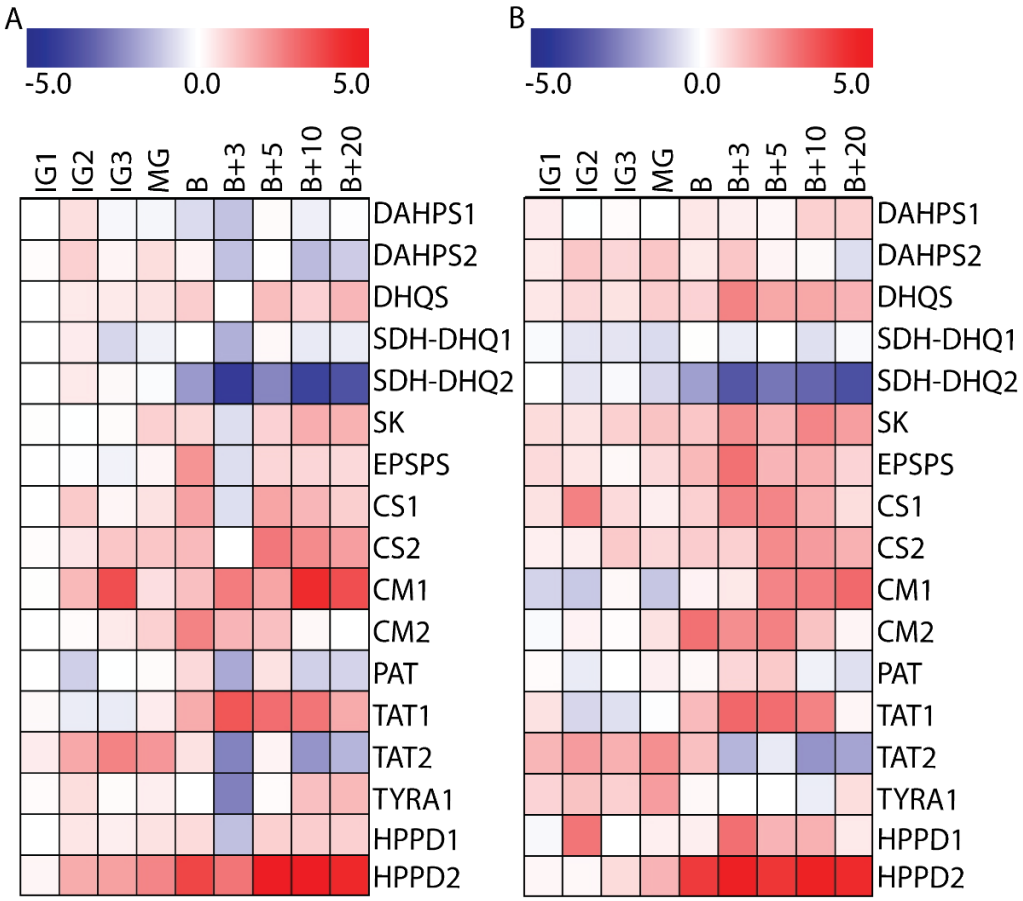


Figure 4-5 Heatmap showing log_2 of relative fold changes of the SK pathway gene expression for nine stages of *S.lycopersicum* fruit ripening and development (n=3). Figure A shows data for the pericarp and B shows data for the epidermis. The scale bar shows log_2 changes from -5.0 to 5.0. The relative expression values for each gene are normalised to IG1 pericarp of tomato fruit development and SlCAC was used as a reference gene.

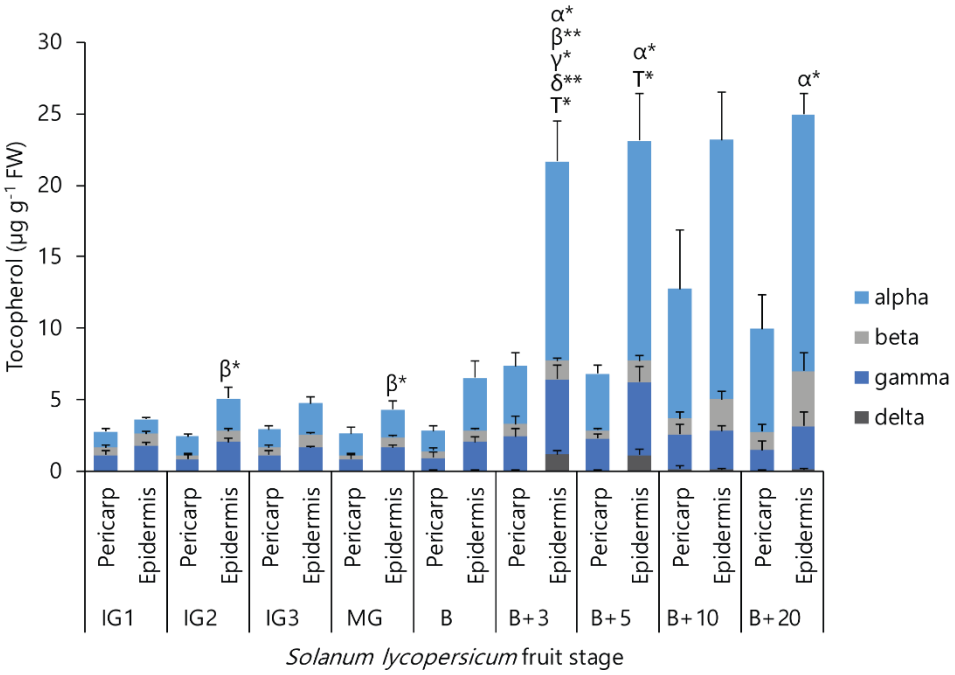


Figure 4-6 Tocopherol contents (µg g^{-1} fresh weight) of the nine S.lycopersicum fruit stages of fruit development and ripening. The tissue of the pericarp and epidermis is shown for each fruit stage. The different forms of tocopherols are depicted using different colours; alpha tocopherol is light blue, beta tocopherol is light grey, gamma tocopherol is dark blue and delta tocopherol is dark grey. The error bars depict standard error of the mean. The stars indicate statistical significance (*) = p value < 0.05, (**) = p < 0.01 for different forms of tocopherols between the pericarp and epidermis, calculated using student t-tests. The different forms of tocopherol are shown using letters α = alpha, β = beta, γ = gamma, δ = delta and T = total. Tukey tests were calculated for statistical significance between the nine stages of S.lycopersicum fruit development and ripening. There was no statistical significance between the nine fruit stages in the tocopherol contents of the pericarp. However, there were statistical significant differences in tocopherol contents in the epidermis, which are shown in appendices table 4-3. Each stage contained (n=4) replicates per sample.

levels in the tissue of the pericarp in the M82 fruit during development. There was a minor increase in overall tocopherol levels from the B+3 stage onwards, which reached a peak of 12.76 μg g^{-1} in B+10 fruit, but, this was not significantly different from levels at other stages. However, total tocopherol levels in the epidermis were significantly increased from the B+3 stage onwards, and total levels rose to 24.94 μg g^{-1}, which was double the total levels observed in the pericarp tissue.

Alpha tocopherol remained the most common form of tocopherol found in the *S.lycopersicum* cv. M82 tomatoes, at all stages and in all tissues (figure 4-6). Gamma (γ) tocopherol was always the second most abundant vitamer, and β- and delta (δ) were always the least abundant vitamers. The increase in tocopherol levels seen in the epidermis at the B+3 stage, was significantly higher than in the pericarp for all forms of tocopherol. This rise occurred concurrently with the increased expression of the MEP and SK pathway genes at the B+3 stage, which was not observed in the pericarp. Total tocopherol levels remained at a similar level throughout the rest of tomato ripening, which is significantly different from the early developmental stages (table 4-3 in appendix). The total tocopherol levels were significantly higher in the epidermis, compared to the pericarp for the B+5 stage. Alpha tocopherol levels were also higher in the epidermis for the B+5 and B+20 stages, in comparison with the contents of the pericarp.

4.3.4 *S.pennellii* fruit demonstrated similar expression patterns for the VTE pathway genes as *S.lycopersicum* cv. M82 fruit

Expression of the VTE pathway genes in developing fruit of *S.pennellii* was more variable that for *S.lycopersicum* cv. M82 (figure 4-7). *VTE2* and *VTE3(2)* showed similar expression patterns (figure 4-3 and figure 4-7), which remained stable throughout fruit maturation. There was an incremental reduction of *VTE2* and *VTE3(2)* transcript levels, but, expression rose in stage 6, to similar levels of expression as in stage 1. The genes; *VTE1, VTE3(1)* and *VTE5,* displayed expression profiles, which were very similar throughout *S.pennellii* fruit maturation. There was a slight increase between stages 2-4 for these genes (figure 4-7), but, there

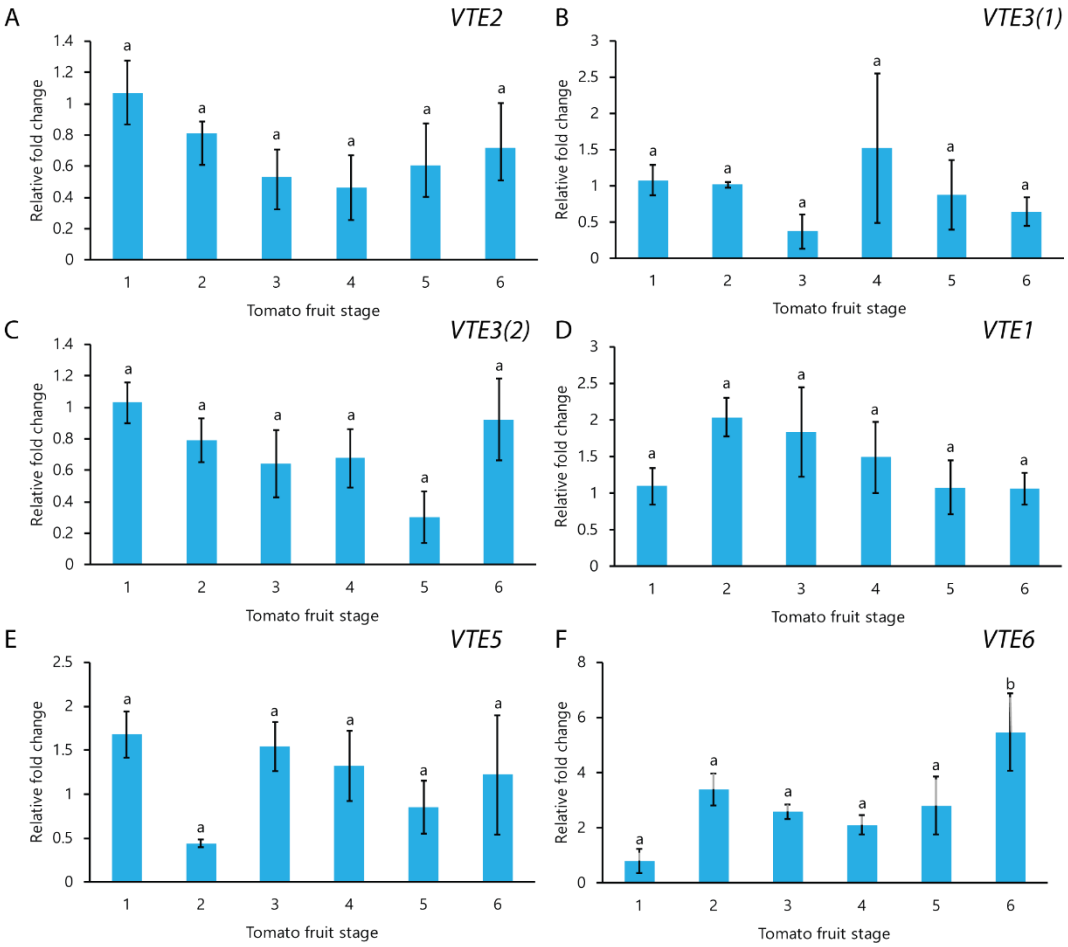


Figure 4-7 Relative fold changes of gene expression of the VTE pathway genes during six stages of *S.pennellii* fruit development. Figure legends represent the following genes: A = VTE2, B = VTE3(1), C = VTE3(2), D = VTE1, E = VTE5, F = VTE6. The error bars show the standard error of the mean. The stars indicate significant differences p value<0.05 = (*), p value <0.01 = (**), between the pericarp and epidermis at a given tomato fruit stage, using student t-tests. Tukey tests for the statistical significance between the different stages of S.lycopersicum for tissues of the pericarp and epidermis are shown in appendix table 4-1. All relative fold changes are normalised to IG1 pericarp stage of tomato fruit development and the reference gene was SlCAC.

were not significant changes of expression. However, *VTE6* was highly expressed and its expression increased throughout the fruit maturation. A similar observation was seen in M82 fruit for *VTE6* (figure 4-3), which also showed increased expression during tomato fruit development and ripening.

4.3.5 Differential expression of the MEP and SK pathway genes during *S. pennellii* fruit development

The MEP pathway genes were highly expressed throughout fruit maturation of *S.pennellii*, particularly; *ISPE* and *GGPS(2)* (figure 4-8). However, *ISPF, IPI(1), IPI(2)* were down regulated during *S.pennellii* maturation, which contrasts with their expression in the pericarp of M82 (figure 4-4). Surprisingly, there was a decrease in gene expression in stage 4 fruit, including genes such as; *ISPF, IPI(1), IPI(2), GGPS(1), GGPS(3).* This was similar to the reduced expression pattern at B+3 of several genes in M82 pericarp. This suggests that stage 4 of fruit of *S.pennellii* could be the equivalent period with transitions in for expression as B+3 in M82. *DXS(1), DXR, HDS*, *HST* and *GGDR* expression displayed the same expression profile during fruit development in *S.pennellii* as a M82 fruit.

Expression of genes in the SK pathway showed that genes encoding enzymes operating earlier in the pathway were induced during the fruit maturation (figure 4-9). However, later genes in the SK pathway displayed reduced expression during fruit maturation. The earlier pathway genes; *DAHPS1, DAHPS2, DAHPS3, SK* and *EPSPS1,* were induced, which was similar to the epidermal gene expression profile, of the SK pathway, in M82. Differential expression of *SDH-DHQS2* was observed in *S.pennellii* compared to the expression of this gene in M82 for the tissues of the pericarp and epidermis. Genes such as; *CS2, CM1, CM2, PAT, TYRA1, HPPD1* and *HPPD2,* showed increased expression in stages 1-3 of *S.pennellii* fruit expression, which was then reduced in the later stages (4-6) of *S.pennellii* fruit maturation. *HPPD1* and *HPPD2* expression showed opposing patterns to M82 fruit for the tissues of the pericarp and the epidermis.

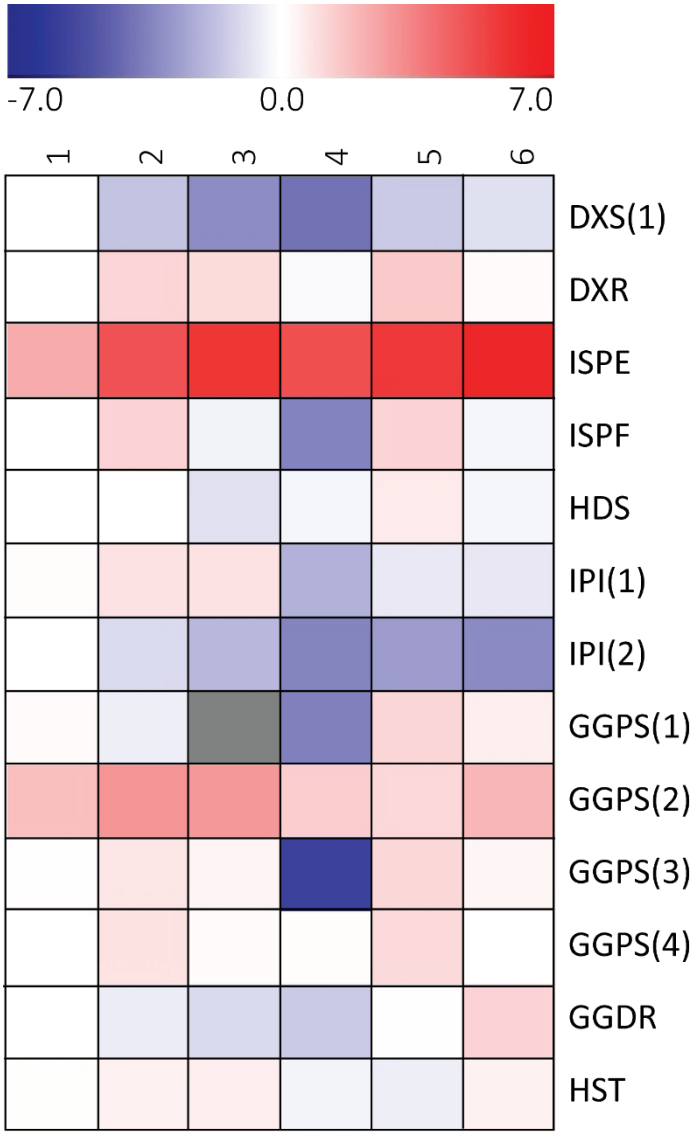


Figure 4-8 Heatmap showing log_2 of relative fold changes of MEP pathway gene expression for six stages of S.pennellii fruit development (n=3). The scale bar shows log_2 changes from -7.0 to 7.0. The relative expression values for each gene is normalised to stage 1 of tomato fruit development and the SlCAC reference gene was used. Grey boxes indicate where expression was not determined.

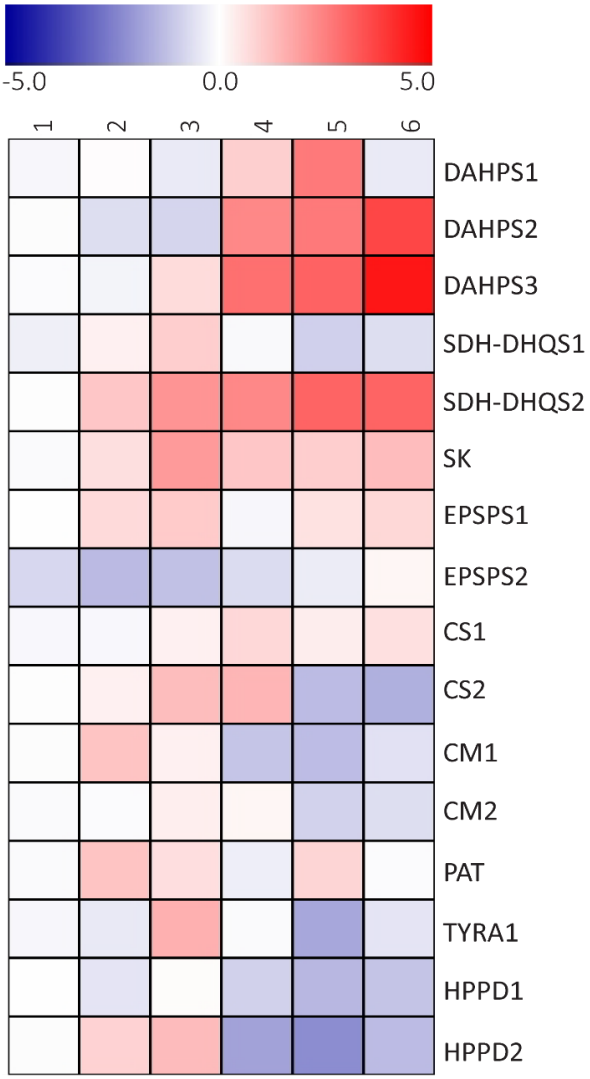


Figure 4-9 Heatmap showing log_2 of relative fold changes of the SK pathway gene expression for six stages of *S.pennellii* fruit development (n=3). The scale bar shows log_2 changes from -5.0 to 5.0. The relative expression values for each gene is normalised to stage 1 of tomato fruit development and the SlCAC reference gene was used.

4.3.6 *S.pennellii* fruit contained more tocochromanols than *S.lycopersicum* cv. M82

The metabolite profile for the *S.pennellii* tomato fruit growth, showed that tocopherols were not the most abundant form of tocopherol, but rather, tocotrienols were most abundant (figure 4-10). Total levels of tocochromanols (tocopherols and tocotrienols), reached a total of 10.75 µg g^{-1} in stage 4 of *S.pennellii* fruit. This was within the range of total levels of tocopherols detected in M82 fruit (figure 4-10). These data differed from the data for M82, as not only were tocotrienols present, but, alpha tocopherol, or, alpha tocotrienol were not the most abundant vitamers within the *S.pennellii* fruit as they are in M82 fruit. Instead, β-tocopherol and β-tocotrienol were the most abundant vitamers in *S.pennellii* fruit. γ-tocopherol and γ-tocotrienol, the precursors of alpha tocopherol and alpha tocotrienol, were the second most abundant vitamers in *S.pennellii* fruit, and δ-tocochromanols were the least abundant.

Total tocopherol levels were significantly different between stage 1 and 2 (1.43 and 2.04 µg g^{-1}, respectively) and the final stage 6 fruit (3.74 µg g^{-1}) (figure 4-10). However, this was not the case for tocotrienols, which showed no significant differences in composition or content during *S.pennellii* fruit development (figure 4-10). Total levels of tocopherols and tocotrienols contents did not alter substantially throughout *S.pennellii* fruit maturation. γ-tocopherol was the only form that altered significantly during *S.pennellii* fruit development (table 4-4 in appendix), between the stages 1-2 and 4-5, compared to stage 6.

4.4 Discussion

4.4.1 Transcriptional regulation of the VTE pathway is essential to maintain VTE levels in *S.lycopersicum* cv. M82 tomato fruit

Transcript abundance of VTE pathway genes remained constant in *S.lycopersicum* cv. M82 in the pericarp of fruit. *VTE2, VTE1, VTE4, VTE5* and *VTE3(1)* expression was slightly decreased until the B+3 stage, but expression was increased in B+5 and B+10, which suggested that co-regulation of these genes

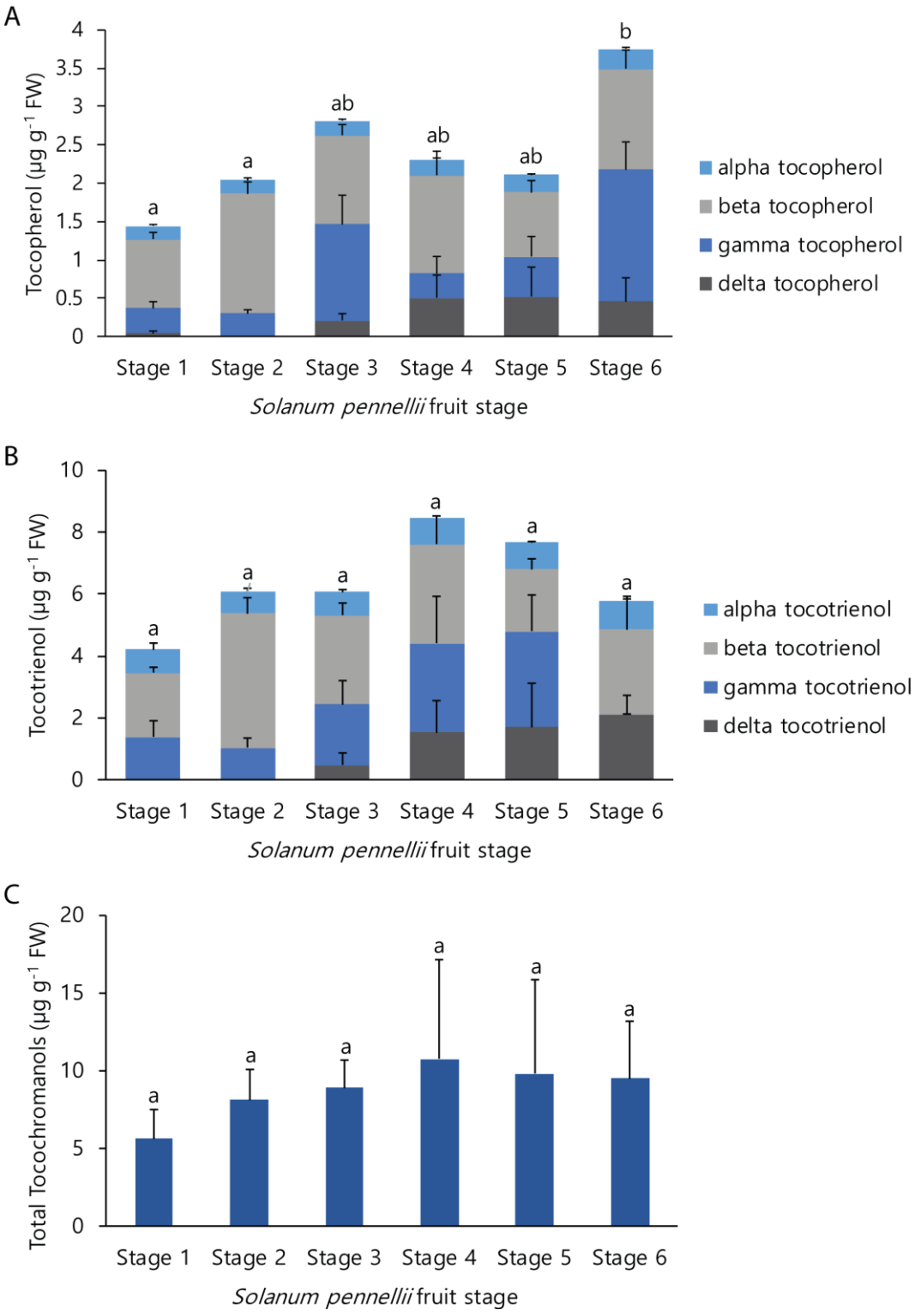

A
Tocopherol (µg g-1 FW)
4
3.5
3
2.5
2
1.5
1
0.5
0
a
a
ab
ab
ab
b
Stage 1
Stage 2
Stage 3
Stage 4
Stage 5
Stage 6
Solanum pennellii fruit stage
alpha tocopherol
beta tocopherol
gamma tocopherol
delta tocopherol
B
Tocotrienol (µg g-1 FW)
10
8
6
4
2
0
a
a
a
a
a
a
Stage 1
Stage 2
Stage 3
Stage 4
Stage 5
Stage 6
Solanum pennellii fruit stage
alpha tocotrienol
beta tocotrienol
gamma tocotrienol
delta tocotrienol
C
Total Tocochromanols (µg g-1 FW)
20
15
10
5
0
a
a
a
a
a
a
Stage 1
Stage 2
Stage 3
Stage 4
Stage 5
Stage 6
Solanum pennellii fruit stage

Figure 4-10 Tocopherol (A), tocotrienol (B) and total tocochromanol (C) contents of the different stages of *S.pennellii* fruit (μg g^{-1} fresh weight). A shows different forms of tocopherol in the different fruit stages, alpha (α) tocopherol is light blue, beta (β) tocopherol is light grey, gamma (γ) tocopherol is dark blue and delta (δ) tocopherol is dark grey. B shows the tocotrienol content in the different fruit stages, α-tocotrienol is light blue, β-tocotrienol is light grey, γ-tocotrienol is dark blue and δ-tocotrienol is dark grey. B shows the tocotrienol content in the fruit stages and the colours. C shows the total tocochromanol content (tocopherols and tocotrienols) for each fruit stage. The error bars show standard error of the mean. The letters indicate the statistical significance ($P<0.05$) of total tocopherol, tocotrienol and tocochromanols, which were calculated using Tukey tests. Tukey tests completed for the different forms of tocopherol are shown in appendices table 4-3. For each stage (n=4), however, each biological replicate consists of a different total number of fruit for each stage (appendix table 4-1). There were no significant changes in compositional tocotrienol content (data not shown).

likely occurs. Although, the genes did not show very large changes in expression, the co-expression of these genes may suggest more nuanced analysis is required to identify transcriptional regulators of this pathway. The changes in relative expression of the VTE pathway genes was reflected in the metabolite data, which showed that tocopherol levels were constant throughout the time course, consistent with literature (Quadrana et al., 2013). All tocopherol forms were detected in tomato fruit in *S.lycopersicum* cv. M82, and, alpha tocopherol was the most abundant, confirming other reports for tomatoes (Abushita et al., 1997, Quadrana et al., 2013).

VTE6 showed an opposing expression profile to the rest of the VTE pathway genes. However, the increase observed of *VTE6* expression occurred concomitant with the degradation of chlorophyll from breaker through to red ripe during tomato fruit ripening. During the chloroplast to chromoplast transition in tomato fruit ripening, chlorophyll is degraded and carotenoids are accumulated (Harris and Spurr, 1969). It is known that the degradation of chlorophyll is needed to generate free phytol for PDP, which is the precursor for tocopherol biosynthesis (Valentin et al., 2006, vom Dorp et al., 2015, Almeida et al., 2016). The *VTE6* enzyme had remained elusive until recently where its activity was demonstrated in *Arabidopsis vte6* knock out lines, in which, tocopherols were abolished (vom Dorp et al., 2015). It is evident that expression of this enzyme is important during *S.lycopersicum* cv. M82 development and ripening of fruit, and its activity may be responsible for the maintenance of VTE levels in ripe fruit. Previous studies have suggested that tomato fruit does not accumulate tocopherols in the late ripening stages due to a lack of availability of the precursor; PDP (Quadrana et al., 2013, Almeida et al., 2015). However, my expression data suggested that although *VTE5* expression decreased, *VTE6* expression increased during fruit ripening. Therefore, balance of transcript abundance between these two genes may maintain tocopherol levels in tomato fruit. *VTE5* could be considered a necessary step in tocopherol biosynthesis, as RNAi lines in tomato show reduced tocopherol levels by 80% (Almeida et al., 2016) and *Arabidopsis vte5* knockouts also have reduced

tocopherols (Valentin et al., 2006). Perhaps, the induction of *VTE6* expression results in an intricate regulation network to maintain VTE biosynthesis by supplementing *VTE5* expression. This is reinforced in the literature which showed that *Arabidopsis vte6* mutant is unable to grow photoautotrophically (vom Dorp et al., 2015). This is probably due to the accumulation of phytyl phosphate, which might impact negatively on chloroplast growth and development. Double *Arabidopsis* mutants of *vte5* and *vte6* were able to complement this phenotype, and mutants were able to grow photoautotrophically and displayed a stay green phenotype (vom Dorp et al., 2015). These data provide new insights into the transcriptional regulation of the VTE pathway.

Differential expression was also observed between the pericarp and epidermis for the VTE pathway genes. Several VTE genes were highly expressed in the epidermis, compared to the pericarp at numerous developmental stages. The expression data is reflected in the metabolite data, which showed that tocopherol accumulation doubled in the epidermis, compared to the pericarp, during ripening. The B+3 stage of tomato fruit showed the greatest differences in expression and metabolite contents between these two tissues. Therefore, this may be a key transition stage for transcriptional regulation of VTE production. This implies that transcriptional regulation between these two tissues could be different. Lu et al (2013) overexpressed *VTE2,* under the control of the RNA operon promoter from tobacco in transgenic tomato lines. The analysis of leaves and fruit of these lines indicated that tissue-specific regulation of biosynthetic gene expression exists, as more tocochromanols accumulated in leaves compared to fruit (Lu et al., 2013). However, these are different organs, and, a recent study showed that in tomato fruit, less than 3% of genes expressed were tissue specific (Matas et al., 2011). However, this generic data can mask specific differences – in this case in expression between tissues.

VTE pathway gene expression was constant throughout *S.lycopersicum* cv. M82 tomato fruit development and ripening. However, the induced expression of *VTE6* has never before been reported, which demonstrated that this gene may be

important to maintain VTE levels in the pericarp. The VTE pathway gene expression was generally higher in the epidermis, compared to the pericarp, which might result in higher VTE levels in the epidermis.

4.4.2 MEP pathway gene expression could modulate VTE levels in *S.lycopersicum* cv. M82 fruit

The MEP pathway gene expression profiles demonstrated that there is a transcriptional shift associated with fruit ripening. Genes within this pathway showed substantial changes in expression, for both pericarp and epidermal tissues. However, expression of MEP pathway genes was higher for the epidermal tissue. This difference is well documented in the literature (Bramley, 2002, Carrari et al., 2006, Alba et al., 2005, Ye et al., 2015, Shinozaki et al., 2018), as the MEP pathway is key for several specialised metabolite pathways, including; carotenoids, abscisic acid, gibberellins, chlorophyll and terpenoid biosynthesis. Therefore, ripening is associated with a vast accumulation of carotenoids, and this transcriptional data suggests it is regulated transcriptionally. Most interestingly, several MEP pathway genes *(DXS(1), ISPE, ISPF, HDS, IPI(1), GGPS(2)* and *GGPS(3))* were induced at B+3, which contrasts with the expression of genes in the VTE pathway, which were generally repressed. The induction of MEP pathway genes in the epidermis, may cause an increase of substrates for VTE synthesis, and consequently, an increase of VTE levels in the tissue of the epidermis. This suggests that there may be crosstalk between the MEP and VTE pathways, possibly involving inducted flux throughout the MEP pathway during ripening, which may trigger the small increases in VTE pathway gene expression. It is known that at the B stage, *PSY1* (which is the first committed step in carotenoid biosynthesis), is expressed (Fraser et al., 1994), resulting in the accumulation of carotenoids. Therefore, the repression of VTE gene expression at B+3 could be a downstream consequence of carotenoid accumulation, as carotenoids 'pull' on the flux of the MEP pathway. This is supported by observations on *PSY1* expression in transgenic *PSY1* overexpression MG fruit, which is associated with decreases of both alpha

and γ-tocopherol contents (Fraser et al., 2007). This implies that transcriptional control is important for VTE biosynthesis, but also, post-translational mechanisms, such as feedback loops, could be important for cross talk between the carotenoid and VTE pathways.

4.4.3 Expression of SK pathway genes was induced early in *S.lycopersicum* cv. M82 fruit

The SK pathway showed similarities with the VTE pathway, as many genes were induced during fruit ripening and development. However, several genes were induced earlier in development, rather than at B+3. Remarkably, at B+3 there was a reduction of expression of numerous SK pathway genes, including; *SK, EPSPS, CS1, CS2, PAT, TAT2, TYRA2* and *HPPD1.* Thus, the B+3 stage appears to be an important stage in transcriptional regulation. This may be a downstream effect of the B stage, which is the transition from mature green fruit to a ripe tomato. *TYRA* and *HPPD* are known to be important regulatory points for the shikimate pathway (Rippert et al., 2004, Zhang et al., 2013). The expression of these two genes was higher in epidermis, compared to pericarp. This provides an indication of why the tocopherol levels were higher in the epidermis, compared to pericarp. The increase in the expression of these genes could result in higher homogentisate pools available for VTE biosynthesis.

4.4.4 Expression of VTE, MEP and SK pathway genes in *S.pennellii* fruit showed smaller changes during fruit development than the same genes in *S.lycopersicum* cv. M82

Expression of VTE pathway genes did not change significantly during tomato fruit growth in *S.pennellii*, but, there were small fluctuations in expression. The exception to this was *VTE6,* which was more was highly expressed in *S.pennellii* throughout fruit growth. In contrast with the *S.lycopersicum* cv. M82 metabolite profile of the pericarp, which showed tocopherols remaining at similar levels throughout the development and ripening stages of fruit, the metabolite profile of the *S.pennellii* fruit showed that tocopherols accumulated cumulatively

during fruit development. Thus, the high expression of *VTE6* demonstrated that this gene may be important for increasing tocopherol levels in *S.pennellii,* rather than maintaining tocopherol levels in *S.lycopersicum* cv. M82.

Several genes in the MEP pathway were down regulated during growth of *S.pennellii* fruit*,* rather than increased. This highlighted the role of *VTE6* in increasing tocopherol levels by regenerating PDP pools. As *S.pennellii* fruit remains green and does not ripen, this suggested that chlorophyll is abundantly available for PDP regeneration from chlorophyll degradation in *S.pennellii* fruit. It is possible that if my time course for sampling *S.pennellii* fruit had included more stages, the tocopherol content might have increased further.

Oxidative stress is associated with ripening in tomato fruit (Jimenez et al., 2002) and tocopherols are potent antioxidants (Schneider, 2005). Therefore, in *S.lycopersicum* cv. M82 fruit, tocopherols might not increase during fruit development and ripening because they might be used as antioxidants. There may be more tocopherol radicals in these fruit, which are not measured in this data. *S.pennellii* fruit do not ripen, therefore tocochromanols may accumulate due to an abundance of chlorophyll for PDP regeneration.

The MEP expression profile in *S.pennellii* fruit was the opposite to the *S.lycopersicum* cv. M82 expression profile, which the latter showed upregulation of the MEP genes during fruit ripening. The increase in MEP expression in *S.lycopersicum* cv. M82 fruit reflects the larger increases in carotenoids, associated with ripening. This is further reinforced by the expression of *DXS(2),* which is an important enzyme for MEP biosynthesis. Expression of this gene in transgenic tomato plants has shown that MEP pathway products were significantly altered, but, tocopherols were not hyperaccumulated (Enfissi et al., 2005). Although *DXS(2)* was not measured here, this is something to consider for future experiments and would elucidate whether the *VTE6* is truly responsible for the higher tocopherol levels by modulating PDP pools in *S.pennellii*, or, whether it is a down-stream effect of increased MEP flux into the VTE pathway. It is likely that *VTE6* expression is highly upregulated to maintain PDP pools in *S.pennellii.* This is

reinforced by the increased levels of tocotrienols, as this would suggest there is more GGDP, therefore the ratio of GGDP:PDP is skewed, resulting in the production of tocotrienols (Zhang et al., 2013).

The SK pathway showed a variable expression profile and the earlier SK pathway genes were upregulated in the later stages 4-6 of fruit development in *S.pennellii*. In contrast, the later SK pathway genes were down regulated in the late stages 4-6. *SDH-DHQS2* expression was different between the two tomato species. This enzyme is known to be more active in *S.pennellii* compared to *S.lycopersicum* fruit (Steinhauser 2010). It is known that several enzymes in primary metabolism are down regulated in *S.lycopersicum* cultivars during ripening, whereas in *S.pennellii* enzyme activities remain stable throughout fruit maturation (Steinhauser et al., 2010). This suggests that primary metabolism changes during fruit development and the SK pathway may be down regulated. Shikimate is known to accumulate to much higher levels in *S.pennellii* fruit compared to fruit of *S.lycopersicum* cultivars (Schauer et al., 2006, Steinhauser et al., 2010). Therefore, the increase in shikimate observed in *S.pennellii* could increase flux throughout the pathway and increase tocochromanol levels in this species via an increase in available homogentisate. This might explain why the genes encoding the late steps of the SK pathway were not upregulated in the late stages of *S.pennellii* fruit development, as there is tight control of tyrosine levels, which ultimately determine the homogentisate pools for tocochromanol biosynthesis (Rippert et al., 2004, Tzin and Galili, 2010a).

It is clear that in *S.pennellii* fruit, each pathway shows differential expression, but also a level of co-regulation. Therefore, it is likely that transcriptional regulators independently control the activity of each pathway. The co-regulation between these pathways may be integrated at the branch points, through post-translational control of pathway flux, responding to levels of GGDP, PDP and homogentisate. Thus, understanding of transcriptional regulation in *S.pennellii* will be complex and likely different from those in *S.lycopersicum* cv. M82, as *S.pennellii* fruit showed different VTE metabolite profiles. This may reflect

major differences in metabolism between red (non-photosynthetic) and green (photosynthetic) fruit.

4.4.5 *S.pennellii* fruit contains a type of tocochromanol that is not normally found in tomato

Tomato species do not contain the homogentisate geranylgeranyl transferase (HGGT) enzyme that is required for tocotrienol biosynthesis. This gene is predominately found in monocot species. Therefore, the presence of tocotrienols in *S.pennellii* was surprising, and they were the most dominant form of tocochromanols in *S.pennellii* fruit. The incidence of high shikimate levels in *S.pennellii* tomato fruit (Steinhauser et al., 2010) suggest that there could be an increase in pathway flux, which resulted in more tocopherols and tocotrienols. Many enzymes involved in primary metabolism in *S.pennellii* are known to increase in activity during tomato maturation, which is in contrast with the transcript abundance of several enzymes in *S.lycopersicum,* which fall during ripening (Steinhauser et al., 2010)). It has been shown that when PDP pools are low, but GGDP is available, VTE2 can use GGDP to produce tocotrienols. (Yang et al., 2011, Zhang et al., 2013). Thus, even though *S.pennellii* does not contain HGGT, retains the potential to produce tocotrienols using VTE2. However, over-expression of *VTE2* alone, does not result in the production of tocotrienols in tomato or other plants (Savidge et al., 2002, Seo et al., 2011).

Homogentisate production is also a limiting factor in tocotrienol biosynthesis (Falk et al., 2003, Rippert et al., 2004). Homogentisate production may be limited by tyrosine, as tyrosine internally inhibits *CM* and *TyrA* (Tzin and Galili, 2010a, Tzin and Galili, 2010b) These are enzymes which provide precursors for tyrosine and homogentisate biosynthesis (figure 4-1). Ten-fold increases of tocochromanols have been achieved by over-expressing a yeast prephenate dehydrogenase (*TYRA1*) and *AtHPPD* in tobacco, and of these the majority of the tocochromanols were tocotrienols (Rippert et al., 2004). Therefore, this has provided promising gene targets to bypass the internal regulation of the SK

pathway to achieve elevated tocotrienol levels. Indeed, combined over-expression of two enzymes; *TyrA* and *HPPD,* in *Arabidopsis* showed that homogentisate is the factor limiting tocotrienol production, as approximately 50% of total tocochromanols were tocotrienols (Zhang et al., 2013). Within these *Arabidopsis* lines, *DXS* and *DXR* were not upregulated (Zhang et al., 2013) . These genes encode key steps for regulation of the MEP pathway, and therefore, they can alter PDP and GGDP pools (Cordoba et al., 2009). De-regulated homogentisate synthesis is an alternative route for tocotrienol synthesis. However, the fruit of *S.pennellii* showed that *TyrA* and *HPPD* were upregulated in the early stages of *S.pennellii* fruit, before they were downregulated in the later fruit stages, 4-6. Although, expression was not consistent throughout *S.pennellii* fruit maturation, the initial induction of these genes could be large enough to increase homogentisate production. Additionally, the dominant tocotrienol composition of *S.pennellii* fruit was also observed in the transgenic *Arabidopsis* lines overexpressing *AtTyrA* and *AtHPPD*, which contained a large proportion of tocotrienols compared to tocopherols (Zhang et al., 2013). Within *S.pennellii* fruit, it is not clear whether there was an increase in GGDP pools, via an induction of *DXS(2)*, or whether there are alternative tocotrienol synthesis routes. These data provided new insights into tocotrienol synthesis in *S.pennellii* tomato fruit which had not previously been identified.

The green fruited *S.pennellii* fruit may contain more tocochromanols than *S.lycopersicum* cv. M82 fruit, as it still a photosynthetic tissue. Chlorophyll degradation is associated with ripening, and the transition between chloroplast and chromoplast in cultivated tomato species (Harris and Spurr, 1969). This does not occur in the *S.pennellii* fruit and the fruit remains green. The process of photosynthesis is concomitant with an increase in hydrogen peroxide (H_2O_2) and singlet oxygen (O_2^-), also known as reactive oxygen species (ROS). Tocochromanols are produced in chloroplasts and they function as preventative antioxidants against lipid peroxidation (Krieger-Liszkay and Trebst, 2006). Therefore, *S.pennellii* fruit tocotrienol biosynthesis might be a result of the high levels of ROS in the

photosynthetically active *S.pennellii* fruit. Shalata and Tal (1998) showed that less lipid peroxidation is observed in salt-stressed *S.pennellii* leaves, compared to salt-stressed *S.lycopersicum* leaves. Several antioxidant enzymes were upregulated in the salt-stressed *S.pennellii* leaves, including; ascorbate peroxidase (APX), superoxide dismutase (SOD) and glutathione reductase (GR) (Shalata and Tal, 1998). These antioxidant enzymes are associated with the ascorbate-glutathione cycle to maintain ROS homeostasis in plants (Smirnoff, 2000), but are also known to reduce the tocopherol radical back to tocopherol to maintain further oxidative stresses (Munne-Bosch and Alegre, 2002). Therefore, tocotrienols might be produced in *S.pennellii* leaves as well, to reduce lipid peroxidation. However, the production of tocotrienols in *S.pennellii* fruit is likely to be a consequence of photosynthesis.

The fact that *S.pennellii* fruit contained tocotrienols was surprising. However, it was not clear whether tocotrienol abundance was due to flux within the MEP pathway, which could cause fluctuations in GGDP:PDP pools sizes, and therefore, affect the substrate available for VTE2 to use for tocochromanol production. The tocotrienols could also be produced via alternative homogentisate biosynthetic routes, which might involve the de-regulation of homogentisate to increase homogentisate pools. Either option could be true, or both could play roles in altering pools sizes of substrates for the biosynthesis of tocotrienols. This requires further experiments on *S.pennellii* to determine the genes responsible for the production of tocotrienols – these are described in more detail in chapter 7.

4.4.6 *S.lycopersicum* cv. M82 showed differential expression of VTE biosynthetic genes and different metabolite profiles to *S.pennellii* fruit

In conclusion, *S.lycopersicum* cv. M82 showed rather constant expression profiles for the VTE pathway genes during fruit development. Genes of the MEP and SK pathways were much more highly expressed in *S.pennellii* fruit in comparison to *S.lycopersicum* cv. M82 fruit. These differences are likely to affect

GGDP:PDP and homogentisate pools, and therefore, determine the type of tocochromanol produced in fruit tissues. The expression profiles of both tomato species demonstrated that cross-talk is likely to play a significant role in the regulation of the VTE pathway. The co-expression of the genes encoding enzymes of this pathway suggested that co-regulation might exist, and transcriptional regulators are likely responsible for the maintenance of VTE levels in tomato fruit.

Chapter 5: Characterisation of SlMYB79 from the trans-eQTL IL9-3-2

Chapter 5: Characterisation of SlMYB79 from the trans-eQTL IL9-3-2

5.1 Introduction

5.1.1 Transcriptional regulation of the VTE, MEP and SK pathways

Very little is known about the transcriptional regulation of Vitamin E (VTE) biosynthesis in tomato or in any other plant species. There have been studies that show differential expression profiles of the genes encoding enzymes of VTE biosynthesis during tomato development, which suggest that transcription factors (TFs) regulate VTE biosynthesis (Quadrana et al., 2013). Promoter analysis of the genes encoding enzymes of the VTE, methyl-erythritol phosphate (MEP) and shikimate (SK) pathways have demonstrated common TF binding motifs (Quadrana et al., 2013). The most common binding motifs found are bound by MYB TFs and bZIP TFs in plants. Therefore, I selected candidates that were based on this prediction during the gene mining of trans-eQTLs affecting VTE biosynthetic genes.

The MEP and SK pathways may share common transcriptional regulators with the VTE pathway as suggested by promoter analyses (Quadrana et al., 2013). Studies have shown that TFs can regulate genes encoding enzymes in the MEP and SK pathway transcriptionally, which can alter VTE content (Verdonk et al., 2005, Enfissi et al., 2010, Lira et al., 2017). This chapter aimed to identify transcriptional regulators of VTE biosynthesis in the trans-eQTL IL9-3-2.

5.1.2 The trans-eQTL analyses revealed two trans-eQTL for candidate gene screening

The trans-eQTL screen (described in chapter 3) elucidated two candidate regions (trans-eQTL IL6-2-2 and trans-eQTL IL9-3-2) from the RNA sequencing data of the *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression lines (ILs) that might contain genes regulating VTE biosynthesis transcriptionally. Genes encoding several candidate TFs were identified in the trans-eQTL IL9-3-2. This trans-eQTL was also identified as a trans-eQTL for lycopene biosynthesis (Li, 2018). I screened the candidate genes encoding TFs in this region using transient silencing

assays to identify transcriptional regulators of VTE biosynthesis. I overexpressed one candidate TF and created CRISPR lines to determine whether this gene regulates VTE biosynthesis transcriptionally.

5.2 Materials and Methods

5.2.1 Candidate gene mining of the trans-eQTL IL9-3-2

The trans-eQTL IL9-3-2 IL region was mined for genes encoding candidate TFs regulating the genes encoding enzymes in the VTE pathway. I used co-expression analysis of the genes encoding the enzymes in the VTE pathway with the expression of candidate TF genes to reduce the number of TFs taken forward for further analysis. A description of how the trans-eQTL IL9-3-2 was identified is provided in chapter 3.

5.2.2 Plant Materials

Microtom tomato plants (T1 generation) expressing *SlMYB79*, under the control of the fruit specific E8 promoter were used for qRT-PCR, trolox equivalent antioxidant capacity (TEAC) assays and metabolic analysis. Fruit at four stages of development were harvested; mature green (MG), breaker (B), breaker + 5 days (B+5) and breaker + 10 days (B+10), seeds were removed, and the pericarp and the epidermis were harvested together. The B+10 tomato fruits were used for TEAC assays.

Moneymaker was used for gene editing of SlMYB79 by CRISPR/Cas9. T1 Moneymaker CRISPR tomatoes were harvested at B+10 stage. The seeds were removed, and the pericarp and the epidermis were harvested together. The tomatoes were used for qRT-PCR and metabolic analysis.

T0 Moneymaker E8:SlMYB71 tomatoes were harvested at B+10 and analysed using metabolite and qRT-PCR analysis. These tomatoes were harvested at B+10, seeds were excised, and the pericarp and epidermis were harvested together.

All tomatoes analysed in this chapter were frozen in liquid nitrogen and ground to a powder. VTE metabolite and qRT-PCR analysis are described in Chapter 2. A full list of primers used in this chapter are available in the appendix.

5.2.3 Viral Induced Gene Silencing (VIGS)

DNA fragments for fruit VIGS of genes of interest (GOI) were cloned into the pTRV2 *Del/Ros* (cloning methods are described in Chapter 2). Once these pTRV2 *Del/Ros* VIGS plasmids were made, they were transformed into *Agrobacterium tumefaciens* strain Agl1. Single colonies were cultured for two days at 28°C. The cultures were washed using Agroinfiltration media (10mM MES, 10mM $MgCl_2$, 200µM acetosyringone, pH5.6) and resuspended to an OD_{600}=0.25 for Moneymaker fruit and OD_{600}=0.1 for Microtom fruit. The tomato fruits were injected at the MG fruit stage and were harvested fourteen days post B. All further RNA extractions, qRT-PCR and metabolite analysis are described in chapter 2.

5.2.4 Plasmid construction for stable transformation for over expression

The SlMYB79 (Solyc09g090790) and SlMYB71 (Solyc05g053150) coding sequences (CDS) were amplified from cDNA using Phusion PCR with primers carrying Gateway® compatible sequences for recombination with Gateway® compatible vectors (Gateway® cloning is described in chapter 2). The CDS with additional Gateway® adapters were recombined into the entry vector pDONR207, using a Gateway® BP reaction. This plasmid was then recombined using the Gateway® LR reaction into pBIN19 E8:GW to generate pBIN19E8:SlMYB79 and pBIN19 E8:SlMYB71. The E8 promoter is a fruit-specific, ethylene induced promoter from tomato fruit (Zhao et al., 2009).

These constructs were transformed into *E.coli* DH5α and extracted using a Qiagen miniprep kit. The construct was transformed into *Agrobacterium tumefaciens* strain Agl1 for tomato transformation. The transformation procedure is described in chapter 2.

5.2.5 Plasmid construction for CRISPR/Cas9 gene editing

Guide RNAs (gRNAs) were designed to target a 20nt target sequence with a -NGG protospacer adjacent motif (PAM) of the genomic SlMYB79 DNA sequence. gRNAs were designed using CRISPR direct (Naito et al., 2015) and were selected based on the least off target sequence matches. These gRNAs were ordered as oligos and recombined into a Golden Gate Level 1 acceptor plasmid, under the control of a U6 promoter by digesting and ligating (Dig-Lig) in the same reaction (described in chapter 2). These Level 1 constructs were recombined into a Level 2 construct containing domesticated Cas9, under the control of a double Cauliflower mosaic virus (CaMV) 35S promoter, using a Dig-Lig reaction for transformation into *E.coli.*

These constructs were propagated in *E.coli* and extracted using a Qiagen miniprep kit. The construct for SlMYB79 editing was transformed into *Agrobacterium tumefaciens* strain Agl1 for tomato transformation. The transformation protocol is described in chapter 2.

5.2.6 Trolox equivalent antioxidant capacity assay

A TEAC assay was adapted from the TEAC assay described by Pellegrini et al (2003). An aqueous solution of 7mmol L^{-1} of ABTS (2,2'-Azino-bis (3-ethylbenzothiazoline-6-sulfonic acid)) with 2.45 mmol L^{-1} potassium persulfate was incubated in the dark at room temperature, overnight. A solution of 50nmol L^{-1} Trolox (6-Hydroxy-2,5,7,8-tetramethylchromane-2-carboxylic acid) was prepared for the standard curve, which was diluted 1:10. This stock was then used to generate a series of 1:2 dilutions for the standard curve. The ABTS solution was diluted to an OD734=0.7±0.02 with ethanol and mixed with the Trolox for the standard curve.

Approximately 100mg of fruit powder was extracted using 600µl of water for the hydrophilic antioxidant fraction. The same sample was then re-extracted with 600 µl of acetone for the lipophilic antioxidant extraction. ABTS solution (1ml) was added to 10µl of extract and incubated for 5 minutes at room temperature.

The OD_{734} of the sample fraction was measured and the results were expressed as TEAC (Trolox equivalent antioxidant capacity) mmol of Trolox per kg fruit (fresh weight) for each sample.

5.2.7 Analysis of candidate R2R3 MYB transcription factors

The method for qRT-PCR of gene expression over the time course of fruit development has been described in chapter 2. The tomatoes used in this analysis were the same as used in chapter 4.

Protein sequences of the candidate genes were used to create a phylogenetic tree against the *Arabidopsis thaliana* R2R3MYB TFs, using IT3F software (Bailey et al., 2008, Bailey et al., 2012). The protein sequences of paralogs of the candidate TFs as well as orthologs from other species were aligned using clustal omega (Sievers et al., 2011) and the promoter regions of these genes were aligned using clustal software (Larkin et al., 2007), and analysed based on TF motifs (Franco-Zorrilla et al., 2014).

5.3 Results

5.3.1 Screening TFs in the trans-eQTL IL9-3-2 interval provided one candidate for further analysis

Gene mining of the trans-eQTL IL9-3-2 showed that there were eleven candidate genes encoding candidate transcriptional regulators (described in chapter 3) (figure 5-1). This region had also been identified as a trans-eQTL for lycopene biosynthesis for tomato fruit (Li, 2018) and for the MEP pathway (chapter 3). The candidate TFs identified in the trans-eQTL IL9-3-2 were reduced to four candidate TFs based on their co-expression pattern with genes encoding enzymes of the VTE pathway (table 5-1). Two TFs (*SlbHLH92* and *SlMYB79)* were down-regulated in the three *S.pennellii* x *S.lycopersicum* cv.M82 ILs (IL9-3, IL9-3-1 and IL9-3-2), together with the down regulation of the expression of genes encoding enzymes of the VTE pathway (table 5-1). One TF candidate *(SlbHLH60)* was upregulated in the three ILs in the trans-eQTL 9-3-2, which is opposite to the general decrease in expression of the genes encoding enzymes of the VTE pathway

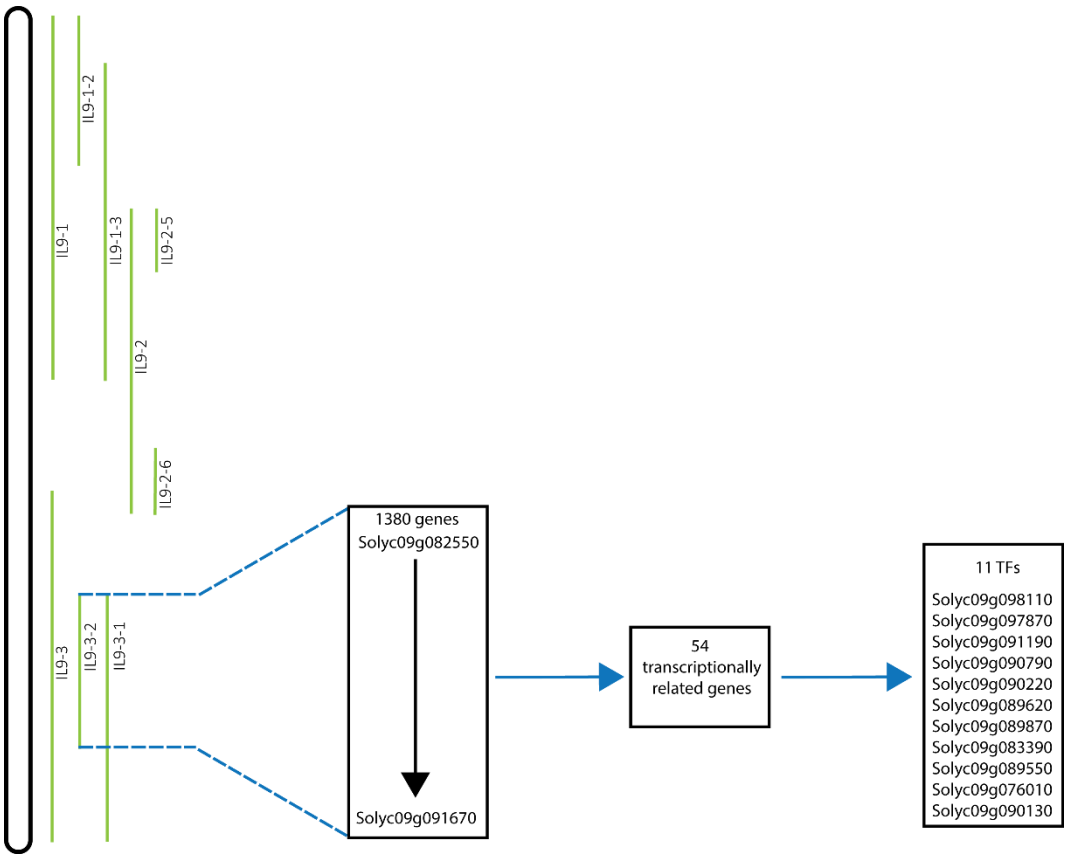


Figure 5-1 Schematic diagram of candidate gene mining for *S.pennellii* x *S.lycopersicum* cv. M82 IL9-3 subgroup. Eleven candidate transcription factors were identified, which showed co-expression patterns with genes encoding enzymes of the VTE pathway in the *S.pennellii* x *S.lycopersicum* cv. M82 fruit RNA sequencing data (Lee and Giovannoni).

Table 5-1 RNA sequencing data showing the relative expression values of genes in VTE biosynthesis in ILs; IL9-3, IL9-3-1 and IL9-3-2, compared to M82. The table shows the expression of genes encoding enzymes of the VTE pathway and candidate TFs that were identified in the three ILs in the trans-eQTL IL9-3.

Gene	**Solgenomics identifier**	***S.pennellii* x *S.lycopersicum* cv.M82 introgression line**			
		M82	**IL9-3**	**IL9-3-1**	**IL9-3-2**
	VTE genes				
VTE2	Solyc07g017770	1	0.87	0.75	0.68
VTE3(1)	Solyc09g065730	1	1.10	0.90	0.98
VTE3(2)	Solyc03g005230	1	0.76	0.69	0.65
VTE1	Solyc08g068570	1	0.87	0.77	0.72
VTE4	Solyc08g076360	1	0.71	1.15	0.99
VTE5	Solyc09g018510	1	0.72	0.56	0.67
VTE6	Solyc07g062180	1	0.92	1.21	1.04
	Candidate TF genes				
SlbHLH60	Solyc09g083220	1	1.21	1.45	1.46
SlbHLH61	Solyc09g097870	1	3.04	3.37	0.79
SlbHLH92	Solyc09g098110	1	0.43	0.27	0.56
SlMYB79	Solyc09g090790	1	0	0	0
	SlMYB79 paralog				
SlMYB71	Solyc05g053150	1	0	0	0

in IL 9-3, IL9-3-1 and IL9-3-2 (table 5-1). *SlbHLH61* was upregulated in IL9-3 and IL9-3-1, but down regulated in IL9-3-2 (table 5-1).

The trans-eQTL IL9-3-2 was compared to the equivalent the *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs (described in chapter 3) (figure 5-2). The equivalent trans-eQTLs to IL9-3-2 in *S.lycopersicoides* x *S.lycopersicum* cv. VF36 were ILs 4270A, 4270B and 4272 and all showed down regulation of genes encoding enzymes of the VTE pathway (table 5-2), which was also observed in the trans eQTL IL9-3-2 in *S.pennellii* x *S.lycopersicum* cv.M82 ILs (table 5-1). The candidate TFs identified in the trans-eQTL reside within the equivalent *S.lycopersicoides* x *S.lycopersicum* cv. VF36 IL 4272 and does not overlap with the other *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs in figure 5-2 and table 5-2. The relative expression of candidate genes encoding TFs in the *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs did not correlate well with the expression of the genes encoding enzymes of the VTE pathway (table 5-2). Therefore, I could not conclude that these candidate TFs are responsible for the expression profile of VTE biosynthetic genes observed in the *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs.

The candidate TFs identified in the IL9-3-2 region were silenced using a transient VIGS assay in *S.lycopersicum* cv. Microtom *Del/Ros* tomato fruit. Table 5-3 showed there was one TF candidate gene (SlMYB79) which showed significantly altered alpha and total tocopherol contents in fruit following VIGS silencing. Consequently, *SlMYB79* was silenced in *S.lycopersicum* cv. Moneymaker *Del/Ros* fruit, as the larger fruit were easier to dissect and there was a larger sample size.

Silencing of *SlMYB79* in Moneymaker *Del/Ros* fruit showed that alpha (α) tocopherol was 1.5-fold higher in pTRV2 *Del/Ros* SlMYB79 silenced sectors compared to pTRV2 *Del/Ros* silenced sectors (figure 5-3). Total tocopherol levels were also 1.6-fold higher in pTRV2 *Del/Ros* SlMYB79 silenced fruit compared to the control (figure 5-3). Expression of *VTE3(1)* and *VTE6* was upregulated in the *SlMYB79* silenced tomato fruit sectors (figure 5-4). Expression of *SlMYB79* was reduced, which indicated that the silencing was effective, although the reduction was not statistically significant compared to the control (figure 5-4). This indicated

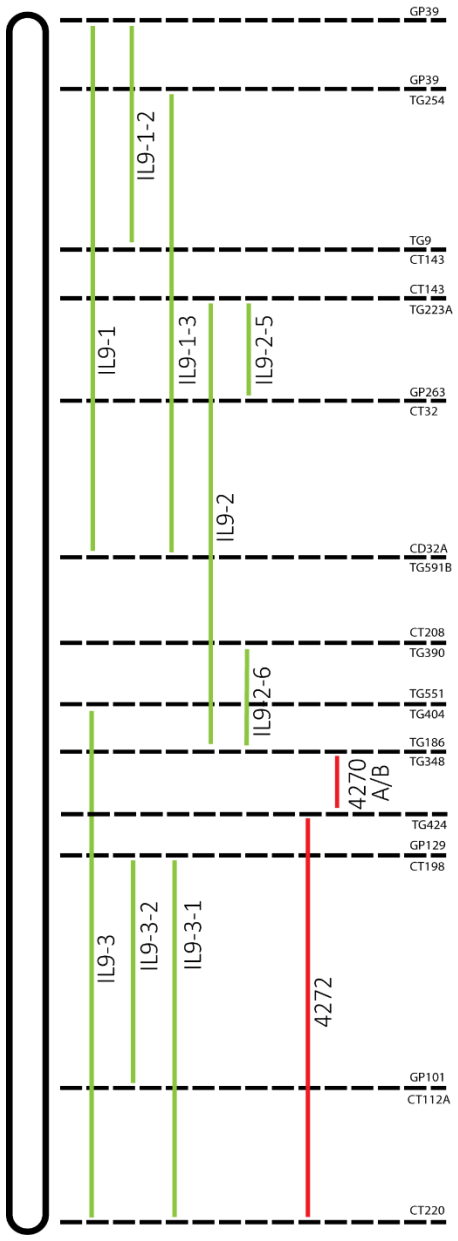


Figure 5-2 Schematic diagram showing *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines (ILs) and the equivalent *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs, corresponding to the trans-eQTL on chromosome 9. The green lines indicate where the *S.pennellii* DNA lies and the red lines show where the *S.lycopersicoides* DNA lies. At the end of each dashed line, the markers that were used to determine the chromosome position are shown.

Table 5-2 RNA sequencing data of *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs showing the relative expression values of genes involved in VTE biosynthesis compared to VF36 (IL parent). The table shows the expression of genes encoding enzymes of the VTE pathway and candidate TFs identified in the three ILs corresponding to the *S.pennellii* trans-eQTL IL9-3-2.

Gene name	**Solgenomics identifier**	**Ratio of expression of *S.lycopersicoides* x *S.lycopersicum* cv. VF36 ILs relative to IL parent (VF36)**			
		VF36	**4270A**	**4270B**	**4272**
VTE genes					
VTE2	Solyc07g017770	1	0.59	0.68	1.29
VTE3(1)	Solyc09g065730	1	1.14	0.82	1.09
VTE3(2)	Solyc03g005230	1	0.90	1.01	1.18
VTE1	Solyc08g068570	1	0.95	1.00	1.70
VTE4	Solyc08g076360	1	0.62	0.74	0.98
VTE5	Solyc09g018510	1	1.21	1.08	1.43
VTE6	Solyc07g062180	1	0.76	0.73	1.05
Candidate TF genes					
SlbHLH60	Solyc09g083220	1	1	1	1
SlbHLH61	Solyc09g097870	1	1	0.48	0.41
SlbHLH92	Solyc09g098110	1	1	0.60	0.02
SlMYB79	Solyc09g090790	1	1	0.80	1
SlMYB79 paralog					
SlMYB71	Solyc05g053150	1	1	1	1

Table 5-3 Summary of the VIGS of candidate genes in Microtom *Del/Ros* fruit. The changes in tocopherols were expressed relative to pTRV2 *Del/Ros* red silenced sectors (n=3). The stars represent statistical significant differences, (*) = $p<0.05$.

SlTF number	TF type	Solgenomics identifier	Statistically significance of tocopherol forms relative to pTRV2 *Del/Ros* red, silenced tomato sectors.				
			Alpha tocopherol	Beta tocopherol	Gamma tocopherol	Delta tocopherol	Total tocopherol
SlbHLH60	bHLH	Solyc09g083220	-	-	-	-	-
SlbHLH61	bHLH	Solyc09g097870	-	-	-	-	-
SlbHLH92	bHLH	Solyc09g098110	-	-	-	-	-
SlMYB79	MYB	Solyc09g090790	*	-	-	-	*

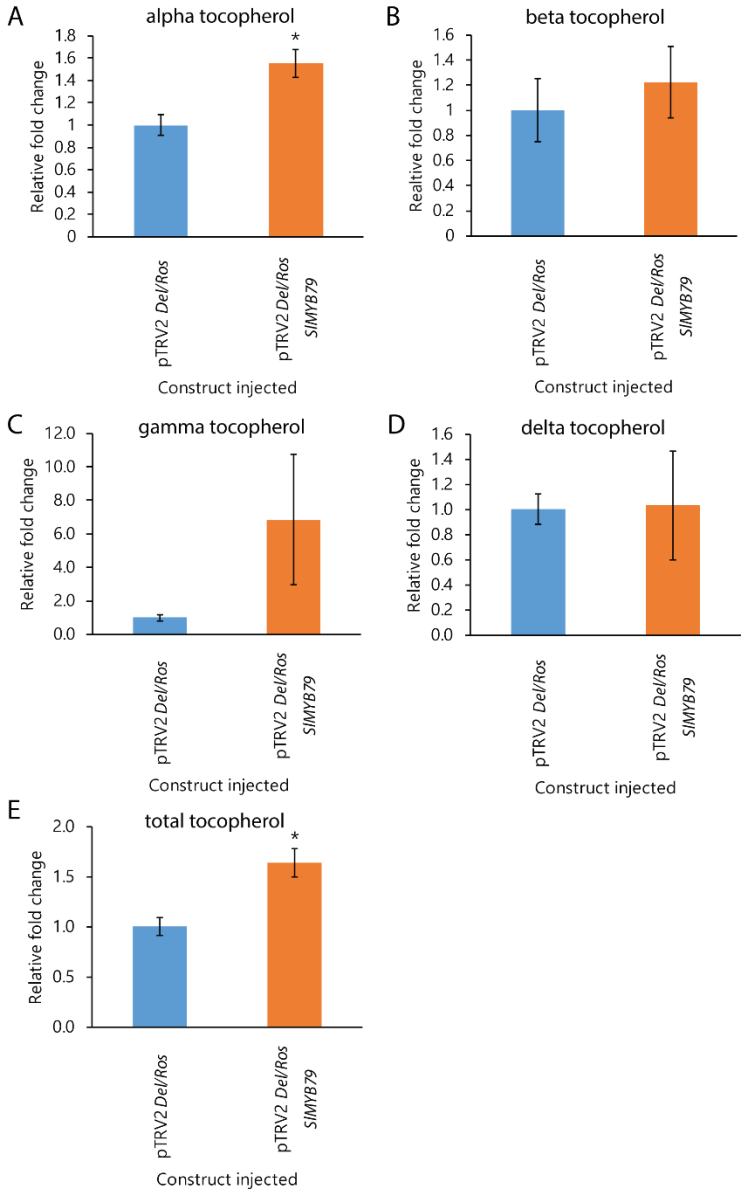


Figure 5-3 Relative fold change of tocopherols in silenced sectors of Moneymaker *Del/Ros* tomato fruits using VIGS. The tocopherol forms are (A) alpha , (B) beta, (C) gamma, (D) delta and (E) total tocopherols. The relative fold changes of pTRV2 *Del/Ros SlMYB79* red silenced sectors are compared to the silenced sectors of pTRV2 *Del/Ros* to generate a ratio. The errors bars depict the standard errors of the mean (n=3). T-tests were used to determine the statistical significance between red silenced sectors of pTRV2 *Del/Ros* and red silenced sectors of pTRV2 *Del/Ros SlMYB79*. The stars represent the significance, (*) = $p<0.05$ and (**) = $p<0.01$.

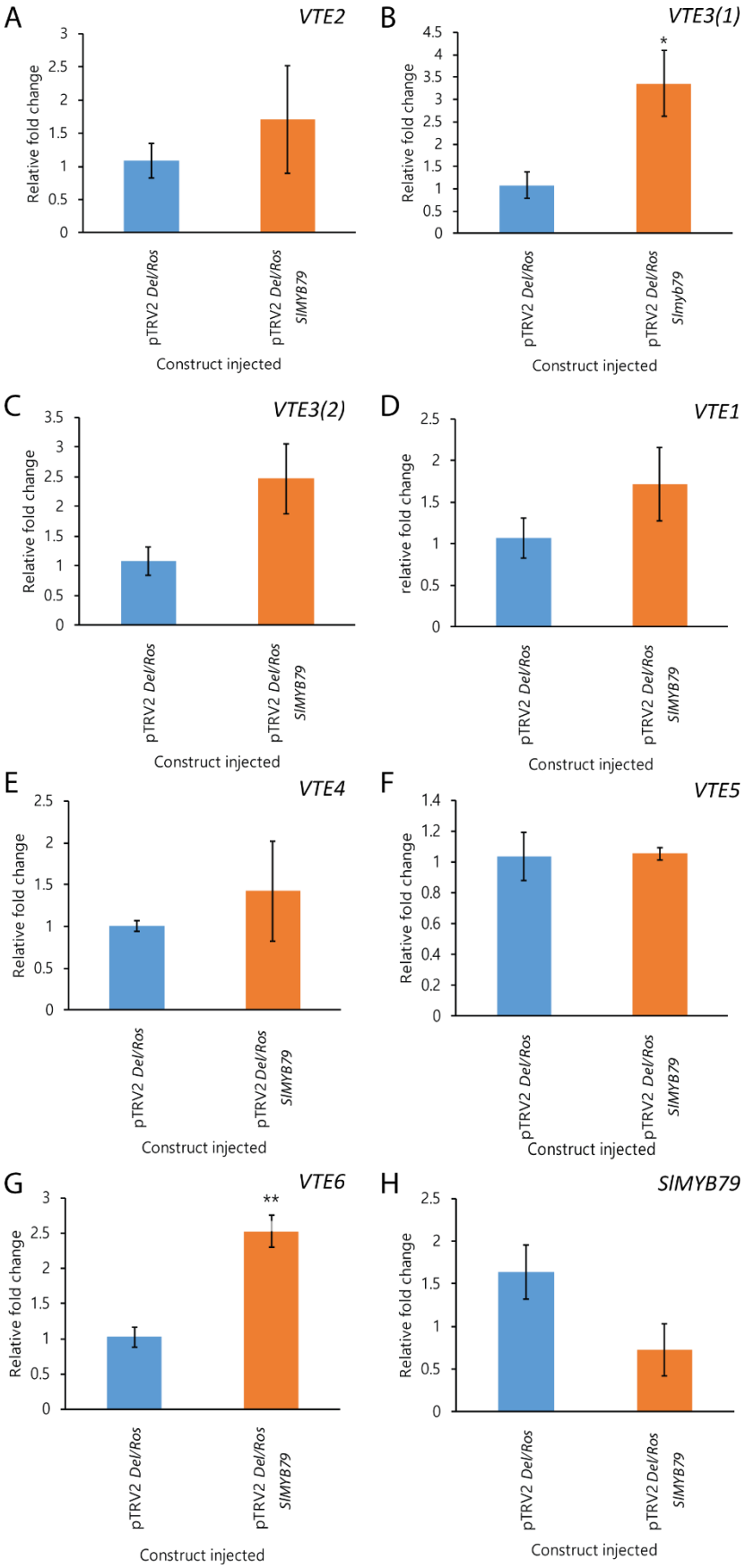

A
VTE2
Relative fold change
3
2.5
2
1.5
1
0.5
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlMYB79
Construct injected
B
VTE3(1)
*
Relative fold change
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
pTRV2 Del/Ros
pTRV2 Del/Ros Slmyb79
Construct injected
C
VTE3(2)
Relative fold change
3.5
3
2.5
2
1.5
1
0.5
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlMYB79
Construct injected
D
VTE1
relative fold change
2.5
2
1.5
1
0.5
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlMYB79
Construct injected
E
VTE4
Relative fold change
2.5
2
1.5
1
0.5
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlMYB79
Construct injected
F
VTE5
Relative fold change
1.4
1.2
1
0.8
0.6
0.4
0.2
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlMYB79
Construct injected
G
VTE6
**
Relative fold change
3
2.5
2
1.5
1
0.5
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlMYB79
Construct injected
H
SlMYB79
Relative fold change
2.5
2
1.5
1
0.5
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlMYB79
Construct injected

Figure 5-4 Relative fold changes of expression of genes encoding enzymes in the VTE pathway and *SlMYB79* in the VIGS assay. (A) *VTE2*, (B) *VTE3(1)*, (C) *VTE3(2)*, (D) *VTE1,* (E) *VTE4*, (F) *VTE5*, (G) *VTE6*, (H) *SlMYB79* are expressed as a ratio of the silenced sectors of pTRV2 *Del/Ros* SlMYB79 tomato fruit, compared to the silenced sectors of pTRV2 *Del/Ros* tomatoes. The errors bars depict the standard errors of the mean (n=3). The relative fold changes in expression were normalised to pTRV2 *Del/Ros* silenced sectors and SlCAC was used as a reference gene. T-tests were used to determine the statistical significance between red silenced sectors of pTRV2 *Del/Ros* and red silenced sectors of pTRV2 *Del/Ros SlMYB79*. The stars represent the significance, (*) = $p<0.05$ and (**) = $p<0.01$.

that *SlMYB79* was a good candidate transcriptional regulator of VTE biosynthesis in tomato fruit and this gene was taken forward for further analysis.

5.3.2 SlMYB79 CRISPR lines suggested that SlMYB79 was a repressor of the genes encoding enzymes of the MEP pathway

The VIGS data suggested that SlMYB79 is a transcriptional repressor, therefore I generated two knock-out lines using genome editing of *SlMYB79* in tomatoes, to determine its function. Guide RNAs (gRNAs) were designed to target sequences near to the start of the *SlMYB79* coding sequence (CDS) near to the ATG, and after the end of the sequence encoding the MYB domain (figure 5-5). Three mutations were found in the analysis of edited lines in the T1 generation (figure 5-5); one mutation was near to the ATG and two mutations were near to the MYB domain. At sgRNA2 there was a three-nucleotide (nt) heterozygous mutation in the CDS after the MYB domain, producing the CRISPR line; *myb79_15* (figure 5-5). This line resulted in a point mutation in the amino acid (AA) sequence (figure 5-6) but was heterozygous with the WT allele, and therefore this line was not taken forward for further analysis. There were two homozygous mutations found; *myb79_16* and *myb79_23*. *Myb79_16* had a four-nt, homozygous deletion near to the start ATG codon at sgRNA1 (figure 5-5). This resulted in a premature stop codon in the AA sequence, which suggested it encoded a protein that had a non-functional MYB domain and therefore was a knock-out allele (figure 5-6). The *myb79_23* CRISPR line had a ten-nt, homozygous deletion of the CDS after the sequence encoding the MYB domain (figure 5-5). The protein of *myb79_23* was predicted to have an intact MYB domain with a premature stop codon shortly after the MYB domain (figure 5-6). This line is also a knock out mutation, as the MYB TF does not have a C-terminal domain, and therefore probably cannot function.

Metabolic analysis of the CRISPR lines showed that tocopherol vitamers were different in the two homozygous lines, but not all were significant (figure 5-7). Fruit of *myb79_23* showed the greatest changes in gamma (γ), delta (δ), and total tocopherol levels, compared to WT fruit (figure 5-7). Fruit of *myb79_16* also

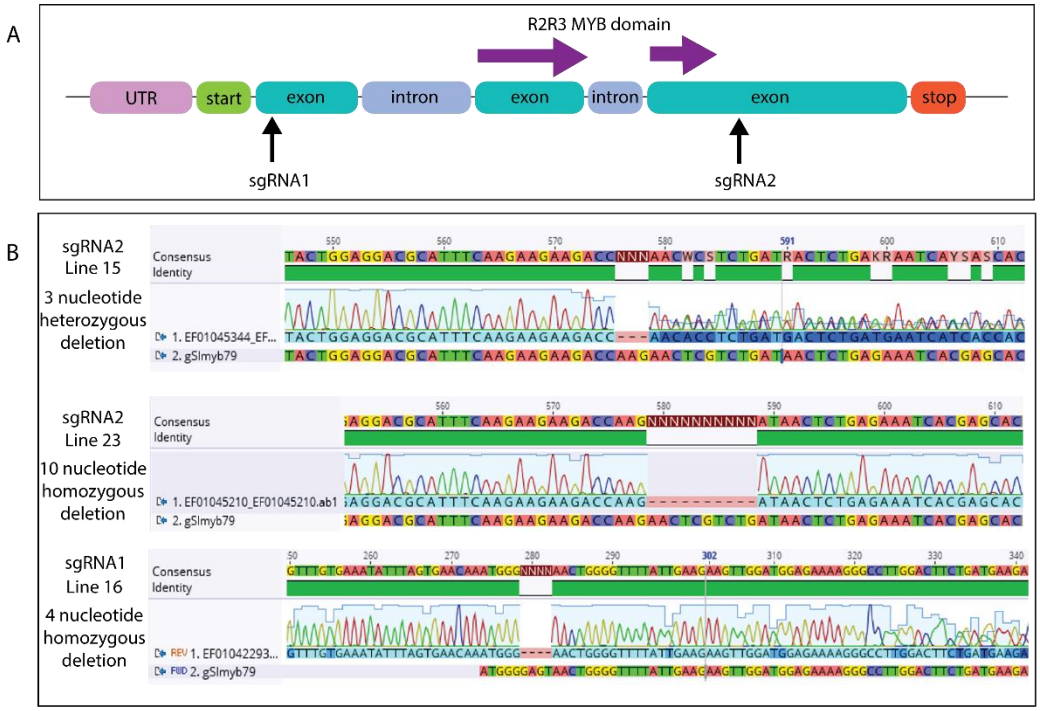


Figure 5-5 Schematic diagram of CRISPR guide RNAs (gRNA) and sequencing chromatograms of the genotypes in the T1 generation of CRISPR plants. (A) gRNA guides in exons of genomic SlMYB79 (gSlMYB79). The purple arrows indicate the R2R3 MYB domain spanning two exons. The sgRNA guides are indicated by the black arrows. (B) Sequencing chromatograms of the gSlMYB79 sequences of the T1 tomato genotypes of the two gRNAs for SlMYB79.

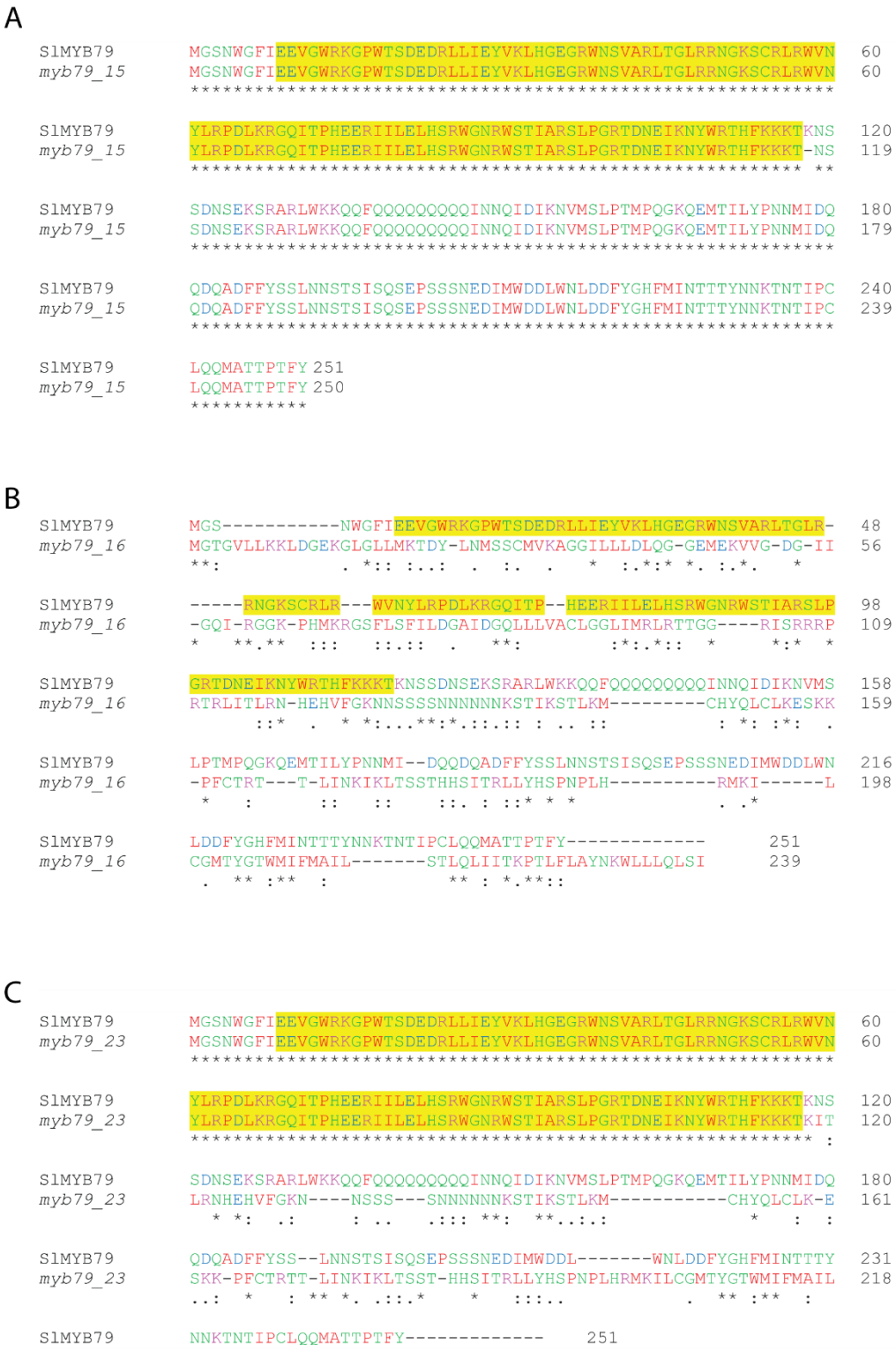


Figure 5-6 Truncated protein alignments of CRISPR lines of *SlMYB79* (A) *myb79_15* (B) *myb79_16* (C) *myb79_23.*

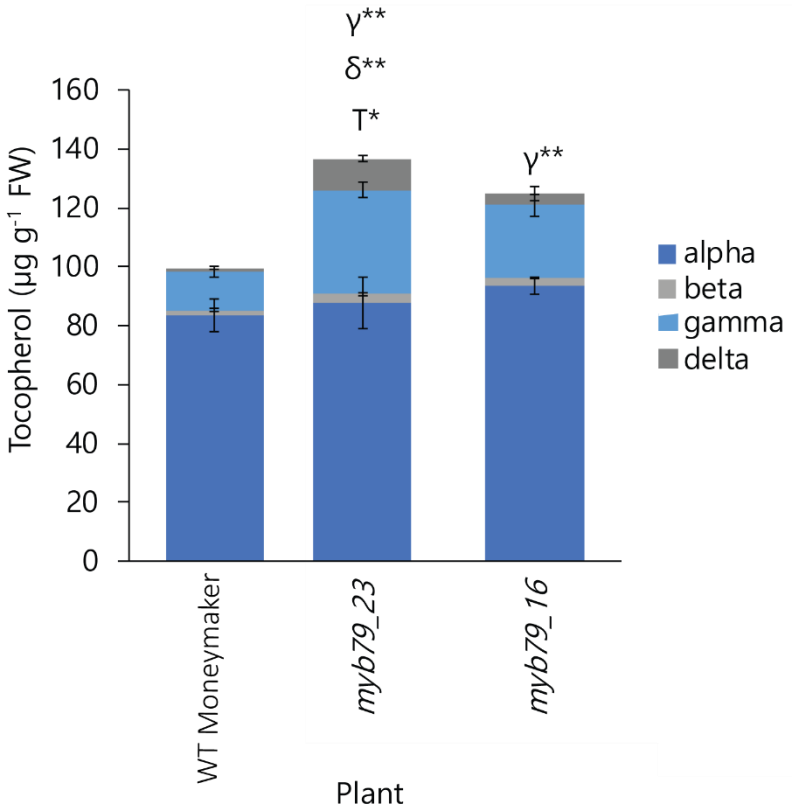


Figure 5-7 Tocopherol contents (µg g^{-1} fresh weight) in the CRISPR knock-out *SlMYB79* lines in B+10 fruit. *myb79_23* has a ten bp deletion and *myb79_16* has a four bp deletion. Tocopherol was measured µg per g of tomato fresh weight (FW) (n=4). Alpha tocopherol is shown in dark blue, beta tocopherol is shown in light grey, gamma tocopherol is shown in light blue and delta tocopherol is shown in dark grey. The errors bars represent standard errors of the mean. The letters represent the statistical significance for the different tocopherol forms α = alpha, β= beta, γ = gamma, δ = delta and T = total. The stars indicate statistical significance measured using students t-tests, (*) = $p < 0.05$, (**) = $p < 0.01$.

showed higher tocopherol levels, although they were not significantly higher, but *myb79_16* had significantly higher γ-tocopherols (figure 5-7). Transcript abundance of *SlMYB79* in these knock-out lines was higher than in WT fruit, but it was not significantly different to WT fruit (figure 5-8).

Expression of genes encoding enzymes of the VTE pathway were not significantly higher in fruit of the CRISPR lines compared to WT fruit (figure 5-9). However, *VTE5* was significantly higher in *myb79_23* fruits compared to WT fruits (figure 5-9).

Expression of genes encoding enzymes of the MEP pathway were generally upregulated in the CRISPR lines, in comparison to WT (figure 5-10). Genes encoding *ISPE, ISPF, GGPS(2)* and *GGPS(3)* were highly expressed in the homozygous CRISPR lines, compared to WT fruit (figure 5-11 and figure 5-12). The GGPS genes encode different isoforms of the GGPS enzymes, and have different expression patterns in tomato plants (figure 5-1 in appendix). *GGPS(1)* is more highly expressed in leaves, than in tomato fruits (appendix figure 5-1). *GGPS(2)* is more highly expressed in tomato fruits than in leaves, and *GGPS(3)* is expressed in both leaves and mature fruits (appendix figure 5-1). Expression of *IPI(1)* and *IPI(2)* was increased in fruit of *myb79_23,* compared to WT (figure 5-11). These data were consistent with the idea that SlMYB79 is a repressor of the genes encoding enzymes of the MEP pathway, which also fits with the VIGS data (figure 5-3 and figure 5-4).

The gene expression profiles for genes encoding enzymes of the SK pathway in both mutants were generally elevated compared to WT levels (figure 5-13). Expression of *DAHPS1* was significantly higher in both knock-out lines, compared to WT (figure 5-14). The genes; *SDH-DHQS1, EPSPS, PAT* and *CM1* were highly expressed in fruit of *myb79_23*, compared to WT (figure 5-14), but were less elevated in fruit of *myb79_16*. Both *DHQS* and *SK* were highly expressed in both knock-out lines, compared to WT (figure 5-14), and although, the differences in expression were not significant, the p-values were very close to $p<0.05$.

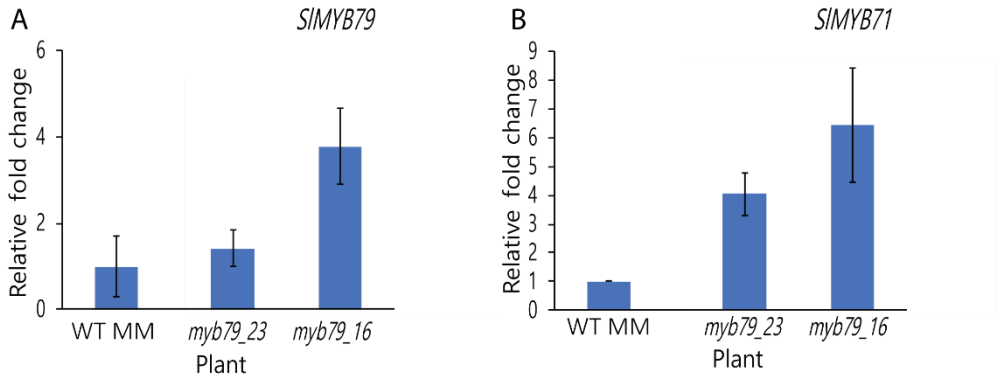


Figure 5-8 Relative fold change in expression of genes encoding TFs in the CRISPR knock-out *SlMYB79* lines in B+10 fruit, (A) *SlMYB79* and (B) *SlMYB71* (n=4). The errors bars depict the standard errors of the means. Statistical significance was calculated using student t-tests, however none of the values were significant compared to WT. All relative values are relative to WT and SlCAC was used as the reference gene.

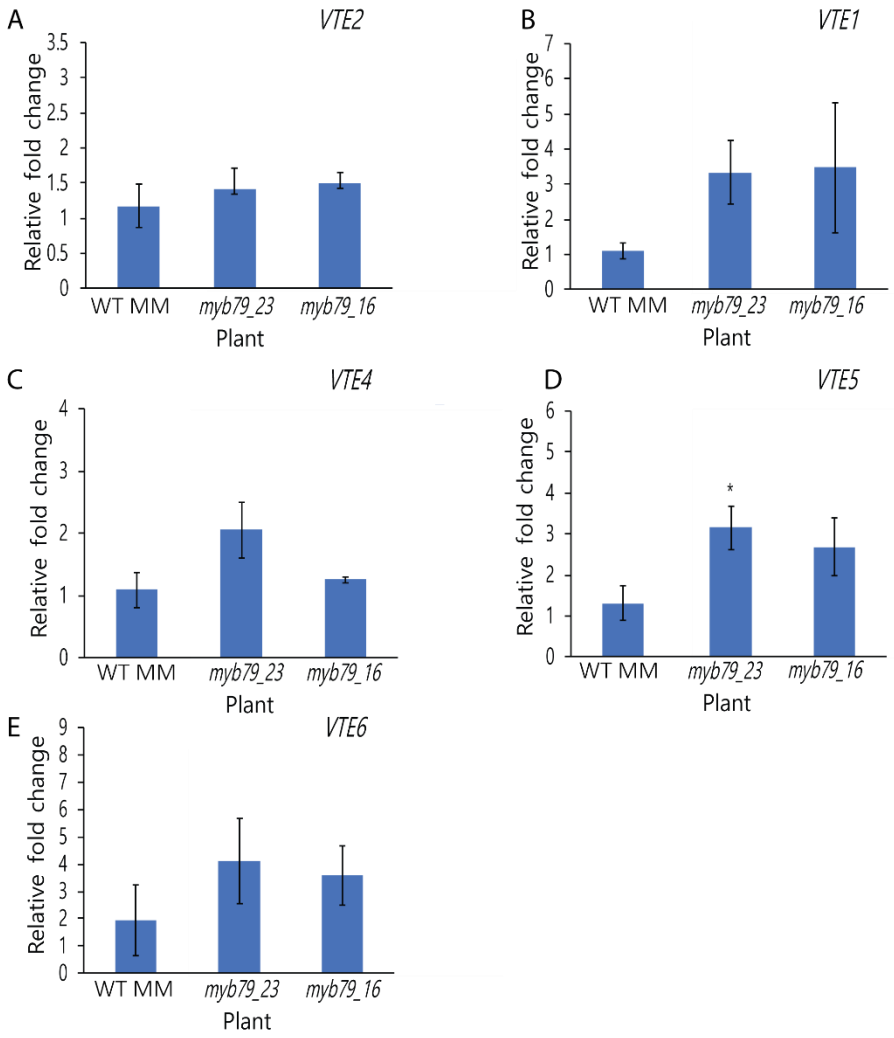


Figure 5-9 Relative fold change in expression of genes encoding enzymes of VTE biosynthesis in the CRISPR knock-out *SlMYB79* lines in B+10 fruit, (A) *VTE2,* (B) *VTE1,* (C) *VTE4,* (D) *VTE5* and (E) *VTE6* (n=4). The errors bars depict the standard errors of the means. The stars show statistical significance, calculated using student t-tests, (*) = $p<0.05$, (**) = $p<0.01$. All relative values are relative to WT and SlCAC was used as the reference gene.

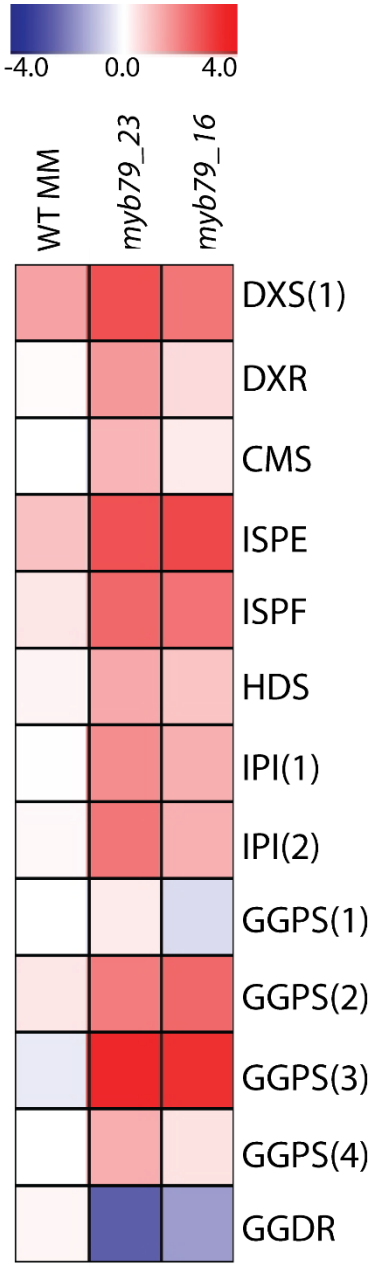


Figure 5-10 Log_2 expression heatmap of genes encoding enzymes in the MEP pathway in the CRISPR knock-out *SlMYB79* lines in B+10 fruit (n=4). The scale bar shows the log_2 values from -4.0 to 4.0. All values are relative to WT and SlCAC was used as the reference gene.

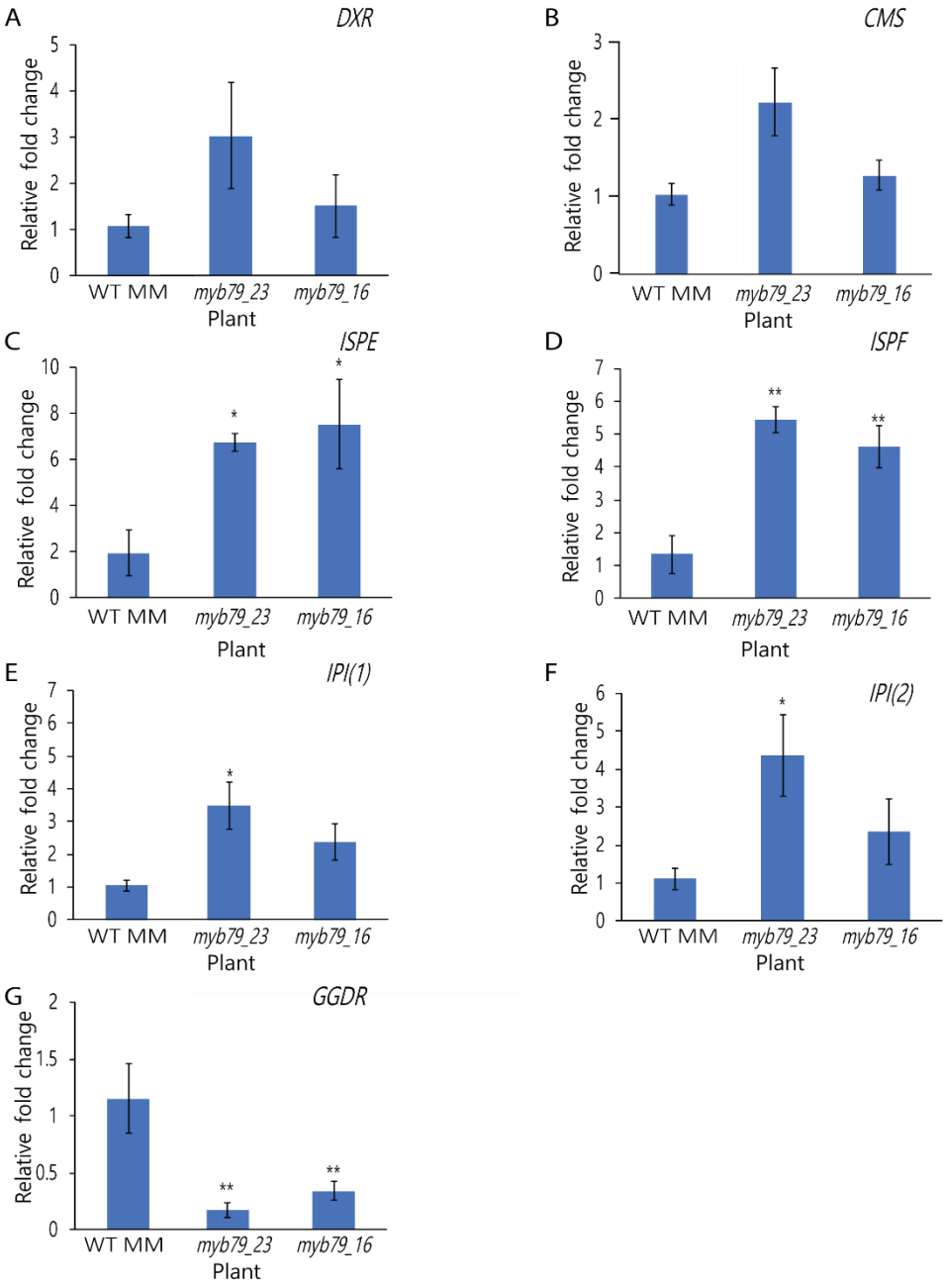


Figure 5-11 Relative fold change in expression of genes encoding enzymes in the MEP pathway in the CRISPR knock-out *SlMYB79* lines in B+10 fruit, (A) *DXR,* (B) *CMS*, (C) *ISPE*, (D) *ISPF*, (E) *IPI(1)*, (F) *IPI(2)* and (G) *GGDR* (n=4). The errors bars depict the standard errors of the means. The stars show statistical significance, calculated using student t-tests, (*) = p<0.05, (**) = p<0.01. All relative values are relative to WT and SlCAC was used as the reference gene.

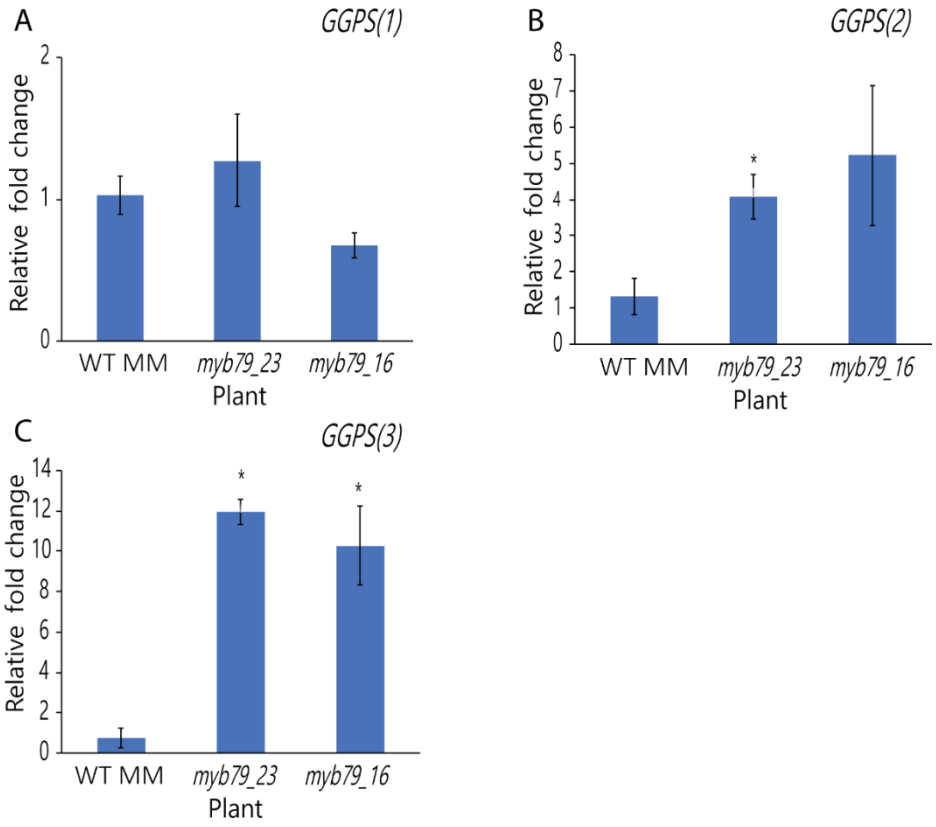


Figure 5-12 Relative fold change in expression of genes encoding enzymes in the MEP in the CRISPR knock-out *SlMYB79* lines in B+10 fruit, (A) *GGPS(1),* (B) *GGPS(2)* and (C) *GGPS(3)* (n=4). The errors bars depict the standard errors of the means. The stars show statistical significance, calculated using student t-tests, (*) = $p<0.05$, (**) = $p<0.01$. All values are expressed relative to WT and SlCAC was used as the reference gene.

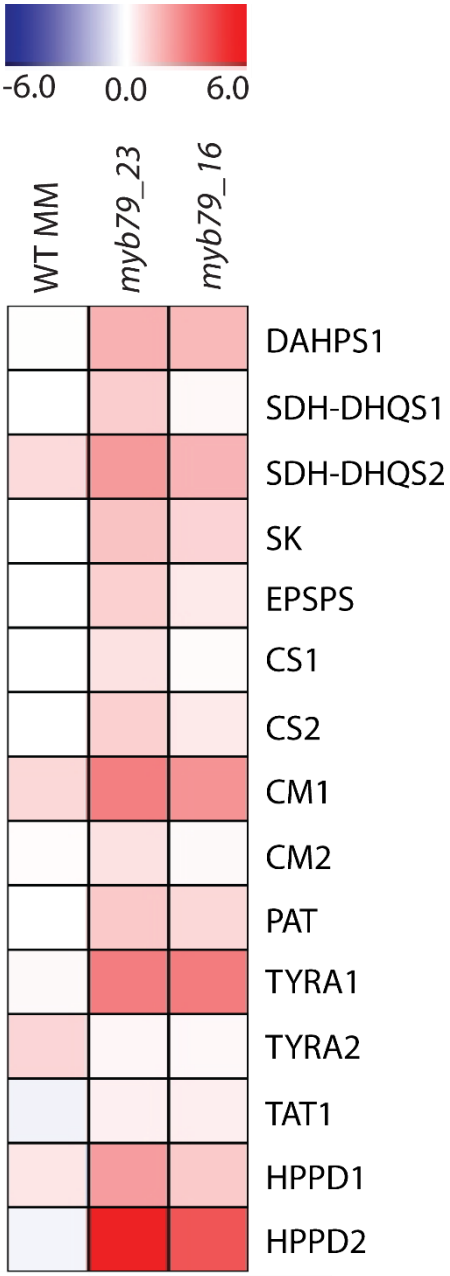


Figure 5-13 Log_2 expression heatmap of genes encoding enzymes in the SK pathway in the CRISPR knock-out *SlMYB79* lines in B+10 fruit (n=4). The scale bar showed the log_2 values from -6.0 to 6.0. All values are expressed relative to WT and SlCAC was used as the reference gene.

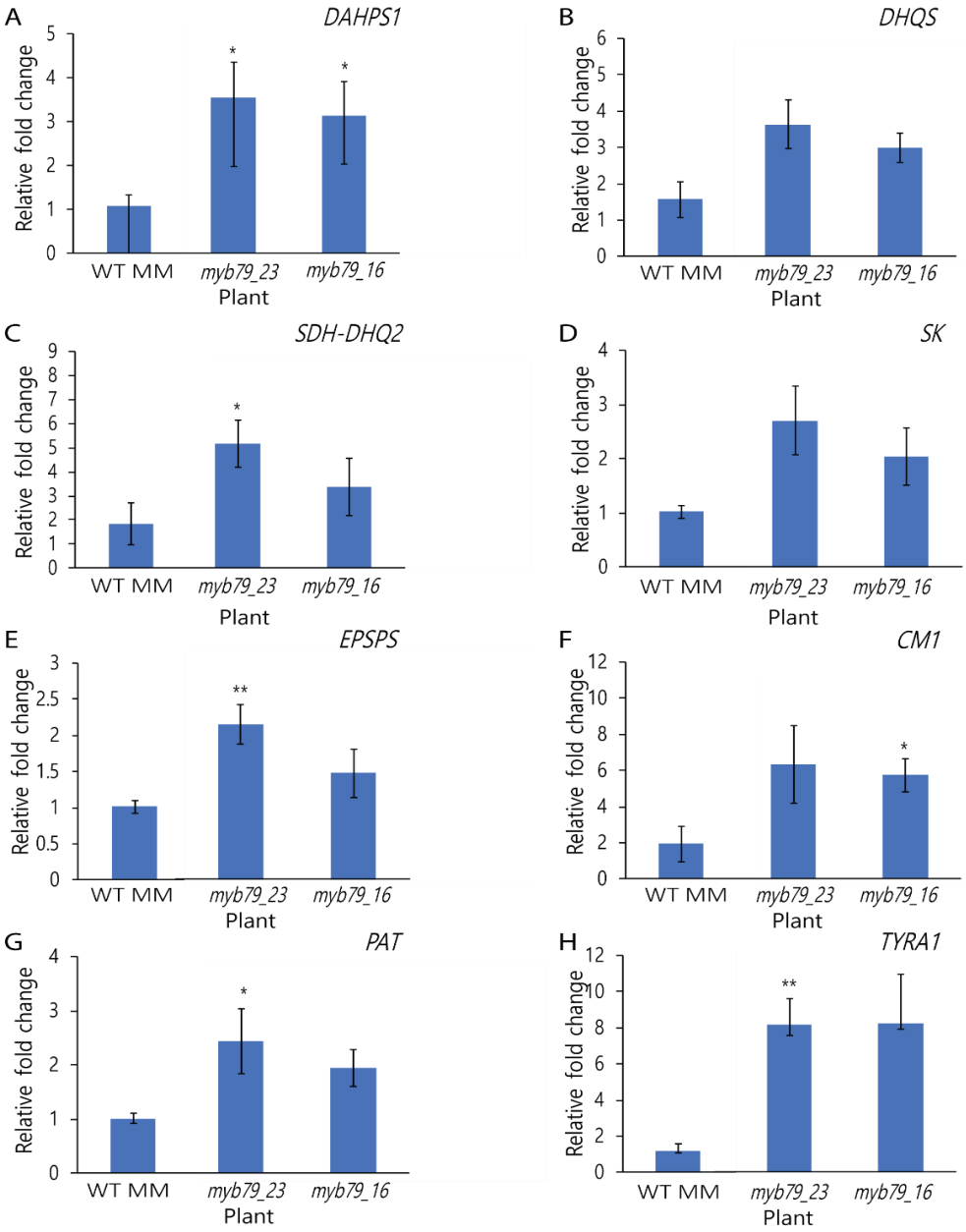


Figure 5-14 Relative fold change in expression of genes encoding enzymes in the SK pathway in the CRISPR knock-out *SlMYB79* lines in B+10 fruit, (A) *DAHPS1,* (B) *DHQS*, (C) *SDH-DHQS2*, (D) *SK*, (E) *EPSPS*, (F) *CM1,* (G) *PAT* and (H) *TYRA1* (n=4). The errors bars depict the standard errors of the means. The stars show statistical significance, calculated using student t-tests, (*) = $p<0.05$, (**) = $p<0.01$. All values are expressed relative to WT. All genes were normalised to the SlCAC reference gene.

These data suggested that SlMYB79 is a transcriptional repressor of the MEP, and possibly also the SK pathway.

5.3.3 Overexpression of SlMYB79 results in higher levels of tocopherols

The CRISPR lines showed that SlMYB79 was a transcriptional repressor, and therefore I overexpressed SlMYB79 under the control of the E8 promoter in Microtom tomato fruit. Analysis of Microtom E8:SlMYB79 transformants showed that the fruit over expressed *SlMYB79* (figure 5-15). Both lines overexpressed SlMYB79 at the B stage (figure 5-15). The overexpression lines accumulated higher levels of tocopherols at the B+10 stage, compared to WT fruit (figure 5-16). E8:SlMYB79 MT1 fruit had 1.5-, 1.22- and 1.41-fold higher total tocopherol levels than WT at the B, B+5 and B+10 stages, respectively (figure 5-16). Relative fold changes of E8:SlMYB79 MT2 were 1.28-, 1.12- and 1.31-fold higher than WT at B, B+5 and B+10 stages, respectively (figure 5-16). Alpha (α) tocopherol constituted 88% of total tocopherols in WT B+10 fruit, however α- tocopherol constituted 92% and 94% in E8:SlMYB79 MT1 and MT2, respectively. Figure 5-16 showed that α-tocopherol levels expressed as a percentage were higher than WT fruit, but generally the relative ratios of α-tocopherol to total tocopherol were similar. This suggested that this TF might work by altering substrate availability of precursors of VTE biosynthesis, rather than by targeting the VTE pathway directly.

Measurements of the relative expression of genes encoding enzymes in the VTE pathway showed that these genes were not differentially expressed in E8:SlMYB79 overexpressing lines in B+10 fruit (figure 5-17). However, *VTE6* was repressed at the B+5 stage in the overexpression lines, compared to WT (figure 5-17). This gene was not highly expressed in the knock-out CRISPR lines of SlMYB79, which suggests that this may not be a direct target of SlMYB79 (figure 5-9).

Genes encoding enzymes in the MEP pathway were differentially expressed in fruit overexpressing *SlMYB79*, compared to WT fruit (figure 5-18). Expression of genes encoding *DXR, ISPE, ISPF, IPI(1), IPI(2), GGPS(1), GGPS(2)* and

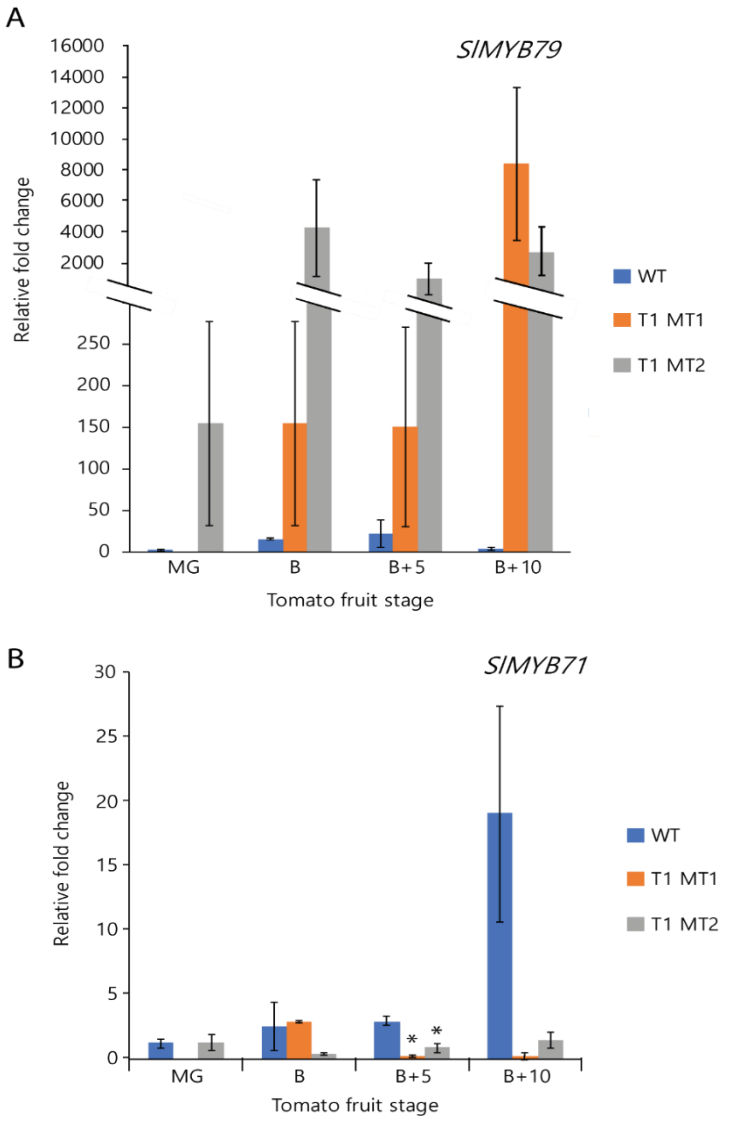


Figure 5-15 Relative fold change of genes encoding TFs, (A) *SlMYB79* and (B) *SlMYB71* for Microtom E8:SlMYB79 transformed tomato fruit. Four tomato fruit stages were analysed for , T1 MT1 E8:SlMYB79 and T1 MT2 E8:SlMYB79 (n=3). MG = mature green, B = Breaker, B+5 = Breaker + 5 days, B+10 = Breaker +10 days. The errors bars depict the standard errors of the means. Statistical significance was calculated using student t-tests, however none of the values were significantly different, compared to WT. All values are relative to WT at the MG stage. All genes were normalised to the SlCAC reference gene.

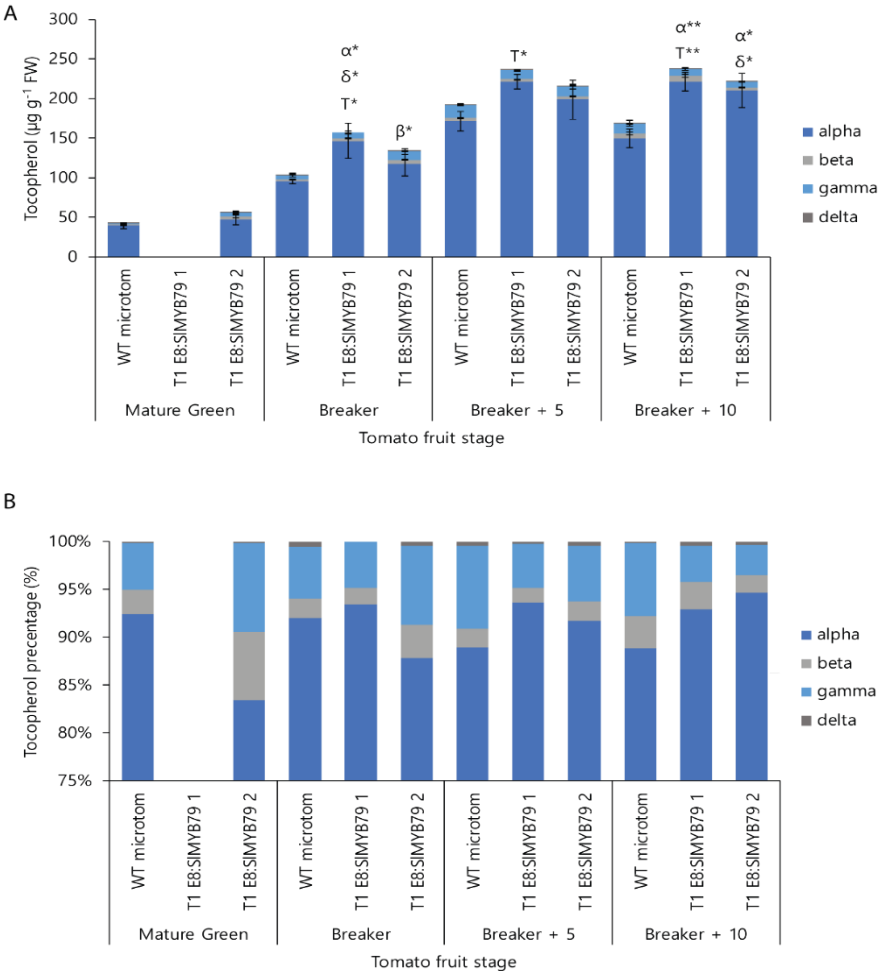


Figure 5-16 (A) Tocopherol contents (µg g^{-1} fresh weight) of Microtom E8:SlMYB79 stable transformant tomato fruit. (B) Tocopherol contents of fruit of two independent T1 Microtom E8:SlMYB79 stable transformants expressed as a percentage of total tocopherols (based on average values). Tocopherol was measured as µg per g of tomato fresh weight (FW) (n=4). Alpha tocopherol is shown in dark blue, beta tocopherol in light grey, gamma tocopherol in light blue and delta tocopherol in dark grey. The errors bars represent standard errors of the mean. The letters represent the statistical significance of the differences between the different tocopherol forms α = alpha, β= beta, γ = gamma, δ = delta and T = total between the different lines. The stars indicate statistical significance measured using students t-tests, (*) = $p < 0.05$, (**) = $p < 0.01$.

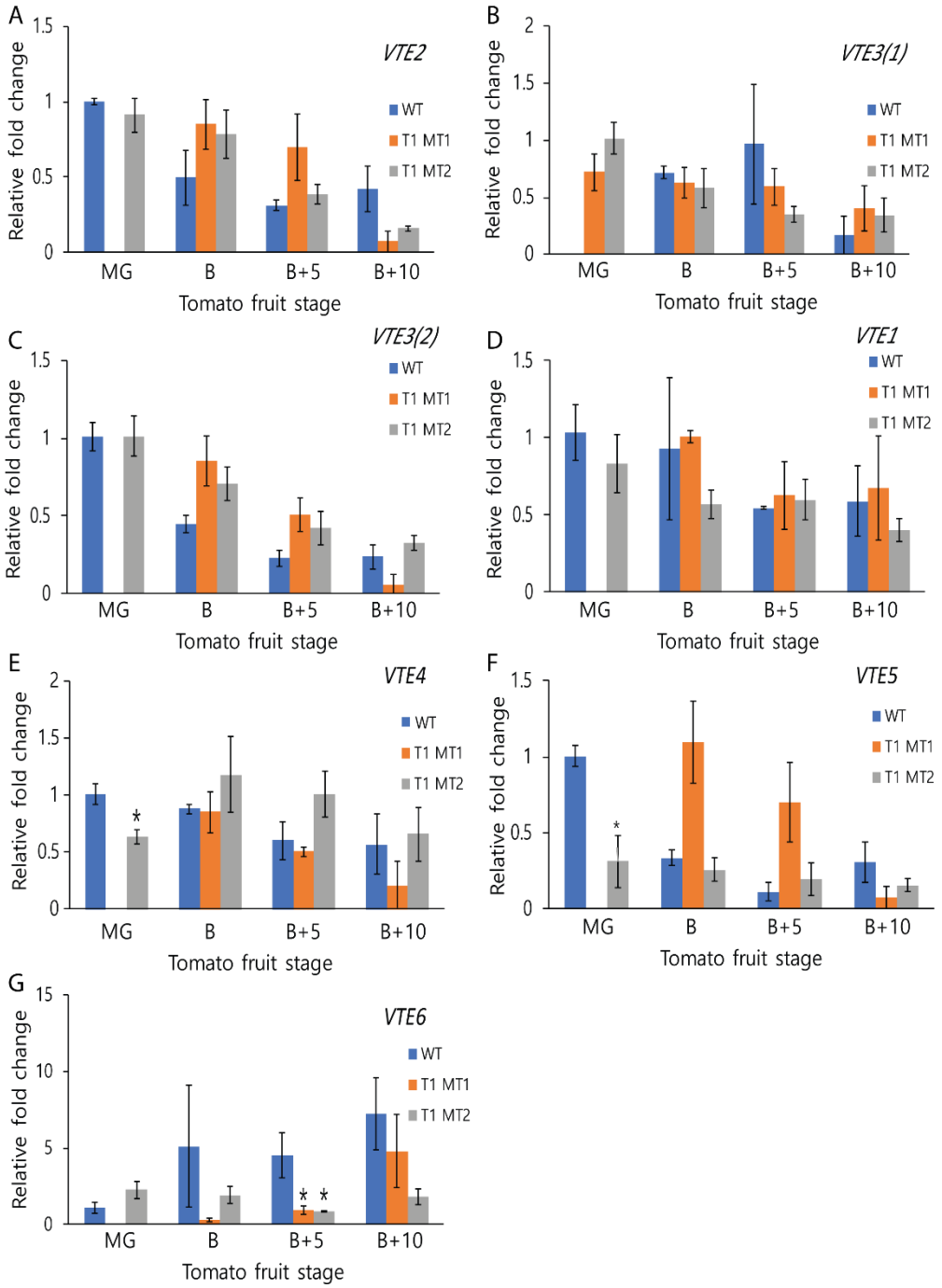


Figure 5-17 Relative fold change in expression of genes encoding enzymes of VTE synthesis; (A) *VTE2,* (B) *VTE3(1),* (C), *VTE3(2)* (D), *VTE1* (E), *VTE4* (F) *VTE5* and (G) *VTE6* for Microtom E8:SlMYB79 transformed tomato fruit. Four tomato fruit stages were analysed for each line, MT1 E8:SlMYB79 and MT2 E8:SlMYB79 (n=3). MG = mature green, B = Breaker, B+5 = Breaker + 5 days, B+10 = Breaker +10 days. The error bars depict the standard errors of the means. The stars show statistically significant differences, calculated using student t-tests, (*) = $p<0.05$, (**) = $p<0.01$. All values are relative to WT at the MG stage. All genes were normalised to the SlCAC reference gene.

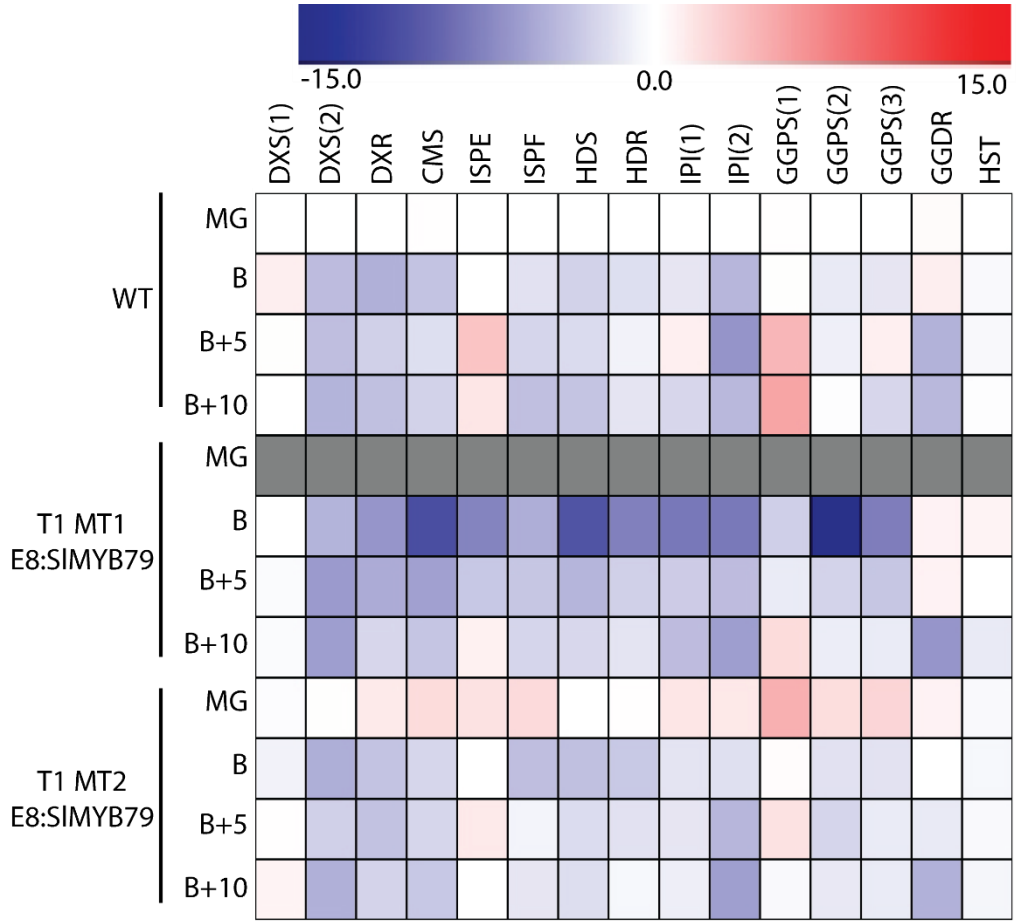


Figure 5-18 Log_2 expression heatmap of genes encoding enzymes in the MEP pathway for Microtom E8:SlMYB79 transformed tomato fruit. Four tomato fruit stages were analysed for each line, T1 MT1 E8:SlMYB79 and T1 MT2 E8:SlMYB79 (n=3). MG = mature green, B = Breaker, B+5 = Breaker + 5 days, B+10 = Breaker +10 days. The grey boxes indicate missing data points. The scale bar shows the log_2 values from -10.0 to 10.0. All values are relative to WT at the MG stage. All genes were normalised to the SlCAC reference gene.

GGPS(3) was upregulated in E8:SlMYB79 MT2 fruit at the MG fruit stage, compared to WT fruit (figure 5-18, figure 5-19 and figure 5-20).

The expression of these genes was reduced significantly at the B stage for *DXR, ISPF* and *IPI(1)* (figure 5-19). *GGPS(1), GGPS(2)* and *GGPS(3)* were the most differentially expressed genes in the overexpression fruit, compared to WT fruit (figure 5-20). The induced *GGPS(1)/(2)/(3)* genes were significantly repressed during tomato ripening, compared to WT (figure 5-20). They were repressed in B and B+5 stages in E8:SlMYB79 MT1 fruit, compared to WT (figure 5-20). Interestingly, expression of *GGDR* was significantly repressed in the CRISPR SlMYB79 knock-out lines compared to WT (figure 5-11), however, this gene was not differentially expressed in fruit overexpressing *E8:SlMYB79* (figure 5-18). Therefore, the repression of *GGDR* in the CRISPR lines may be a down-stream effect of the regulation of the MEP pathway because of an increase in flux along the MEP pathway (figure 5-18).

The genes encoding enzymes of the SK pathway showed more variable expression profiles (figure 5-21). Genes encoding enzymes of the SK pathway in the fruit of overexpression lines were not significantly different when compared to WT fruit (figure 5-21). The most differentially expressed genes in the overexpression lines of genes encoding enzymes in the SK pathway were; *CM2,* which was increased at B+10, and *TAT1*, which was repressed at B and B+5 stage, compared to WT fruit (figure 5-21), but these changes in expression were not significant. However, in the CRISPR knock-outs, several genes encoding enzymes of the SK pathway were highly expressed (figure 5-14). SlMYB79 may not be a repressor of these genes, and rather their up-regulation may be due to cross-talk between the pathways.

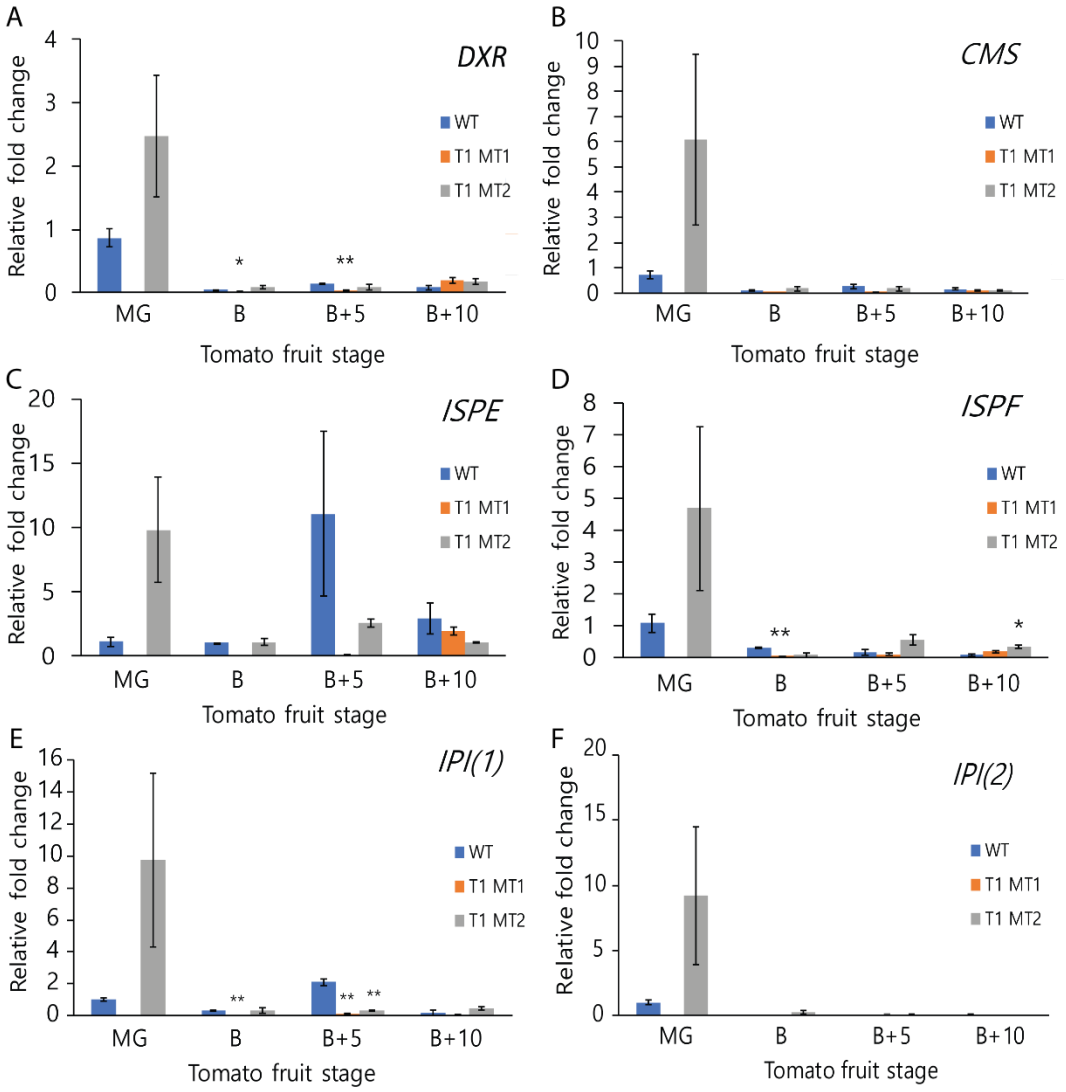


Figure 5-19 Relative fold change in expression of genes encoding enzymes in the MEP pathway for Microtom E8:SlMYB79 transformed tomato fruit, (A) *DXR,* (B) *CMS*, (C) *ISPE*, (D) *ISPF*, (E) *IPI(1)* and (F) *IPI(2)*. Four tomato fruit stages were analysed for each line, MT1 E8:SlMYB79 and MT2 E8:SlMYB79 (n=3). MG = mature green, B = Breaker, B+5 = Breaker + 5 days, B+10 = Breaker +10 days. The errors bars depict the standard errors of the means. The stars show statistical significance, calculated using student t-tests, (*) = $p<0.05$, (**) = $p<0.01$. All values are relative to WT at the MG stage. All genes were normalised to the SlCAC reference gene.

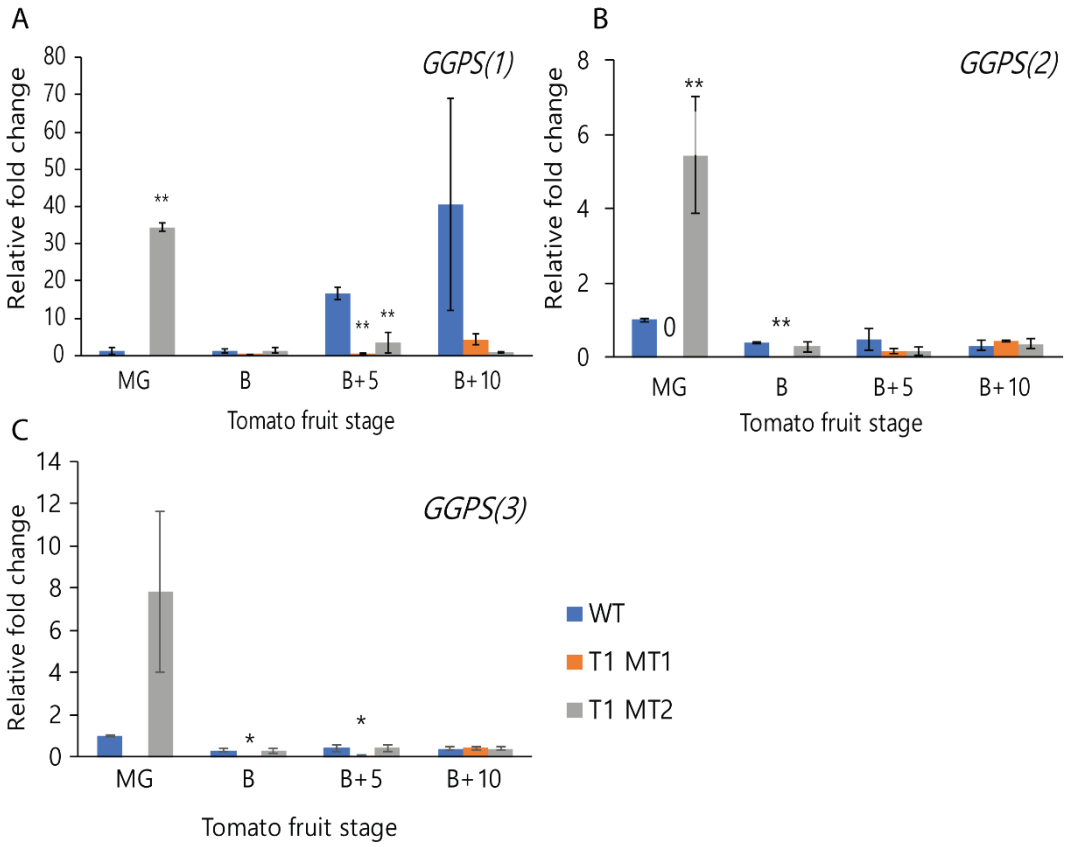


Figure 5-20 Relative fold change of genes encoding enzymes in the MEP pathway for the Microtom E8:SlMYB79 transformed tomato fruit, (A) *GGPS(1),* (B) *GGPS(2)* and (C) *GGPS(3)*. Four tomato fruit stages were analysed for each line, MT1 E8:SlMYB79 and MT2 E8:SlMYB79 (n=3). MG = mature green, B = Breaker, B+5 = Breaker + 5 days, B+10 = Breaker +10 days. The errors bars depict the standard errors of the means. The stars show statistical significance, calculated using student t-tests, (*) = $p<0.05$, (**) = $p<0.01$. All values are relative to WT at the MG stage. All genes were normalised to the SlCAC reference gene.

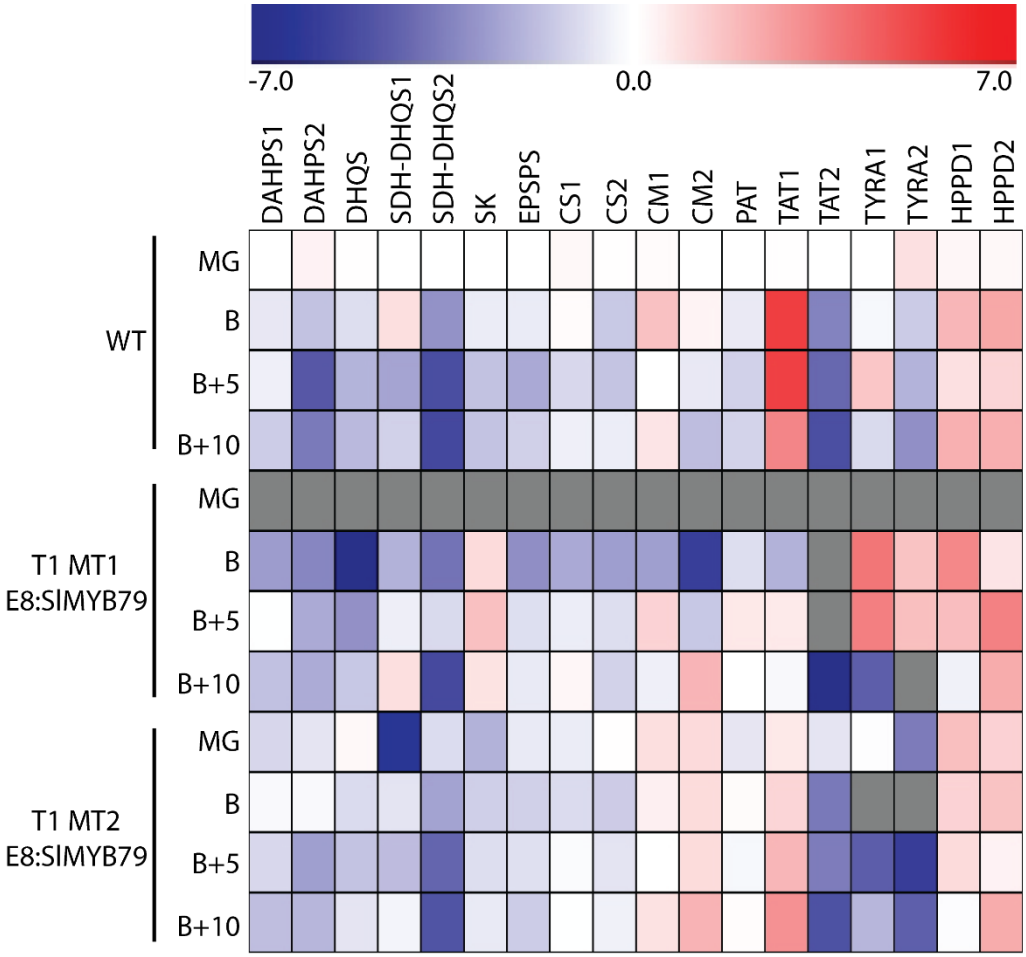


Figure 5-21 Log_2 expression heatmap of genes encoding enzymes in the SK pathway for Microtom E8:SlMYB79 transformed tomato fruit. Four tomato fruit stages were analysed for each line, T1 MT1 E8:SlMYB79 and T1 MT2 E8:SlMYB79 (n=3). MG = mature green, B = Breaker, B+5 = Breaker + 5 days, B+10 = Breaker +10 days. The grey boxes missing data points. The scale bar shows the log_2 values from -7.0 to 7.0. All values are relative to WT at the MG stage. All genes were normalised to the SlCAC reference gene.

5.3.4 Tomatoes overexpressing SlMYB79 had a higher antioxidant capacity than controls

The water fraction of tomato fruits overexpressing *SlMYB79* had a higher antioxidant capacity than WT (figure 5-22). However, only E8:SlMYB79 MT1 tomato fruits had significantly higher antioxidant capacity than WT. Acetone fractions (containing lipophilic antioxidants) were lower than WT levels (figure 5-22). However, acetone may not be the optimum extraction method for lipophilic antioxidants, since tocopherols (which are lipophilic) were significantly higher in these tomato fruits (figure 5-16).

5.3.5 Overexpression of SlMYB71 in Moneymaker tomatoes

The tomato genome encodes a paralog of SlMYB79; SlMYB71, which is phylogenetically closely related to SlMYB79 (figure 5-23), and, to date, neither gene has a reported function. The phylogenetic tree shown in figure 5-23 suggests that SlMYB71 is likely an ortholog of AtMYB79, and SlMYB79 is an ortholog of AtMYB71. When the tomato genes were annotated, they were probably annotated incorrectly on the Solgenomics network (Fernandez-Pozo et al., 2015a). The *Arabidopsis* genes encoding AtMYB79 and AtMYB71 probably shared a common ancestor, together with *SlMYB79* and *SlMYB71*. This suggests that the duplication of these genes in *Arabidopsis* and *Solanum* species may have occurred early in the ancestry of land plants, because *Arabidopsis* belongs to the Rosids family, and tomato to the Asterid clade.

SlMYB79 and SMYB71 are R2R3 MYB TFs, which have conserved MYB DNA binding domains and relatively closely related C-terminal domains (figure 5-24), suggesting that they may share very similar DNA binding motifs. Protein alignments of these paralogs showed that their C-terminal activation domains are similar, and they both contain the glutamine rich and acidic domains associated with transcriptional activation (figure 5-24). The SlMYB71 protein is a truncated

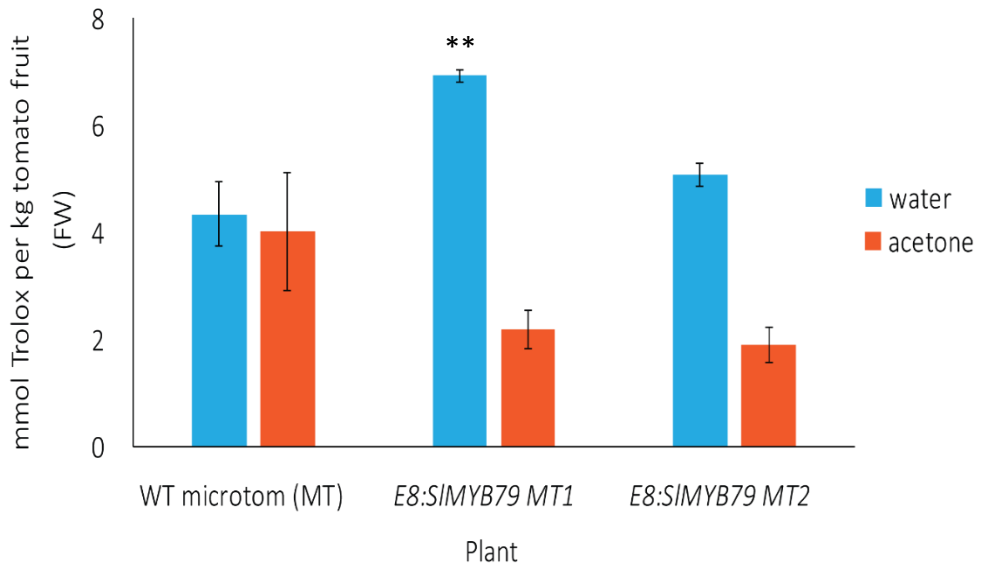


Figure 5-22 Trolox-equivalent antioxidant capacity assay of Microtom E8:SlMYB79 transformed tomato fruit. Fruit were analysed at B+10 (n=4). The antioxidant capacity is expressed as mmol Trolox per kg of tomato fruit (fresh weight - FW). Water and acetone extracts of the tomato fruit were analysed. The errors bars depict the standard errors of the means. The stars show statistical significance, calculated using student t-tests between WT and the overexpression line, (*) = $p<0.05$, (**) = $p<0.01$.

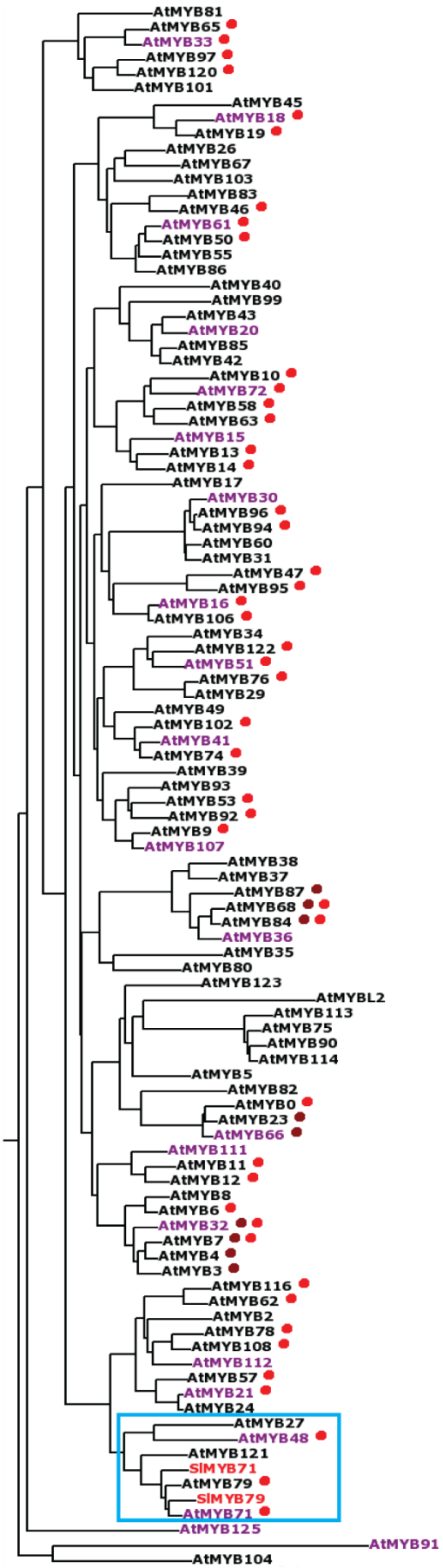

AtMYB81
AtMYB65
AtMYB33
AtMYB97
AtMYB120
AtMYB101
AtMYB45
AtMYB18
AtMYB19
AtMYB26
AtMYB67
AtMYB103
AtMYB83
AtMYB46
AtMYB61
AtMYB50
AtMYB55
AtMYB86
AtMYB40
AtMYB99
AtMYB43
AtMYB20
AtMYB85
AtMYB42
AtMYB10
AtMYB72
AtMYB58
AtMYB63
AtMYB15
AtMYB13
AtMYB14
AtMYB17
AtMYB30
AtMYB96
AtMYB94
AtMYB60
AtMYB31
AtMYB47
AtMYB95
AtMYB16
AtMYB106
AtMYB34
AtMYB122
AtMYB51
AtMYB76
AtMYB29
AtMYB49
AtMYB102
AtMYB41
AtMYB74
AtMYB39
AtMYB93
AtMYB53
AtMYB92
AtMYB9
AtMYB107
AtMYB38
AtMYB37
AtMYB87
AtMYB68
AtMYB84
AtMYB36
AtMYB35
AtMYB80
AtMYB123
AtMYBL2
AtMYB113
AtMYB75
AtMYB90
AtMYB114
AtMYB5
AtMYB82
AtMYB0
AtMYB23
AtMYB66
AtMYB111
AtMYB11
AtMYB12
AtMYB8
AtMYB6
AtMYB32
AtMYB7
AtMYB4
AtMYB3
AtMYB116
AtMYB62
AtMYB2
AtMYB78
AtMYB108
AtMYB112
AtMYB57
AtMYB21
AtMYB24
AtMYB27
AtMYB48
AtMYB121
SlMYB71
AtMYB79
SlMYB79
AtMYB71
AtMYB125
AtMYB91
AtMYB104

Figure 5-23 Phylogenetic tree showing SlMYB79 and SlMYB71 aligned with all *Arabidopsis thaliana* R2R3 MYBs. This phylogenetic tree was adapted from the tree generated from IT3F (Bailey et al., 2008, Bailey et al., 2012). The blue box shows the clade in which SlMYB79 and SlMYB71 reside. The spots depict gene duplications; dark red = older duplications, red = newer duplications.

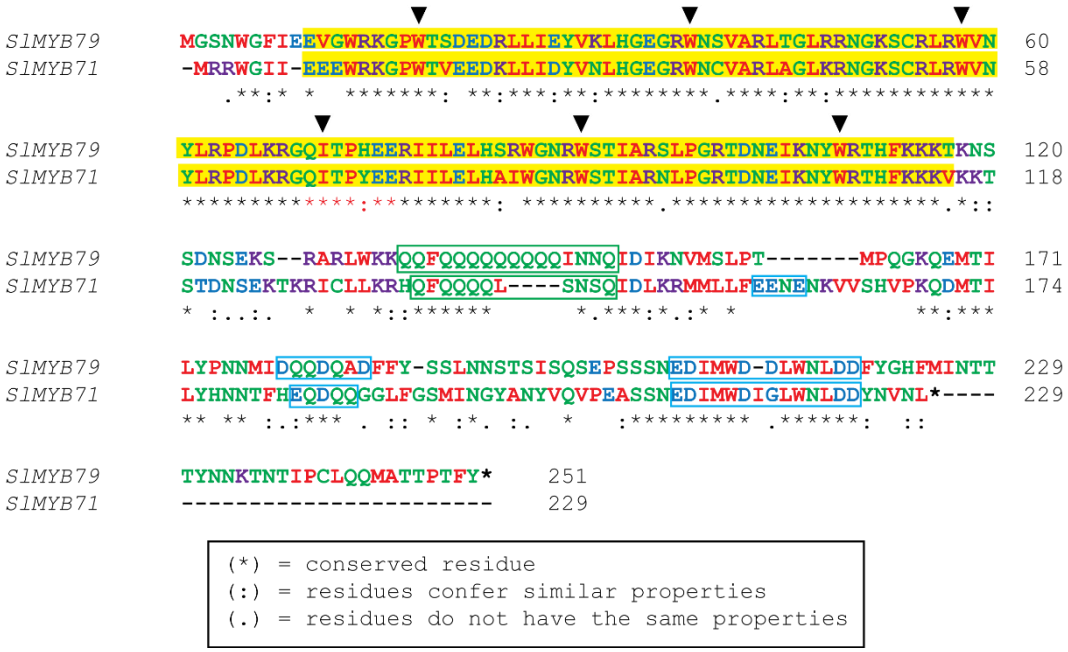


Figure 5-24 Protein alignments of SlMYB79 and SlMYB71, generated using clustal omega (Larkin et al., 2007). The yellow box highlights the MYB DNA binding domain. The black arrow heads show the conserved tryptophan residues, of which three are associated with each MYB repeat. The first hydrophobic reside in R3 is I, rather than W in plant R2R3 MYB proteins (Martin and Paz-Arez, 1997). The amino acid residues colours are as follows; red = small and hydrophobic, including tyrosine (Y), blue = acidic, magenta = basic histidine (H), green = hydroxyl, sulfhydryl, and amine, including glycine (G). Transcriptional activation domains in MYB TFs are associated with amino acids (proline and glutamine), or with acidic areas (Jin and Martin, 1999). The glutamine rich area is shown in the green box and the acidic areas are shown in blue boxes.

version of SlMYB79 and there are 22 AA missing (figure 5-24). The protein sequence alignments suggest that these proteins might be functionally redundant or have overlapping functionalities. Figure 5-25 shows SlMYB79 and SlMYB71 proteins aligned with AtMYB79 and AtMYB71. The alignment shows that the MYB domains are very similar, and the C-terminal activation domains are also similar. Therefore, it is likely that in *Arabidopsis* the proteins encoded by these genes might have similar functions. The fact that AtMYB71/SlMYB79 and AtMYB79/SlMYB71 have maintained distinct structures over long evolutionary time suggests that they have conserved, but distinct functionalities. AtMYB79, AtMYB71 and SlMYB79 share an additional C-terminal domain which is absent in SlMYB71, near to a conserved acidic domain and is characteristic of an activation domain. The absence of this domain in SlMYB71 could indicate that SlMYB71 has different transcriptional regulatory properties to SlMYB79 and the two proteins in *Arabidopsis*.

Expression of *SlMYB71* was elevated in the *SlMYB79* CRISPR knock-out lines, but the increase in expression was not statistically significant compared to transcript levels in WT (figure 5-8). In tomatoes overexpressing *SlMYB79*, expression of *SlMYB71* was repressed significantly at B+5 (figure 5-15). This suggested that SlMYB79 may negatively regulate expression of the gene encoding SlMYB71. Therefore, I overexpressed in Moneymaker tomatoes and analysed the fruit of one line in the T0 generation. Expression of *SlMYB71* was high in the overexpression line (figure 5-26). Total tocopherols were significantly increased in these overexpressing tomato fruits by 1.7-fold in E8:SlMYB71-1, compared to fruit of WT (figure 5-27). The expression of *SlMYB79* was elevated in the E8:SlMYB71 fruit (figure 5-26). The expression of the genes encoding enzymes in the MEP, SK and VTE pathway were generally increased in E8:SlMYB71-1 fruit (figure 5-26). The analysis was carried out on T0 fruit, therefore further analysis of T1 fruit, examination of more independent transformants and a larger analysis of genes encoding enzymes in the MEP and SK pathways would clarify the targets for SlMYB71. Expression of *SlMYB79* was elevated in tomato fruit overexpressing

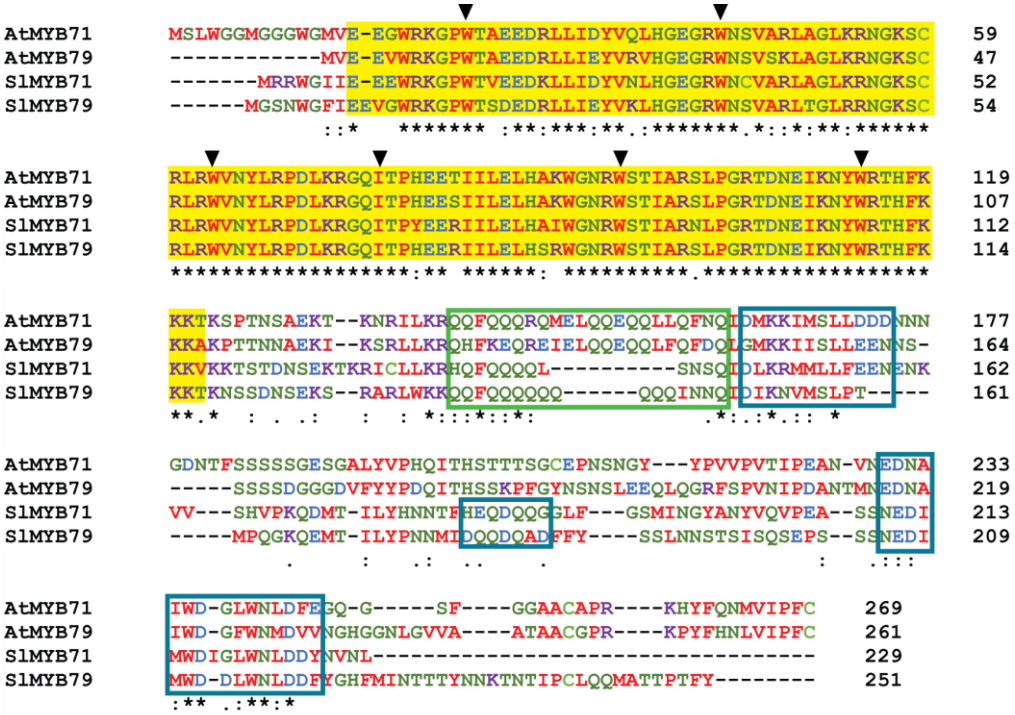


Figure 5-25 Protein alignments of SlMYB79, SlMYB71 and the orthologs AtMYB71 and AtMYB79, generated using clustal omega (Larkin et al., 2007). The yellow box highlights the MYB DNA binding domain. The black arrow heads show the conserved tryptophan residues, of which three are associated with each MYB repeat. The first hydrophobic reside in R3 is I, rather than W in plant R2R3 MYB proteins (Martin and Paz-Arez, 1997). The amino acid residues colours are as follows; red = small and hydrophobic, including tyrosine (Y), blue = acidic, magenta = basic histidine (H), green = hydroxyl, sulfhydryl, and amine, including glycine (G). Transcriptional activation domains in MYB TFs are associated with amino acids (proline and glutamine), or with acidic areas (Jin and Martin, 1999). The glutamine rich areas are shown in the green boxes and the acidic areas are shown in blue boxes.

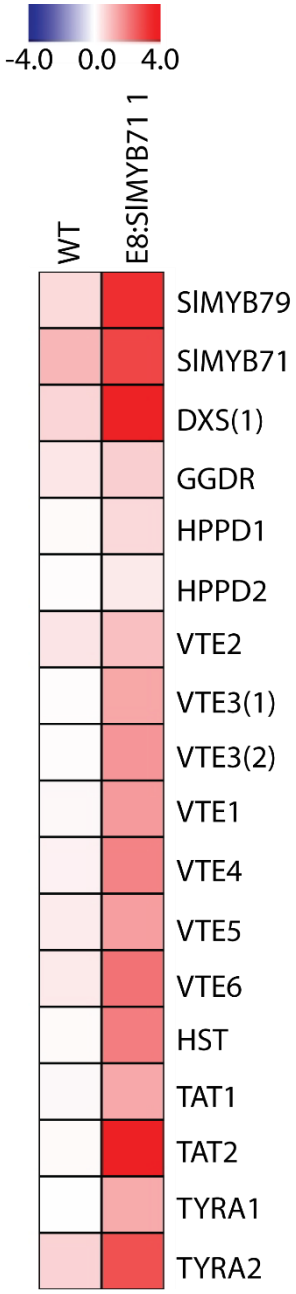


Figure 5-26 Log_2 expression heatmap of genes encoding enzymes in the VTE, MEP and SK pathways, *SlMYB71* and *SlMYB79* for the E8:SlMYB71 overexpressing lines (n=3). The scale bar shows the log_2 values from 0.0 to 12.0. All values are relative to WT at the MG stage. All genes were normalised to the SlCAC reference gene.

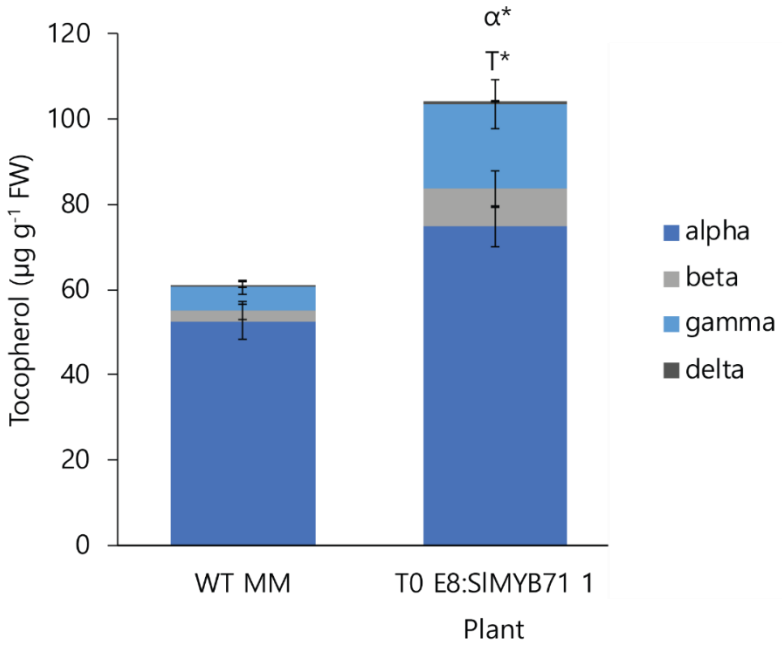


Figure 5-27 Tocopherol contents (µg g^{-1} fresh weight) of Moneymaker T0 E8:SlMY71 stable transformant tomato fruit. Tocopherol was measured µg per g of tomato fresh weight (FW) (n=4). Alpha tocopherol is dark blue, beta tocopherol is light grey, gamma tocopherol is light blue and delta tocopherol is dark grey. The errors bars represent standard errors of the mean. The letters represent the statistical significance for the different tocopherol forms α = alpha, β= beta, γ = gamma, δ = delta and T = total. The stars indicate statistically significant differences measured using students t-tests, (*) = $p < 0.05$, (**) = $p < 0.01$

SlMYB71 (figure 5-26), and *SlMYB71* expression in fruits overexpressing SlMYB79 was repressed (figure 5-15). This suggests that SlMYB79 may be a repressor of SlMYB71 and these two TFs might both regulate expression of their encoding genes.

5.3.6 Expression of SlMYB79 and SlMYB71 during tomato fruit development and ripening

SlMYB79 was not expressed in the trans-eQTL IL9-3-2 (table 5-1). The paralog of SlMYB79; *SlMYB71*, was also not expressed in the trans-eQTL IL9-3-2 (table 5-1). These genes are expressed in roots, based on eFP browser data (appendix figure 5-2). However, in the eFP browser data, neither *SlMYB79* nor *SlMYB71* are reported as expressed to high levels in tomato fruit. Therefore, I analysed the expression of these genes in tomatoes of the IL parent *S.lycopersicum* cv. M82 during development and ripening to determine their patterns of expression.

SlMYB79 was highly expressed in the pericarp tissue, and it was induced at the B stage in *S.lycopersicum* cv. M82 tomato fruit (figure 5-28). In the epidermis, expression of *SlMYB79* was low in the MG and B stages and then increased in B+5 fruit stage (figure 5-28). These changes in expression were not significantly different, but with further biological replicates the reproducibility of this increase might be clarified. Expression of *SlMYB71* peaked at B stage in the pericarp (figure 5-28). There was a peak in expression of *SlMYB71* in the epidermis at the B stage (figure 5-28). The expression of *SlMYB71* then was reduced in the later stages of fruit ripening (figure 5-28).

The genes encoding enzymes of the VTE pathway were down-regulated compared to M82 (table 5-1). This suggested that SlMYB79 could be a transcriptional activator of the genes encoding enzymes of the VTE pathway. However, transient silencing of *SlMYB79* and CRISPR knock-out lines suggested that this transcriptional regulator increased tocopherol production. My analyses

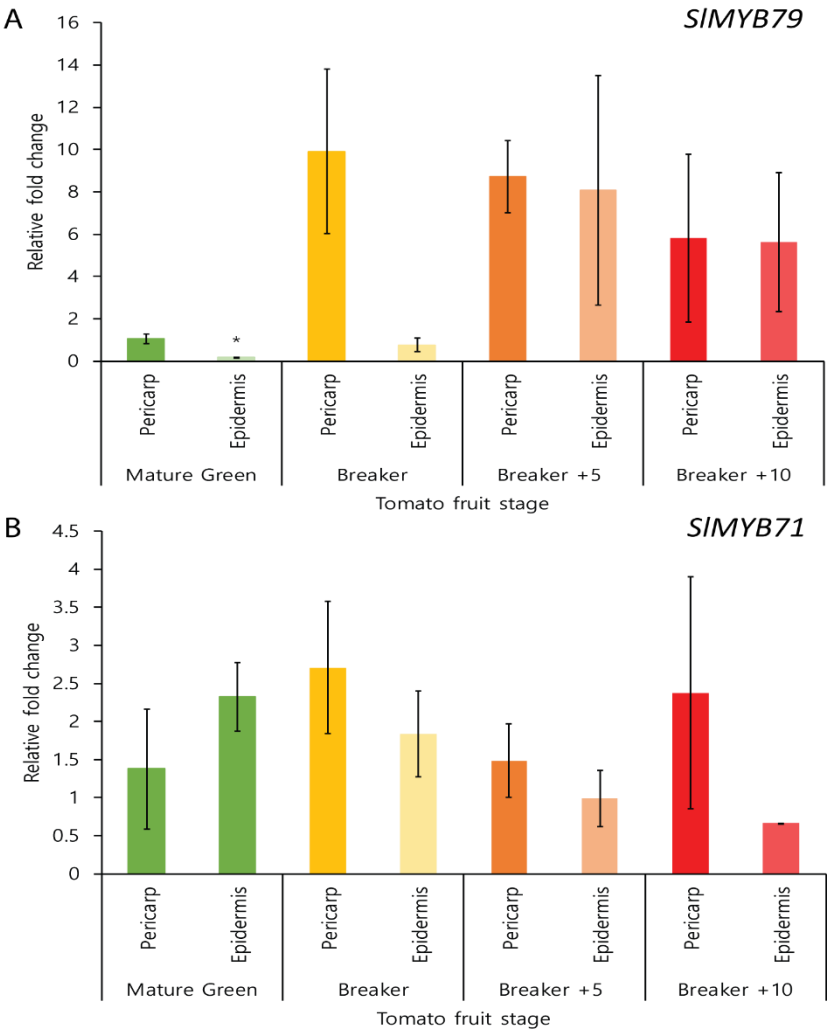


Figure 5-28 Relative expression of (A) *SlMYB79* and (B) *SlMYB71* during *S.lycopersicum* cv. M82 fruit development. Tissues of the pericarp and epidermis were analysed at 4 stages of tomato fruit development and ripening, which included; mature green, breaker, breaker +5 days and breaker +10 days (n=3). The errors bars represent the standard errors of the mean. The relative fold changes of expression at each stage were compared to the mature green stage for normalisation. Statistical significance was calculated using students t-tests and represented by the stars; (*) = $p < 0.05$, Tukey tests showed no significant differences across tomato development, between any tissue of any stage for either gene.

suggested that SlMYB79 may negatively regulate the expression of genes encoding enzymes of the MEP pathway and possibly *VTE6*. RNA sequencing data of *S.pennellii* fruit confirmed that *SlMYB79* was not expressed (table 5-4). This coupled with the data from the CRISPR knock-out lines of *SlMYB79* and *E8:SlMYB79* overexpression tomatoes suggested that SlMYB79 is not likely to be the causal TF resulting in the downregulation of the genes encoding enzymes of the VTE pathway in the trans-eQTL IL9-3-2. *SlMYB79* is expressed in *S.lycopersicum* cv. M82, but *SpMYB79* is not expressed in *S.pennellii* (table 5-4). Therefore, it is unlikely that the differences between transcript levels of VTE biosynthetic genes in the trans-eQTL IL9-3-2 and M82 are due to differences in SpMYB79 expression or activity.

5.3.7 Analysis of protein alignments between SlMYB79 and SlMYB71

Alignments of SlMYB79 and SpMYB79 proteins indicated that the MYB DNA binding domain is identical in these proteins (figure 5-29). There is a two AA deletion in the glutamine-rich region of SpMYB79 in the C-terminus, compared to SlMYB79 (figure 5-29). However, this difference is very small and the other acidic domains in the C terminus are identical. Therefore, this analysis suggested that *SlMYB79* is probably differentially expressed between *S.lycopersicum* cv. M82 and *S.pennellii* (table 5-4). SlMYB79 and SpMYB79 are structurally very similar and therefore any differences in their activity in fruit between *S.lycopersicum* and *S.pennellii* are likely to be due to differences in the regulation of their expression, rather than differences in protein structure.

Protein alignments between SlMYB79 and SlMYB71 showed that they have a fairly conserved MYB domain with some conservative substitutions in the C-terminal (figure 5-24). Generally, the protein alignments of SlMYB79 and SlMYB71 show that the transcriptional activation domains (TADs) are similar, which might suggest they have similar functions. However, my data has suggested that SlMYB79 is a repressor of SlMYB71, but there are no obvious changes in TADs between the protein sequences. Additionally, protein alignments of the *Solanum*

Table 5-4 RNA sequencing data of fruit of *S.lycopersicum* cv. M82 and *S.pennellii* showing the RPKM-normalised values compared. The table shows the expression of *SlMYB79* and *SlMYB71* and the statistical significance (displayed as a p value) between the two tomato species. These RNA sequencing data were publicly available (Koenig et al., 2013).

Solgenomics identifier	Gene name	RKPM normalized value		P value between *S.lycopersicum* cv. M82 and *S.pennellii*
		S.lycopersicum cv. M82	*S.pennellii*	
Solyc09g090790	SlMYB79	3.51	0	4.04×10^{-6}
Solyc05g053150	SlMYB71	0	0	0

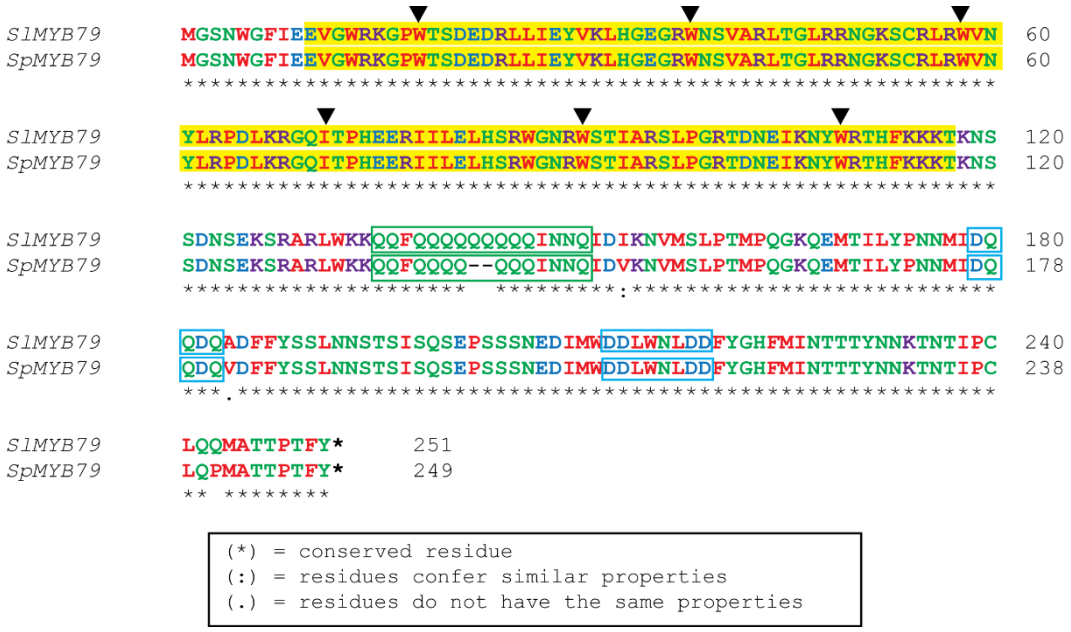


Figure 5-29 Protein alignments of R2R3 MYB TFs, SlMYB79 and SpMYB79, generated using clustal omega (Larkin et al., 2007). The yellow box highlights the MYB DNA binding domain. The black arrows heads show the conserved tryptophan residues, of which three are associated with each MYB repeat. The first hydrophobic reside in R3 is I, rather than W in plant R2R3 MYB proteins (Martin and Paz-Arez, 1997). The amino acid residues colours are as follows; red = small and hydrophobic, including tyrosine (Y), blue = acidic, magenta = basic histidine (H), green = hydroxyl, sulfhydryl, and amine, including glycine (G). Transcriptional activation domains in MYB TFs are associated with amino acids (proline and glutamine), or with acidic areas (Jin and Martin, 1999). The glutamine rich area is shown in the green box and the acidic areas are shown in blue boxes.

paralogs with the *Arabidopsis* orthologs show that they all share very similar MYB domains (figure 5-25). The C-terminal domains of the *Arabidopsis* and *Solanum* homologs show that there are several regions that are either glutamine rich or contain acidic domains, which are associated with TADs of MYB TFs (Jin and Martin, 1999, Martin and Paz-Arez, 1997). There are no clear differences between the TADs of the orthologs of SlMYB71 and AtMYB79 compared to SlMYB79 and AtMYB71 (figure 5-25). In some mammalian MYB TFs, mutations in the C-terminus have resulted in deletions of TADs, which can alter the function of MYB TFs (Chen and Lipsick, 1993). The TADs of SlMYB79 and SlMYB71 are similar and therefore any differences in TADs sequences are unlikely to be the reason why SlMYB79 represses SlMYB71. Therefore, other regulatory elements might bind to the AAs in the C-terminus, which have a role in altering TF activity.

Protein alignments of SlMYB71 and SpMYB71 showed that they have identical MYB DNA binding domains (figure 5-30). The C terminal activation domains are also very similar. However, there is a 16 AA difference in SpMYB71 compared to SlMYB71 in the C-terminal domain (figure 5-30), but this difference does not affect the acidic C-terminal activation domains. Therefore, SlMYB71 and SpMYB71 may be differentially expressed. *SpMYB71* is not expressed in the trans-eQTL IL9-3-2 (table 5-1) and was not expressed in RNA sequencing data of *S.pennellii* or *S.lycopersicum* cv. M82 (table 5-4). However, it is clear from my data that *SlMYB71* is expressed in tomato fruit (figure 5-4), although expression was not observed in B+10 tomatoes in the RNA sequencing data (table 5-1 and table 5-4). Overexpression of SMYB71 in tomato fruit, resulted in higher total tocopherol contents (figure 5-27) and higher expression of *SlMYB79* (figure 5-26), which suggested that SlMYB71 may activate expression of SlMYB79.

5.3.8 Analysis of promoters of orthologs of SlMYB79 and SlMYB71

All available data suggest that *SpMYB79* is not expressed in *S.pennellii* fruit. Whereas, *SlMYB79* is expressed in *S.lycopersicum* fruit, and this differential

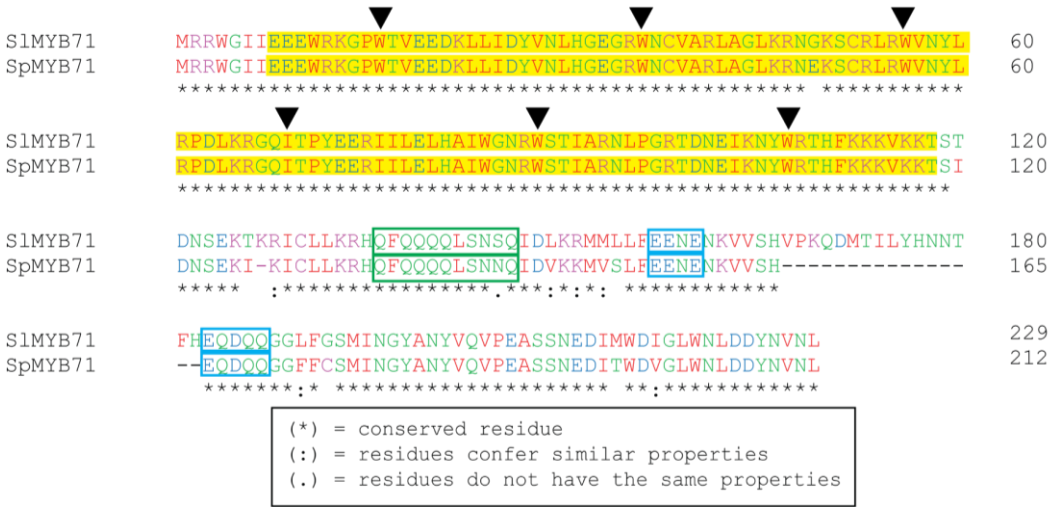


Figure 5-30 Protein alignments of R2R3 MYB TFs, SlMYB71 and SpMYB71, generated using clustal omega (Larkin et al., 2007). The yellow box highlights the MYB DNA binding domain. The black arrows heads show the conserved tryptophan residues, of which three are associated with each MYB repeat. The first hydrophobic reside in R3 is I, rather than W in plant R2R3 MYB proteins (Martin and Paz-Arez, 1997). The amino acid residues colours are as follows; red = small and hydrophobic, including tyrosine (Y), blue = acidic, magenta = basic histidine (H), green = hydroxyl, sulfhydryl, and amine, including glycine (G). Transcriptional activation domains in MYB TFs are associated with amino acids (proline and glutamine), or with acidic areas (Jin and Martin, 1999). The glutamine rich area is shown in the green box and the acidic areas are shown in blue boxes.

expression may impact the activity of SlMYB79 between the two species. Promoter comparisons between SlMYB79 and SpMYB79 showed that there were several differences in the region 500bp upstream of the ATG start codon (figure 5-31). The most prominent difference is a 10bp insertion in the SpMYB79 promoter at the -343 position, which includes a GARP binding motif near to a GATA TF binding motif observed in the SlMYB79 promoter. There were also two other TF binding motifs (**T**eosinte branched 1, **C**ycloidea and **P**CF (TCP) TF domain and **H**omeo**d**omain leucine **zip**per (HD-ZIP) domain) (figure 5-31). This suggests that differential regulation might exist between *S.lycopersicum* cv. M82 and *S.pennellii.*

Promoter alignments of 500bp upstream of the ATG start codon in SlMYB71 and SpMYB71 showed that there were a few differences in the DNA sequences (figure 5-32). There are several differing TF binding motifs between the SlMYB71 and SpMYB71 promoters, including; a GARP TF binding domain near to a GATA TF binding motif in *S.lycopersicum* that is not present in *S.pennellii* and a **N**o-**A**pical **M**eristem, ***A****rabidopsis* transcriptional activation factor and Cup-shaped cotyledon (NAC/NAM) binding motif in *S.pennellii,* which is not present in the *S.lycopersicum* promoter (figure 5-32). It is possible that, both SlMYB71 and SlMYB79 may be differentially regulated by different TFs, which affect their activity within different tomato species.

These data suggest that *SlMYB79* and *SlMYB71* are differentially regulated from their orthologs in *S.pennellii.* This coupled with the fact that SlMYB79 acts as a negative regulator of SlMYB71 suggests that a model can be proposed (figure 5-33). The model shows that in *S.lycopersicum* fruit, SlMYB79 might be regulated by different TFs than in *S.pennellii* fruit. SlMYB79 acts as a negative regulator of the MEP pathway and SlMYB71, which results in higher tocopherol levels in the CRISPR *SlMYB79* knock-out lines (figure 5-7). In *S.pennellii,* it is not clear if SpMYB79 could regulate the expression of *SpMYB71* since *SpMYB79* is not expressed in fruit. The model suggests that the differences in binding motifs in the promoter of SpMYB79,

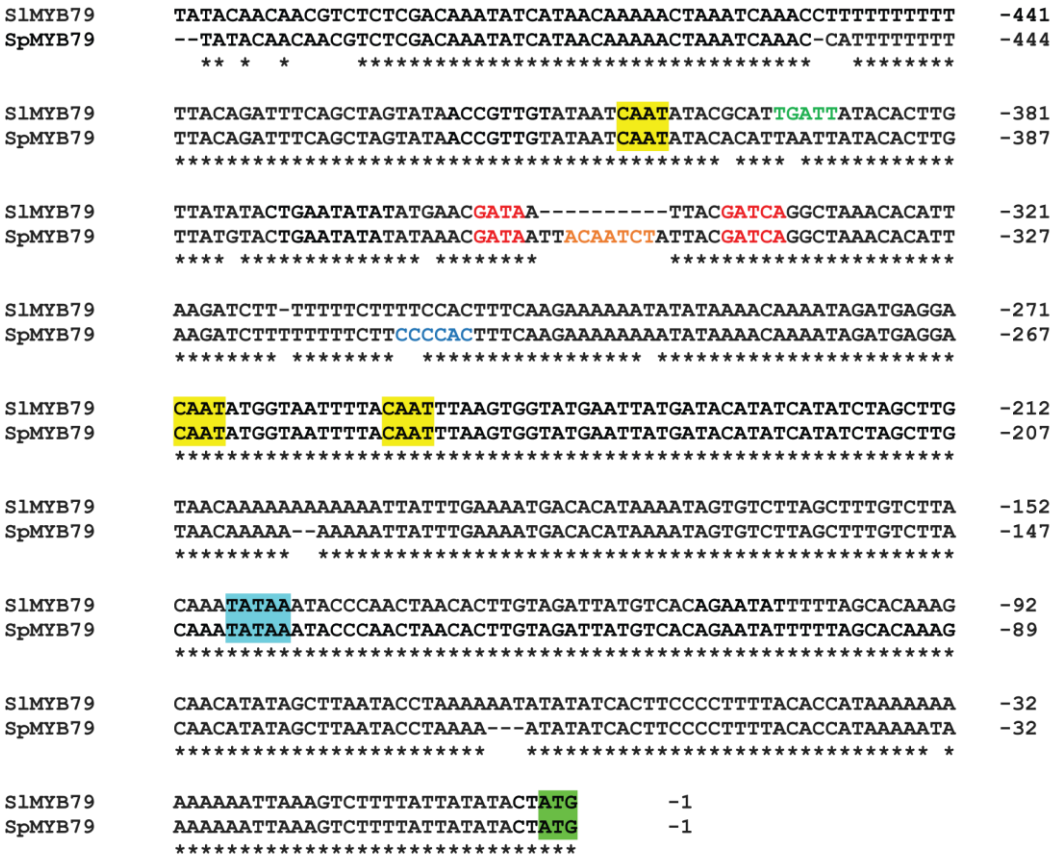


Figure 5-31 Promoter alignments of SlMYB79 and SpMYB79. The blue box represents the TATA (TATAA) box, which lies upstream of the start codon (ATG) shown in the green box. The yellow boxes show the CAAT boxes associated with transcription. The coloured DNA motifs represent the following: orange = GARP TF binding motif, red = GATA TF binding motif, green = HD-ZIP TF binding motifs and blue = TCP TF binding motifs. These motifs were identified based on the motifs identified by Franco-Zorilla et al. (2014).

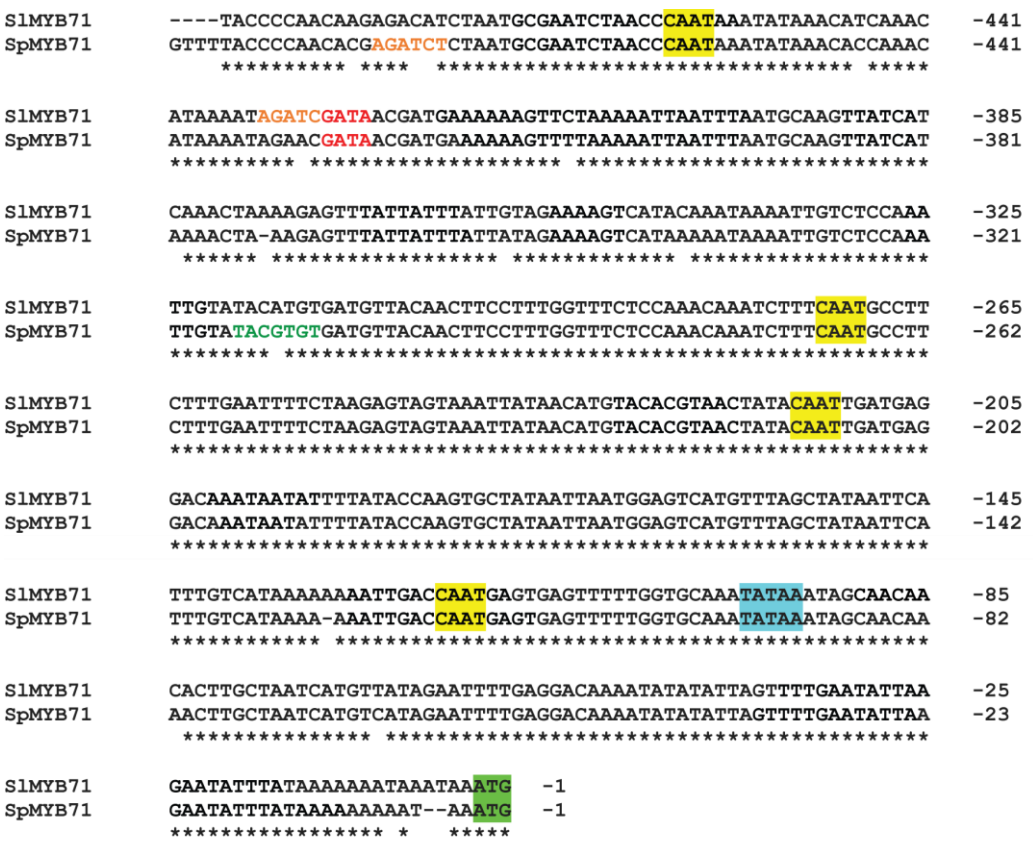


Figure 5-32 Promoter alignments of SlMYB71 and SpMYB71. The blue box represents the TATA (TATAA) box, which is upstream of the start codon (ATG) in the green box. The yellow boxes show the CAAT boxes associated with transcription. The coloured DNA motifs represent the following: orange = GARP TF binding motif, red = GATA TF binding motif and green = NAC TF binding motifs. These motifs were identified based on the motifs identified by Franco-Zorilla et al. (2014).

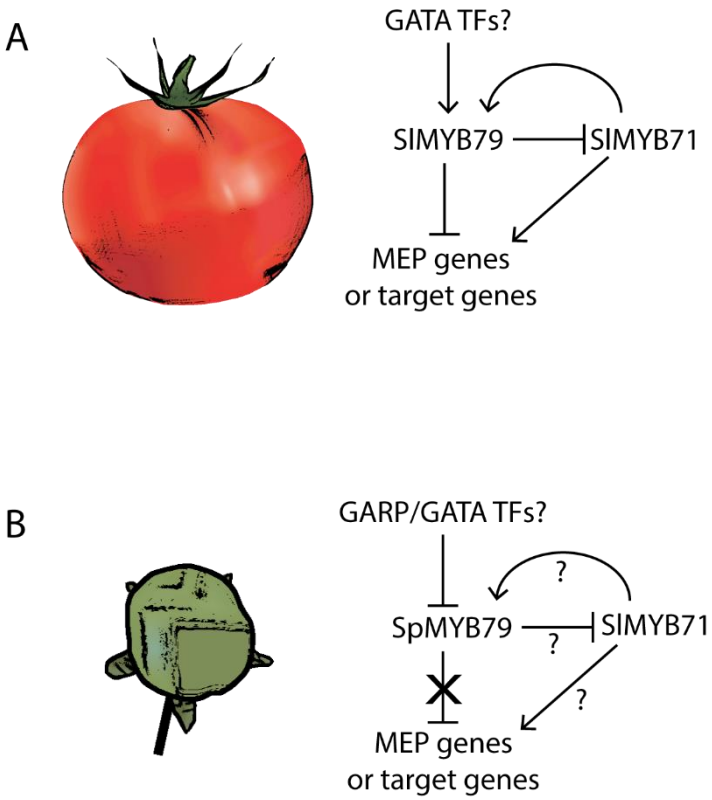


Figure 5-33 Model of possible SlMYB79 regulatory mechanism of genes encoding enzymes of the MEP pathway, or other target genes. The model shows the possible regulation of SlMYB79 in (A) *S.lycopersicum* and SpMYB79 in (B) *S.pennellii* fruits. Tomato fruits overexpressing SlMYB79/SlMYB71 and knockout lines of SlMYB79 have suggested possible interactions between SlMYB79 and SlMYB71. It is not clear if MYB79 regulates the genes encoding enzymes of the MEP pathway directly. The arrow heads indicate possible activation by TFs, and the T bars indicate possible negative regulation by TFs. The question marks represent unknown mechanisms.

compared to SlMYB79 might result in the lack of *SpMYB79* expression in fruit. As a consequence, the genes encoding enzymes of the MEP pathway are not repressed in fruit at ripening and this may result in the high levels of tocopherols observed in *S.pennellii* fruits (described in chapter 4).

5.4 Discussion

5.4.1 Gene mining of the trans-eQTL IL9-3-2 and transient assays revealed one gene encoding a MYB TF that might regulate VTE biosynthesis transcriptionally

Gene mining revealed eleven genes encoding candidate TFs that were identified from the trans-eQTL IL9-3-2. Out of these eleven genes, four genes encoding candidate TFs showed a similar or inverse co-expression pattern with the genes encoding enzymes of the VTE pathway. These four genes encoding candidate TFs were transiently silenced using VIGS. The VIGS data suggested that SlMYB79 was a candidate transcriptional regulator of the VTE biosynthetic pathway, as expression of several enzymes were upregulated in the silenced sectors of the VIGS fruit, and α- and total tocopherol levels were increased. These data suggested that SlMYB79 may be a transcriptional repressor of VTE biosynthesis.

The expression of genes encoding enzymes of the VTE pathway was down-regulated in the trans-eQTL IL9-3-2. The RNA sequencing data showed that *SlMYB79* was not expressed in this IL region, compared to other ILs and the *S.pennellii* x *S.lycopersicum* cv. M82 IL parent, *S.lycopersicum* cv. M82. This suggested that SlMYB79 should be a transcriptional activator of the pathway, however my VIGS data suggested otherwise. Therefore, *SpMYB79* is probably not the gene causing the down-regulation of the genes encoding enzymes of the VTE pathway in the trans-eQTL IL9-3-2. This trans-eQTL was also identified as a trans-eQTL of lycopene biosynthesis (Li, 2018) and for the MEP pathway (chapter 3). Genes encoding enzymes of the MEP pathway were highly expressed in this region *(DXS(1)* and *GGPS(1)),* which is described in chapter 3.

5.4.2 The SlMYB79 CRISPR knock-out lines suggested that SlMYB79 is a transcriptional repressor

Two knock-out CRISPR lines were generated for SlMYB79. *Myb79_23* carries a 10 nt deletion after the MYB domain which would result in a premature termination of the SlMYB79 protein lacking the entire C-terminus of the TF. I predict this would be a knock-out mutation of SlMYB79. *Myb79_16* has a four nt deletion near to the start codon, which resulted in a predicted protein with no MYB domain and is therefore likely to be a knock-out mutation. Both lines showed that compositional tocopherol levels in fruit were different from WT fruit. However, only *myb79_23* had significantly higher total tocopherol levels. Total tocopherol levels did not reach as high levels as seen in the SlMYB79 overexpression lines in Microtom fruit. However, both the CRISPR lines in Moneymaker fruit had higher tocopherol levels, which suggests that SlMYB79 is an effective transcriptional repressor of VTE biosynthesis. These data fit with the observed phenotype seen in the VIGS data, which also suggested that SlMYB79 was a transcriptional repressor.

Expression of *ISPE, ISPF, IPI(1), IPI(2), GGPS(2)* and *GGPS(3)* were upregulated in all the CRISPR lines. Most interestingly, expression of *DAHPS1, SH-DHQS2, EPSPS, PAT, CM1* and *TYRA1* were also upregulated in the CRISPR lines. This suggests that knocking-out SlMYB79 function by CRISPR altered gene expression profiles for the MEP, VTE and SK pathways in fruit. If SlMYB79 is a repressor, the knock-out lines could increase the expression of genes in the MEP and SK pathways, or the induction of SK pathway genes could be due to cross talk between the pathways. An increase in the flux of the MEP pathway might cause an increase in flux along the SK pathway, as a response through metabolic control of transcript levels to maintain tocopherol biosynthesis in tomato fruit.

It is clear in the *myb79_23* knock-out that *VTE5* was upregulated. Tomato fruits modulate *VTE5* expression for phytyl diphosphate (PDP) regeneration from free phytol derived from the chlorophyll degradation pathway (Valentin et al., 2006). This suggests PDP levels are important for tocopherol biosynthesis in these

tomato fruits. Tocopherol is an important antioxidant that scavenges singlet oxygen produced from photosynthesis and prevents lipid peroxidation of poly-unsaturated fatty acids (PUFAs) in maintenance of membrane integrity (Sattler et al., 2004, Havaux et al., 2005). Therefore, it is essential that tocopherol is synthesised to reduce oxidative stresses in photosynthetic tissues.

5.4.3 Tomato fruit overexpressing SlMYB79 have higher tocopherol contents

Overexpression of *SlMYB79*, under the control of the E8 promoter, in Microtom tomato fruit increased tocopherol levels significantly. α-tocopherol was the most abundant form of tocopherol found in these fruits and was increased 1.41 and 1.31-fold in MT1 and MT2 B+10 fruit, respectively, compared to WT. α-tocopherol is the most bioactive form of tocopherol in humans because it is preferentially retained by the α-tocopherol transfer protein (αTTP) (Schneider, 2005, Hosomi et al., 1997), therefore these tomatoes can be considered to be nutritionally enhanced. The composition of tocopherols in the WT fruit and fruit overexpressing SlMYB79 was similar when they were expressed as percentages. This suggested that SlMYB79 may alter flux of VTE substrates that are available for VTE biosynthesis, rather than target the genes encoding enzymes of the VTE pathway directly. However, expression of *VTE6* in the B+5 tomato fruits overexpressing *SlMYB79* was repressed compared to WT. This implies that SlMYB79 might be transcriptional repressor of *VTE6.* This gene encodes for an enzyme that catalyses the phosphorylation of phytyl phosphate to PDP. This is part of the chlorophyll degradation pathway for regeneration of free phytol for VTE biosynthesis. This step is important as tocopherols are nearly completely abolished in *Arabidopsis vte6* knock-out mutants (vom Dorp et al., 2015). *Arabidopsis vte6* mutants struggle to grow photo autotrophically. This phenotype can be rescued by crossing with *Arabidopsis vte5* mutants, which suggests that plants cannot accumulate phytyl phosphate (vom Dorp et al., 2015) and therefore, *VTE5* and *VTE6* expression is tightly controlled (Quadrana et al., 2013). In the ripe tomato fruits overexpressing *SlMYB79,* expression of *VTE5* was not significantly

different from the WT fruits, which suggests that SlMYB79 does not regulate *VTE5* expression.

Generally, the composition of the tocopherol vitamers were similar in WT and *SlMYB79* over expression lines. Therefore, changes in expression of genes encoding enzymes of pathways upstream of VTE biosynthesis might explain the increase in tocopherol observed in the fruit. Expression of genes encoding enzymes of the MEP pathway supported this hypothesis as expression of *DXR, ISPE, ISPF, IPI(1), IPI(2), GGPS(1), GGPS(2)* and *GGPS(3)* was repressed in E8:SlMYB79 MT1 and MT2 fruit at the B fruit stage, compared to WT fruit. This supports the idea that SlMYB79 is a transcriptional repressor of the genes encoding enzymes of the MEP pathway as these genes were upregulated in the CRISPR knock-out lines. However, the expression of these genes was higher in MG fruit and then declined rapidly. The expression of these genes was not measured in MG fruit in E8:SlMYB79 MT1 as the fruit were not ready at the time of this analysis. However, considering that the expression profiles for the rest of the tomato fruit stages and the metabolite profiles are similar in both lines, it is likely that these potential target genes would also be upregulated in MT1 at the MG stage.

Similarly, the transcript abundance of the *GGPS(1), GGPS(2)* and *GGPS(3)* was significantly repressed in B and B+5 fruits overexpressing SlMYB79, compared to WT. This suggested that these genes may be targets of SlMYB79. The GGPS enzymes described are different isoforms of the same enzyme. The eFP browser data indicate that *GGPS(2)* and *GGPS(3)* are expressed in tomato fruits, whereas *GGPS(1)* is expressed more highly in leaves. The repression of all *GGPS* isoforms supports the idea that SlMYB79 may be transcriptional regulator of the MEP pathway in different tissues. Genes encoding enzymes of the SK pathway were not differentially expressed, but several were upregulated in the CRISPR knock-out lines. Also, this fits with the idea that SlMYB79 is not the causal candidate TF responsible for the downregulation of genes encoding enzymes of the VTE pathway in the trans-eQTL IL9-3-2. This trans-eQTL was also shown to be a trans-

eQTL for the MEP pathway and two genes were highly expressed in this region *(DXS(1)* and *GGPS(1)),* which is described in chapter 3. Expression of *GGPS(1)* was also highly repressed in the tomato fruit overexpressing *SlMYB79*. Therefore, the absence of *SlMYB79* expression in this trans-eQTL region might result in tomato fruits expressing genes encoding enzymes of the MEP pathway to high levels.

Increased expression of the same genes encoding enzymes in the MEP pathway was observed in the CRISPR lines and the same genes were repressed in fruit overexpressing *SlMYB79*, suggesting that SlMYB79 has a role in regulating the MEP pathway. Of course, other TFs probably also regulate this pathway transcriptionally. SlMYB79 likely affects pathway flux through modulation of transcript levels of genes encoding MEP pathway enzymes or higher pathway enzymes that alter substrate flux in the MEP pathway for VTE biosynthesis and other pathways. This may affect substrate availability of PDP, which is the substrate for VTE biosynthesis (Cahoon et al., 2003, Yang et al., 2011). It is not clear from my data, whether SlMYB79 is a direct or indirect negative regulator of expression of genes encoding enzymes of the MEP pathway. The activity of SlMYB79 as a direct repressor of the genes encoding enzymes of the MEP pathway and *VTE6* needs to be tested in future experiments.

The TEAC assay results supported the hypothesis that SlMYB79 may have other targets that impact the activity of other metabolic pathways that are not related to the VTE or MEP pathways, or increases in pathway flux may alter other metabolic biosynthetic pathways. Fruit overexpressing *SlMYB79* had a significantly higher water-soluble antioxidant capacity than WT fruit. The products of the MEP and VTE biosynthesis are lipophilic. Therefore, there could be other biosynthetic pathways affected by overexpression of *SlMYB79*. Further untargeted mass spectrometry (MS) (untargeted MS-MS) would reveal other possible metabolite profiles which might be affected to determine additional possible targets of SlMYB79.

5.4.4 The transcription factors SlMYB79 and SlMYB71 might regulate VTE biosynthesis indirectly

The majority of the work of my thesis has focused on SlMYB79. However, the protein alignments of SlMYB79 with its paralog (SlMYB71) showed that SlMYB71 has a truncated C terminus, compared to SlMYB79. The orthologs of SlMYB79 and SlMYB71 in *Arabidopsis* (AtMYB79 and AtMYB71) were described by Kranz et al. (1998). These authors suggested that AtMYB79 has a 106 AA deletion in its C-terminus, compared to AtMYB71. However, this was probably the result of a mis-annotation at the time of the analysis as the protein alignments of SlMYB79 and SlMYB71 with the *Arabidopsis* ortholog showed AtMYB79 and AtMYB71 have conserved MYB binding domains, and although there were several differences in the C-terminal domains, no large deletion between in AtMYB79 was apparent from comparing the proteins. The predicted MYB domains of SlMYB79 and SlMYB71 are very similar, and their C-terminal activation domains are fairly similar. This suggests that SlMYB71 likely also modulates VTE biosynthesis and should be studied to establish its role in VTE biosynthesis.

Overexpression and CRISPR lines of SlMYB79 showed that they accumulated more tocopherols in fruit, than in WT. Interestingly, the paralog of SlMYB79; *SlMYB71*, was highly expressed in the CRISPR lines, compared to WT fruit. Overexpression of *SlMYB71,* under the control of the E8 promoter in Moneymaker tomatoes, resulted in higher levels of tocopherols. Analysis of *E8:SlMYB71* fruit showed high transcript levels of *SlMYB79*, suggesting that SlMYB71 might activate expression of *SlMYB79.* The expression profiles of the genes encoding enzymes of the VTE, MEP and SK pathways were higher in the *SlMYB71* overexpression line. However, this analysis was carried out on only one line, therefore further analysis of other independent lines is needed. Expression analysis during tomato development and ripening in the next generation (T1) should reveal the relationship between SlMYB71 expression and VTE biosynthesis as well as its interaction with SlMYB79.

Protein alignments have shown that SlMYB71 is truncated relative to SlMYB79. The homolog of SlMYB71, AtMYB79 was suggested to be a dominant inhibitor of AtMYB71 (Kranz et al., 1998). Similarly, SlMYB79 may be a negative regulator of SlMYB71. This idea is reinforced by my data which showed that expression of *SlMYB71* was induced when *SlMYB79* was knocked out by CRISPR/Cas9 genome editing. Fruits overexpressing *SlMYB79* showed that expression of *SlMYB71* was significantly repressed in B+5 fruits, and was also repressed at B+10. Possibly, SlMYB79 may be a repressor of SlMYB71, but SlMYB71 may be a positive regulator of *SlMYB79* because expression of *SlMYB79* was induced when *SlMYB71* was overexpressed in tomato fruit, as summarised in the model (figure 5-33). It is also possible that these TFs may compete for the same promoter binding sites. Dual luciferase assays with promoters of the genes encoding enzymes of the MEP pathway and promoters of the paralogs would determine whether SlMYB79 and SlMYB71 can bind to MEP promoters and each other's promoters, to determine whether both these TFs regulate the genes encoding enzymes of MEP pathway or higher pathways transcriptionally.

The protein alignments of SlMYB79 and SlMYB71 showed that these proteins are very similar. Their C-terminal domains show very similar TADs, but other AAs do differ, most notably the 22AA shorter C-terminal domain of SlMYB71. Phosphorylation and ubiquitin have been shown to be important regulators of transcriptional activity of MYB TFs (Morse et al., 2009, Salghetti et al., 2001, Salghetti et al., 2000). Serine residues in C-terminal domains of MYB TFs are important AAs that are phosphorylated to increase transcriptional activation functions in pine species (Morse et al., 2009). Additionally, transcriptional activity of MYB TFs can be enhanced by ubiquitination of lysine residues near to a TAD, before the protein is degraded through the ubiquitin-proteasome pathway (Salghetti et al., 2001, Salghetti et al., 2000). Therefore, other AA residues could alter transcriptional activity of MYB TFs. This could also explain the different functionalities of SlMYB71 as an activator and SlMYB79 as a repressor of transcription in tomato fruits.

5.4.5 SlMYB79 may be a repressor in non-photosynthetic tissues

SlMYB79 was not expressed in the RNA sequencing data of the trans-eQTL IL9-3-2, therefore I checked the expression of *SlMYB79,* and its paralog *(SlMYB71)*, in fruit of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs parent *(S.lycopersicum* cv. M82) during tomato development and ripening. The expression of *SlMYB79* peaked at B stage in the pericarp tissue of *S.lycopersicum* cv. M82 fruit, whereas, in the epidermis, *SlMYB79* expression was induced in the epidermis at B+5. *SlMYB71* expression peaked at B stage in *S.lycopersicum* cv. M82 fruit and then was reduced during tomato development. The induction of expression of *SlMYB79* at the B stage, suggested that VTE levels might decline in the pericarp if SlMYB79 is a transcriptional repressor. However, tocopherol contents in the pericarp remained constant throughout tomato development and ripening (described in chapter 4), which suggested that *SlMYB79* is not the only transcriptional regulator that affects production of VTE.

Both SlMYB79 and its paralog (SlMYB71) are R2R3 type MYB TFs, which can be grouped into phylogenetic sub-groups which commonly share biological functions (Dubos et al., 2010), however SlMYB79 and SlMYB71 and their homologs in *Arabidopsis* have not been classified into a functional sub-group. From the eFP browser, these TFs may play more prominent roles in roots, as genes encoding these TFs are highly expressed in *S.lycopersicum* and *S.pennellii* root tissues. However, these data are not sensitive as the qRT-PCR data. The eFP browser data suggested that these genes are not expressed in fruit, yet my qRT-PCR data suggested that *SlMYB79* and *SlMYB71* are expressed in fruit during development and ripening. My data suggested that *SlMYB79* was not expressed to high levels in MG fruit. Expression of *SlMYB79* was not observed in the RNA sequencing data of fruit of *S.pennellii.* This coupled with the eFP browser data suggested that this repressor may be most active in non-photosynthetic tissues. Kranz et al. (1998) measured expression of all the *A.thaliana* MYB TFs, using Northern blot analysis and AtMYB71 and AtMYB79 were not expressed in response to several hormones, sucrose, nitrogen, and infections with *Pseudomonas syringae.* Therefore, the

expression of SlMYB79 has not yet been aligned with any specific physiological functions. These data suggest that SlMYB79 is a repressor in non-photosynthetic tissues, although further validation of expression is needed in IL9-3-2 to confirm this conclusion.

5.4.6 *MYB79* and *MYB71* might be differentially expressed in *S.pennellii* fruits, compared to *S.lycopersicum* cv. M82

Neither *SpMYB79* nor *SlMYB71* were expressed in the RNA sequencing data of the trans eQTL IL9-3-2, which suggested that orthologs of these genes might be differentially expressed. Additionally, RNA sequencing data of the B+10 fruit of the *S.pennellii* x *S.lycopersicum* cv.M82 IL parents revealed that *SlMYB79* was expressed in *S.lycopersicum* cv. M82, but *SpMYB79* was not expressed in *S.pennellii* fruit. *MYB71* was not expressed in fruit in either *S.pennellii* x *S.lycopersicum* cv.M82 IL parents. This suggested that *MYB79* may be differentially expressed between the IL parents. The protein alignments of the SlMYB79 and SpMYB79 proteins showed that they had very similar MYB DNA binding domains and C-terminal activation domains. However, the promoter sequences of SlMY79 and SpMYB79 were not identical, which means that the genes might be differentially expressed. They differed specifically by the presence of a GARP binding motif, close to a GATA binding motif in the promoter of SpMYB79 which was missing in the promoter of SlMYB79.

Promoter alignments, approximately 500bp upstream of the ATG start codon of SlMYB79 and SpMYB79 showed that there was a 10 bp insertion of a GARP binding motif near to a GATA TF binding motif in the *S.pennellii* SpMYB79 promoter. The insertion may affect binding of a TF to the GATA-box present in *S.lycopersicum.* Both GARP TFs and GATA TFs have been associated with photosynthesis (Bi et al., 2005, Waters et al., 2009). In *Arabidopsis,* GARP TFs called *GOLDEN2-like (GLK)* are associated with chloroplast development. *Arabidopsis glk1 and glk2* mutants are unable to form grana of the thylakoid membranes in chloroplasts, which are the site for light-dependent reactions of

photosynthesis (Waters et al., 2008, Yasumura et al., 2005). Analysis of the double mutant, *glk1 glk2,* showed that the GARP TFs bind to promoters of many genes encoding photosynthetic proteins, such as proteins involved in the light-harvesting complex (Waters et al., 2009). In *Arabidopsis,* GATA TFs are type IV zinc finger proteins that have specific DNA domain ($CX_2CX_{17-20}CX_2C$) and a basic DNA binding C-terminus (Behringer and Schwechheimer, 2015). These TFs were originally identified as regulators of genes encoding enzymes of nitrogen metabolism during ammonium deprivation in plants (Reyes et al., 2004). GATA TF binding motifs have been identified in promoters of light-responsive genes; encoding the small subunit of ribulose-1,5-biphosphate carboxylase-oxygenase (RbcS) and chlorophyll a/b binding proteins (cab) (Gilmartin et al., 1990). These nuclear genes encode for enzymes and proteins that are involved in photosynthesis. RbcS encodes the small subunit of the enzyme RbcS, also known as RuBisCo, which fixes carbon during photosynthesis. Cab proteins are membrane-bound proteins that form a larger part of the photosystem II, which generates nicotinamide adenine dinucleotide phosphate (NADPH) and adenosine triphosphate (ATP). These proteins are important for the fixation of carbon during photosynthesis. A GATA TF called GATA factor, nitrate-inducible, carbon metabolism-involved (GNC) is involved in chlorophyll biosynthesis and is induced by nitrate. The *Arabidopsis gnc* knock-out mutant shows reduced chlorophyll in leaves (Bi et al., 2005). GNC is also regulated by phytochrome interacting factor 3 (PIF3), which is a light-responsive bHLH TF (Richter et al., 2010). PIFs are also known to be light-responsive regulators of the MEP pathway (Toledo-Ortiz et al., 2014, Gangappa and Kumar, 2017). Chlorophyll degradation is an important part of the VTE salvage pathway, to regenerate PDP from free phytol released during chlorophyll degradation. Therefore, the insertion of the GARP binding motif near to the GATA TF binding site in the *S.pennellii* could provide a plausible link between VTE biosynthesis and TFs regulating chlorophyll synthesis. *S.pennellii* fruits do not ripen and accumulate carotenoids like *S.lycopersicum* fruit and remain green. Therefore, *S.pennellii* fruit are photosynthetically active and they produce more tocochromanols than

S.lycopersicum cv. M82 fruit (described in chapter 4). Possibly, SlMYB79 is induced by GATA TFs in *S.lycopersicum* cv. M82 but SpMYB79 is not induced due to the disruption of the GATA box, or negative regulation by GARP TFs in *S.pennellii.* My data suggest that SlMYB79 is a repressor of genes encoding enzymes of the MEP pathway and VTE6 and it is not expressed in green *S.pennellii* fruit. This suggests that in *S.pennellii* fruit are able to accumulate more tocochromanols because SpMYB79 is not induced. This also supports the idea that SlMYB79 is a repressor operating in non-photosynthetic tissues.

Protein alignments of SlMYB71 and SpMYB71 showed that these proteins share highly similar MYB DNA binding domains and C-terminal domains. Promoter alignments of SlMYB71 show that there are several point mutations between the promoters, but no large differences, equivalent to those between the SlMYB79 and the SpMYB79 promoters. Neither *SlMYB71* nor *SpMYB71* were expressed in *S.pennellii* or *S.lycopersicum* cv. M82 fruit at the time point that these fruits were harvested for the RNA sequencing data. However, my own data showed that *SlMYB71* is expressed in *S.lycopersicum* cv. M82 fruit during tomato development and ripening, and this might mean that *SpMYB71* is also expressed in *S.pennellii* fruit. I need to confirm the expression of *SlMYB71* in *S.pennellii* experimentally. It is also possible that differences in the binding motifs between the promoters of *SlMYB71* and *SpMYB71* result in differential expression between *S.lycopersicum* cv. M82 and *S.pennellii.* Differential expression of SlMYB71 could also result in altered tocopherol levels, as overexpression of *SlMYB71* in fruit resulted in higher total tocopherol levels and induction of the genes in the MEP and SK pathways. It is clear from these data that *MYB79* and *MYB71* are likely to be differentially expressed between *S.lycopersicum* and *S.pennellii* which results in altered tocopherol contents between *S.lycopersicum* and *S.pennellii* fruit.

Overall, my data suggest that SlMYB79 is a repressor of *SlMYB71* and the genes encoding enzymes in the MEP pathway, which is summarised in the model (figure 5-33).

5.4.7 Metabolic engineering of SMYB79 leads to nutritionally enhanced tomatoes

Both fruit-specific overexpression and CRISPR lines of *SlMYB79*, overexpression lines of *SlMYB79* and overexpression lines of *SlMYB71* have shown that tocopherol contents can be manipulated, however, these genes are probably part of a wider transcriptional network that controls VTE biosynthesis and/or the supply of precursors for VTE biosynthesis.

SlMYB79 was overexpressed in Microtom fruit and the *SlMYB79* CRISPR genome editing was completed in Moneymaker fruit. Therefore, these stable transformations were completed in different tomato backgrounds. Further analysis of SlMYB79 needs to be completed in a Moneymaker background to ensure that the tomato background is not affecting the effects of *SlMYB79* on tocopherol levels.

α-tocopherol was the most abundant form of tocopherol observed in Microtom fruit overexpressing *SlMYB79* and the tomatoes are nutritionally enhanced. α-tocopherol is preferentially retained by the human body by the hepatic α-tocopherol transfer protein (Lim and Traber, 2007, Hosomi et al., 1997). Therefore, α-tocopherol is often mistaken for being the most potent antioxidant in humans, but actually this is because it is more bioavailable than other tocopherol vitamers (Traber and Atkinson, 2007). High levels of VTE in these nutritionally enhanced tomatoes might provide alternative sources of α-tocopherol in the human diet.

The levels of tocopherols are higher in fruit overexpressing *SlMYB79* than in tomato fruits of *SlORESARA1 (ORE1)* RNAi lines (Lira et al., 2017). SlORE1 is a NAC TF that regulates senescence and therefore RNAi lines have prolonged photosynthesis in leaves (Lira et al., 2017), which is important as chlorophyll breakdown increases the availability of PDP for VTE biosynthesis (vom Dorp et al., 2015). My data suggested that tocopherol levels were higher in lines with engineered TF levels, than achieved using RNAi approaches to increase VTE levels by manipulating chlorophyll breakdown.

This chapter has identified a transcriptional regulator which probably alters substrate availability of precursors for VTE biosynthesis. It is not clear from these data whether SlMYB79 regulates the MEP pathway directly, or indirectly via modulation of other transcriptional regulators, or by directly altering activity of other pathways, for example those involved in supplying substrates. Therefore, to understand the role of SlMYB79 in VTE biosynthesis, untargeted MS-MS needs to be carried out to determine whether any other pathways are affected. Dual luciferase assays would determine whether SlMYB79 and SlMYB71 regulate genes encoding enzymes of the MEP pathway and VTE biosynthesis, directly. Dual luciferase assays with SlMYB79 and SlMYB71 promoters with SlMYB71 and SlMYB79, respectively would also elucidate whether these TFs control each other's expression. Furthermore, chromatin immunoprecipitation (ChIP)-sequencing would allow the identification of direct targets of SlMYB79. Consequently, my trans-eQTL analysis has identified two transcriptional regulators that can be engineered to increase VTE levels in tomatoes to provide nutritionally enhanced tomatoes. However, the details of their mechanisms of action require further investigation.

Chapter 6:

Transient screening of candidate transcription factors from the trans-eQTL IL6-2-2

Chapter 6: Transient screening of candidate transcription factors from the trans-eQTL IL6-2-2

6.1 Introduction

6.1.1 The trans-eQTL analyses elucidated two regions for candidate gene screening

The trans-eQTL analyses of the *Solanum pennellii* x *Solanum lycopersicum* cv. M82 introgression lines (ILs) identified two trans-eQTLs for further investigation (trans-eQTL IL9-3 and trans-eQTL IL6-2-2) (as described in chapter 3). This chapter describes the characterisation of candidate transcription factors (TFs) lying in the trans-eQTL 6-2-2 region, using co-expression analyses of TFs with genes encoding enzymes in the Vitamin E (VTE) pathway. Using transient assays, I have characterised the three TFs regulating the VTE pathway transcriptionally.

6.1.2 Transcriptional regulation of the VTE pathway

Transcriptional regulators of VTE biosynthesis have not yet been identified. Studies have suggested that overexpression of TFs that regulate the pathways that supply substrates to the VTE pathway can alter VTE contents (Toledo-Ortiz et al., 2010, Enfissi et al., 2010, Enfissi et al., 2005, Davuluri et al., 2005). Therefore, transcriptional regulation of VTE biosynthesis is likely complex, and requires further characterisation. This chapter aimed to characterise candidate TFs that have been predicted to transcriptionally regulate the VTE pathway, using transient assays.

6.2 Materials and Methods

6.2.1 Candidate gene mining of the trans-eQTL IL6-2-2

The trans-eQTL IL6-2-2 IL region was mined for candidate TFs putatively regulating the VTE pathway genes. I used co-expression analysis of the genes encoding the enzymes in the VTE pathway with the candidate TF genes to reduce the number of TFs taken forward for further analysis. A full description of how the trans-eQTL IL6-2-2 was identified is provided in chapter 3.

6.2.2 Viral Induced Gene Silencing (VIGS)

VIGS fragments were approximately 300 nucleotides in length and designed using the Solgenomics network (SGN) VIGS tool (Fernandez-Pozo et al., 2015b). VIGS fragments were cloned using Phusion PCR and transformed into the pTRV2 *Del/Ros* vector using Gateway® reactions. These plasmids were transformed into the *Agrobacterium tumefaciens* strain Agl1 and injected into mature green tomatoes. The tomatoes were harvested between ten and fourteen days post breaker for analysis. This protocol is described fully in chapter 2. All further RNA extractions, qRT-PCR and tocochromanol extraction were completed, as described in chapter 2. A full list of primers used in this experiment are included in appendix.

6.2.3 Transient over expression of genes in tomato fruit

The transient viral over-expression system (pTRV2) was created, and kindly provided, by Dr Vera Thole. Coding sequences (CDS) of the candidate genes were cloned using Phusion PCR using primers with restriction enzyme adapters (described in chapter 2). The PCR products were ligated into the pTRV2 overexpression plasmid. The plasmids were transformed into the *Agrobacterium tumefaciens* stain Agl1 and injected into ripe tomatoes (breaker + 10 days). These tomatoes were harvested five days post injection for analysis. The full protocol is described in chapter 2. All further RNA extractions, qRT-PCR and tocochromanol extractions were completed, as described in chapter 2. A full list of primers used in this experiment are included in the appendix.

6.2.4 Phylogenetic analysis

Using the IT3F phylogenetic tool (Bailey et al., 2008) protein sequences of candidate TFs were added to an existing phylogenetic tree of *Arabidopsis thaliana* MYB and bZIP TFs to generate trees displaying their phylogenetic positions.

6.3 Results

6.3.1 Screening TFs in the trans-eQTL IL6-2-2 interval provided three candidate TFs for further analysis

Twenty six candidate genes were identified as candidate transcriptional regulators of the VTE pathway from the *S.pennellii* trans-eQTL IL6-2 (full details in chapter 3). Figure 6-1 shows the trans-eQTL IL6-2-2 and the candidate gene mining used to identify the TFs in this region. Genes encoding candidate TFs were reduced to seven candidate TFs for cloning and VIGS, which were chosen based on their co-expression patterns with genes encoding the enzymes of VTE pathway (table 6-1). The candidate TFs; *SlTF2, SlTF4, SlTF7* and *SlTF10* were upregulated in the two ILs identified in the trans-eQTL6-2-2, which showed a similar which co-expression profile as the genes encoding enzymes of the VTE pathway (table 6-1). Interestingly, these genes were all highly expressed in fruit of *S.pennellii* compared to fruit of *S.lycopersicum* cv. M82 (table 6-2). However, only expression of *SlTF4, SlTF7* and *SlTF10* were statistically different between the IL parents (table 6-2). There were three candidate TFs (*SlTF9, SlTF11* and *SlTF12)* which were downregulated in the ILs identified in the trans-eQTL IL6-2-2 (table 6-1). These genes showed the opposite expression profile to genes encoding enzymes of the VTE pathway in the ILs. These genes were also more highly expressed in fruit of *S.lycopersicum* cv. M82 compared to fruit of *S.pennellii* (table 6-2), however they were not statistically different.

The trans-eQTL L6-2-2 did not correlate to the *S.lycopersicoides* fruit RNA sequencing data in chapter 3. However, this trans-eQTL did correlate with the leaf RNA sequencing data (Chitwood et al., 2013), identified in chapter 3. Therefore, I checked the expression of the candidate TFs in the leaf RNA sequencing data for co-expression (table 6-3). *SlTF2* was lowly expressed in the leaf RNA sequencing data, which correlated with the expression of the VTE pathway genes in the leaf RNA sequencing data in chapter 3 (table 6-3). *SlTF12* was highly expressed in this region, whereas SlTF10 showed a more variable expression profile (table 6-3). The other candidate TFs were not differentially expressed, or the expression of some

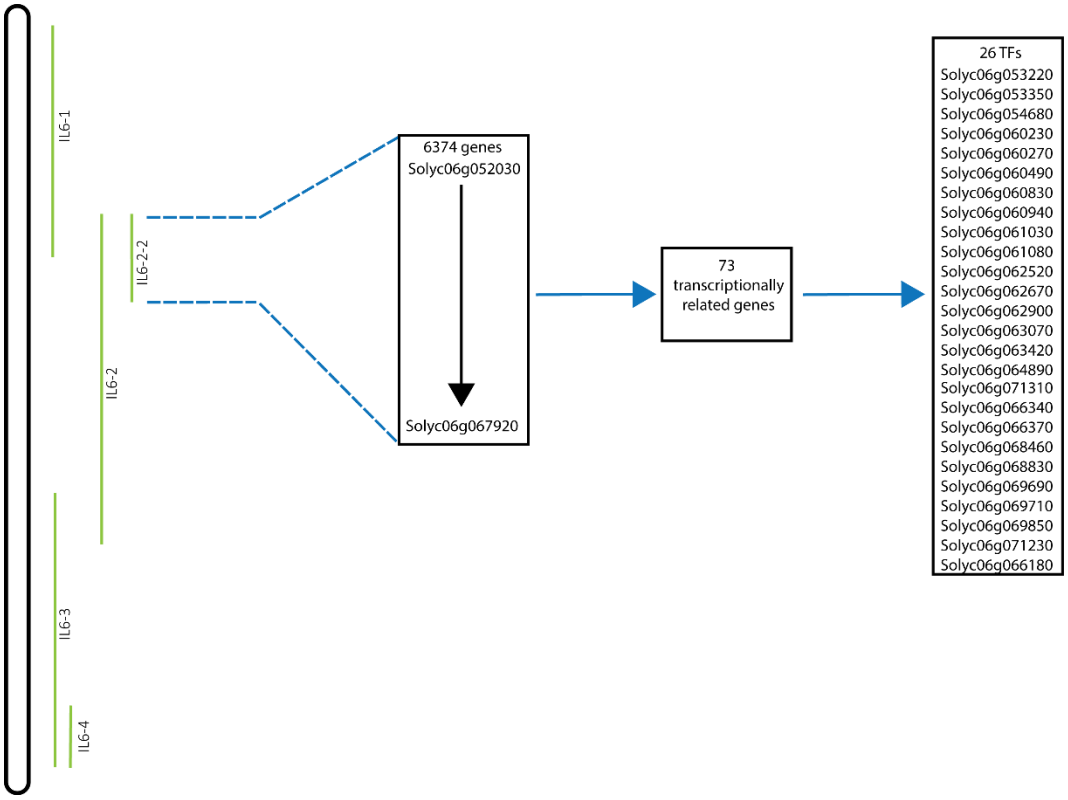


Figure 6-1. Schematic diagram of candidate gene mining for the trans-eQTL IL6-2-2. Twenty-six candidate transcription factors were identified, which showed co-expression patterns with genes encoding enzymes of the VTE pathway in the *S.pennellii* x *S.lycopersicum* cv. M82 fruit RNA sequencing data (Lee and Giovannoni).

Table 6-1 *S.pennellii* x *S.lycopersicum* IL fruit RNA sequencing data showing the relative expression values compared to M82 (=1). The table shows the two ILs in the trans-eQTL IL6-2 and the expression of VTE genes and candidate TFs which were analysed.

Gene	**Solgenomics identifier**	**Relative expression of genes in *S.pennellii* x S.lycopersicum cv. M82 introgression line**		
		M82	**IL6-2**	**IL6-2-2**
VTE genes				
VTE2	Solyc07g017770	1	1.53	2.78
VTE3(1)	Solyc09g065730	1	0.99	1.63
VTE3(2)	Solyc03g005230	1	0.78	1.31
VTE1	Solyc08g068570	1	0.57	1.49
VTE4	Solyc08g076360	1	0.95	1.89
VTE5	Solyc09g018510	1	0.66	0.96
VTE6	Solyc07g062180	1	1.59	1.09
Candidate TF genes				
SlTF2	Solyc06g060490	1	1.50	1.80
SlTF4	Solyc06g066180	1	20.55	39.97
SlTF7	Solyc06g071230	1	3.21	1.55
SlTF9	Solyc06g069710	1	0.50	0.38
SlTF10	Solyc06g069850	1	3.69	1.92
SlTF11	Solyc06g060230	1	0.25	0.42
SlTF12	Solyc06g061080	1	0.20	0.17

Table 6-2 Ratio of means of the *S.lycopersicum* cv.M82 and *S.pennellii* fruit RNA sequencing data. The statistical significance is shown by the p value, which p value <0.05 it is statistically significant. The letters represent which IL parent, the expression of the gene was highest in.

Gene	**Solgenomics identifier**	**Ratio of means between *S.lycopersicum* cv.M82 and *S.pennellii***	**Statistical significance shown as a p value**	**Higher expression in *S.lycopersicum* cv. M82 (L) or *S.pennellii* (P)**
SlTF2	Solyc06g060490	1.03	0.48	P
SlTF4	Solyc06g066180	11.99	1.26×10^{-6}	P
SlTF7	Solyc06g071230	2.95	3.28×10^{-10}	P
SlTF9	Solyc06g069710	0.31	0.48	L
SlTF10	Solyc06g069850	12.39	7.57×10^{-6}	P
SlTF11	Solyc06g060230	0.82	1	L
SlTF12	Solyc06g061080	0.22	0.14	L

Table 6-3 Relative values of RNA sequencing data of leaves of the *S.pennellii* x *S.lycopersicum* introgression lines (ILs) (Chitwood et al., 2013). The values are shown for the ILs that reside in the fruit trans-eQTL IL6-2-2 that was identified in chapter 3. The relative expression values are relative to the average value of the gene across all the *S.pennellii* x *S.lycopersicum* cv. M82 ILs. The letters (N.D.) represent genes that were not detected in the RNA sequencing data.

Gene name	**Solgenomics identifier**	**Relative expression in the *S.pennellii* x *S.lycopersicum* cv. M82 ILs**		
		IL6-1	**IL6-2**	**IL6-2-2**
SlTF2	Solyc06g060490	0.92	0.34	0.40
SlTF4	Solyc06g066180	N.D	N.D	N.D
SlTF7	Solyc06g071230	1.07	0.90	1.10
SlTF9	Solyc06g069710	N.D	N.D	N.D
SlTF10	Solyc06g069850	0.69	1.31	0.91
SlTF11	Solyc06g060230	0.88	0.98	0.97
SlTF12	Solyc06g061080	0.54	1.78	2.6

TFs candidates were not determined in the ILs of the trans-eQTL IL6-2-2 (table 6-3).

These candidate TFs were transiently silenced using VIGS in Microtom *Del/Ros* transformed fruit, and as shown in table 6-4 there were three TFs which showed significantly altered VTE contents in their silenced lines. *SlTF4* is a R1R2 type MYB TF (1R MYB), and this gene was transiently silenced in tomato fruit. Transiently silenced *SlTF4* tomato sectors showed significantly higher alpha (α), beta (β) and total tocopherol levels, relative to pTRV2 *Del/Ros* silenced sectors (table 6-4). Another 1R MYB TF *(SlTF7)* also showed higher α- and β-tocopherol contents in transiently silenced *SlTF7* tomato sectors. A NAC TF (*SlTF11)* showed that, when silenced, α- tocopherol was significantly altered. Silencing of *SlTF2* also showed altered tocopherol levels.

Three TFs (*SlTF2, SlTF4* and *SlTF7)* were taken for further VIGS analysis in Moneymaker fruit for qRT-PCR analysis, as well as transient over expression assays, because the Microtom fruit were very small and hard to dissect sectors. Consequently, therefore the transient assay was repeated in larger *S.lycopersicum* cv. Moneymaker *Del/Ros* fruit so that red, silenced sectors and purple, non-silenced sectors were easier to dissect.

6.3.2 Transient silencing of the candidate TFs in *S.lycopersicum* cv. Moneymaker *Del/Ros* fruit

SlTF2, *SlTF4* and *SlTF7* were transiently silenced using VIGS in *S.lycopersicum* cv. Moneymaker *Del/Ros* tomato fruit. Figure 6-2 shows that *SlTF2* did not significantly alter relative levels of any form of tocopherol or total tocopherols. Silencing of *SlTF2* was effective as *SlTF2* expression was reduced in the silenced sectors (figure 6-3).

Table 6-4 Summary of VIGS screening in Microtom *Del/Ros* tomato fruit. The table shows statistical significance of metabolite contents of the silenced sectors of a *Del/Ros* and TF silenced fruit compared to *Del/Ros* silenced fruit. The stars represent statistically significance (*) = $p < 0.05$, (**) = $p < 0.01$.

SlTF number	TF type	Solgenomics identifier	alpha tocopherol	beta tocopherol	gamma tocopherol	total tocopherol
2	bZIP	Solyc06g060490	-	-	-	-
4	MYB	Solyc06g066180	**	**	-	**
7	MYB	Soly06g071230	*	*	-	-
9	NAC	Solyc06g069710	-	-	-	-
10	MYB	Solyc06g069850	-	-	-	-
11	NAC	Solyc06g060230	*	-	-	-
12	NAC	Solyc06g061080	-	-	-	-

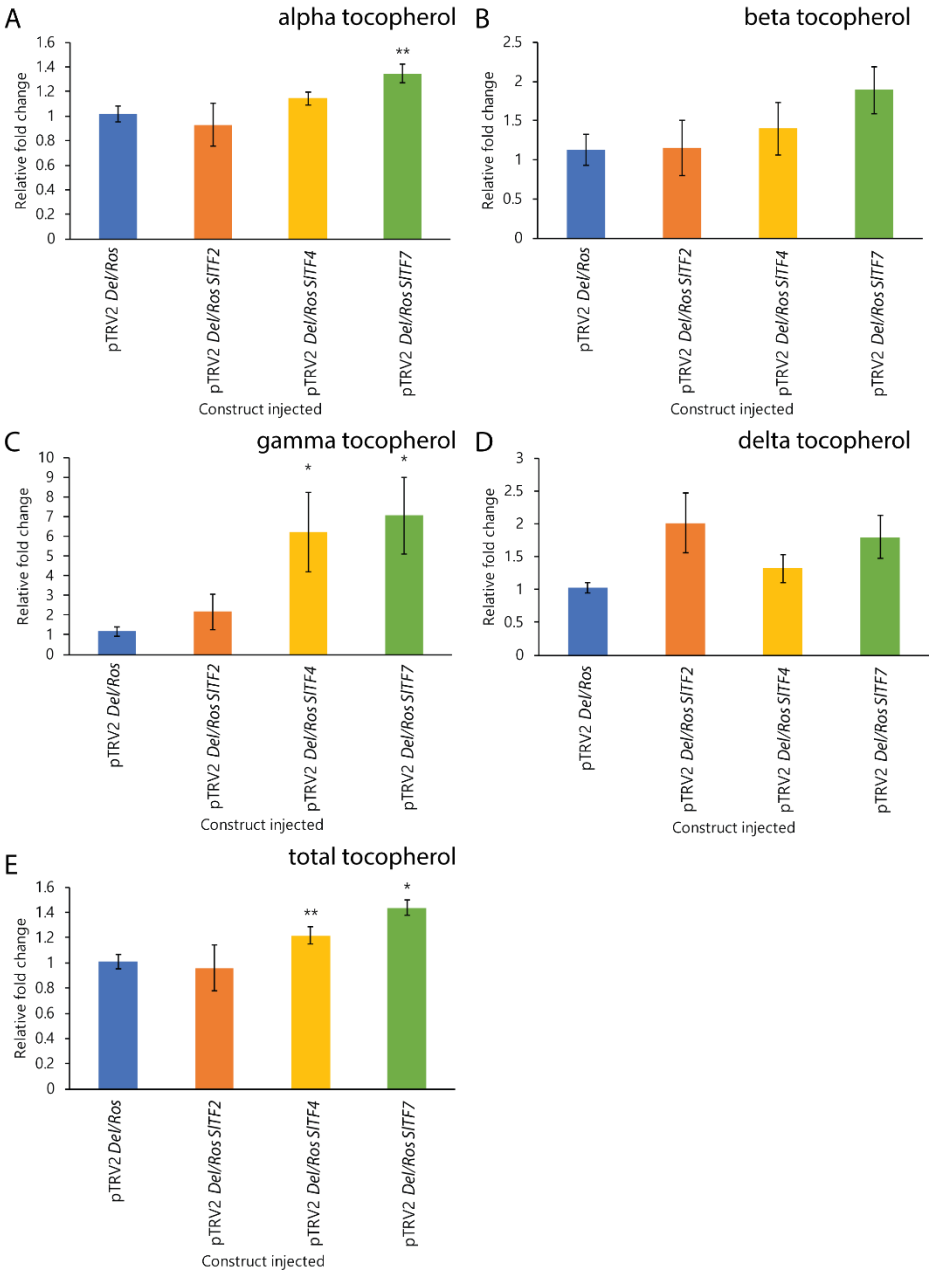


Figure 6-2 Relative fold change of tocopherols of injected VIGS Moneymaker tomato fruit; (A) alpha tocopherol (B) beta tocopherol (C) gamma tocopherol (D) delta tocopherol (E) total tocopherol. These data were analysed from red silenced sectors of injected fruit and the relative fold change measurements were determined by comparing against the control; pTRV2 *Del/Ros*. The error bars are standard error and the stars represent statistical significance; (*) = $p<0.05$, (**) = $p<0.01$, which were calculated using t-tests.

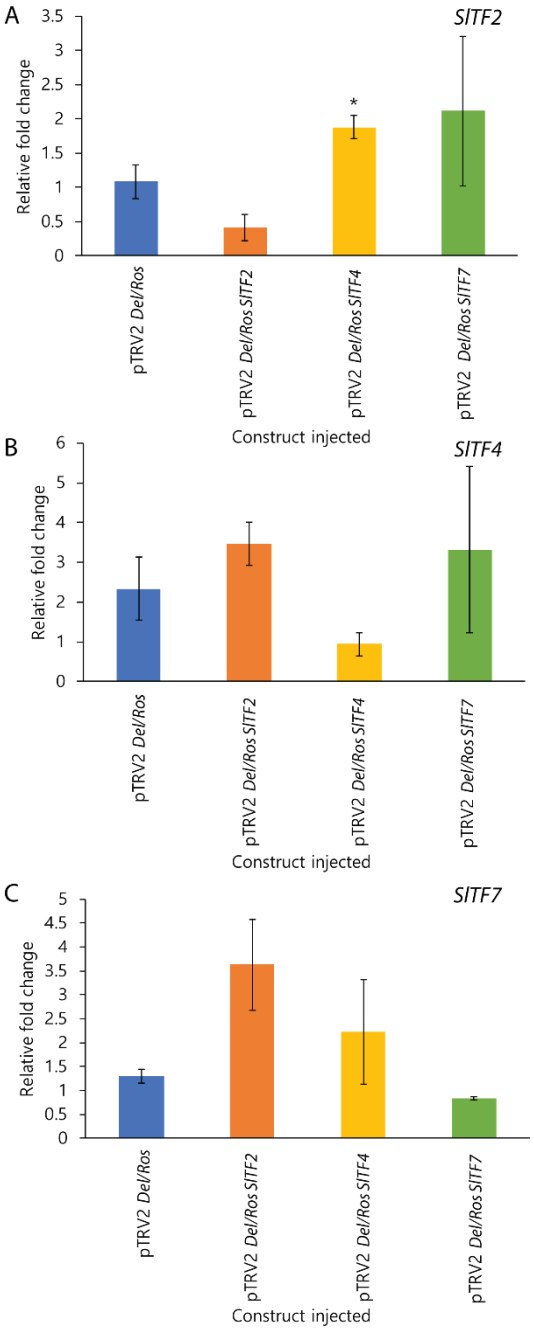


Figure 6-3 Relative expression of the candidate genes of injected VIGS Moneymaker tomato fruit; (A) *SlTF2,* (B) *SlTF4* and (C) *SlTF7*. These data were analysed from red silenced sectors of injected fruit and the relative fold change measurements were determined by comparing against the control; pTRV2 *Del/Ros*. The error bars are standard error and the stars represent statistical significance; (*) = $p<0.05$, (**) = $p<0.01$, which were calculated using t-tests. SlCAC was used as a reference gene.

Silenced *SlTF4* and *SlTF7* tomato sectors showed significant differences in the forms of tocopherol in fruit (figure 6-2). Transiently silenced *SlTF4 Del/Ros* fruit showed that *SlTF4* silenced sectors had significantly higher γ-tocopherol and total tocopherol contents, relative to the silenced pTRV2 *Del/Ros* control sectors. Therefore, the silencing of *SlTF4* resulted in an increase in tocopherol, which implies that *SlTF4* could be a candidate repressor. Expression of *SlTF4* was reduced in the transiently silenced fruit (figure 6-3), demonstrating that the VIGS had been successful. Relative fold changes of genes encoding VTE enzymes showed that *VTE6* expression was increased (figure 6-4), but that expression of other VTE genes was not significantly altered.

SlTF7 was transiently silenced in *S.lycopersicum* cv. Moneymaker *Del/Ros* tomato fruit, and α-tocopherol was significantly higher in the silenced *SlTF7* sectors, relative to the pTRV2 *Del/Ros* silenced control sectors (figure 6-2). γ-tocopherol and total tocopherol levels are also significantly higher than the pTRV2 *Del/Ros* injected sectors. Expression of *SlTF7* was reduced in the respective silenced sectors (figure 6-3). Surprisingly, expression of *SlTF7* was higher in the pTRV2 *SlTF2* and pTRV2 *SlTF4* injected tomatoes, suggesting that there may be interactions between these three genes. The expression of genes encoding enzymes of the VTE pathway showed significant changes in pTRV2 *SlTF7* silenced sectors compared to the pTRV2 *Del/Ros* silenced control sectors. *VTE2, VTE3(1)* and *VTE3(2)* expression was higher in pTRV2 SlTF7 silenced sectors compared to controls. These data implied that *SlTF7* could be a repressor of the VTE pathway.

Expression of *VTE4* (figure 6-4), a methyl-transferase responsible for the synthesis of α- and β-tocopherol, was reduced in transiently silenced *SlTF2* tomatoes. These data were unexpected as tocopherol content was not altered. The pTRV2 *Del/Ros* SlTF4 injected tomatoes, which had silenced *SlTF4* showed increased expression of *SlTF2,* compared to pTRV2 *Del/Ros* injected fruit.

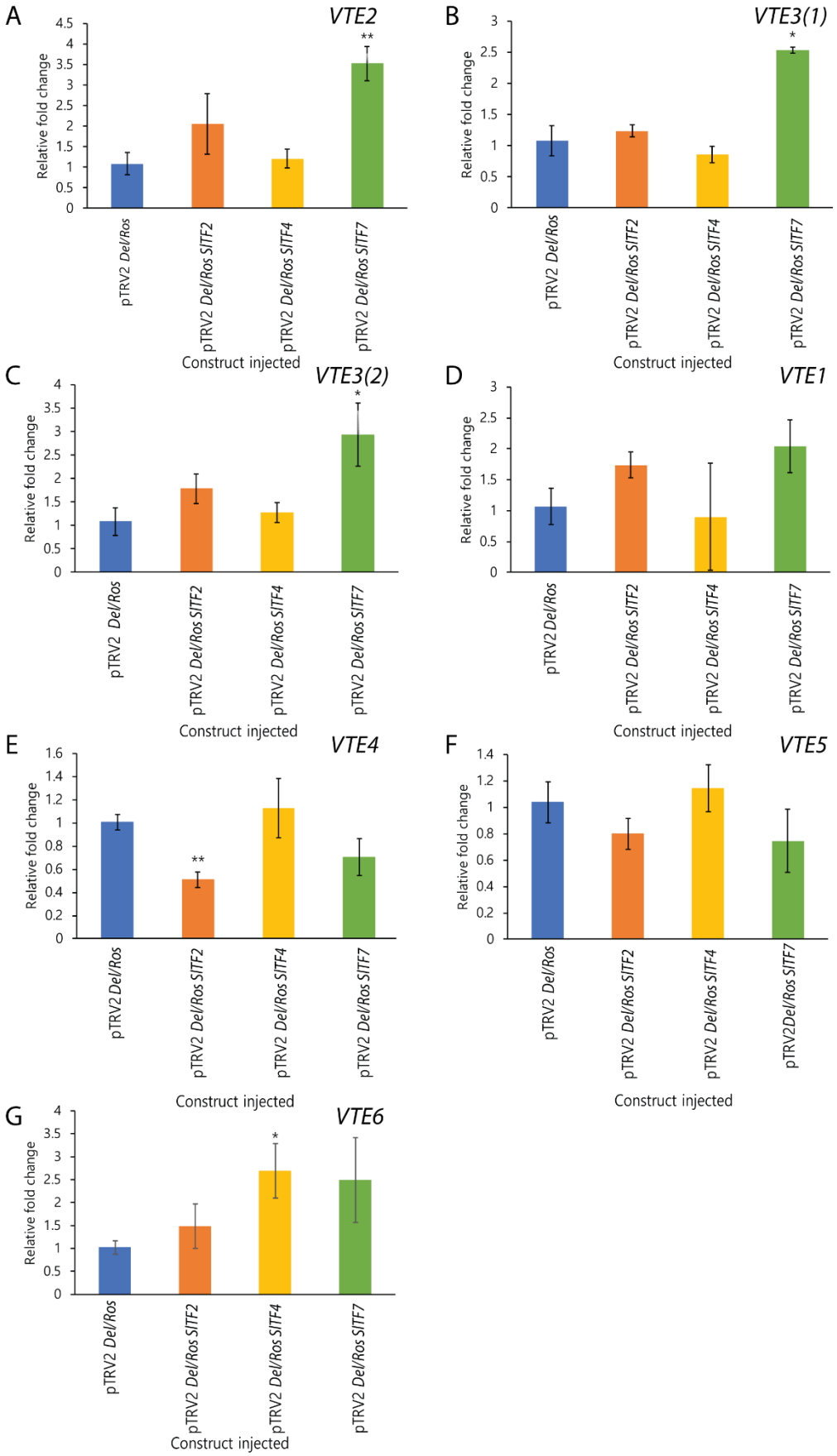

A
VTE2
**
Relative fold change
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlTF2
pTRV2 Del/Ros SlTF4
pTRV2 Del/Ros SlTF7
Construct injected
B
VTE3(1)
*
Relative fold change
3
2.5
2
1.5
1
0.5
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlTF2
pTRV2 Del/Ros SlTF4
pTRV2 Del/Ros SlTF7
Construct injected
C
VTE3(2)
*
Relative fold change
4
3.5
3
2.5
2
1.5
1
0.5
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlTF2
pTRV2 Del/Ros SlTF4
pTRV2 Del/Ros SlTF7
Construct injected
D
VTE1
Relative fold change
3
2.5
2
1.5
1
0.5
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlTF2
pTRV2 Del/Ros SlTF4
pTRV2 Del/Ros SlTF7
Construct injected
E
VTE4
**
Relative fold change
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlTF2
pTRV2 Del/Ros SlTF4
pTRV2 Del/Ros SlTF7
Construct injected
F
VTE5
Relative fold change
1.4
1.2
1
0.8
0.6
0.4
0.2
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlTF2
pTRV2 Del/Ros SlTF4
pTRV2Del/Ros SlTF7
Construct injected
G
VTE6
*
Relative fold change
4
3.5
3
2.5
2
1.5
1
0.5
0
pTRV2 Del/Ros
pTRV2 Del/Ros SlTF2
pTRV2 Del/Ros SlTF4
pTRV2 Del/Ros SlTF7
Construct injected

Figure 6-4 Relative expression of the VTE pathway genes of injected VIGS Moneymaker tomato fruit; (A) *VTE2* (B) *VTE3(1)* (C) *VTE3(2)* (D) *VTE1* (E) *VTE4* (F) *VTE5* (G) *VTE6*. These data were analysed from red silenced sectors of injected fruit and the relative fold change measurements were determined by comparing against the control; pTRV2 *Del/Ros*. The error bars are standard error and the stars represent statistical significance; (*) = $p<0.05$, (**) = $p<0.01$, which were calculated using t-tests. SlCAC was used as a reference gene.

6.3.3 The candidate transcriptional regulators are differentially expressed during *S.lycopersicum* cv. M82 tomato development and ripening

To understand the potential roles of these three TFs in VTE biosynthesis, the expression of the candidate genes was checked during *S.lycopersicum* cv. M82 tomato fruit development and ripening. Expression of *SlTF2* was highest at breaker (figure 6-5), in the pericarp and epidermis. The expression of *SlTF2* declined at the B+5 and B+10 stages. In contrast, expression of *SlTF4* was low in the mature green and breaker stages of tomato fruit in the epidermis (figure 6-5). The expression of *SlTF4* peaked at the B+5 stage, but, expression did not change significantly during tomato ripening. Within the epidermis, *SlTF4* expression was very low in MG and B fruit stages, and expression was induced in the B+5 and B+10 stages. *SlTF7* expression was highest at the MG stage in the pericarp, but *SlTF7* expression remained constant during fruit ripening (figure 6-5). Expression of *SlTF7* in the epidermis was constant during fruit development and ripening.

6.3.4 Transient over expression of candidate TFs in *S.lycopersicum* cv Moneymaker tomato fruit suggested that they are repressors of VTE biosynthesis

Transient over expression of *SlTF2* using the pTRV2 over expression system resulted in a decrease in several tocopherol forms (figure 6-6). β-, gamma (γ) and delta (δ) tocopherols were significantly reduced, although, these forms are minor forms of tocopherols in tomato, and did not alter total tocopherol levels. Figure 6-7 showed that *SlTF2* was significantly over expressed in fruit injected with the pTRV2 OE *SlTF2* construct. Genes encoding enzymes of the VTE pathway were not differentially expressed in tomatoes overexpressing *SlTF2*, compared to WT (figure 6-8). Similarly, *SlTF4* expression was significantly higher in the pTRV2 OE SlTF4 injected tomatoes, compared to the pTRV2 OE injected control fruit (figure 6-7). γ- and δ- tocopherol forms were significantly reduced in *SlTF4* overexpression fruit, in comparison to the control (figure 6-6), but, these forms are minor forms

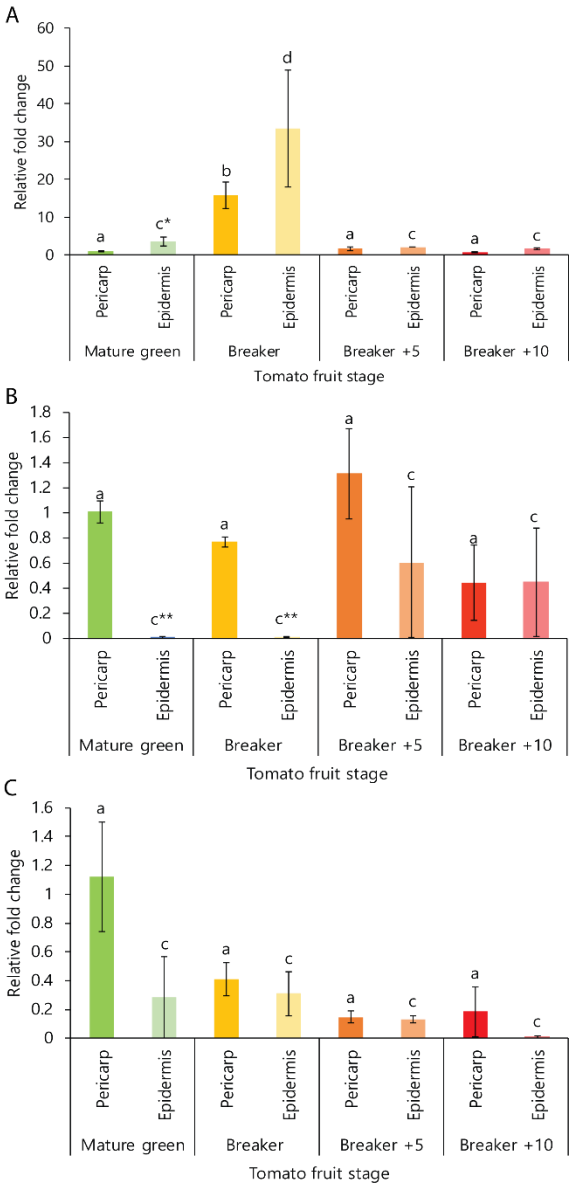


Figure 6-5 Candidate transcription factor gene expression during *S.lycopersicum* cv.M82 tomato development and ripening; (A) *SlTF2* (B) *SlTF4* and (C) *SlTF7*. Relative fold change of the candidates was normalised to mature green pericarp. The error bars are standard error and the stars represent t-tests between pericarp and epidermis; (*) = $p<0.05$, (**) = $p<0.01$. The letters represent tukey tests and show the statistical significant of a tissue and were analysed during the time course. The tissue of the pericarp tukey tests are represented by the letters a and b. The tukey tests between the tissue of the epidermis are represented by the letters c and d. The tukey tests do not compare across tissues, and instead, the t-tests represent significance at one stage of tomato for both tissues. SlCAC was used as a reference gene.

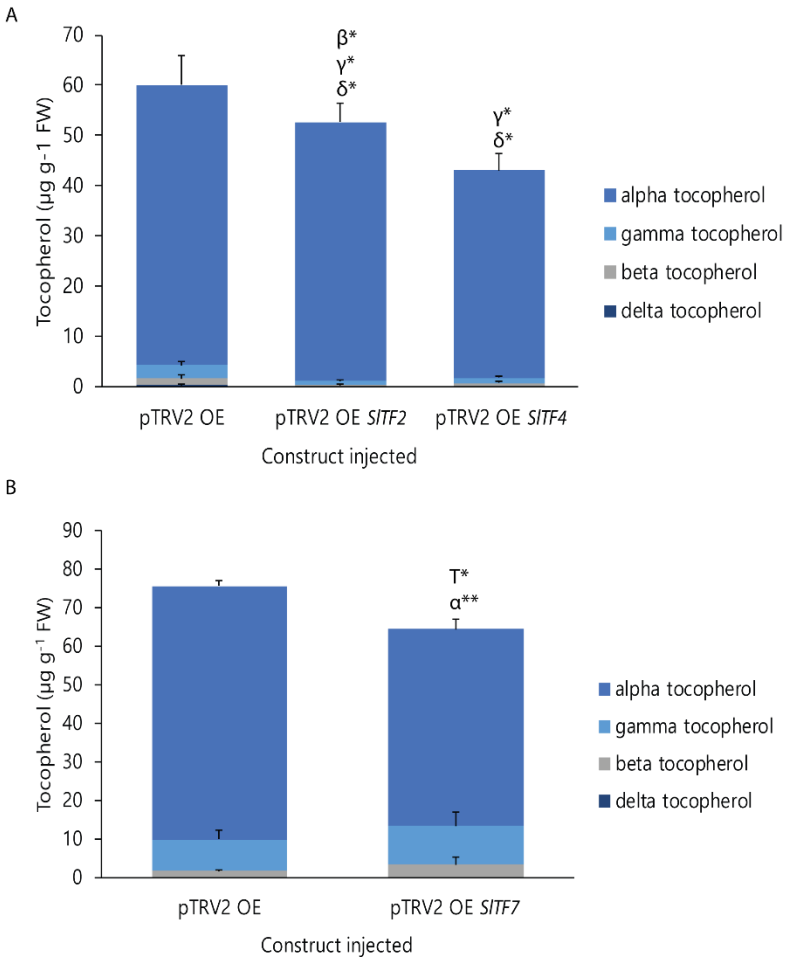


Figure 6-6 Metabolite contents of transiently over expressed constructs (A) pTRV2 OE SlTF2 and pTRV2 OE TF4, in comparison with the pTRV2 OE control. (B) pTRV2 SlTF7, in comparison with the pTRV2 OE control. All constructs were injected into WT *S.lycopersicum* cv.M82 breaker +10 days fruit. The error bars are standard error. The letters represent the different forms of tocopherol; α = alpha, β = beta, γ = gamma, δ = delta, T= total. The stars next to the letters represent whether they are statistically significant from the control; (*) = $p<0.05$, (**) = $p<0.01$.

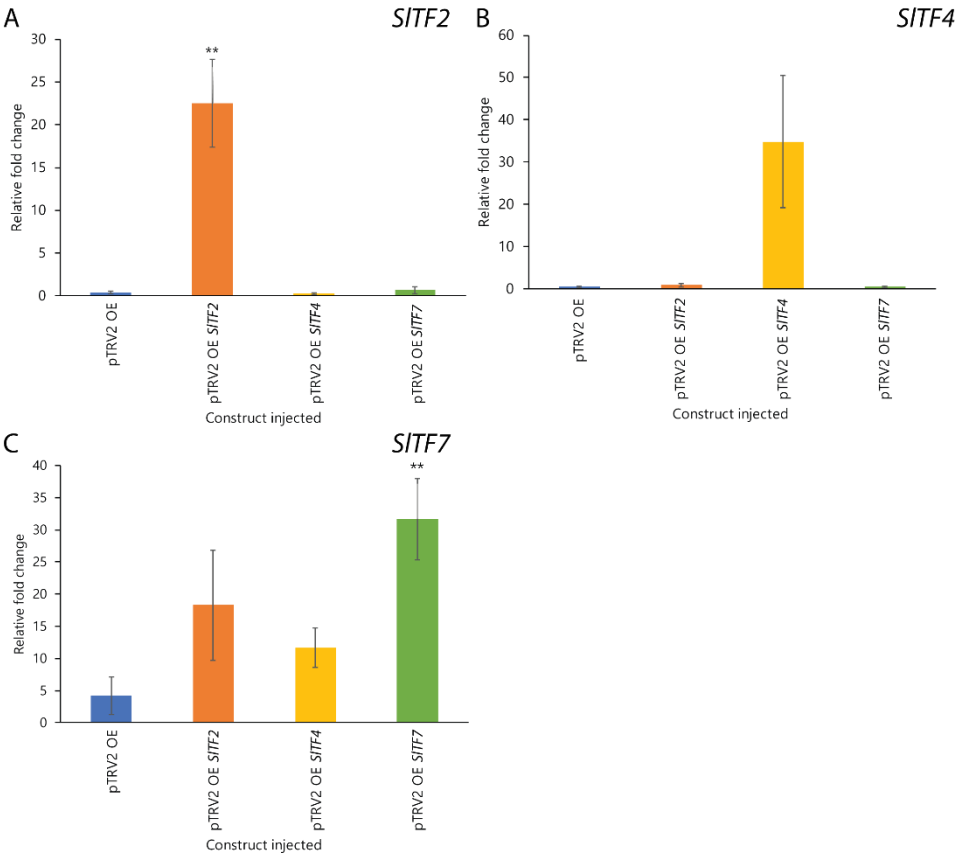


Figure 6-7 Relative fold change of expression of the candidate TFs in tomatoes injected with different transient overexpression constructs; (A) *SlTF2,* (B) *SlTF4* and (C) *SlTF7.* The error bars were standard error. The stars represent t-tests between the injected constructs with the gene of interest compared to the pTRV2 OE control; (*) = $p<0.05$, (**) = $p<0.01$. SlCAC was used as a reference gene.

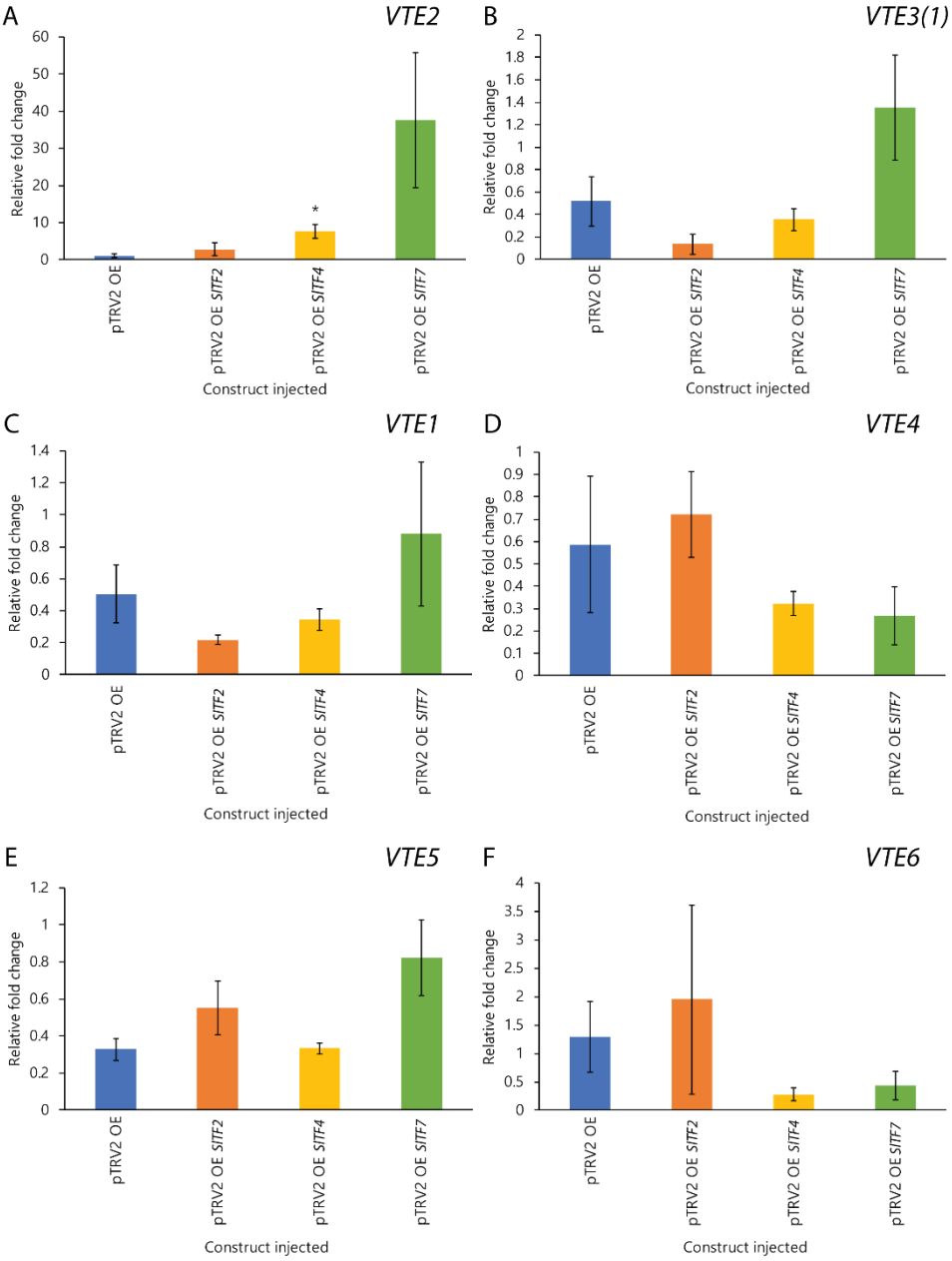


Figure 6-8 Relative fold change of expression of VTE pathway genes in WT *S.lycopersicum* cv. Moneymaker tomatoes, which were injected with different pTRV2 over expression constructs; (A) VTE2 (B) VTE3(1) (C) VTE1 (D) VTE4 (E) VTE5 (F) VTE6. The stars represent t-tests between the injected constructs with the gene of interest compared to the pTRV2 OE control; (*) = $p<0.05$, (**) = $p<0.01$. SlCAC was used as a reference gene.

of tocopherol. *VTE2* was highly expressed in tomatoes transiently overexpressing *SlTF4* (figure 6-8).

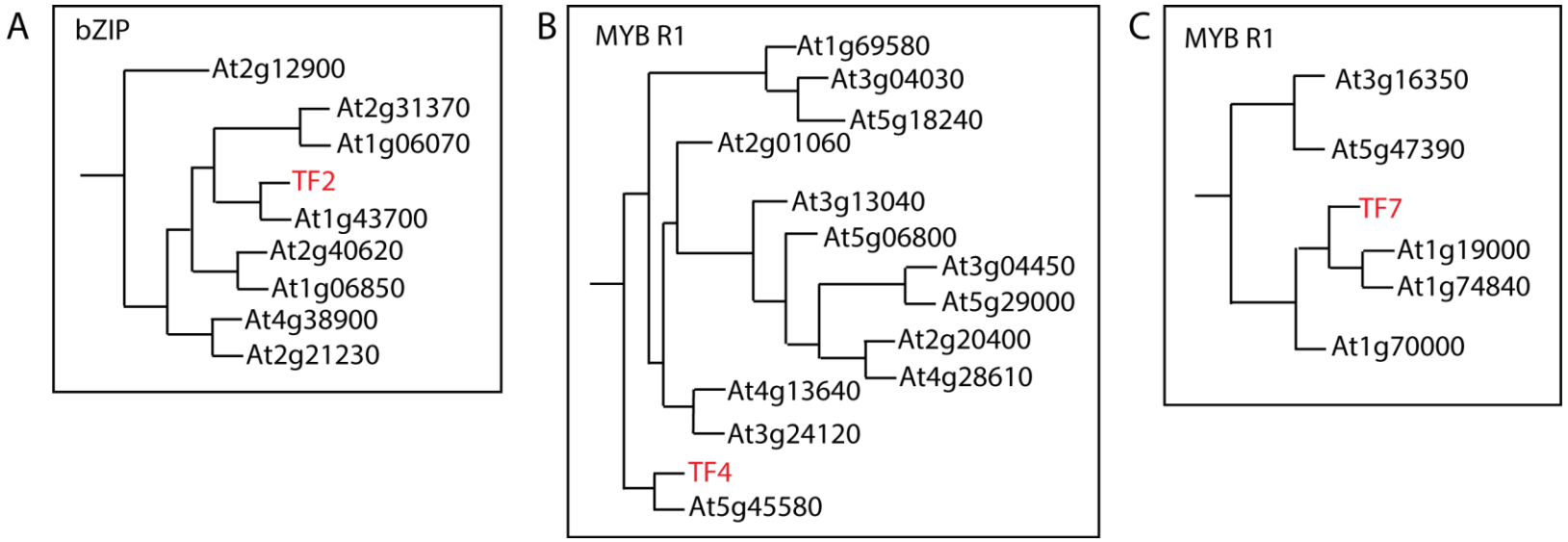
Tomatoes that transiently expressed *SlTF7* showed reduced α- and total tocopherol contents (figure 6-6). α-tocopherol and total tocopherol levels were reduced by 22% and 14%, respectively. β- and γ- tocopherol levels were low in comparison with α-tocopherol levels, but were increased by 93% and 23%, respectively, although these increases were not statistically significant. Expression of *SlTF7* was significantly higher in transient over-expression *SlTF7* tomato fruit, compared to the pTRV2 control fruit (figure 6-7). Genes encoding VTE enzymes, including *VTE2, VTE3(1), VTE1* and *VTE5,* were more highly expressed, but, these increases were not statistically significant (figure 6-8).

6.3.5 Transient assays that silenced and overexpressed the candidate TFs suggested that they regulate the VTE biosynthesis

The data from the VIGS and overexpression assays suggested that the candidate TFs regulate the VTE pathway. From the VIGS data, *SlTF2* positively regulates *VTE4* expression. *SlTF2* and *SlTF4* may positively regulate *SlTF7* expression, as *SlTF7* expression was increased in the pTRV2 *SlTF2* and *SlTF4* over expression fruits. Therefore, *SlTF2* and *SlTF4* may have an indirect effect on repressing the expression of genes encoding enzymes in the VTE pathway. VIGS of *SlTF7* showed that transient silencing induced expression of several VTE genes encoding enzymes, including; *VTE3(1)* and *VTE3(2).* Additionally, *SlTF4* may bypass *SlTF7* and directly regulate the VTE pathway genes, as modified *SlTF4* expression was shown to affect *VTE6* and *VTE2* expression in VIGS and over expression assays, respectively.

Phylogenetic analyses of these candidate TFs showed that they were not phylogenetically related (figure 6-9). The bZIP TF (SlTF2) shares homology with AtVIP1 (At1g43700), which is an important protein for *Agrobacterium* infection (Maes et al., 2014). SlTF4 and SlTF7 both encode for MYB R1 TFs, although they do not share extensive homology, as shown in figure 6-9. SlTF4 shares amino acid

Figure 6-9 Phylogenetic trees of candidate TFs transiently assessed (A) TF2 (B) TF4 and (C) TF7. TF2 is a bZIP TF and TF4 and 7 are MYBR1 TFs. The trees were adapted from outputs from the IT3F website (Bailey et al., 2008) and show subgroups of a larger phylogenetic tree.



homology with the protein product of At5g45580, which has no known function in *Arabidopsis.* Additionally, SlTF7 shares homology with *AtMYBD* (At1g19000), which is a positive regulator of light-mediated anthocyanin production (Nguyen et al., 2015).

6.4 Discussion

6.4.1 Screening of the trans-eQTL IL6-2 TF candidates resulted in three putative TFs of the VTE pathway.

The initial screen of candidate TFs using VIGS in *S.lycopersicum* cv. Microtom *Del/Ros* fruit was an effective way to reduce the number of candidate TFs for further analysis. The VIGS screen identified three candidate TFs which were reanalysed by VIGS in *S.lycopersicum* cv. Moneymaker *Del/Ros* fruit, to complete expression analysis of the VTE pathway genes. *SlTF2, SlTF4* and *SlTF7* candidate TFs showed that silenced tomato sectors, injected with pTRV2 *Del/Ros SlTF4* and pTRV2 *Del/Ros SlTF7,* altered the tocopherol composition and increased expression of some VTE biosynthetic genes. In contrast, pTRV2 *Del/Ros SlTF2* altered expression of only *VTE4.* These candidate TFs were taken forward for further transient over expression assays.

The transient over expression assay demonstrated that *SlTF2* and *SlTF4* altered the composition of tocopherols in fruit. Transient over expression of *SlTF7* resulted in tomato fruit with reduced α- and total tocopherol contents, which suggested that this is a repressor of VTE synthesis. This was consistent with the increase in expression of *VTE2* when *SlTF7* was transiently silenced.

6.4.2 Transient assays of *SlTF2* showed that it may have an indirect role in transcriptional regulation of the VTE pathway

SlTF2 is a bZIP TF and its homolog in *Arabidopsis* is *AtVIP1*. This bZIP resides in group I of the bZIP family, which defines proteins involved in stress responses in *Arabidopsis* (Droge-Laser et al., 2018). The VIP1 protein interacts with the *Agrobacterium* DNA binding protein; VirE2, which is important for *Agrobacterium* infection (Maes et al., 2014, Tsugama et al., 2016). Over expression of *35S:AtVIP1*

in *Arabidopsis* showed that transgenic lines had increased ABA sensitivity compared to the WT seeds. The loss of function G-protein mutant, *agb1-2,* also shows ABA sensitivity. G-proteins reside on plasma membranes and are important for signal transduction in *Arabidopsis*. *Agb1-2* mutants treated with ABA or drought stressed have higher *AtVIP1* transcript levels, compared to WT (Xu et al., 2015). This implies that *AtVIP1* has a role in regulating ABA signalling, although, the extent of its role in ABA signalling is not yet clear.

Substrates for ABA synthesis derive from the MEP pathway, therefore it is possible that there is competition between the ABA and VTE biosynthesis pathways for substrates. This could explain why compositional tocopherol contents were altered in the transient over expression assay for *SlTF2*. If *SlTF2* functions in a similar way to its homolog, it may cause the upregulation of *VTE4* as observed in the transient over expression assay. *SlTF2* may not directly regulate the VTE pathway, but might be responsible for cross-talk between the ABA and VTE pathways. Transcriptional regulation between these two pathways is likely to be linked and complex, therefore an increase in *VTE4* expression in response to silencing of *SlTF2* may be due to indirect transcriptional regulation of the ABA signalling pathway.

Expression of *SlTF2* was highest at the B stage during tomato development and ripening in the epidermis. This correlates well with the data from chapter 4, which showed that VTE levels were highest at the B+3 stage. Therefore, this gene may be 'ripening related', and may affect expression of VTE enzymes only indirectly. *SlTF2* might play a role in regulating VTE synthesis, although my transient assays did not provide full functional characterisation of this gene. Stable transformations would provide further clarification of the role of this TF in the regulation of VTE biosynthesis and possible cross talk with ABA signalling.

6.4.3 *SlTF4* is a possible negative regulator of the VTE pathway

SlTF4 is a MYB 1R TF and its closest *Arabidopsis* homolog has no known function. This is a novel TF that putatively regulates VTE biosynthesis, but it is not clear whether this is direct or indirect regulation.

Transiently silenced pTRV2 *Del/Ros SlTF4* tomatoes had significantly higher tocopherol contents relative to the control. Thus, it is likely that stable transformations would provide further clarification of how *SlTF4* interacts the VTE pathway.

Transient over expression analysis of *SlTF4* showed that it upregulated *VTE2* expression and down regulated expression of *VTE1*, *VTE3(1)*, *VTE4* and *VTE6*, although not to the levels of statistical significance. *VTE6* expression was also reduced in transiently *SlTF4* overexpressing fruit, although this was not significant. Total tocopherol contents were not significantly altered in tomato fruits overexpressing *SlTF4*, although γ- and δ- tocopherol levels were significantly different. These forms of tocopherol contribute only a very small proportion of total tocochromanols. VIGS silencing of *SlTF4* showed a significant increase in transcript levels of *VTE6*, despite measured increases in levels of total tocopherols and α-tocopherol in particular Perhaps, SlTF4 negatively regulates *VTE6* expression and so fruits tocopherol biosynthesis by limiting the supply of precursors from chlorophyll breakdown in fruit development. *SlTF4* could also mediate VTE biosynthesis by regulating *SlTF7*, as this gene was highly expressed in *SlTF4* overexpression assays.

Expression of *SlTF4* in the M82 tomatoes was highest at the B stage during tomato development and ripening. Previously, in chapter 4, the metabolite and expression analysis of genes encoding enzymes of the VTE biosynthetic pathway, in tomato fruit showed that at B+3, several VTE genes were induced. Metabolite contents in the epidermis were also increased in the pericarp of M82 fruit at the B+3 stage (chapter 4). Therefore, there is a link between the induction of *SlTF4* expression and the resulting increase of expression of VTE genes and metabolite content at the B+3 stage (described in chapter 4). It is possible that *SlTF4* might be

the cause for the increase in VTE gene expression and metabolite content, however, stable transformations would provide further insight into *SlTF4's* regulatory role of VTE biosynthesis.

6.4.4 *SlTF7* encodes a novel repressor of the VTE pathway, the expression of which may be modulated by light

Transient overexpression of *SlTF7* caused a decrease in α- and total tocopherols, which suggested that this TF might be a repressor of VTE biosynthesis. The expression of some VTE pathway genes (*VTE2, VTE3(1)* and *VTE3(2)*) were increased in the transiently silenced pTRV2 *Del/Ros SlTF7* tomatoes, reinforcing the idea that this TF is a repressor. However, these VTE pathway genes were also upregulated in the *SlTF7* transiently overexpressing tomatoes, although these changes in transcript levels were not significant. The *Arabidopsis* homolog of *SlTF7* is a *AtMYB-like domain (AtMYBD)* TF, which is a small MYB TF. *AtMYBD* negatively regulates *AtMYB-like 2 (AtMYBL2)* - a negative regulator of anthocyanin biosynthesis (Dubos et al., 2008, Nguyen et al., 2015). Expression of *AtMYBD* in *Arabidopsis,* increases in response to cytokinins and light and it is regulated directly by *AtHY5* – a bZIP TF (Nguyen et al., 2015). *AtHY5* is vital for light responses and has previously been identified as a regulator of chalcone synthase *(CHS),* which is involved in anthocyanin biosynthesis (Shin et al., 2007). This suggests that if *SlTF7* had a function similar to its *Arabidopsis* homolog, it could affect anthocyanin biosynthesis and also be regulated by light.

PIF TFs have been shown to interact with *HY5* (Toledo-Ortiz et al., 2014). PIFs are transcriptional regulators of PHYs that perceive red and far-red light. *AtPIF3* and *AtHY5* regulate an anthocyanin biosynthesis gene F3-H (flavanone 3-hydroxylase) in *Arabidopsis,* but, *HY5* must be present for *PIF3* to bind (Shin et al., 2007). PIFs also play a role in carotenoid production and *SlPIF1a* has been identified as a direct regulator of *PSY1* expression (Toledo-Ortiz et al., 2010). Carotenoid and tocopherols share the same substrates from the MEP pathway. Therefore, if *SlTF7* shares a similar function to its *Arabidopsis* homolog, *SlTF7*

might negatively regulate the earlier steps of the MEP pathway as well as anthocyanin biosynthesis. This would alter flux of substrates for tocopherol biosynthesis. It is likely that the MEP and VTE pathways share transcriptional regulators as promoter sequence analysis of the MEP pathway and VTE pathways showed the genes encoding enzymes share common TF binding motifs (Quadrana et al., 2013).

The transient assays of *SlTF7* support a role as a novel negative transcriptional regulator of the VTE pathway and possibly other pathways, such as carotenoid biosynthesis. This could be mediated through *HY5*, like its *Arabidopsis* homolog; *AtMYBD,* and suggests the important role of light regulation for tocochromanol biosynthesis.

6.4.5 The three candidate TFs may interact to control VTE biosynthesis and genes encoding enzymes of VTE biosynthesis

Transient over expression assays showed that *SlTF7* was up regulated by the pTRV2 OE *SlTF2* and pTRV2 OE *SlTF4* constructs, but, these changes in transcript levels were not significant. These data demonstrated that these three TFs may interact in regulating the VTE pathway. The data suggested that *SlTF2* and *SlTF4* induce *SlTF7* expression. Interestingly, the *SlTF2* and *SlTF4* show differential expression during tomato fruit development and ripening. *SlTF2* expression is induced at breaker in the tissues of the pericarp and epidermis. This induction of *SlTF2* expression in the epidermis may cause the upregulation of *SlTF4* at breaker +5 stage in the epidermis. These TFs could possibly be regulating each other. The expression of *SlTF7* was induced in the pericarp at the mature green stage, but then, the expression remained constant throughout further tomato ripening. These interpretations imply that *SlTF2* and *SlTF4* do not regulate *SlTF7*, although the transient over expression data showed the opposite to these results. Further stable transformations are needed to understand these data more fully.

The candidate TFs that were transiently silenced were identified as transcriptional repressors of VTE biosynthesis. However, the candidate TFs were

highly expressed in the trans-eQTL IL6-2-2 from the RNA sequencing data (described in chapter 3). The genes encoding enzymes of VTE biosynthesis were highly expressed in trans-eQTL IL6-2-2. Therefore, it is likely that these genes are not the causal TFs causing the changes in expression in the trans-eQTL IL6-2-2. *SlTF4* and *SlTF7* were highly expressed in *S.pennellii* fruit, compared to *S.lycopersicum* cv. M82 in the RNA sequencing data. If they were candidate TFs regulating the expression of genes encoding enzymes of the VTE pathway, they should be highly expressed in *S.lycopersicum* and lowly expressed in *S.pennellii.*

These transient assays have provided insight into other transcriptional mechanisms that may regulate the VTE pathway. It is likely that light regulation will impact VTE biosynthesis and these assays have suggested *SlTF7* as a strong candidate for negatively regulating VTE biosynthesis. Further clarifications are required to define transcriptional regulation of the VTE pathway and the roles that these TFs play. *SlTF2* and *SlTF4* may indirectly regulate the VTE pathway, either through *SlTF7,* or, possibly through other mechanisms. Stable over expression and CRISPR tomato transformations should provide further evidence to support these putative roles for regulating the VTE pathway.

Chapter 7: Tocotrienol biosynthesis in *Solanum pennellii* tomato fruit

Chapter 7: Tocotrienol biosynthesis in *Solanum pennellii* tomato fruit

7.1 Introduction

7.1.1 *Solanum pennellii* fruits contain tocotrienols

Metabolic and expression analysis of fruit during development and ripening of fruit of the species, *Solanum pennellii* and *Solanum lycopersicum* revealed that *S.pennellii* fruit synthesised tocotrienols (described in chapter 4), whereas fruit of *S.lycopersicum* cv.M82 did not. Tocotrienols are not normally produced in tomato fruit, as tomatoes lack the gene encoding homogentisate geranylgeranyl transferase (HGGT) responsible for tocotrienol synthesis (Yang et al., 2011). In this chapter, I screened fruit from fifty six of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs to identify the genomic region and the putative candidate genes responsible for tocotrienol synthesis in *S.pennellii.*

7.1.2 Tocotrienol biosynthesis

Dicotyledonous plants lack the gene encoding HGGT for synthesis of tocotrienols (Yang et al., 2011). Tocotrienol biosynthesis in dicotyledonous plants has been achieved by over-expressing a barley *HGGT* in *Arabidopsis* (Cahoon et al., 2003), which also lacks an endogenous HGGT. It has been suggested that VTE2 is responsible for the trace levels of tocotrienols that are found in tomato fruit (Chun et al., 2006). Homogentisate prenyltransferase (HPT - VTE2) is the enzyme responsible for tocopherol synthesis which uses phytyl diphosphate (PDP) as a substrate to synthesise tocopherols. VTE2 has been shown to be able to use GGDP as a substrate and to synthesise tocotrienols when PDP pools are low and GGDP pools are high (Yang et al., 2011). This study also showed HGGT primarily uses GGDP as a substrate but can use PDP when GGDP pools are low. Therefore, fruit of *S.pennellii* possibly do not need HGGT to synthesise tocotrienols.

7.1.3 The Vitamin E salvage pathway

S.pennellii is a wild relative of cultivated *S.lycopersicum* tomato and it does not synthesise carotenoids, such as lycopene in its fruit. Carotenoid accumulation is associated with ripening in *S.lycopersicum* fruit (Fraser et al., 1994). However, in

S.pennellii fruit carotenoids are not synthesised in tomato due to a non-expressed phytoene synthase 1 (PSY1) gene (Bolger et al., 2014), which is the first committed step in carotenoid biosynthesis in fruit. Tomato fruit ripening is also associated with the transition of chloroplasts to chromoplasts and the breakdown of chlorophyll (Hörtensteiner, 2013). *S.pennellii* fruit do not do through the ripening process and they remain green and retain chlorophyll (Perez-Fons et al., 2014).

The breakdown of chlorophyll is important for the regeneration of free phytol for VTE biosynthesis (Spicher et al., 2017). First, the Mg^{2+} ion is removed from the chlorophyll and pheophytin a is produced. The pheophorbide a oxygenase (PAO) pathway then catabolises pheophytin a into red chlorophyll catabolites. This process involves the dephytylation of pheophytin a and free phytol is produced (Hortensteiner and Krautler, 2011). Genes encoding enzymes involved in VTE biosynthesis (VTE5 and VTE6) use the free phytol to regenerate PDP for VTE biosynthesis (Valentin et al., 2006, Mach, 2015, vom Dorp et al., 2015).

7.2 Materials and Methods

7.2.1 Plant Materials

Tomatoes used in the *S.pennellii* x *S.lycopersicum* cv. M82 IL screen were harvested at B+10, seeds were removed, and the pericarp and epidermis were harvested together. These tomatoes were harvested by Dr Vincenzo D'Amelia and only one tomato fruit was used for analysis of each IL line. Fruit from fifty-six ILs out of the seventy-six lines of the *S.pennellii x S.lycopersicum* IL population were analysed (a full list is provided in the appendix), because fruit were not available for all the *S.pennellii* x *S.lycopersicum* cv. M82 ILs. However, fruit from the ILs analysed provided good genome coverage. Tomato fruits were ground to a powder using liquid nitrogen and stored for analysis. All other metabolite and qRT-PCR analyses carried out on the fruit are described in chapter 2.

For the analysis of the backcrossed inbred lines (BILs), seven lines were grown, which included; M82, IL6-4, IL6-3, 6601, 6603, 6343, 6344. Tomatoes were harvested at B+10 days for analysis, seeds were removed, and the pericarp and

epidermis were harvested together. Tomato fruits were frozen and ground to a powder using liquid nitrogen and stored before further analysis. All further metabolite and qRT-PCR analyses carried out on the fruit were described fully in chapter 2. All primers used in this chapter are available in the appendix.

7.2.2 RNA sequencing data of *S.pennellii, S.lycopersicum* and *S.pennellii* x *S.lycopersicum* cv. M82 ILs

RNA-sequencing data from *S.pennellii, S.lycopersicum* cv. M82 and *S.pennellii* x *S.lycopersicum* ILs are publicly available (Lee and Giovannoni, Koenig et al., 2013). Tomato fruits were harvested when 80% of the fruit were ripe and only the pericarp was used for extraction. The same RNA sequencing data were used for trans-eQTL analysis in chapter 3.

7.2.3 SNP analysis and genotyping of the BILs

BILs were generated, genotyped, and kindly provided by Dr Iati Ofner and Prof. Dani Zamir (Ofner et al., 2016). These lines are the products of a cross between the *S.pennellii* and *S.lycopersicum* cv. M82. The progeny from this cross were then backcrossed to the *S.lycopersicum* cv. M82 parents and they provide finer resolution of traits. Each line was shown to have, on average, 2.7 introgressed regions per line (Ofner et al., 2016).

BIL6601 has three introgressed fragments on chromosomes 5, 6 and 10. BIL6603 has four introgressed fragments on chromosomes 5, 6, 10 and 11. BIL6343 has three introgressed fragments on chromosomes 4, 6 and 11. BIL6344 has four introgressed fragments on chromosomes 6, 10, and two on chromosome 4. SNP genotyping for these other introgressed fragments are not shown.

The SNP genotyping of the BILs used in this thesis does not show clearly where the introgressed fragment lies on chromosome 6. The *S.pennellii* introgressed fragment is near to the end of the chromosome, but the SNP genotyping does not define where the introgressed fragment starts in the IL6-4 interval.

7.2.4 Candidate gene mining of IL6-4

The *S.pennellii* x *S.lycopersicum* IL (IL6-4) was mined for candidate genes that might affect tocotrienol biosynthesis. Protein sequences of the candidate genes from *S.lycopersicum* were compared to the predicted protein sequences in *S.pennellii* in the IL6-4 interval using the blastp programs available through the Solgenomics network (Fernandez-Pozo et al., 2015a), to determine their possible protein functions. To identify functional protein domains, protein sequences were annotated using the Interpro database (Finn et al., 2017).

Protein sequences of candidate genes were taken from the Solgenomics website (Fernandez-Pozo et al., 2015a) and the *S.lycopersicum* and *S.pennellii* protein sequences were aligned using Clustal Omega software (Larkin et al., 2007, Goujon et al., 2010).

7.3 Results

7.3.1 Screening of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs and BILs revealed that tocotrienols are produced in IL6-4

A time course analyses (described in chapter 4) of fruit (*S.lycopersicum* cv. M82 and *S.pennellii)* showed that *S.pennellii* fruit contained tocotrienols, which were the dominant form of tocochromanols found in these fruits (figure 7-1). The tocotrienol levels remained constant throughout *S.pennellii* fruit development and beta (β) tocotrienols were the most abundant tocotrienol vitamer in *S.pennellii* fruits (figure 7-1). These data suggested the ability to synthesise tocotrienols in fruit would segregate in the ILs and the gene(s) responsible for tocotrienol biosynthesis could be identified. Therefore, I completed a screen of fifty-six *S.pennellii* x *S.lycopersicum* cv. M82 IL fruit (a full list is in the appendix). Just one of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs contained tocotrienols (figure 7-2). M82 fruit did not produce tocotrienols, but IL 6-4 produced all four tocotrienol vitamers (figure 7-2). Total tocotrienol levels reached 4.97 μg g^{-1} in IL6-4, but there were no detectable tocotrienols in M82 fruit (figure 7-2C).

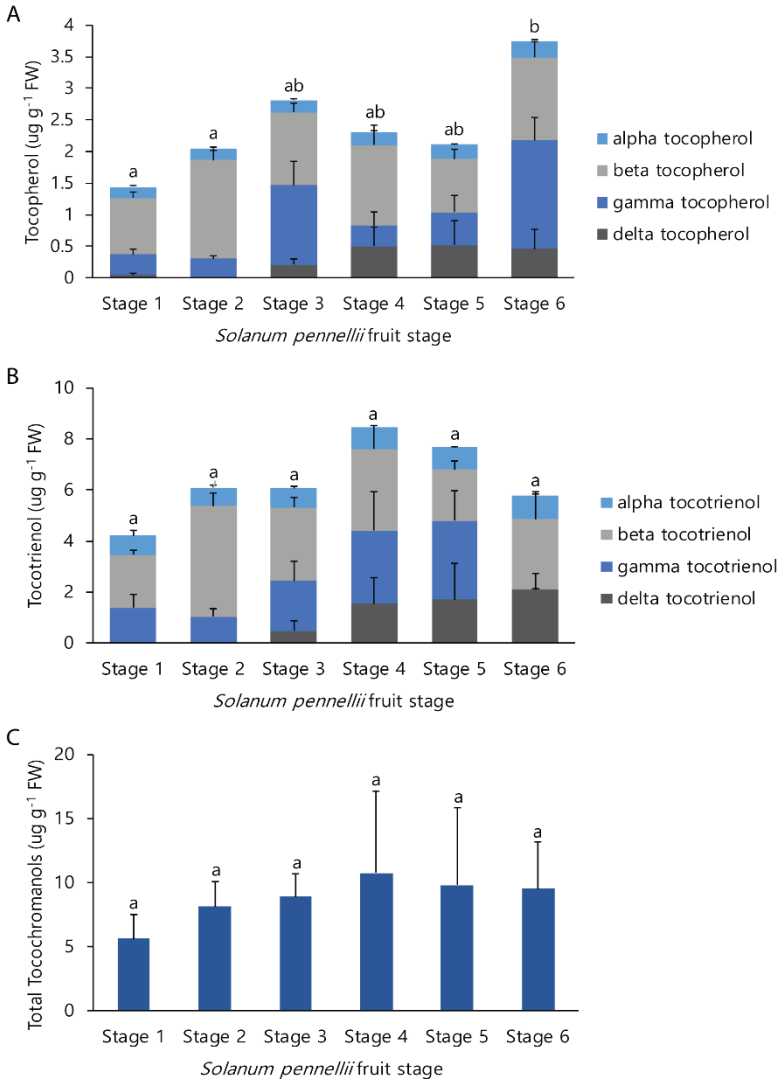
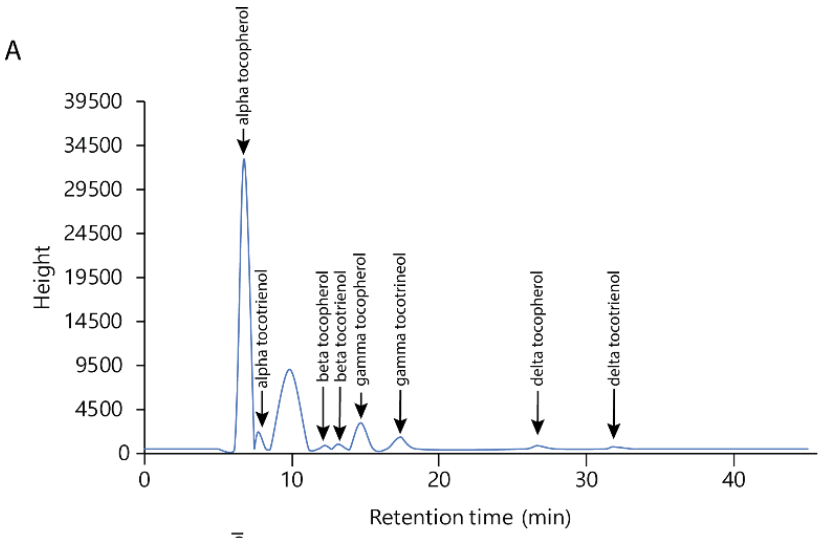
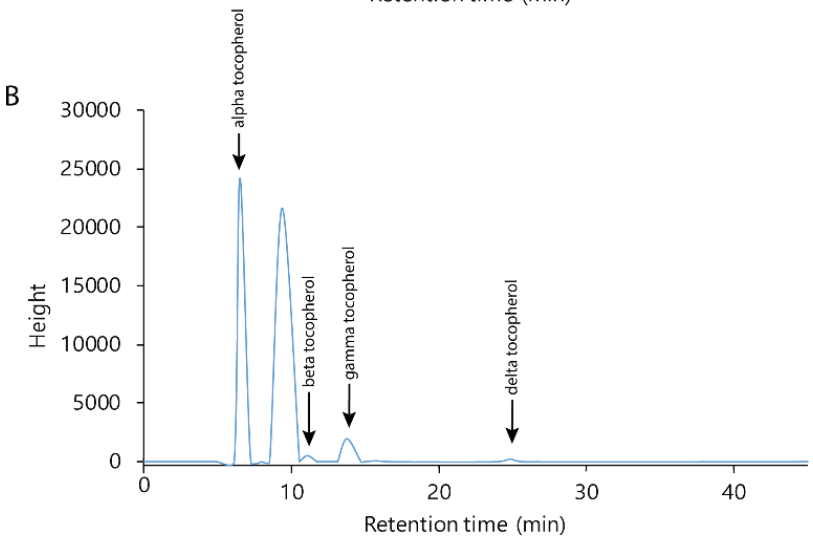


Figure 7-1 Tocopherol (A), tocotrienol (B) and total tocochromanol (C) contents of fruit at the different stages of development in *S.pennellii* (A) shows the levels of different forms of tocopherol in the different fruit stages, alpha (α) tocopherol in light blue, beta (β) tocopherol in light grey, gamma (γ) tocopherol in dark blue and delta (δ) tocopherol in dark grey. (B) shows the tocotrienol content in fruit at the different developmental stages, α-tocotrienol in light blue, β-tocotrienol in light grey, γ-tocotrienol in dark blue and δ-tocotrienol in dark grey. (C) shows the total tocochromanol content (tocopherols and tocotrienols) for each fruit stage. The error bars show standard errors of the mean. The letters indicate the statistical significance ($P<0.05$) of total tocopherol, tocotrienol and tocochromanols levels, which were calculated using Tukey tests. Tukey tests for differences in the levels of the different forms of tocopherol are shown in appendices table 4-3. For each stage (n=4), however, each biological replicate consists of a different total number of fruit for each stage (appendices table 4-1). There were no significant changes in compositional tocotrienol content during fruit development (data not shown).


A
Height
39500
34500
29500
24500
19500
14500
9500
4500
0
alpha tocopherol
alpha tocotrienol
beta tocopherol
beta tocotrienol
gamma tocopherol
gamma tocotrineol
delta tocopherol
delta tocotrienol
0
10
20
30
40
Retention time (min)


B
Height
30000
25000
20000
15000
10000
5000
0
alpha tocopherol
beta tocopherol
gamma tocopherol
delta tocopherol
0
10
20
30
40
Retention time (min)

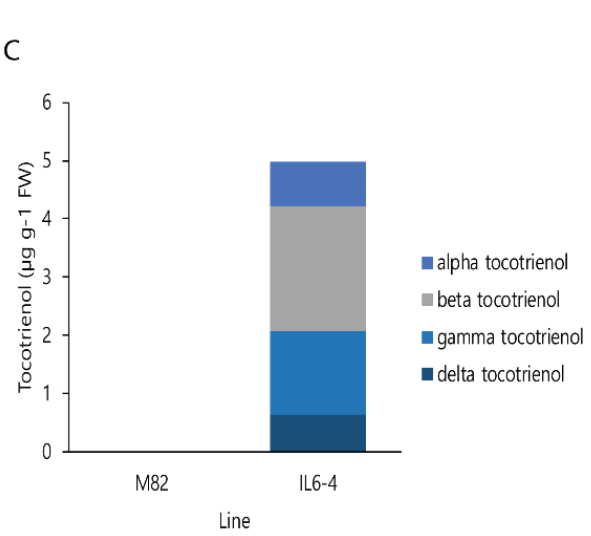

C
Tocotrienol (µg g-1 FW)
6
5
4
3
2
1
0
M82
IL6-4
Line
alpha tocotrienol
beta tocotrienol
gamma tocotrienol
delta tocotrienol

Figure 7-2 Screen of fruit of *S.pennellii* x *S.lycopersicum* cv. M82 ILs for tocotrienols. (A) IL6-4 LC-FDA chromatogram (B) M82 LC-FDA chromatogram (C) tocotrienol contents of M82 and IL6-4. (A) and (B) show the LC-FDA chromatograms that contain alpha (α), beta (β), gamma (γ) and delta (δ) tocopherol and α, β, γ and δ tocotrienols. All tocochromanols are marked by arrows with their name and the chromatograms are based on peak heights. The unmarked peak at approximately ~10 minutes is butylated hydroxytoluene (BHT), which was added to every sample. (C) shows the quantified tocotrienol contents of fruit of M82 and IL6-4 (n=1). α tocotrienol is shown in blue, β tocotrienol in grey, γ tocotrienol in light blue and δ tocotrienol in dark blue. A full list of tomatoes analysed can be found in the appendix.

The LC-FDA chromatograms showed the presence of other tocopherol vitamers in both M82 and IL6-4 fruit (figures 7-2A and B). The introgressed segment of IL6-4 overlaps with IL6-3 (figure 7-3). However, IL6-3 did not contain tocotrienols (data not shown). Therefore, the gene responsible for tocotrienol biosynthesis likely resides in IL6-4.

7.3.2 SNP genotyping of the *S.pennellii* backcrossed ILs

To reduce the size of the region for candidate gene mining, I screened four backcrossed ILs (BILs), which contained the introgressed region of IL6-4 (figure 7-4). Table 7-1 showed the SNP genotyping for the BILs. The BILs (6603, 6343 and 6344) appeared to carry a region of DNA near to the end of the chromosome from *S.pennellii* in a *S.lycopersicum* cv. M82 background (table 7-1). BIL6601 may also carry the same region as the other BILs, although this was not clear as the SNP marker; solcap_snp_31671 from the *S.pennellii* IL parent failed (table 7-1). All the SNP markers shown in table 7-1 reside within the *S.pennellii* IL6-4 region. It is not clear from table 7-1 where the *S.pennellii* introgressed fragment resides in the BILs. The introgressed fragment resides in the IL6-4 and is located towards the end of chromosome 6, but it is not clear from the SNP markers where it begins in the IL6-4 interval.

The SNP genotyping indicated that the BILs all carried the same region of DNA which resided in the IL6-4 region (figure 7-4). However, BIL6601 and BIL6603 had yellow and orange fruit phenotypes (respectively) (figure 7-5), unlike the red fruit phenotype observed in M82 and IL6-4 fruit (figure 7-5). BIL6343 and BIL6344 had red fruit phenotypes (figure 7-5). The orange fruit phenotype was also observed in IL6-3 (figure 7-5), and has been documented in the literature, as due to high activity of CYCβ in ripe fruit resulting in elevated β-carotene levels (Ronen et al., 2000, Ofner et al., 2016). CYCβ encodes for the enzyme responsible for β-carotene biosynthesis from lycopene in chromoplasts (figure 7-6). BIL6601 and BIL6603 SNP genotyping revealed that CYCβ in these BILs was the same allele as in

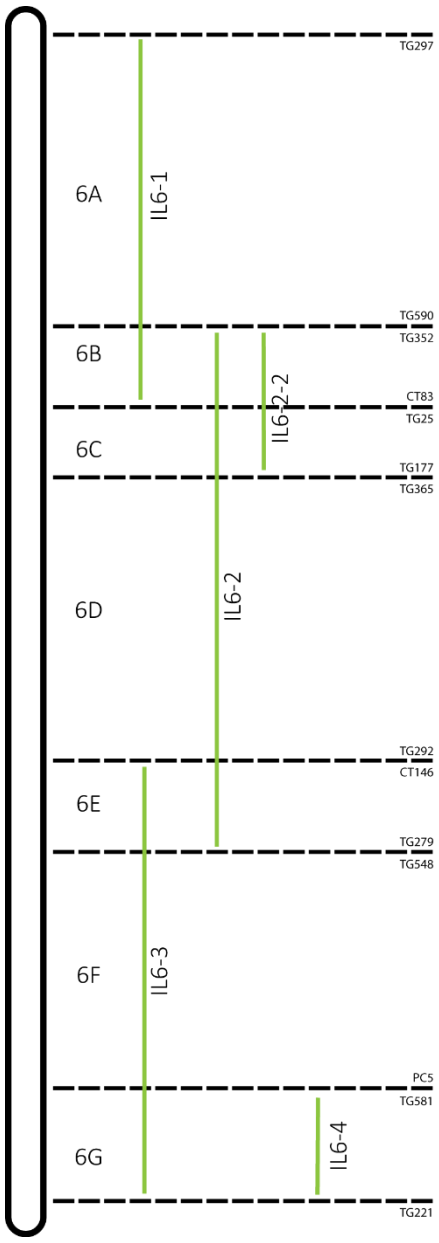


Figure 7-3 Schematic diagram showing *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines (ILs) on chromosome 6. 6A-G indicate bin mapping positions, which are indicated by the dashed lines. At the end of each dashed line, the markers that were used to determine the chromosome position are shown. The green lines delimiting IL6-1 to IL6-4 are physical intervals from *S.pennellii* in the ILs mapped and used for analysis.

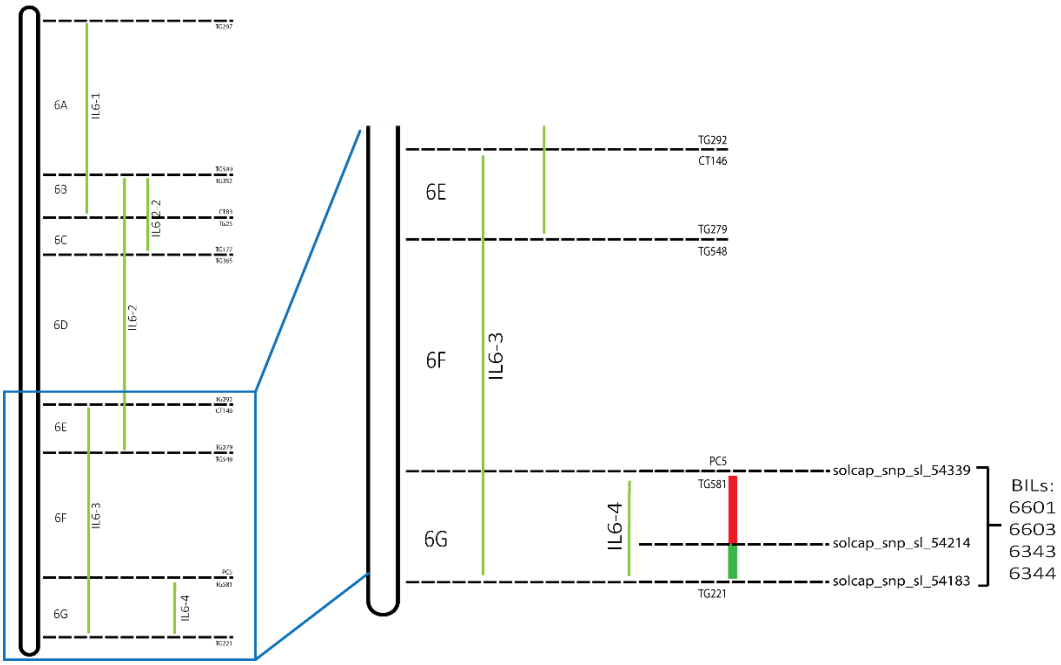


Figure 7-4 Schematic diagram of the *S.pennellii* x *S.lycopersicum* cv. M82 backcrossed inbred lines (BILs) containing the IL6-4 region on chromosome 6. 6A-G indicate bin mapping positions, which are indicated by the dashed lines. At the end of each dashed line, the markers that were used to determine the chromosome position are shown. The green lines delimiting IL6-1 to IL6-4 are physical introgression lines mapped and used for analysis. The BILs result from a cross between *S.pennellii* and *S.lycopersicum* cv. M82. These lines are represented by the red and green lines in the IL6-4 region. The solcap numbers on the dashed lines represent the SNPs that were used to identify these BILs.

Table 7-1 SNP analysis of the *S.pennellii* x *S.lycopersicum* cv. M82 backcrossed inbred lines (BILs). The markers and chromosome positions they define are shown. The *S.lycopersicum* cv. M82 SNPs are shown in red, the *S.pennellii* SNPs are shown in green. The yellow boxes represent SNPs which could be from either parent.

Marker	Chromosome	Position (bp)	Introgression line parents		BILS			
			S.lycopersicum cv.M82	*S.pennellii*	6601	6603	6343	6344
			IL6-4 start					
solcap_snp_sl_54339	ch06	48740122	G	T	G	G	G	G
CL015625-0444	ch06	48745275	G	G	G	G	G	G
solcap_snp_sl_24250	ch06	48746460	C	G	C	C	C	C
solcap_snp_sl_19445	ch06	48780631	G	G	G	G	G	G
U146140_177c	ch06	48787325	C	T	C	C	C	C
U146140_369c	ch06	48787516	T	C	T	T	T	T
U146140_402c	ch06	48787549	C	C	C	C	C	C
U146140_479c	ch06	48787626	T	G	T	T	T	T
solcap_snp_sl_31676	ch06	48838772	A	G	A	A	A	A
solcap_snp_sl_54304	ch06	48846487	G	G	G	G	G	G
solcap_snp_sl_31671	ch06	48940051	C	failed	failed	C	C	C
solcap_snp_sl_54282	ch06	48958590	A	A	A	A	A	A
solcap_snp_sl_54259	ch06	49096703	T	T	T	T	T	T
solcap_snp_sl_54244	ch06	49262285	A	A	A	A	A	A
solcap_snp_sl_54214	ch06	49469915	A	C	C	C	C	C
solcap_snp_sl_43238	ch06	49482087	C	C	C	C	C	C
solcap_snp_sl_54183	ch06	49631131	T	T	T	T	T	T
			IL6-4 end					

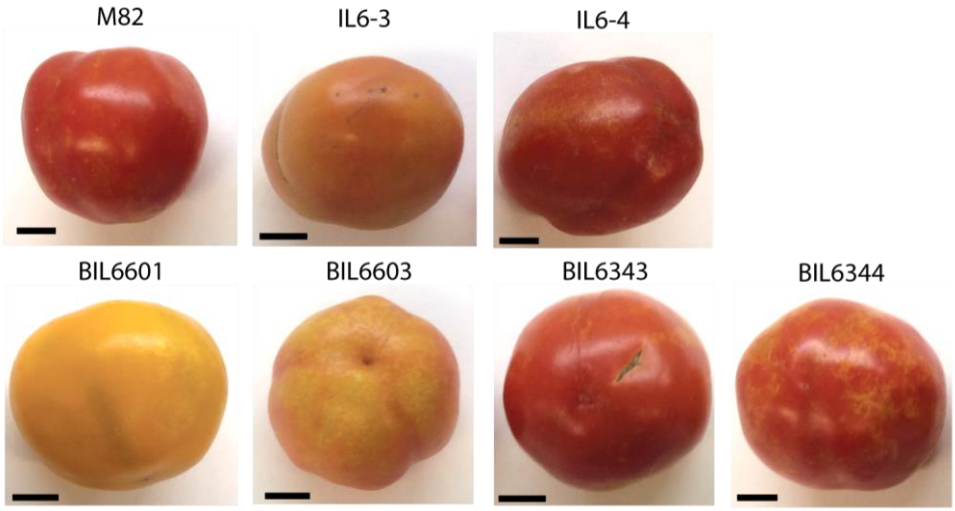


Figure 7-5 Fruit of M82, *S.pennellii* x *S.lycopersicum* cv. M82 ILs and *S.pennellii* x *S.lycopersicum* cv. M82 backcrossed inbred lines (BILs) to show their phenotypes. The black scale bars are 1cm.

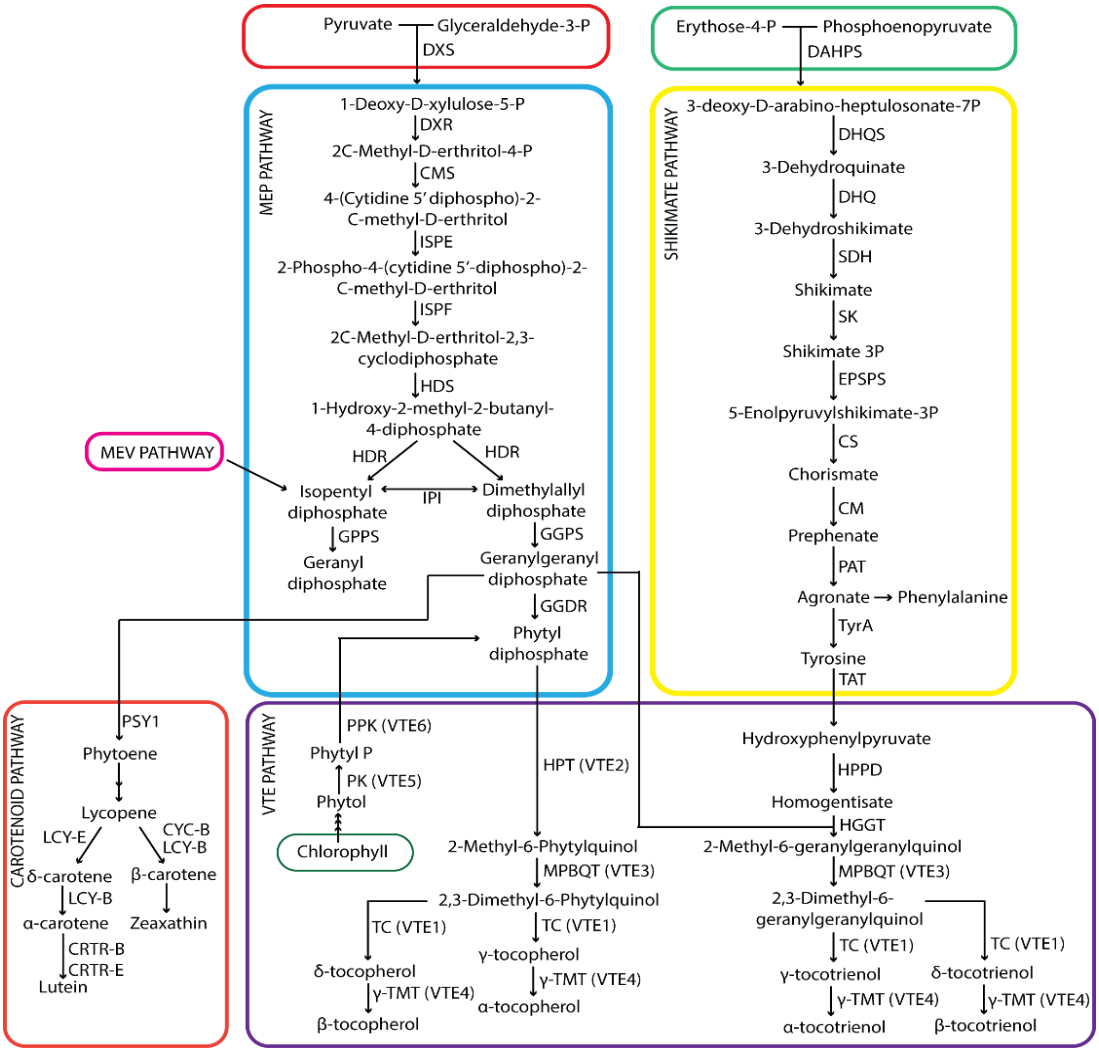

Pyruvate
Glyceraldehyde-3-P
DXS
Erythose-4-P
Phosphoenopyruvate
DAHPS
MEP PATHWAY
1-Deoxy-D-xylulose-5-P
DXR
2C-Methyl-D-erthritol-4-P
CMS
4-(Cytidine 5′ diphospho)-2-C-methyl-D-erthritol
ISPE
2-Phospho-4-(cytidine 5′-diphospho)-2-C-methyl-D-erthritol
ISPF
2C-Methyl-D-erthritol-2,3-cyclodiphosphate
HDS
1-Hydroxy-2-methyl-2-butanyl-4-diphosphate
HDR
HDR
MEV PATHWAY
Isopentyl diphosphate
IPI
Dimethylallyl diphosphate
GPPS
Geranyl diphosphate
GGPS
Geranylgeranyl diphosphate
GGDR
Phytyl diphosphate
SHIKIMATE PATHWAY
3-deoxy-D-arabino-heptulosonate-7P
DHQS
3-Dehydroquinate
DHQ
3-Dehydroshikimate
SDH
Shikimate
SK
Shikimate 3P
EPSPS
5-Enolpyruvylshikimate-3P
CS
Chorismate
CM
Prephenate
PAT
Agronate → Phenylalanine
TyrA
Tyrosine
TAT
Hydroxyphenylpyruvate
HPPD
Homogentisate
HGGT
CAROTENOID PATHWAY
PSY1
Phytoene
Lycopene
LCY-E
CYC-B
LCY-B
δ-carotene
β-carotene
LCY-B
α-carotene
Zeaxathin
CRTR-B
CRTR-E
Lutein
VTE PATHWAY
PPK (VTE6)
Phytyl P
PK (VTE5)
Phytol
Chlorophyll
HPT (VTE2)
2-Methyl-6-Phytylquinol
MPBQT (VTE3)
2,3-Dimethyl-6-Phytylquinol
TC (VTE1)
TC (VTE1)
δ-tocopherol
γ-tocopherol
γ-TMT (VTE4)
γ-TMT (VTE4)
β-tocopherol
α-tocopherol
2-Methyl-6-geranylgeranylquinol
MPBQT (VTE3)
2,3-Dimethyl-6-geranylgeranylquinol
TC (VTE1)
TC (VTE1)
γ-tocotrienol
δ-tocotrienol
γ-TMT (VTE4)
γ-TMT (VTE4)
α-tocotrienol
β-tocotrienol

Figure 7-6 Outline of the tocochromanol pathway. MEP, SK, carotenoid and VTE pathway in blue, yellow, orange, and purple, respectively. Enzyme names are as follows: **DXS**; 1-Deoxy-D-xylulose-5-P synthase, **DXR**; 2C-Methyl-D-erythritol-4-phosphate synthase, **CMS**; 2C-Methyl-D-erythritol-4-phosphate cytidyltransferase, **ISPE**; 4-2-CMethyl-D-erythritol kinase, **ISPF**; 2C-Methyl-D-erythritol-2-3-cyclodiphosphate synthase, **HDS;** 4-Hydroxy-3-methylbut-2-enyl-diphospahte synthase, **HDR**; 4-Hydroxy-3-methylbut-2-enyl-diphosphate reductase, **IPI**; Isopentyl diphosphate δ isomerase, **GPPS;** Geranyl pyrophosphate synthase, **GGPS**; Geranylgeranyl pyrophosphate synthase **GDDR**; Geranylgeranyl reductase, **PSY**; Phytoene synthase, **DAHPS**; 3-Deoxy-D-arabino-hepulosonate, **DHQS**; 3-Dehydroquianate synthase**, SDH-DHQS1**; 3-Dehydroquianate dehydratase**, SDH-DHQS2**; Shikimate 5-dehydrogenase**, SK**; Shikimate kinase**, EPSPS**; 5-Enolpyruvylshikimate-3-P-synthase, **CS**; Chorismate synthase, **CM**; Chorismate mutase, **PAT**; Prephenate aminotransferase, **TyrA**; Arogenate dehydrogenase, **TAT**; Tyrosine aminotransferase, **HPPD**; 4-Hydroxyphenylpyruvate dioxygenase, **HPT (VTE2)**; Homogentisate phytyl transferase, **MPBQMT (VTE3)** Dimethyl-phytylquinol methyl transferase, **TC (VTE1)**; Tocopherol cyclase, **γ-TMT (VTE4)**; γ-Tocopherol C-methyl transferase, **PK (VTE5)**; Phytol kinase, **PPK (VTE6)**; Phytyl-phosphate kinase, **PSY1**; Phytoene synthase 1, **CYC-B**; chromoplast-specific lycopene cyclase, **LYC-B**; lycopene β cyclase , **LYC-E** lycopene ε cyclase , **CRTR-B**; β-ring hydroxylase , **CRTR-E**; ε-ring hydroxylase.

the *S.lycopersicum* cv. M82 parent, and therefore did not carry the *S.pennellii* allele of CYCβ associated with IL6-3 (data not shown).

However, BIL6601 and BIL6603 carried the same allele for LCYβ as the *S.pennellii* parent in another introgressed fragment on chromosome 10 (data not shown). LCYβ encodes for an enzyme that converts β-carotene to lycopene in chloroplasts, and has been mapped to the *orange (og)* mutant in tomato (Ronen et al., 2000). The SNP genotyping revealed that each BIL has several introgressed fragments on several chromosomes, which may affect synthesis of many metabolites in tomato fruit. Therefore, this explained why the tomatoes had different phenotypes.

7.3.3 Analysis of the backcrossed inbred lines revealed different metabolite and expression profiles

IL6-3 and IL6-4 showed higher total tocopherol contents compared to M82 (figure 7-7). However, IL6-4 produced tocotrienols which reached a total level of 28.66 μg g^{-1}, which was higher than the total tocopherol content in this line (25.18 μg g^{-1}) (figure 7-7). This reinforced my preliminary screening of the *S.pennellii* x *S.lycopersicum* cv.M82 ILs, which also showed that IL6-4 produced tocotrienols (figure 7-2). Figure 7-7 showed that very low levels of tocotrienols were present in fruit of IL6-3.

In contrast to the *S.pennellii* x *S.lycopersicum* cv.M82 ILs analysed, the BILs showed different results. Total tocopherols were significantly higher in fruit of BIL 6601, compared to fruit of M82 (figure 7-7). However, fruit of BIL6603, BIL6343 and BIL6344 did not show significantly higher tocopherol contents. Tocotrienols were observed in BIL 6603 and reached levels up to 23.36 μg g^{-1}, but tocotrienols were not observed in the fruit of the other BILs (figure 7-7). δ-tocotrienol was the dominant tocotrienol vitamer in BIL6603. β-tocotrienol was the dominant form of tocotrienols in *S.pennellii* fruit (figure 7-1 and figure 7-7). Fruit of BIL6601 showed significantly lower lycopene and β-carotene contents than M82 (figure 7-7), which was consistent with its yellow phenotype (figure 7-5). Lycopene contents were not

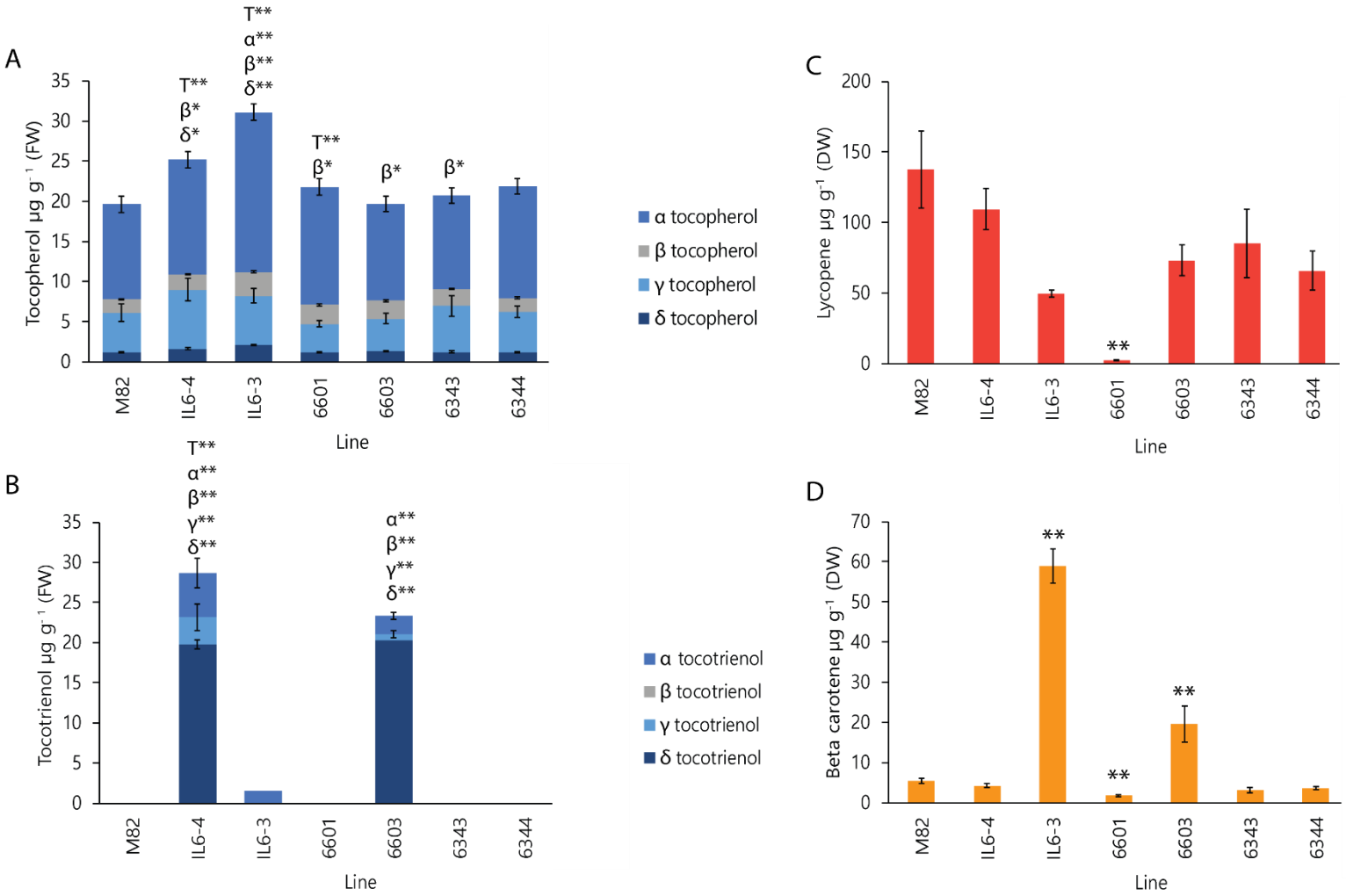

A
Tocopherol µg g⁻¹ (FW)
T**
β*
δ*
T**
α**
β**
δ**
T**
β*
β*
β*
M82
IL6-4
IL6-3
6601
6603
6343
6344
Line
α tocopherol
β tocopherol
γ tocopherol
δ tocopherol
B
Tocotrienol µg g⁻¹ (FW)
T**
α**
β**
γ**
δ**
α**
β**
γ**
δ**
Line
α tocotrienol
β tocotrienol
γ tocotrienol
δ tocotrienol
C
Lycopene µg g⁻¹ (DW)
**
Line
D
Beta carotene µg g⁻¹ (DW)
**
**
**
Line

Figure 7-7 Tocopherol (A), tocotrienol (B), lycopene (C) and beta (β) carotene (D) contents of M82, IL6-4, IL6-3 and BILs. (A) shows tocopherol contents of the BILs analysed; alpha (α) tocopherol is shown in blue, beta (β) tocopherol in grey, gamma (γ) tocopherol in light blue and delta (δ) tocopherol in dark blue. (B) shows tocotrienols contents of the BILs analysed; the alpha (α) tocotrienol in blue, beta (β) tocotrienol in grey, gamma (γ) tocotrienol in light blue and delta (δ) tocotrienol in dark blue. The tocochromanols were measured in μ g g^{-1} fresh weight (FW) and the carotenoids were measured in μg g^{-1} dry weight (DW). The error bars depict standard errors of the mean (n=4) (B) shows that fruit of IL6-3 contained a small amount of α-tocotrienol, but α-tocotrienol was found in only one fruit (n=1). The stars indicate statistical significance calculated using Student's t-test, (*) = $p<0.05$ and (**) = $p<0.01$. The letters represent the statistical significance for different tocopherol and tocotrienol vitamers, (α) alpha (β) beta, (γ) gamma, (δ) delta and (T) total.

significantly different from M82 fruit in the other ILs or BILs analysed. Consistent with the LYCβ allele from *S.pennellii,* fruit of IL6-3 and BIL6603 showed significantly higher β- carotene contents compared to M82 (figure 7-7), which was reflected in their orange phenotypes (figure 7-5).

Analysis of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs and BILs indicated that the BILs had distinct gene expression profiles. Genes that encoded enzymes in the VTE pathway (*VTE2, VTE3(1), VTE1, VTE4, VTE5* and *VTE6)* were down regulated in IL6-4, BIL6601 and BIL6344, compared to M82 (figure 7-8). *VTE4* and *VTE5* were slightly upregulated in IL6-3, BIL6603 and BIL6343, in comparison to M82, although expression was variable for each line (figure 7-8). In IL6-4, expression of many genes encoding enzymes in the carotenoid, MEP and SK pathways were down regulated, or remained at similar levels to M82. This suggested that these genes/enzymes were not good candidates for tocotrienol biosynthesis.

BIL6603 produced tocotrienols and *TAT1* was upregulated and *VTE2* and *VTE1* were down regulated in BIL6603 fruit compared to fruit of the M82 control (figure 7-8). These genes showed similar expression patterns in IL6-4 (figure 7-8), which suggested that regulation of these genes might be important for tocotrienol biosynthesis. *CYCβ* was more highly expressed in IL6-3 and BIL6603 than in other lines, consistent with the increase in β-carotene contents (figure 7-7 and 7-8) (Ronen et al., 2000). BIL 6343 showed many genes encoding enzymes were upregulated slightly, and *TAT1* showed the largest changes in expression, but fruit of the BIL6343 did not produce tocotrienols (figure 7-7 and 7-8). This suggested that *TAT1* was not a strong candidate for tocotrienol biosynthesis. BIL6344 showed a general decrease of *DXS(1), CYCβ, HST, HPPD1, HPPD2, VTE2, VTE1, VTE4, VTE5* and *VTE6* transcript levels in fruit compared to M82 (figure 7-8). However, these changes in VTE gene expression was not reflected in differences in the metabolite data for BIL6344 (figure 7-7).

My analyses of BILs showed conflicting results and I cannot conclude that the gene responsible for tocotrienols resides in the regions of *S.pennellii* DNA

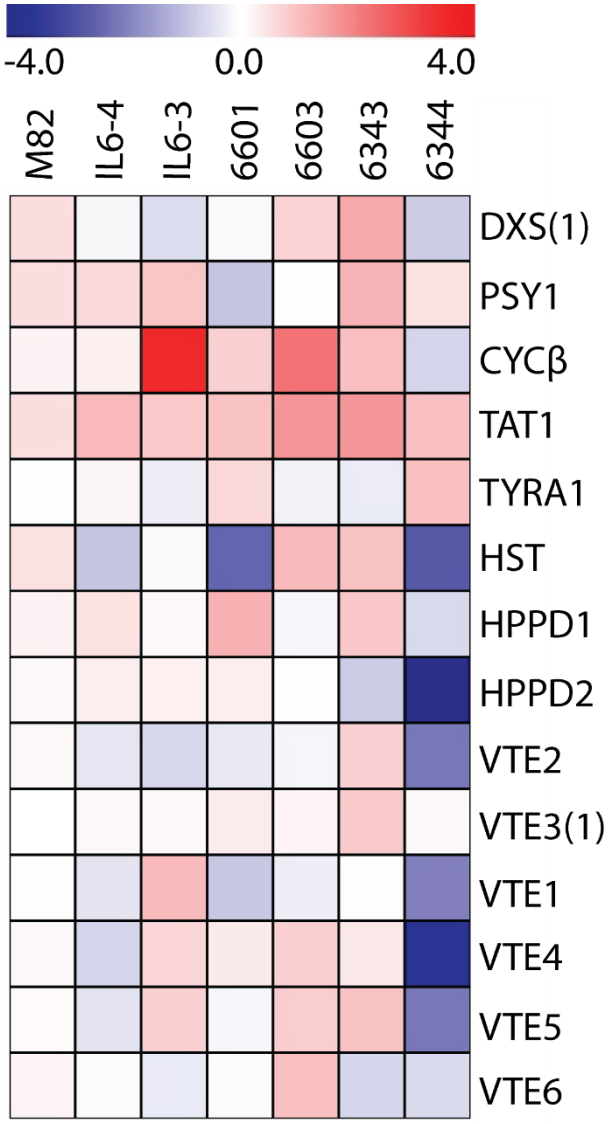


Figure 7-8 Relative expression of genes in MEP, SK, VTE and carotenoid biosynthesis pathways for the IL and BIL analysis (n=3). The relative fold changes are relative to M82. SlCAC was used as a reference gene. The scale bar shows relative expression on a scale of 0.0-5.0, from blue to red.

present in the BILs as they had many introgressed fragments on other chromosomes, which affected the production of carotenoids and possibly tocochromanols in fruit. Fruit of IL6-4 contained tocotrienols, thus it was likely that the candidate gene does reside in this region. Therefore, the whole IL6-4 region was mined to identify candidate gene(s) from *S.pennellii* determining tocotrienol biosynthesis.

7.3.4 The gene(s) responsible for tocotrienol biosynthesis in *S.pennellii* fruit was not 'gained' or 'lost' from *S.pennellii* or *S.lycopersicum*

The 'gene' responsible for tocotrienol biosynthesis might be a new gene that was introgressed from the *S.pennellii* parent into the *S.lycopersicum* cv. M82 background when they were crossed to generate the ILs. Table 7-2 shows the 'candidates' that were introgressed as new genes from *S.pennellii* and were not present in *S.lycopersicum* cv. M82 background. Therefore, these genes were 'gained' by IL6-4.

Genes that were identified in this region included; two caffeic 3-O-methyltransferases (Sopen06g03810 and Sopen06g03820), two Tudor proteins (Sopen06g035620 and Sopen06g035630) and many proteins with no known function (Sopen06g034870, Sopen06g035060, Sopen06g035080, Sopen06g035090 and Sopen06g035420) (table 7-2). None of these genes were likely to be involved in tocotrienol biosynthesis as they cannot catalyse the reaction needed to synthesise tocotrienols.

The IL6-4 donor fragment from *S.pennellii* might have replaced genes that were in the *S.lycopersicum* cv. M82 region that inhibited tocotrienol biosynthesis. Therefore, tocotrienol production in tomato fruit could be due to loss of a gene from the *S.lycopersicum* cv. M82 background. Table 7-3 shows the *S.lycopersicum* genes that were replaced by the *S.pennellii* fragment. Many of the genes 'lost' in the *S.lycopersicum* region had no known function and the protein sequences had no predicted functional domains. The genes with predicted functions (Solyc06g083340, Solyc06g083400, Solyc084490, Solyc06g084760) were unlikely

Table 7-2 Genes that lie within the IL6-4 *S.pennellii* region that has been introgressed into *S.lycopersicum* background in IL6-4 and are not present in the equivalent interval in *S.lycopersicum*. The Solgenomics identifier, coding sequence (CDS) and predicted protein functions are shown.

***S.pennellii* genes in IL6-4 that are not present in the *S.lycopersicum* region**	**CDS?**	**Predicted function based on protein domains**
Sopen06g03810	CDS	Caffeic acid 3-O-methyltransferase
Sopen06g03820	CDS	Caffeic acid 3-O-methyltransferase
Sopen06g034870	CDS	unknown protein function which has ribonuclease H domain
Sopen06g035060	CDS	unknown protein function and no predicted protein domain
Sopen06g035080	CDS	unknown protein function and no predicted protein domain
Sopen06g035090	CDS	unknown protein function and no predicted protein domain
Sopen06g035420	CDS	Unknown protein function and no predicted protein domain
Sopen06g035620	CDS	Tudor protein agenet domain
Sopen06g035630	CDS	Tudor protein agenet domain

Table 7-3 Genes that lie in the region of *S.lycopersicum* that are not present in the introgressed segment from *S.pennellii* in IL6-4. The Solgenomics identifier, CDS and predicted protein functions are shown.

Genes that are present in *S.lycopersicum* region but not in *S.pennellii* region	**CDS?**	**Predicted function based on protein domains**
Solyc06g083240	no CDS	unknown function and no predicted protein domain
Solyc06g083340	CDS	Wound induced basic protein
Solyc06g083400	CDS	Xyloglucan endotransglucosylase/hydrolase
Solyc06g083740	CDS	unknown function which has a non-cytoplasmic domain
Solyc06g084490	CDS	transport inhibitor response
Solyc06g084560	CDS	unknown function and no predicted protein domain
Solyc06g084660	no CDS	unknown function and no predicted protein domain
Solyc06g084760	CDS	cDNA cytochrome P450 monooxygenase
Solyc06g084770	no CDS	unknown function and no predicted protein domain

to affect tocotrienol biosynthesis, as these genes are involved in wound healing or cell wall biosynthesis based on their annotation. These data indicated that tocotrienol production was unlikely to be due to 'gain' or 'loss' of a specific gene function. These data suggested that tocotrienol biosynthesis was due to a gene which was differentially expressed, or an ortholog which functions differently between the *S.pennellii* and *S.lycopersicum.*

7.3.5 Gene mining of IL6-4 revealed that a gene encoding a transcription factor might affect flux of available substrates into the VTE pathway

The gene responsible for tocotrienol biosynthesis in fruit of *S.pennellii* was likely to be present in *S.lycopersicum* cv. M82 and *S.pennellii* genomes. The candidate gene could encode an enzyme, or a transcription factor (TF). Orthologs of the candidate enzyme or TF might function differently or might be differentially expressed in the two species, either of which could result in tocotrienol biosynthesis.

To identify whether an enzyme or TF was responsible for the observed phenotype of fruit of IL6-4, I identified the eleven enzymes encoded by genes that reside in IL6-4 (table 7-4). These enzymes are involved in cell wall expansion and biosynthesis (Van Sandt et al., 2007, Marin-Rodriguez et al., 2002), ascorbate antioxidant recycling (Packer et al., 2001), lignin biosynthesis (Zubieta et al., 2002) and gibberellin biosynthesis (Koksal et al., 2011). Several of these enzymes were differentially expressed in the *S.pennellii* x *S.lycopersicum* cv. M82 IL; IL6-4 and in *S.pennellii* (table 7-4), compared to *S.lycopersicum* cv. M82. However, none of the enzymes identified were predicted to be able to catalyse the reaction between GGDP and homogentisate to produce tocotrienols. Some of these enzymes have a tentative link to VTE biosynthesis, such as shikimate dehydrogenase (SDH) which is involved in shikimate biosynthesis (figure 7-6). Glutathione transferase is involved in ascorbate recycling which interacts with the VTE antioxidant recycling (Packer et al., 2001). However, the protein alignments of the orthologs of these

Table 7-4 Relative expression of candidate genes from RNA sequencing data from fruit of *S.pennellii* x *S.lycopersicum* cv. M82 introgression lines (ILs), *S.pennellii* and *S.lycopersicum.* The transcript levels of candidate genes that lie within IL6-4 are shown relative to these of fruit of M82. This table shows the ratio of expression between the *S.pennellii* IL parents (*S.lycopersicum* cv. M82 : *S.pennellii)* and the statistical significance between the parents.

Solgenomics identifier	Gene description	Relative fold change of expression compared to M82.			Ratio of expression between *S.lycopersicum* and *S.pennellii*	Statistically significant, $p < 0.05$ (Y/N)
		M82	IL6-3	IL6-4		
Solyc06g083310	Polygalacturonate 4-alpha-galacturonosyltransferase	1	1.1	1.34	2.51	Y
Solyc06g083380	Acyl-acyl carrier protein (ACP) thioesterase	1	2.61	2.52	3.48	Y
Solyc06g083400	Xyloglucan endotransglucosylase	1	0	0	0.33	N
Solyc06g083450	Caffeic acid 3-O-methyltransferase (COMT)	1	0	0	18.33	N
Solyc06g083580	Pectate lyase family protein	1	4.36	8.78	94.76	Y
Solyc06g083650	GDSL esterase/lipase	1	0	0	1	N
Solyc06g083770	Glutathione transferase	1	1.37	0.91	0.37	N
Solyc06g084030	S-adenosyl-L-methionine-dependent methyltransferase	1	5.5	9	187.63	Y
Solyc06g084240	Ent-copalyl diphosphate synthase	1	0	0	1	N
Solyc06g084460	Shikimate 5-dehydrogenase	1	0.78	0.59	4.81	Y
Solyc06g084570	S-adenosyl-L-methionine-dependent methyltransferase	1	0.76	1.1	0.83	N

enzymes showed that there were very few differences between them in the two species (data not shown). This analysis suggested that it was unlikely that a gene encoding an enzyme was responsible for tocotrienol biosynthesis in *S.pennellii*, and rather a TF, which might alter flux into the VTE biosynthetic pathway was a more likely candidate responsible for tocotrienol production.

Further evidence supporting this hypothesis is presented in table 7-5, which shows the relative expression of the genes encoding VTE pathway enzymes in the *S.pennellii* x *S.lycopersicum* cv. M82 ILs; IL6-3 and IL6-4. Many of the VTE genes were not differentially expressed in the RNA sequencing data of IL6-4 compared to IL 6-3 (table 7-5), although *VTE2* and *VTE6* were upregulated. This differs from the BIL and IL expression data, which showed that VTE genes were down regulated in IL6-4 (figure 7-8) and qRT-PCR is more reliable than RNA sequencing data. *VTE6* is responsible for the VTE salvage pathway, to regenerate PDP from phytol (figure 7-6). However, GGDP, not PDP, is the substrate for tocotrienol biosynthesis (Yang et al., 2011). *VTE2* has been proposed to use GGDP as a substrate to synthesise tocotrienols when PDP pools are low (Yang et al., 2011), therefore a change in *VTE2* expression might alter flux affecting GGDP and PDP pools. Consequently, a TF that altered the expression of enzymes in the MEP pathway could alter the availability of substrates for *VTE2,* and so affect the synthesis tocotrienols from GGDP. However, a TF that affects both *VTE2* and *VTE6* expression would be unlikely to alter substrate availability and would probably not result in tocotrienol biosynthesis, because *VTE6* is responsible for the regeneration of PDP for tocopherol biosynthesis.

The RNA sequencing data showed that expression of genes encoding enzymes in the MEP pathway (figure 7-9) (*ISPF, HDS* and *HDR)* were reduced in IL6-3 and IL6-4, compared to M82. I did not measure the expression of these genes in my BIL and IL analysis, therefore I have used the RNA sequencing data to understand the possible cause of tocotrienol biosynthesis in *S.pennellii* fruit. Increased *GGPS* expression might result in increases in the pool of GGDP (figure 7-6 and 7-9). However, the *GGPS* isoforms were similarly expressed in IL6-4 and

Table 7-5 Relative fold change in expression from *S.pennellii* x *S.lycopersicum* cv. M82 introgression line (IL) RNA sequencing data, compared to M82. The table shows the genes encoding enzymes in the Vitamin E pathway and their relative expression in IL6-3 and IL6-4.

Gene	Solgenomics identifier	Relative fold change of expression compared to M82		
		M82	IL6-3	IL6-4
VTE2	Solyc07g017770	1	0.74	1.28
VTE3(1)	Solyc09g065730	1	0.95	0.91
VTE3(2)	Solyc03g005230	1	0.89	0.83
VTE1	Solyc08g068570	1	1.45	0.75
VTE4	Solyc08g076360	1	0.99	0.96
VTE5	Solyc09g018510	1	0.65	0.75
VTE6	Solyc07g062180	1	0.87	1.36

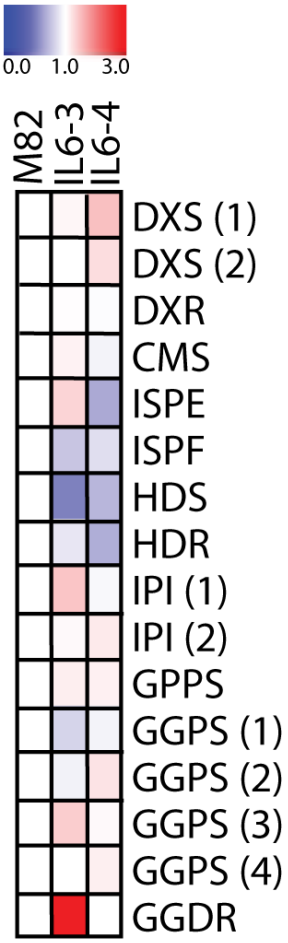


Figure 7-9 Heatmap of relative gene expression calculated from the *S.pennellii* x *S.lycopersicum* cv. M82 fruit RNA sequencing data for genes encoding enzymes of the MEP pathway for IIL6-3 and IL6-4 (Lee and Giovannoni). The values are shown relative to M82, which equals 1. The scale bar ranges from 0 to 3-fold changes in expression.

M82. Interestingly, expression of *GGDR* was highly upregulated in IL6-3, but was not differentially expressed in IL6-4. This gene catalyses the reduction of GGDP to PDP (figure 7-6), therefore the low GGDR expression in IL6-4 might result in higher GGDP pools, relative to PDP pools, which could skew VTE biosynthesis towards tocotrienols. This suggested that the candidate gene controlling tocotrienol biosynthesis might encode a TF that represses GGDR expression.

7.3.6 Gene mining of IL6-4 revealed several candidate transcription factors

Eight genes encoding candidate TFs were identified in IL6-4 (table 7-6). The candidate TFs identified in IL6-4 are involved in a myriad of roles, which include; responses to light (Wan et al., 2015), seed germination and dormancy (Willmann et al., 2011), ABA signalling (Bassel et al., 2006, Delmas et al., 2013), floral stigma development (Gao et al., 2018) and jasmonate synthesis (Fonseca et al., 2014, Huang et al., 2018). Three of the candidate TFs (Solyc06g083170, Solyc06g083430 and Solyc06g083900) are highly expressed in *S.pennellii* fruit (table 7-6), which suggested that they might be transcriptional activators of the MEP pathway. Also, more than one TF may be responsible for tocotrienol biosynthesis. *GGDR* expression was not altered in IL6-4, unlike in IL6-3 (figure 7-9), which implied that a transcriptional repressor might inhibit GGDR expression or a repressor might inhibit PDP synthesis, which might result in larger GGDP pools, relative to PDP pools. This suggested that the gene would be highly expressed in *S.pennellii* compared to *S.lycopersicum,* and there were three candidate TFs (Solyc06g083170, Solyc06g083430 and Solyc06g083900) which were highly expressed in *S.pennellii.* Further characterisation of these candidate genes is needed to determine whether any of the candidate genes are responsible for tocotrienol synthesis in *S.pennellii*.

7.4 Discussion

7.4.1 Tocotrienols are present in *S.pennellii* and were identified in one IL, IL6-4

S.pennellii fruit contained high levels of tocotrienols (as described in chapter 4), even though tomato does not contain a gene encoding HGGT

Table 7-6 Candidate transcription factors identified in *S.pennellii* IL 6-4. The relative expression of the candidate genes in IL6-3 and IL6-4 are shown in comparison with M82. RNA sequencing data of *S.pennellii* and *S.lycopersicum* are shown as a ratio between the two. If the ratio is statistically significant, the p value <0.05.

Gene name	Solgenomics identifier	Gene annotation	Relative expression in *S.pennellii* RNA sequencing data			Ratio of expression between *S.pennellii* and *S.lycopersicum* RNA sequencing data		
			M82	IL6-3	IL6-4	Ratio	p value	Highest expression in *S.lycopersicum* (L) or *S.pennellii* (P)
SlbHLH62	Solyc06g083170	bHLH	1	1.6	4.8	6.53	0.19	P
SlTrhelix1	Solyc06g083430	Trihelix	1	0.81	0.77	1.21	0.2	P
SlABI3	Solyc06g083590	bZIP TF	1	0	0	7.62	0.67	L
SlABI3	Solyc06g083600	bZIP TF	1	0	0	3.33	0.86	L
SlNAC74	Solyc06g083840	NAC domain TF	1	0	0	0.04	0.98	L
SlNAC74	Solyc06g083850	NAC domain TF	1	0.05	0	0.08	0.23	L
SlMYB13	Solyc06g083900	R2R3 MYB	1	0	0	121.33	0.001	P
SlbHLH3	Solyc06g083980	bHLH	1	0.62	0.51	0.4	0.002	L

(Cahoon et al., 2003). To identify the region responsible for tocotrienol production in *S.pennellii* fruit, I completed a screen of tomato fruit of a subset of the *S.pennellii* ILs. IL6-4 produced all four tocotrienol vitamers, unlike *S.lycopersicum* cv.M82 which did not contain any tocotrienols.

Further analysis of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs and BILs revealed that fruit of IL6-4 and BIL6603 produced tocotrienols. The BIL and IL screen showed that tocotrienols were accumulated in IL6-4 fruit, but not in IL6-3. The four BIL lines that were analysed showed that fruit of BIL6603 also produced tocotrienols, although unfortunately this BIL had the most introgressed fragments on other chromosomes compared to the other BILs. The BIL expression data suggested that *VTE2* and *VTE1* could be important enzymes contributing to tocotrienol biosynthesis in tocotrienol producing fruit. The RNA sequencing data showed that *VTE2* expression was upregulated in IL6-4. However, the RNA sequencing data was carried out when 80% of the fruit were ripe. Therefore, a mix of fruit stages may have affected the expression data, which might explain why this is different from my expression data. VTE2 can use GGDP as a substrate to produce tocotrienols (Yang et al., 2011). Down regulation of *VTE2* was observed in my expression data of IL6-4 and BIL6603, which may be a result of tocotrienol biosynthesis. These contradictory results need to be resolved before a reliable conclusion can be made to explain why tocotrienols are synthesised in *S.pennellii* and IL6-4. A time course throughout fruit development and ripening of *S.pennellii* and IL6-4 would elucidate the key regulatory points during tocotrienol biosynthesis.

IL6-3 fruit contained α-tocotrienol, but this was observed in a single fruit, so this cannot be considered a reliable result. If the presence of tocotrienols in IL6-3 can be confirmed, the gene responsible could lie in the overlapping segment between IL6-3 and IL6-4, but the absence of tocotrienols in most IL6-3 fruit suggested that this was not a valid interpretation. In conclusion, tocotrienols were observed in IL6-4 and in one BIL6603, but it was not clear where the gene(s) responsible resided. Therefore, IL6-4 was mined for candidate genes.

7.4.2 Tocotrienol biosynthesis correlated with beta carotene contents in the BIL tomato fruit

BIL6601 and BIL6603 had yellow and orange fruit phenotypes (respectively), which was similar to the IL6-3 orange phenotype reported by Ronen et al (2000). Fruit of BIL6603 produced tocotrienols and had higher levels of β-carotene, compared to fruit of M82. Fruit of BIL6601 did not have significantly altered β-carotene contents nor tocotrienols, and instead their lycopene contents were significantly reduced. This coupled with the fact the BILs have more than one introgressed fragment, suggests that the metabolite contents of the fruit might be altered. Fruit of IL6-4 produced tocotrienols but did not have higher levels of β-carotene. This shows that the BILs are not a useful tool to identify candidate genes for tocotrienol biosynthesis, and that any apparent association between β-carotene content and tocotrienol production is unlikely to be reliable.

Transgenic sorghum lines overexpressing *SB-AKAF::HGGT, SB-AKAF::DXS, SB-AKAF::PSY1* and *SB-AKAF::CRT1* showed that β-carotene was increased over 18-fold and the seeds were visibly orange, in comparison with the wild type (WT) (Che et al., 2016). These lines overexpress several carotenoid biosynthetic genes as well as HGGT, resulting in six fold increases in tocotrienols in these transgenic lines (Che et al., 2016). These authors showed that tocotrienols prevent the oxidation of β-carotene during storage tests, compared to WT. Fruit of BIL6603 also contained tocotrienols and had higher β-carotene contents compared to WT fruit, which implied that the presence of tocotrienols might also prevent the oxidation of β-carotene in tomato fruit. However, *CYCβ* expression was increased in IL6-3 and BIL6603. Elevated *CYCβ* expression in IL6-3 is expected as the gene encoding CYCβ maps to this IL (Ronen et al., 2000). The SNPs genotyping showed that BIL6603 carried the *S.lycopersicum* allele of CYCβ and the *S.pennellii* allele of LYCβ (data not shown). Therefore, the *S.pennellii* LYCβ allele suggested why β-carotene levels were elevated in fruit of BIL6603 and IL6-3. Fruit of IL6-4 produced tocotrienols but did not have higher β-carotene content. This line also did not show increased

CYCβ expression, therefore it is likely that higher β-carotene levels are due to the LYCβ allele and not due to the presence of tocotrienols in BIL6603. The production of tocotrienols in both IL6-4 and BIL6603 suggests tocotrienol production occurs independently of β-carotene and LYCβ/CYCβ expression. By analogy to the report of Che et al, (2016) the production of tocotrienols could prevent oxidation of β-carotene and consequently give rise to higher levels of β-carotene in BIL6603, but increased β-carotene was not observed in IL6-4. Fruit with increased β-carotene may have other applications for nutritional quality of tomato fruit. However, these data are not conclusive due to the problems with the BIL genotyping and further markers need to be used to determine the introgressed fragments.

7.4.3 *S.pennellii* fruits are still photosynthetically active and might synthesise tocotrienols to maintain ROS homeostasis.

VTE prevents lipid peroxidation and scavenges singlet oxygen (O_2^-) species produced by photosynthesis. Scavenging O_2^- is the main function of VTE in photosynthetic organisms (Havaux et al., 2005, Sattler et al., 2006). *S.pennellii* fruits do not break down chlorophyll during ripening. Therefore, the fruits remain green and are still photosynthetically active, unlike its red-fruited *S.lycopersicum* relative. Chlorophyll breakdown is an important part of the VTE salvage pathway, which regenerates PDP from free phytol derived from chlorophyll breakdown (Valentin et al., 2006, vom Dorp et al., 2015). This suggests that there is demand in *S.pennellii* fruit to maintain reactive oxygen species (ROS) homeostasis, as fruit remain photosynthetically active and tocotrienols may be produced in *S.pennellii* fruit in response to high ROS levels. An absence of chlorophyll breakdown may result in low PDP pools, and higher GGDP pools and could skew VTE biosynthesis towards tocotrienols, depending on substrate availability (Yang et al., 2011). This provides a verifiable hypothesis to explain why tocotrienols are produced in *S.pennellii* and IL6-4 fruit.

Further investigation is required to determine whether the candidate *S.pennellii* genes or alleles in IL6-4 might be responsible for tocotrienol

biosynthesis. GGDP and PDP pools would need to be measured in IL6-4 and *S.pennellii* fruit to ensure VTE2 is indeed the enzyme responsible for tocotrienol biosynthesis. ROS assays that establish the antioxidant capacity of *S.pennellii* fruit would provide an explanation for why tocotrienols are synthesised in fruit. Transient assays to silence and over-express these candidate genes tomato fruit would allow for the identification of the gene responsible. Further stable transformations and CRISPR techniques could provide functional characterisation. Further analysis of other green fruited tomato relatives might provide additional insights into tocotrienol biosynthesis and show whether tocotrienol production is a common feature of photosynthetically-active tomato fruit.

The stay green tomato mutant; *green-flesh (gf),* does not degrade chlorophyll, but accumulates carotenoids during ripening (Barry et al., 2008). Therefore, the *gf* mutant is brown at the ripe fruit stage (Barry et al., 2008). This mutant contains a mutation of an invariant amino acid of the STAY-GREEN protein (Barry et al., 2008), which is thought to de-stabilise the light harvesting complex pigments during senescence (Sakuraba et al., 2014). The *gf* tomato mutant should produce less PDP, which is normally regenerated from free phytol from chlorophyll degradation, and so GGDP pools should be higher. Analysis of the *gf* mutant showed that all tocopherol vitamers and total tocopherols were higher in ripe fruit of the *gf* mutant, compared to mature green *gf* fruit (Almeida et al., 2015). Also, the composition of β-, γ- and δ-tocopherols levels were higher in *gf* ripe fruit, compared to fruit of the WT. This was reflected in the expression data, which showed that genes encoding the enzymes of VTE biosynthesis were differentially expressed (Almeida et al., 2015). Genes encoding enzymes of the MEP pathway were upregulated, which suggested that these genes were expressed in response to the lack of PDP regeneration in ripe fruit (Almeida et al., 2015). However, these authors did not analyse *gf* fruits for tocotrienol accumulation. Therefore, further work is needed to determine if fruits of *gf* accumulate tocotrienols. This would support the hypothesis that tocotrienols accumulate because of higher GGDP

levels, compared to PDP. This would provide further support for the accumulation of tocotrienols in *S.pennellii* fruit.

In conclusion, the data presented suggest that tocotrienols may be produced in *S.pennellii* fruit to meet the demands of the photosynthetically active fruit for scavenging for ROS. The *S.pennellii* x *S.lycopersicum* cv. M82 ILs have identified several TFs as potential targets for engineering VTE biosynthesis for tocotrienol production in tomato fruit which might be beneficial for increasing antioxidant potential in tomato fruit. However, tocotrienols are not preferentially retained by the α-tocopherol transfer protein in the liver of humans, compared to α-tocopherol (Hosomi et al., 1997). Therefore, they have a lower bioavailability in humans. Engineering tocotrienol-enriched tomato fruit may not provide nutritionally enhanced fruit for humans. Tocotrienols are probably produced at the expense of α-tocopherol biosynthesis, but they may still have specific beneficial effects in plant tissues.

Chapter 8: General Discussion

Chapter 8: General discussion and outlook

8.1.1 General summary

Vitamin E (VTE) is a lipophilic antioxidant that is synthesised in photosynthetic organisms. Humans do not synthesise VTE and obtain all their VTE from dietary sources. Studies suggest that increasing VTE in the diet can aid in the prevention of diseases (Fitzpatrick et al., 2012, Kris-Etherton et al., 2002). Therefore, there are potential benefits of eating VTE rich foods. There have been many attempts to biofortify foods to increase their VTE content although these have usually relied on manipulating expression of genes encoding enzymes of the VTE pathway. Consequently, there is a limit on substrate availability and pathway flux, which means that VTE composition is often skewed towards producing less bioactive vitamers (Shintani and DellaPenna, 1998, Schledz et al., 2001, Collakova and DellaPenna, 2001, Savidge et al., 2002, Shintani et al., 2002, Cheng et al., 2003). The aim of my thesis was to identify transcriptional regulators of VTE biosynthesis that might be used to enhance VTE levels in tomato. Metabolic engineering of transcriptional regulators would overcome the limitations imposed on pathway flux by individual biosynthetic steps, which could result in nutritionally enhanced crops.

There were no known transcription factors that regulate VTE biosynthesis transcriptionally at the start of my PhD project. My data have shown that VTE biosynthesis in tomato fruit is regulated transcriptionally in two species of tomatoes. *Solanum lycopersicum* cv. M82 and *Solanum pennellii* showed differential expression of VTE biosynthetic genes and metabolite profiles during fruit development, which suggest that these tomato species have biologically distinct regulatory mechanisms controlling VTE biosynthesis. The eQTL analysis elucidated two trans-eQTLs that probably contained transcriptional regulators of VTE biosynthesis. My data have provided new evidence of an R2R3 MYB TF (SlMYB79) that may regulate the genes encoding enzymes of the MEP pathway transcriptionally. The overexpression of *SlMYB79* in tomato fruit, has resulted in a 1.41- and 1.31-fold increase in total tocopherols, of which the majority was the

most bioavailable vitamer, alpha (α) tocopherol. My overexpression lines have resulted in nutritionally enhanced tomatoes, which contain more than the recommended daily allowance (RDA) of VTE (15mg per day (NIH, 2016)). The overexpression of SlMYB79 led to increased VTE levels, and contain between 22.16-23.83mg per 100g of fresh tomatoes. These values seem much higher than the range of values reported for tomatoes by the USDA (table 1-2), howe ver these measurements are well within the range of values reported in the literature for tocopherol contents of tomato (Lira et al., 2017, Lira et al., 2016). Transient assays have provided other candidate TFs that might be direct regulators of VTE biosynthesis. In this thesis I have also shown that tocotrienols are produced in *S.pennellii* fruits, and I have identified candidate genes which may be responsible.

8.1.2 Expression Quantitative Trait Loci analysis

Quantitative trait loci (QTLs) are loci which can have pleiotropic effects on phenotypes, and they are difficult to identify. Typically, introgression lines (ILs) have been used to segregate QTLs transgressively and enable the molecular characterisation of these loci (Fernie et al., 2006, Lippman et al., 2007). However, the development of high throughput gene expression profiling allows more complex traits to be identified. The combination of gene expression and QTLs has allowed for the identification of expression QTLs (eQTLs) (Alberts et al., 2007, Cubillos et al., 2012, Hansen et al., 2008).

I have used eQTL and co-expression analyses of RNA sequencing data of tomato fruit of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs to elucidate candidate transcriptional regulators of VTE biosynthesis. eQTLs can be classified into cis- or trans- eQTLs which act locally or distally to a locus of interest, respectively. Using the *S.pennellii S.lycopersicum* cv. M82 ILs (figure 3-1), I have identified several regions which contain trans-eQTLs by using co-expression analyses of the genes encoding enzymes of the VTE pathway (figure 3-5). Two candidate regions were identified as trans-eQTLs for the VTE biosynthetic genes; trans-eQTL IL6-2-2 and trans-eQTL IL9-3-2 (figure 3-5).

The co-expression analyses of the trans-eQTL IL9-3-2 (figure 3-6) showed that genes encoding enzymes of the VTE pathway were down-regulated by an allele from *S.pennellii* in this region (figure 3-5). Co-expression analyses of the expression profile of genes encoding enzymes of the methyl erythritol pathway (MEP) were downregulated in this trans-eQTL (figure 3-7). The MEP pathway provides the phytyl diphosphate (PDP) for VTE biosynthesis (figure 3-2). This suggested that this region could contain a candidate TF that regulates the VTE and MEP pathway transcriptionally. Additionally, the corresponding ILs in the *S.lycopersicoides* IL population showed that the ILs contain the DNA corresponding to the *S.pennellii* region in IL9-3-2 showed reduced expression of genes encoding the enzymes of the VTE pathway (table 3-2 and figure 3-10). This suggested that this region might contain transcriptional regulators that can regulate VTE biosynthesis in more than one species.

The trans-eQTL IL6-2-2 (figure 3-13) showed that the genes encoding enzymes of the VTE pathway were upregulated compared to M82 (figure 3-5). This trans-eQTL also appeared in the RNA sequencing data of leaves of the IL (figure 3-9), which suggested that this region may contain transcriptional regulators that regulate VTE biosynthesis in other organs. The trans-eQTL identified in the fruit RNA sequencing data also showed upregulation of genes encoding enzymes of the shikimate (SK) pathway (figure 3-8), suggesting that the upregulation of the VTE pathway could be a downstream effect of the increase in expression of genes encoding enzymes of the SK pathway. Therefore, this region may contain candidate TFs that regulate both the SK and/or the VTE pathway.

Candidate transcription factors were identified from the two trans-eQTLs (figure 3-12 and figure 3-15). Expression of the candidate TFs was correlated to the co-expression profile of the genes encoding enzymes of the VTE pathway. These genes were transiently silenced using viral induced gene silencing (VIGS) in chapters 5 and 6, to determine their function. Twenty-six candidate TFs were identified in the trans-eQTL IL6-2-2 (figure 3-15) and eleven candidate TFs were identified in the trans-eQTL IL9-3-2 (figure 3-12). The trans-eQTL screen and co-

expression analyses provide powerful tools for identifying TFs controlling subtle traits. Therefore, the analyses used in this thesis have enabled the identification of several candidate TFs that may regulate VTE biosynthesis transcriptionally. This approach can easily be applied to many other biosynthetic pathways.

8.1.3 Transcriptional regulation of VTE biosynthesis in tomato fruits

The understanding of transcriptional regulation of VTE biosynthesis is limited. There are some studies which have suggested that transcriptional regulation exists during tomato development and ripening (Quadrana et al., 2013). However, these studies failed to identify transcriptional regulators. I analysed tomato fruits of *S.lycopersicum* cv. M82 and *S.pennellii* during tomato development and ripening, to understand the transcriptional environment for VTE biosynthesis.

In *S.lycopersicum* cv. M82 tomato fruits, VTE levels were constant in the pericarp tissue and the gene expression profiles of genes encoding enzymes of the VTE pathway were constant throughout tomato development and ripening (figure 4-3 and figure 4-6). However, in the epidermis, there was an increase in VTE levels at the breaker + 3 days (B+3) stage (figure 4-6). The expression profiles of the genes encoding enzymes of VTE biosynthesis of this tissue suggested that transcriptional control is probably different from the pericarp as they showed more highly expressed profiles than the pericarp (figure 4-3). Changes in expression were subtle, suggesting that more nuanced analysis was required for identification of transcriptional regulators, such as trans-eQTL analysis. In *S.lycopersicum* cv. M82 fruit, the expression profile of genes encoding enzymes of the MEP and SK pathways was more variable during fruit development than the genes encoding enzymes of VTE biosynthesis (figure 4-4 and figure 4-5), which suggested that these pathways might be important regulatory points altering substrate flux into the VTE pathway.

The expression levels of genes encoding enzymes of the VTE, MEP and SK pathways in *S.pennellii* fruit generally were much higher than in *S.lycopersicum* cv.

M82 fruit (figure 4-7, figure 4-8 and figure 4-9). This is consistent with the notion that many genes encoding enzymes that are involved in primary metabolism are also highly expressed in *S.pennellii* (Steinhauser et al., 2010). This was reflected in the metabolite data, which showed that *S.pennellii* tomato fruit accumulated much higher levels of tocochromanols than *S.lycopersicum* cv. M82 fruit (figure 4-10). Additionally, *S.pennellii* fruit contained both tocopherols and tocotrienols (figure 4-10).

The expression profiles of genes encoding enzymes of VTE biosynthesis, suggested that these genes were likely to be co-regulated as there were no substantial changes in expression during tomato development and ripening for either species (figure 4-3 and figure 4-7). Therefore, the higher VTE levels in *S.pennellii* were probably due to the variation in expression of the genes encoding enzymes of the MEP and SK pathways (figure 4-4, figure 4-5, figure 4-8 and figure 4-9). The differential expression profiles of the MEP and SK pathways of *S.lycopersicum* cv. M82 and *S.pennellii* tomato fruits showed that production of VTE might be controlled by different transcriptional regulators in the two-tomato species. Therefore, eQTL analysis offered a powerful tool to identify transcriptional regulators that affect VTE biosynthesis, because the changes in gene expression are subtle.

8.1.4 VTE biosynthesis might be regulated by several TFs

Two trans-eQTL regions (trans-eQTL IL9-3-2 and trans-eQTL IL6-2-2) were mined for candidate genes encoding TFs that were co-expressed with the expression profiles of genes encoding enzymes of the VTE pathway. Previously, Quadrana et al (2013) showed that there were several common TF binding motifs in the promoters of the VTE, MEP and SK pathways. The most common motifs found were those bound by R2R3 MYB and bZIP TFs. Therefore, I selected candidate TFs based on this promoter analysis.

8.1.4.1 The trans-eQTL IL9-3-2 revealed a putative transcriptional repressor that regulates genes encoding enzymes of the MEP pathway

The trans-eQTL screen suggested that the trans-eQTL IL9-3-2 might contain a transcriptional regulator of VTE biosynthesis (figure 5-1). I identified eleven TFs that co-expressed with the expression profiles of genes encoding enzymes of the VTE pathway. I transiently silenced four of these candidate genes and found SlMYB79 acted as a transcriptional repressor of VTE biosynthesis (table 5-3, figure 5-3 and figure 5-4).

SlMYB79 was not expressed in the trans-eQTL IL9-3-2 or in *S.pennellii* fruit, but it was expressed in *S.lycopersicum* cv. M82 fruit (table 5-1 and table 5-4). Therefore, if SlMYB79 is a transcriptional repressor of VTE biosynthesis, this may explain why *S.pennellii* fruit contain more tocopherols (figure 4-10) than *S.lycopersicum* cv. M82 (figure 4-6). However, its expression cannot explain why VTE biosynthetic genes were down regulated in the trans-eQTL IL9-3-2, unless this affects higher pathways than the MEP pathway due to increasing expression of *DXS* or *GGPS*. Additionally, protein alignments of SlMYB79 and SpMYB79 suggest that the MYB DNA binding domain and C-terminal activation domains of the two TFs are very similar, suggesting that they may have the same function in both IL parents (figure 5-29). Promoter alignments of SlMYB79 and SpMYB79 showed an insertion of 10 base pairs in the SpMYB79 promoter which includes a GARP TF binding domain near to a GATA TF binding motif, which was not present in the promoter of SlMYB79 (figure 5-31). Therefore, this gene may be differentially regulated between the two parents. This insertion may result in the absence of *SpMYB79* expression in fruit, which enables *S.pennellii* fruit to accumulate more tocochromanols, as shown in the model (figure 4-10 and figure 5-33).

Two homozygous knock-out lines of *SlMYB79* were generated using CRISPR/Cas9 genome editing. The expression of genes encoding enzymes of the MEP pathway were significantly upregulated in these lines (figure 5-11 and figure 5-12), as well as some genes encoding enzymes of the SK pathway (figure 5-4). These lines showed changes in tocopherol composition (figure 5-7). It is not clear

from these data whether SlMYB79 binds the repressed genes directly, or whether affects their levels of expression through an indirect mechanism. The changes in tocopherol levels were inconsistent with the expression profiles of genes of MEP, SK and VTE pathways in both the overexpression and CRISPR lines.

Overexpression of *SlMYB79*, under the control of an E8 promoter, in Microtom fruit showed that several genes encoding enzymes of the MEP pathway were significantly repressed (figure 5-19 and figure 5-20). However, the overexpression of *SlMYB79* resulted in higher tocopherol levels, and in particular α-tocopherol was increased (figure 5-16). Additionally, when the paralog of SlMYB79, *SlMYB71,* was overexpressed in Moneymaker tomatoes, total tocopherols were increased, but the expression profiles of genes encoding enzymes of VTE biosynthesis, MEP and SK pathways were increased (figure 5-26 and figure 5-27).Therefore, it is unclear how SlMYB79 affects VTE levels, but it does seem to act as a transcriptional repressor of the MEP pathway.

8.1.4.2 Transient screening of candidate TFs of the trans-eQTL IL6-2-2 revealed two negative regulators of VTE biosynthesis

Gene mining of the trans-eQTL IL6-2-2 provided twenty-six candidate TFs that might regulate VTE biosynthesis transcriptionally (figure 6-1). I transiently silenced seven candidate TFs. Three of these candidates were taken forward for further analysis, and they altered tocopherol contents in tomato when silenced (table 6-4). The viral induced gene silencing and transient overexpression of three candidate TFs *(SlTF2, SlTF4 and SlTF7*) showed that tocopherol composition was significantly altered (figures 6-3, figure 6-4 and figure 6-6). Both SlTF4 and SlTF7 acted like transcriptional repressors of VTE biosynthesis in the VIGS and transient overexpression assays (figure 6-6 and figure 6-8). SlTF4 and SlTF7 were MYB 1R type TFs, the activity of which may be modulated by light. The homolog of SlTF7 is AtMYBD, which is a negative regulator of a transcriptional repressor of anthocyanin biosynthesis (Nguyen et al., 2015). AtMYBD is regulated by HY5, which is a key light responsive transcriptional regulator (Nguyen et al., 2015,

Hardtke et al., 2000). This candidate TF may provide a link between light regulation and VTE biosynthesis. HY5 is known to modulate bHLH TFs known as phytochrome interacting factors (PIFs) and regulate carotenoid and chlorophyll levels in tomato fruit (Toledo-Ortiz et al., 2014, Toledo-Ortiz et al., 2010). Therefore, this TF could be part of a larger regulatory network that controls VTE biosynthesis.

8.1.5 Tocotrienols

The metabolic and expression analysis of *S.pennellii* fruit (chapter 4) showed that these fruits contained tocotrienols (figure 4-10). Tocotrienols are not normally produced in tomatoes, as they like all dicotyledonous plants lack the gene that encodes for homogentisate geranylgeranyl transferase (HGGT). This gene encodes for an enzyme that catalyses the reaction between homogentisate (supplied from the SK pathway) and geranylgeranyl diphosphate (GGDP) (the product of the MEP pathway), to synthesise tocotrienols. Tocotrienols have been suggested to have a higher antioxidant capacity than tocopherols (Khanna et al., 2005, Schneider, 2005, Serbinova et al., 1991). *S.pennellii* remains green and does not accumulate carotenoids, therefore it is photosynthetically active, unlike like *S.lycopersicum.* Tocochromanols are important antioxidants that scavenge singlet oxygen produced during photosynthesis (Kruk et al., 2005), and prevent lipid peroxidation of poly unsaturated fatty acids (PUFAs) (Sattler et al., 2004). Therefore, tocotrienols may be produced in photosynthetically active *S.pennellii* fruit to reduce oxidative stress associated with photosynthesis (Havaux et al., 2005, Kruk et al., 2005, Sattler et al., 2004, Sattler et al., 2006).

Analysis of the fruit of the *S.pennellii* x *S.lycopersicum* cv. M82 ILs showed that one IL contained tocotrienols (IL 6-4) (figure 7-2). Gene mining of this region suggested that the gene responsible was not likely to be an enzyme, because none of the enzymes identified in this region were able to catalyse the reaction to synthesise tocotrienols (table 7-4). Therefore, it is likely that the gene responsible alters flux of pathway products that alter availability of substrates for VTE biosynthesis. The identification of a gene affecting tocotrienol biosynthesis would

be beneficial. Tocotrienols are more potent antioxidants than tocopherols (Munné-Bosch and Alegre, 2002, Serbinova et al., 1991) and tocotrienols may serve to prevent oxidation of other phytonutrients, such as beta-carotene (Che et al., 2016).

8.1.6 Future outlook

Much more research is required to address transcriptional regulation of VTE biosynthesis adequately. I aimed to identify the TFs that regulate this pathway during this PhD project. Using the trans-eQTL analyses I proved that this is a useful tool for the identification of transcriptional regulators and I identified three TFs which may regulate VTE biosynthesis. However, it is likely that the TFs identified alter substrate availability for VTE biosynthesis, to result in enhanced VTE levels, rather than regulating the expression of VTE biosynthetic genes directly.

SlMYB79 is a candidate transcriptional regulator of the MEP pathway and further tests are required to determine whether this binds directly or indirectly to the gene targets. Additionally, the effects of the different genetic backgrounds need to be addressed in the *SlMYB79* overexpression and CRISPR knockout lines. It would be interesting to understand the mechanisms underlying the differential expression of *SpMYB79* in *S.pennellii.* Two other MYB 1R type TFs were also identified as repressors of the VTE pathway (SlTF4 and SlTF7). These genes need to be overexpressed in stable transformations and knock-out lines using CRISPR need to be produced to assess the function of these genes at the molecular level.

Tocotrienols were produced in *S.pennellii* fruits and the candidate gene controlling this trait needs to be identified. Overall, a combination of these methods has furthered the understanding of the transcriptional regulation of VTE biosynthesis and has resulted in the production of nutritionally enhanced tomatoes, but there remains a long way to go for full understanding of the control of VTE production in tomatoes.

Appendix

Appendix

List of plasmids

Table 1 List of plasmids used in this thesis. The table shows the plasmid name, resistance, and description of its use.

Plasmid name	Resistance marker	Description
pDONR207 Gateway®	Gentamycin, Chloramphenicol	Used as an entry plasmid and contains a Gateway® cassette
pDONR207 Gateway® SlMYB79 VIGS	Gentamycin	Entry plasmid containing SlMYB79 VIGS fragment, Gateway® compatible
pDONR207 Gateway® SlbHLH60 VIGS	Gentamycin	Entry plasmid containing SlbHLH60 VIGS fragment, Gateway® compatible
pDONR207 Gateway® SlbHLH61 VIGS	Gentamycin	Entry plasmid containing SlbHLH61 VIGS fragment, Gateway® compatible
pDONR207 Gateway® SlbHLH92 VIGS	Gentamycin	Entry plasmid containing SlbHLH92 VIGS fragment, Gateway® compatible

Table 1 continued.

pDONR207 Gateway® SlTF2 VIGS	Gentamycin	Entry plasmid containing SlTF2 VIGS fragment, Gateway® compatible
pDONR207 Gateway® SlTF4 VIGS	Gentamycin	Entry plasmid containing SlTF4 VIGS fragment, Gateway® compatible
pDONR207 Gateway® SlTF7 VIGS	Gentamycin	Entry plasmid containing SlTF7 VIGS fragment, Gateway® compatible
pDONR207 Gateway® SlTF10 VIGS	Gentamycin	Entry plasmid containing SlTF10 VIGS fragment, Gateway® compatible
pDONR207 Gateway® SlTF11 VIGS	Gentamycin	Entry plasmid containing SlTF11 VIGS fragment, Gateway® compatible
pDONR207 Gateway® SlTF12 VIGS	Gentamycin	Entry plasmid containing SlTF12 VIGS fragment, Gateway® compatible
pDONR207 Gateway® SlMYB79 CDS	Gentamycin	Entry plasmid containing SlMYB79 coding sequence, Gateway® compatible
pDONR207 Gateway® SlMYB71 CDS	Gentamycin	Entry plasmid containing SlMYB71 coding sequence, Gateway® compatible

Table 1 continued.

pDONR207 Gateway® SlTF2 CDS	Gentamycin	Entry plasmid containing SlTF2 coding sequence, Gateway® compatible
pDONR207 Gateway® SlTF4 CDS	Gentamycin	Entry plasmid containing SlTF4 coding sequence, Gateway® compatible
pDONR207 Gateway® SlTF7 CDS	Gentamycin	Entry plasmid containing SlTF7 coding sequence, Gateway® compatible
pTRV2 *Del/Ros*	Kanamycin	Used as a control for viral induced gene silencing
pTRV2 *Del/Ros* Gateway®	Kanamycin, Chloramphenicol	Used as a destination plasmid for viral induced gene silencing, Gateway ® compatible
pTRV1	Kanamycin	Used for viral induced gene silencing
pTRV2 *Del/Ros* Gateway® SlMYB79 VIGS	Kanamycin	Used for viral induced gene silencing of SlMYB79
pTRV2 *Del/Ros* Gateway® SlbHLH60	Kanamycin	Used for viral induced gene silencing of SlbHLH60
pTRV2 *Del/Ros* Gateway® SlbHLH61	Kanamycin	Used for viral induced gene silencing of SlbHLH61
pTRV2 *Del/Ros* Gateway® SlbHLH92	Kanamycin	Used for viral induced gene silencing of SlbHLH92

Table 1 continued.

pTRV2 *Del/Ros* Gateway® SlTF2	Kanamycin	Used for viral induced gene silencing of SlTF2
pTRV2 *Del/Ros* Gateway® SlTF4	Kanamycin	Used for viral induced gene silencing of SlTF4
pTRV2 *Del/Ros* Gateway® SlTF7	Kanamycin	Used for viral induced gene silencing of SlTF7
pTRV2 *Del/Ros* Gateway® SlTF10	Kanamycin	Used for viral induced gene silencing of SlTF10
pTRV2 *Del/Ros* Gateway® SlTF11	Kanamycin	Used for viral induced gene silencing of SlTF11
pTRV2 *Del/Ros* Gateway® SlTF12	Kanamycin	Used for viral induced gene silencing of SlTF12
pBIN19 E8 Gateway®	Kanamycin, Chloramphenicol	Used as a destination plasmid for stable overexpression, under the control of the E8 promoter, Gateway® compatible
pBIN19 E8 Gateway® SlMYB79	Kanamycin	Used as a destination plasmid for stable overexpression of SlMYB79, Gateway® compatible
pBIN19 E8 Gateway® SlMYB71	Kanamycin	Used as a destination plasmid for stable overexpression of SlMYB71, Gateway® compatible

Table 1 continued.

pTRV2 overexpression	Kanamycin	Used as an expression vector to overexpress candidate genes, under the control of the pea early browning virus (PEBV) promoter
pTRV1 overexpression	Kanamycin	Used for the overexpression vector (pTRV2) to overexpress candidate genes
pTRV2 overexpression SlTF2 CDS	Kanamycin	Overexpression pTRV2 vector containing the coding sequence of SlTF2
pTRV2 overexpression SlTF7 CDS	Kanamycin	Overexpression pTRV2 vector containing the coding sequence of SlTF7
pTRV2 overexpression SlTF4 CDS	Kanamycin	Overexpression pTRV2 vector containing the coding sequence of SlTF4
pICH47732 gRNA1	Ampicillin	Level 1 Golden gate plasmid containing gRNA1 for SlMYB79, under the control of the U6 promoter
pICH47742 gRNA2	Ampicillin	Level 1 Golden gate plasmid containing gRNA2 for SlMYB79, under the control of the U6 promoter
pICSL002203 gRNA 1/2	Kanamycin	Level 2 Golden gate plasmid containing gRNA1/2 for SlMYB79
pICH41744 (linker)	Streptomycin	Linker plasmid for Golden gate construction

Table 2 List of primers used in thesis. Primer name, corresponding Solgenomics identifier, use, primer sequence, amplicon length and origin are shown

Gene name	Solgenomics identifier	Primer name	Use of primer	F primer/R primer sequence	Amplicon length (bp)	Origin (EF designed or specific paper)
SlMYB79	Solyc09g090790	SlMYB79_GW_F1	CDS Gateway® primers	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGGGAGTAACTGGGGTTTTA	817	EF
		SlMYB79_GW_F1		GGGGACCACTTTGTACAAGAAAGCTGGGTATTAATAGAAAGTTGGAGT		
		SlMYB79_F1	CDS primers	ATGGGGAGTAACTGGGGTTTTA	756	EF
		SlMYB79_R1		GCGTTAATAGAAAGTTGGAGT		
		SlMYB79_VIGS_GW_F1	VIGS Gateway® primers	GGGGACAAGTTTGTACAAAAAAGCAGGCTTA TCACGAGCACGTCTTTGGAAAA	295	EF
		SlMYB79_VIGS_GW_R1		GGGGACCACTTTGTACAAGAAAGCTGGGTATGAAGAGGGTTCGGACTG		
		SlMYB79_VIGS_F1	VIGS primers	TCACGAGCACGTCTTTGGAAAA	234	EF
		SlMYB79_VIGS_R1		TGAAGAGGGTTCGGACTG		
		SlMYB79_qRT-PCR_F20	qRT-PCR primers	TTTGAGGCCAGATTTGAAGAGG	89	EF

Table 2 continued.

SlMYB79	Solyc09g090790	SlMYB79_qRT-PCR_R20	qRT-PCR primers	ACCATCTATTGCCCCATCTAG		
		MYB79_R	Sequencing reverse primer near to ATG	GCTTTCCTTGAGGCATAGTTGG	N/A	EF
		SlMYB79_UTR_F	Sequencing forward primer in the UTR	CACTTCCCCTTTTACACC	N/A	EF
		SlMYB79_N1	Nested forward sequencing primer for end of SlMYB79	GAGAAATCACGAGCACGTC	N/A	EF
		SlMYB79_N2	Nested reverse sequencing primer for beginning of SlMYB79	CCCTCTTCAAATCTGGCCT	N/A	EF

Table 2 continued.

SlMYB71	Solyc05g053150	SlMYB71_GW_F1	CDS Gateway® primers	GGGGACAAGTTTGTACAAAAAAGC AGGCTTAATGAGAAGGTGGGGAATT	122	EF
		SlMYB71_GW_F1		GGGGACCACTTTGTACAAGAAAGC TGGGTATTAAAGATTGACATTATA		
		SlMYB71_F1	CDS primers	ATGAGAAGGTGGGGAATT	1165	EF
		SlMYB71_R1		GGGTATTAAAGATTGACATTATA		
		SlMYB71_qRT-PCR_F1	qRT-PCR primers	AATTGTGTTGCTAGGCTTGCTG	80	EF
		SlMYB71_qRT-PCR_R1		CGATGATAGGTTTGTTCAAGATGGA		
TF2	Solyc06g060490	Solyc06g060490 – TF2_NotI_F	CDS restriction enzyme primers	TATGCGGCCGCATGGACGCTAAGTTC GCCGGAAAG	723	EF
		Solyc06g060490 – TF2_ApaI_R		CCCGCAGGGCCCTCAGTTCTTTGTTTT CAGCATTTAGTCCAG		
		622_11	CDS primers	ATGGACGCTAAGTTCGCC	697	EF
		622_12		TCAGTTCTTTGTTTTCAGC		
		622_13	VIGS Gateway® primers	GGGGACAAGTTTGTACAAAAAAGCAG GCTTAATGGACGCTAAGTTCGCC	342	EF
		622_14		GGGGACCACTTTGTACAAGAAAGCTG GGTAGCCTGAACATCGAGACTG		
		622_15	VIGS primers	622_11	200	EF
		622_16		GCCTGAACATCGAGACTG		
		622_TF2_3	qRT-PCR primers	GACAATCTGCTGCACGTTCC	104	EF
		622_TF2_4		CTGAGCGGATAGAGTGGTGG		

Table 2 continued.

TF4	Solyc06g066180	Solyc06g066180 - TF4_NotI_F	CDS restriction enzyme primers	TGCGCGGCGGCCGCATG TATAATGCAATTGATAGGATG	602	EF
		Solyc06g066180 - TF4_ApaI_R		CACTAAGGGCCCTAACT AGACGTTGGTAGCTCTATG		
		622_27	CDS primers	ATGTATAATGCAATTGATAG	576	EF
		622_28		GCGTAACTAGACGTTGGTAGC		
		622_29	VIGS Gateway® primers	GGGGACAAGTTTGTACAA AAAAGCAGGCTTAGAAGT ATAGACTTGGACA	361	EF
		622_30		GGGGACCACTTTGTACAAG AAAGCTGGGTATTAAATTTG ACAGAGGTG		
		622_31	VIGS primers	GAAGTATAGACTTGGACAGC	300	EF
		622_32		TTAAATTTGACAGAGGTGGC		
		622_TF4_23	qRT-PCR primers	ATGGGCTTGAAGGGTTTGACA	99	EF
		622_TF4_24		GTTTGCCTCGGCTGCATTTT		
TF7	Solyc06g071230	Solyc06g071230 – TF7_NotI_F	CDS restriction enzyme primers	TATATGCGGCCGCATGTCG AGCGTTTGCAGTGATAAG	935	EF
		Solyc06g071230 – TF7_ApaI_R		TTATAAGGGCCCTCATGCC ACACGGATGATGCTTTC		
		622_59	CDS primers	ATGTCGAGCGTTTGCAGT	909	EF
		622_60		TCATGCCACACGGATGAT		

Table 2 continued.

TF7	Solyc06g071230	622_61	VIGS Gateway® primers	GGGGACAAGTTTGTACAAAAAAGCAG GCTTAACCATCACACATGTCTCA	361	EF
		622_62		GGGGACCACTTTGTACAAGAAAGCTG GGTATCATAGGATCCACTTTTA		
		622_63	VIGS primers	ACCATCACACATGTCTCAGC	300	EF
		622_64		TCATAGGATCCACTTTTAGC		
		622_225	qRT-PCR primers	ACCAATGCATTTTCGGTGGC	116	EF
		622_226		AGCATCGATGAATTCGCATGA		
TF9	Solyc06g069710	622_205	VIGS Gateway® primers	GGGGACAAGTTTGTACAAAAAAGCA GGCTTACAGAGAGATTTTCAGGGG	361	EF
		622_206		GGGGACCACTTTGTACAAGAAAGCTG GGTAGTAGCAGTAGCAGAATCT		
		622_207	VIGS primers	CAGAGAGATTTTCAGGGG	300	EF
		622_208		GTAGCAGTAGCAGAATCTGC		
		622_TF9_3	qRT-PCR primers	TCCAGCTTCATTTGGGGTTCA	113	EF
		622_TF9_4		CCCCTGTTTCTTGGGGTACA		
TF10	Solyc06g069850	622_213	VIGS Gateway® primers	GGGGACAAGTTTGTACAAAAAAGCAG GCTTAGCTGGGATTTATCAAGTT	361	EF
		622_214		GGGGACCACTTTGTACAAGAAAGCTG GGTATTCATGTTCGGTCAAATC		
		622_215	VIGS primers	GCTGGGATTTATCAAGTT	300	EF
		622_216		TTCATGTTCGGTCAAATC		
		622_TF10_9	qRT-PCR primers	GCTTCAAGTGCTGAAAACATTGC	112	EF
		622_TF10_10		TTGTTGTTCAAGGATGACATAGG		

Table 2 continued.

TF11	Solyc06g060230	622_221	VIGS Gateway® primers	GGGGACAAGTTTGTACAAAAAAGC AGGCTTAAGCTGATGTTGATCGATC	361	EF
		622_222		GGGGACCACTTTGTACAAGAAAGC TGGGTAGTGAATTCGGGTGAGAGC		
		622_223	VIGS primers	AGCTGATGTTGATCGATC	300	EF
		622_224		GTAGTGAATTCGGGTGAGAGC		
		622_TF11_7	qRT-PCR primers	GTGGGGCCACCAATGGATTA	88	EF
		622_TF11_8		TATTCTACTACGGCGGCACG		
TF12	Solyc06g061080	622_125	VIGS Gateway® primers	GGGGACAAGTTTGTACAAAAAAGC AGGCTTATCTCTGTGATGATGTTTC	361	EF
		622_126		GGGGACCACTTTGTACAAGAAAGC TGGGTAGAGTCATAGTCCTCTATT		
		622_127	VIGS primers	TCTCTGTGATGATGTTTC	300	EF
		622_128		AGTCATAGTCCTCTATT		
		622_TF12_5	qRT-PCR primers	TCTGCAAACCTCTGGATGCC	118	EF
		622_TF12_6		GCTGCCATGAAGTTGAGTTTGA		

Table 2 continued.

DXR	Solyc03g114340	DXR_PENNELLII_F1	qRT-PCR primer for *S.pennellii*	ACAGCTAGGGTGGCCTGATA	101	EF
		DXR_PENNELLII_R1	qRT-PCR primer for *S.pennellii*	AAAGATCAAGCCGAGGCCAA		
		DXR_F1	qRT-PCR primer for *S.lycopersicum*	GCTCACTACCTTTTCGGAGCTG	198	EF
		DXR_R1	qRT-PCR primer for *S.lycopersicum*	AAGCCGAGGCCAAGTAATCTC		
CMS	Solyc01g102820	CMS_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CCCAAGAATGTATTGCCTTC	207	(Quadrana et al., 2013)
		CMS_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CCCCTTTCCTCCAGCAAG		
ISPE	Solyc01g009010	ISPE_PENNELLII_F1	qRT-PCR primer for *S.pennellii*	GAACCTCCCGCCTTTGAAGT	80	EF
		ISPE_PENNELLII_R1	qRT-PCR primer for *S.pennellii*	TCATATTGTCCTCGGCCAGC		
		ISPE_F1	qRT-PCR primer for *S.lycopersicum*	GTAATGCTGCAACAACTC	208	(Quadrana et al., 2013)
		ISPE_R1	qRT-PCR primer for *S.lycopersicum*	GGCTTTATGAGGACCATTGG		

Table 2 continued.

ISPF	Solyc08g081570	ISPE_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CCACTCCGTTGAAGTCTC	195	(Quadrana et al., 2013)
		ISPE_R1		CAATATCAGGAAGCCCCAGAGC		
HDS	Solyc11g069380	HDS_PENNELLII_F1	qRT-PCR primer for *S.pennellii*	GAGGGTTAAGTTCCGGCGAA	88	EF
		HDS_PENNELLII_R1	qRT-PCR primer for *S.pennellii*	TTCGGATGCAGGCTGAAGTT		
		HDS_F1	qRT-PCR primer for *S.lycopersicum*	GAAGTATGGACGTGCAATGC	198	(Quadrana et al., 2013)
		HDS_R1	qRT-PCR primer for *S.lycopersicum*	CGATACGCCTGAACCATAAC		
IPI(1)	Solyc04g056390	IPI(1)_PENNELLII_F1	qRT-PCR primer for *S.pennellii*	CTTGCTGCAGCCATCCACTA	90	EF
		IPI(1)_PENNELLII_R1	qRT-PCR primer for *S.pennellii*	CGAGAAGCTTCCTTTGTGCAG		
		IPI(1)_F1	qRT-PCR primer for *S.lycopersicum*	CTTCTCATTCTTCACTCCG	207	(Quadrana et al., 2013)
		IPI(1)_R1	qRT-PCR primer for *S.lycopersicum*	CATGGTCATTCTCATCCACC		

Table 2 continued.

IPI(2)	Solyc05g055760	IPI(2)_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CTGATGGAAAGTGGGGAG	240	(Quadrana et al., 2013)
		IPI(2)_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GGATGGTCCCCTTCTCAAC		
GGDR	Solyc03g115980	GGDR_PENNELLII_F1	qRT-PCR primer for *S.pennellii*	CACCGGCGGAAAAATCATCC	102	EF
		GGDR_PENNELLII_R1	qRT-PCR primer for *S.pennellii*	CCTGCTGCATCCCCAACTAA		
		GGDR_F1	qRT-PCR primer for *S.lycopersicum*	CAGAGACGCTCGCTAAGG	201	(Quadrana et al., 2013)
		GGDR_R1	qRT-PCR primer for *S.lycopersicum*	GCTTCAGAGTCTGTCCGATATC		
GGPS(1)	Solyc11g011240	GGPS(1)_PENNELLII_F2	qRT-PCR primer for *S.pennellii*	TAGGTGCATTGGGCTGTTGT	92	EF
		GGPS(1)_PENNELLII_R2	qRT-PCR primer for *S.pennellii*	GGTCCTTACCCGCTGTCTTT		
		GGPS(1)_F1	qRT-PCR primer for *S.lycopersicum*	CACCAAGCCAATCTTACAG	196	(Quadrana et al., 2013)
		GGPS(1)_R1	qRT-PCR primer for *S.lycopersicum*	GGTATTGCTTCATCTAGTGC		

Table 2 continued.

GGPS(2)	Solyc04g079960	GGPS(2)_PENNELLII_F4	qRT-PCR primer for *S.pennellii*	GTTGGGGGAAACCAAGGGAA	110	EF
		GGPS(2)_PENNELLII_R4	qRT-PCR primer for *S.pennellii*	CGGAGATCGTCGTCATCCAT		
GGPS(3)	Solyc02g085700	GGPS(3)_PENNELLII_F1	qRT-PCR primer for *S.pennellii*	GTATAGCTGCTTGTGAA	220	EF
		GGPS(3)_F1	qRT-PCR primer for *S.lycopersicum*	GTGTATAGCTGCTTGTGAG		(Quadrana et al., 2013)
		GGPS(3) _R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GTGCTCAAAGGCTAATGC		
GGPS(4)	Solyc09g008920	GGPS(4)_PENNELLII_F4	qRT-PCR primer for *S.pennellii*	CGATGATCTACGTCGTGGCA	79	EF
		GGPS(4)_PENNELLII_R4	qRT-PCR primer for *S.pennellii*	AAGTGCATCCCCTGCAAGAA		
		GGPS(4)_F1	qRT-PCR primer for *S.lycopersicum*	CACTGCCTGTGCCTTAGAG	202	(Quadrana et al., 2013)
		GGPS(4)_R1	qRT-PCR primer for *S.lycopersicum*	GGGAACGAGATCACTTGG		

Table 2 continued.

DAHPS1	Solyc11g009080	DAHPS1_PENNELLII_F1	qRT-PCR primer for *S.pennellii*	TGGAGACTGTGCTGAGAG	200	EF
		DAHPS1_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GTAACTTGGTAGCTTCACTC		(Quadrana et al., 2013)
		DAHPS1_F1	qRT-PCR primer for *S.lycopersicum*	GGGAGACTGTGCTGAGAG		(Quadrana et al., 2013)
DAHPS2	Solyc01g105390	DAHPS2_PENNELLII_F2	qRT-PCR primer for *S.pennellii*	TCAGGGCTGAGGTAAGAGCA	97	EF
		DAHPS2_PENNELLII_R2	qRT-PCR primer for *S.pennellii*	TCTGTGACGTTTTGGCCTGT		
		DAHPS2_F1	qRT-PCR primer for *S.lycopersicum*	GTGCTGTTCTTATGTTTGGTG	191	(Quadrana et al., 2013)
		DAHPS2_R1	qRT-PCR primer for *S.lycopersicum*	GATCAGCCTCTGAGGGTC		(Quadrana et al., 2013)
DAHPS3	Solyc04g05390	DAHPS3_PENNELLII_F3	qRT-PCR primer for *S.pennellii*	TTGCTGGTGAGGCTAGAAGC	87	EF
		DAHPS3_PENNELLII_R3	qRT-PCR primer for *S.pennellii*	CGGCACAATCTCCTCCTTGT		

Table 2 continued.

DHQS	Solyc02g083590	DHQS_PENNELLII_F2	qRT-PCR primer for *S.pennellii*	TGCAGTTGAAACTGGCGTTG	101	EF
		DHQS_PENNELLII_R2	qRT-PCR primer for *S.pennellii*	TCCAACCAAGACGGTGTGAC		
		DHQS_F1	qRT-PCR primer for *S.lycopersicum*	CACGATTGAAGGTGTTGG	196	(Quadrana et al., 2013)
		DHQS_R1	qRT-PCR primer for *S.lycopersicum*	GTAACGACAAGGACTCTC		
SDH/DQHS1	Solyc01g067750	SDH/DQHS1_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GAGGATCTTGGCAATCTAG	181	(Quadrana et al., 2013)
		SDH/DQHS1_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GACAAAGTATCCGAGACATC		
SDH/DQHS2	Solyc06g084460	SDH/DQHS2_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GTATTAGTTGGGGCAGGTG	223	(Quadrana et al., 2013)
		SDH/DQHS2_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GATCAGACTTTGGCTGCATG		
SK	Solyc04g051860	SK_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GCCTTCAGTTGTCCTCATG	229	(Quadrana et al., 2013)
		SK_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GATGCAGAGAAACATCCTG		

Table 2 continued.

EPSPS 1	Solyc01g091190	EPSPS1_PENNELLII_F2	qRT-PCR primer for *S.pennellii*	AGGTCCTCGAGAAGATGGGG	105	EF
		EPSPS1_PENNELLII_R2	qRT-PCR primer for *S.pennellii*	CAACGGCACGCAAATGTTTC		
		EPSPS1_F1	qRT-PCR primer for *S.lycopersicum*	CGTATTCTCCTTCTTGCTG	224	Quadrana et al. (2014)
		EPSPS1_R1	qRT-PCR primer for *S.lycopersicum*	GTTCCTGCATTTCCAAGG		Quadrana et al. (2014)
EPSP2	Solyc05g050980	EPSPS2_PENNELLII_F1	qRT-PCR primer for *S.pennellii*	TGCCGGTAAAGAGTCCAAGG	83	EF
		EPSPS2_PENNELLII_R1	qRT-PCR primer for *S.pennellii*	CAACTGCTGCTGTTAGTGGC		
CS1	Solyc04g049350	CS1_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CTGTGACAAGAGACAAACACG	202	Quadrana et al. (2014)
		CS1_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GCTGATTCAGGTGATGAC		Quadrana et al. (2014)
CS2	Solyc04g009620	CS2_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CCATCCAAGCTCTGTTCTTC	186	Quadrana et al. (2014)
		CS2_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CTCTGAGAGTGGGAGTCG		Quadrana et al. (2014)

Table 2 continued.

CM1	Solyc02g088460	CM1_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CAAACAACAATTACCTCCTTCTTCC	169	Quadrana et al. (2014)
		CM1_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	TGTGCATTTGGTAAGTGGTATGA		Quadrana et al. (2014)
CM2	Solyc11g017240	CM2_PENNELLII_F1	qRT-PCR primer for *S.pennellii*	CATCGATCACAGCCCAAACG	101/202, for *S.pennellii* and *S.lycopersicum*, respectively	EF, Quadrana et al. (2014) for *S.lycopersicum* primer
		CM2_F1	qRT-PCR primer for *S.lycopersicum*	GCTCAGTACTGTTTCAATGCGGA		
		CM2_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	ACACCACCACCGTGAGATTC		
PAT	Solyc04g054710	PAT_PENNELLII_F1	qRT-PCR primer for *S.pennellii*	TGACTGGTTGGAGACTTGGC	84	EF
		PAT_PENNELLII_R1	qRT-PCR primer for *S.pennellii*	CTGACGTGAACTGGCTCTGT		
		PAT_F1	qRT-PCR primer for *S.lycopersicum*	CGGAGTAGAGGTTGATGGATTTG	150	Quadrana et al. (2014)
		PAT_R1	qRT-PCR primer for *S.lycopersicum*	GTCGAAAGCGATGCTGCATA		

Table 2 continued.

TAT1	SGN-U577103	TAT1_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CTACTGTGGGACTTCCTC	216	(Quadrana et al., 2013)
		TAT1_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CGATGTTTCTAAATGCAGCAC		
TAT2	SGN-U563404	TAT2_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GCATCCTTCCTGCTAGAC	249	(Quadrana et al., 2013)
		TAT2_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CCTCCCACTCCTTCTCTG		
TYRA 1	Solyc07g007590	TYRA1_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CTCTGTCTCTTCTCCGTC	225	(Quadrana et al., 2013)
		TYRA1_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GTTCTTGAATGAGCCAAC		
HPPD 1	Solyc07g045050	HPPD1_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CCAGGGCAGGGGATATACTG	209	(Quadrana et al., 2013)
		HPPD1_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CTCCTTCCTCGTTTTTCAGC		

Table 2 continued.

HPPD 2	Solyc05g041200	HPPD2_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CCAGGCGTGTGAAGAATTG	214	(Quadrana et al., 2013)
		HPPD2_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CGATCTAAACAGCTCAGAG		
HST	Solyc03g051810	HST_PENNELLII_F1	qRT-PCR primer for *S.pennellii*	CCAGATGTGGAGGGTGATCG	88	EF
		HST_PENNELLII_R1	qRT-PCR primer for *S.pennellii*	GACCAGAGCCGAGAAATGCT		
		HST_F1	qRT-PCR primer for *S.lycopersicum*	GCTGCTAACTTGGTGCTC	230	(Quadrana et al., 2013)
		HST_R1	qRT-PCR primer for *S.lycopersicum*	GATCCTAGCACAGTCCCACG		
VTE2	Solyc07g017770	VTE2_PENNELLII_F4	qRT-PCR primer for *S.pennellii*	CGCGGCCCCATACCATAATA	117	EF
		VTE2_PENNELLII_R4	qRT-PCR primer for *S.pennellii*	CAATGGCCTCTAACACCCCA		
		VTE2_F1	qRT-PCR primer for *S.lycopersicum*	CAATTCCAGTTCCTGCTGAG	154	(Quadrana et al., 2013)
		VTE2_R1	qRT-PCR primer for *S.lycopersicum*	CCTCCAACATGCTCTTGCGTG		

Table 2 continued.

VTE3 (1)	Solyc09g065730	VTE3(1)_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CTTGACCAATCTCCTCATC	210	(Quadrana et al., 2013)
		VTE3(1)_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	GCACGCCTTTCCTCCAGG		
VTE3 (2)	Solyc03g005230	VTE3(2)_PENNELLII_F1	qRT-PCR primer for *S.pennellii*	AGCCATATGCTGGGGAATCG		EF
		VTE3(2)_PENNELLII_R1	qRT-PCR primer for *S.pennellii*	TATGCCAAGCAGGAATCGCA		
		VTE3(2)_F1	qRT-PCR primer for *S.lycopersicum*	GCTAAGGCTAGGCAGAAGGAG	185	(Quadrana et al., 2013)
		VTE3(2)_R1	qRT-PCR primer for *S.lycopersicum*	CAGGCAACCCCACCTATGG		
VTE5	Solyc09g018510	VTE5_PENNELLII_F2	qRT-PCR primer for *S.pennellii*	TTGGCAGGAGGTTCGGAAAA	96	EF
		VTE5_PENNELLII_R2	qRT-PCR primer for *S.pennellii*	TGGATGCTAAAAGACCGGCA		
		VTE5_F1	qRT-PCR primer for *S.lycopersicum*	CGTATCAGGACGGGCTCGC	246	(Quadrana et al., 2013)
		VTE5_R1	qRT-PCR primer for *S.lycopersicum*	TCACCACCACACATCATTGCTAATG		

Table 2 continued.

VTE6	Solyc07g062180	VTE6_PENNELLII_F2	qRT-PCR primer for *S.pennellii*	CCTTAGCTGGGCTTCTTGCT	89	EF
		VTE6_PENNELLII_R2	qRT-PCR primer for *S.pennellii*	TACACATATCGCAGCCCCAG		
		VTE6_F1	qRT-PCR primer for *S.lycopersicum*	TGTGCAAATTACCCCAACAAAGA	80	EF
		VTE6_R1	qRT-PCR primer for *S.lycopersicum*	AGCTTCTCACGTAAGACTCCG		
VTE1	Solyc08g068570	VTE1_PENNELLII_F2	qRT-PCR primer for *S.pennellii*	GAACTGGGGTGGTTCCTTCC	105	EF
		VTE1_PENNELLII_R2	qRT-PCR primer for *S.pennellii*	AGCCTTAATCCACCACCAGC		
		VTE1_F1	qRT-PCR primer for *S.lycopersicum*	CGAACTCCTCATAGCGGGTATC	189	(Quadrana et al., 2013)
		VTE1_R1	qRT-PCR primer for *S.lycopersicum*	CACGCCAGTAAACCGAGGC		
VTE4	Solyc08g076360	VTE4_F1	qRT-PCR primer for *S.lycopersicum*	GAGTGGAGAACACATGCCGA	80	EF
		VTE4_R1	qRT-PCR primer for *S.lycopersicum*	GGATGATTGTGCCTCCTGGT		

Table 2 continued.

SlCAC	Solyc08g006960	SlCAC_F1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CCTCCGTTGTGATGTAACTGG	173	(Exposito-Rodriguez et al., 2008)
		SlCAC_R1	qRT-PCR primer for *S.pennellii* and *S.lycopersicum*	CGGGACTTGAGTTGTGACTCTT		
PSY1	Solyc03g031860	PSY1_F1	qRT-PCR primer for *S.lycopersicum*	TGGCCCAAACGCATCATATA	92	(Li, 2018)
		PSY1_R1	qRT-PCR primer for *S.lycopersicum*	CACCATCGAGCATGTCAAATG		
CYCβ	Solyc06g074240	CYCβ_F1	qRT-PCR primer for *S.lycopersicum*	GATCCTAAATACTGGCAAGGGTTCC	103	(Li, 2018)
		CYCβ_R1	qRT-PCR primer for *S.lycopersicum*	ACCTAGTCATGTTTGAGCCATGTCC		
pDONR207	N/A	pDONR_F1	sequencing primer for pDONR207 Gateway® plasmids	TCGCGTTAACGCTAGCATGGATCTC	N/A	Gateway® manual
		pDONR_R1		GTGTCTCAAAATCTCTGATGTTAC	N/A	Gateway® manual
pTRV2	N/A	pTRV2_VIGS_F1	sequencing primer for pTRV2 VIGS vector	CTCAAGGAAGCACGATGAG	N/A	Eugenio Butelli
		pTRV2_VIGS_R1	sequencing primer for pTRV2 VIGS vector and pTRV2 overexpression vector	CGATCAATCAAGATCAGTCGAG	N/A	
		pTRV2_OE_F1	sequencing primer for pTRV2 overexpression vector	CGCAAGAACAGGCTGAGCATC	N/A	Vera Thole

Table 2 continued.

pBINE8:GW	N/A	E8_F1	sequencing primer for E8 promoter into pBINE8:GW Gateway® site	GTGAAACAACAACGTTTTGG	N/A	Eugenio Butelli
		E8_R1	sequencing primer for CaMV terminator into pBINE8:GW Gateway® site	ATGCTTCCGGCTCGTATGTT	N/A	
Cas9	N/A	Cas9_F1	Cas9 primer for genotyping	TATCGGCACAAACAGCGTC	N/A	Fabio D'Orso
		Cas9_R1	Cas9 primer for genotyping	CGAAAAGCTGATTGTAAGTCTG	N/A	
Left border primer	N/A	LB_F1	seq primer for Left border of Level 1 Golden gate	CTTAATAACACATTGCGGACG	N/A	Fabio D'Orso
Right border primer	N/A	RB_F1	seq primer for Right border primer of Level 2 Golden Gate	CCCGCCAATATATCCTGTC	N/A	
gRNA1	Solyc09g090790	Slmyb79_sgRNA1	CRISPR primer for cloning of gRNA1, with *Arabidopsis* U6 promoter	TGTGGTCTCAATTGTGAACAAATGGGGAGTAACGTTTTAGAGCTAGAAATAGCAAG	N/A	EF
gRNA2	Solyc09g090790	Slmyb79_sgRNA2	CRISPR primer for cloning of gRNA2, with *Arabidopsis* U6 promoter	TGTGGTCTCAATTGGAGTTATCAGACGAGTTCTGTTTTAGAGCTAGAAATAGCAAG	N/A	EF

Table 2 continued.

SlbHLH60	Solyc09g083220	SlbHLH60_F1/R1	qRT-PCR primers	GTCCAGCCTTTTCAAGGCCCA	105	EF
		SlbHLH60_F1/R1		ATCTGTGGCTTGTCCCCTCC		EF
		SlbHLH60_VIGS_GW_F1	VIGS Gateway® primers	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAGAGGGGGACAGTGGTGAA	262	EF
		SlbHLH60_VIGS_GW_R1		GGGGACCACTTTGTACAAGAAAGCTGGGTAGGACTGGTCATCCGGTTG		EF
SlbHLH61	Solyc09g097870	SlbHLH61_F1/R1	qRT-PCR primers	AACCCAGGATTTGAGTGCTGT	144	EF
		SlbHLH61_F1/R1		CTCTCACGGCAAGCTTTGTT		EF
		SlbHLH61_VIGS_GW_F1	VIGS Gateway® primers	GGGGACAAGTTTGTACAAAAAAGCAGGCTTATTGCACTGAGGGCTGACA	275	EF
		SlbHLH61_VIGS_GW_R1		GGGGACCACTTTGTACAAGAAAGCTGGGTACAAGCAGCAGGAGGTCTG		EF
SlbHLH92	Solyc09g098110	SlbHLH92_F1/R1	qRT-PCR primers	TACTGTTTCAATGGGGAGAGG	102	EF
		SlbHLH92_F1/R1		TCTTGCGTTGTGAAGCTGAG		EF
		SlbHLH92_VIGS_GW_F1	VIGS Gateway® primers	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAGTTTCGCGATGGGTTCTT	296	EF
		SlbHLH92_VIGS_GW_R1		GGGGACCACTTTGTACAAGAAAGCTGGGTATGAGTAGCACAACCACGC		EF

Media recipes

Luria-Bertani (LB) Broth

	(1L)
Tryptone	10g
Yeast extract	5g
NaCl	10g

pH7.0
(For LB agar, agar was added at a final concentration of 1.5% (w/v))

Super Optimal broth with Catabolite repression (SOC)

	(1L)
Tryptone	20g
Yeast extract	5g
NaCl	0.5g
KCl	0.186g

pH7.0
The medium was autoclaved before addition of 20mM glucose and 2mM $MgCl_2$ (final concentrations)

Murashige and Skoog (MS)
1/2 MS salts

	(1L)
Myo-inositol	50mg
Thiamine	0.5mg
Pyroxidine	0.25mg
Nicotinic acid	0.25mg
MES	0.25g

KOH added to pH 5.7
(Agar was added for MS agar to a final concentration of 0.8% (w/v))

Tryptone yeast (TY)

	(1L)
Peptone from casein (Tryptone)	5g
Yeast Extract	3g
Calcium chloride hexahydrate	1.32g

TBD

	(500ml)
100mM $RbCl_2$	6g
45mM $MnCl_2.4H_2O$	4.45g
10mM $CaCl_2.2H_2O$	0.56g
35mM KAcetate	1.71g
15% Glycerol	75ml
ddH_2O	up to 500ml

Freezing buffer

	(200ml)
10mM $RbCl_2$	0.25g
75mM $CaCl_2.2H_2O$	1.67g
MOPS	0.4g
15% Glycerol	30ml
ddH_2O	up to 200ml

Add KOH to pH to 6.8

Seed Germination

	(1L)
MS and vitamins	1x (4.4g)
Agarose	6g

pH 5.8 (KOH)

Cell Suspension

	(1L)
MS and vitamins	1x (4.4g)
Sucrose	10g
Agarose	6g

pH 5.7 (KOH)

Filter sterilise and add after autoclaving:

2,4-D (in ethanol)	0.5mg

Regeneration

	(1L)
MS and vitamins	1x (4.4g)
Sucrose	20g
Agar gel	4g

pH 6.0 (KOH)

Filter sterilise and add after autoclaving:

Zeatin Riboside (trans isomer)	2mg
Ticarcillin	320mg
Kanamycin (or the selection antibiotic)	100mg
Cefotaxime	250mg

Rooting

	(1L)
MS and vitamins	0.5x (2.2g)
Sucrose	5g
Gelrite	2.25g
pH 6.0 (KOH)	
Autoclave	
Filter sterilise and add after autoclaving:	
Ticarcillin	320mg
Kanamycin (or the selection antibiotic)	100mg

Chapter 1 appendix

List of *S.pennellii* x *S.lycopersicum* introgression lines

IL1-1
IL1-1-2
IL1-1-3
IL1-2
IL1-3
IL1-4
IL1-4-18
IL2-1
IL2-1-1
IL2-2
IL2-3
IL2-4
IL2-5
IL2-6
IL2-6-5
IL3-1
IL3-2
IL3-3
IL3-4
IL3-5
IL4-1
IL4-1-1
IL4-2
IL4-3
IL4-3-2
IL4-4
IL5-1
IL5-2
IL5-3
IL5-4
IL5-5
IL6-1
IL6-2
IL6-2-2
IL6-3
IL6-4
IL7-1
IL7-2
IL7-3
IL7-4
IL7-4-1
IL7-5
IL7-5-5
IL8-1
IL8-1-1
IL8-1-5
IL8-2
IL8-2-1
IL8-3
IL8-3-1
IL9-1
IL9-1-2
IL9-1-3
IL9-2
IL9-2-5
IL9-2-6
IL9-3
IL9-3-1
IL9-3-2
IL10-1
IL10-1-1
IL10-2
IL10-2-2
IL10-3
IL11-1
IL11-2
IL11-3
IL11-4
IL11-4-1
IL12-1
IL12-1-1
IL12-2
IL12-3
IL12-3-1
IL12-4
IL12-4-1

Chapter 4 appendix

Table 4-1 Table of S.pennellii stages for the time course, showing the weight range for fruit for each stage, the number of fruit per replicate and the number of replicates for analysis.

S.pennellii stage	1	2	3	4	5	6
Weight (g)	0.05-0.3	0.31-0.75	0.76-1.25	1.26-2.0	2.01-2.5	2.5+
Fruit per replicate	10	5	3	3	3	3
No. of replicates	4	4	4	4	4	4

Table 4-2 Table showing statistical significance of *S.lycopersicum* expression data for nine fruit stages. A = VTE2 pericarp, B = VTE2 epidermis, C = VTE1 pericarp, D = VTE1 epidermis, E = VTE3(1) pericarp, F = VTE3(1) epidermis, G = VTE4 pericarp, H = VTE4 epidermis, I = VTE5 pericarp, J = VTE5 epidermis, K = VTE6 pericarp, L = VTE6 epidermis, M = VTE3(2) pericarp, N = VTE3(2) epidermis. The dashed line (-) depict no statistical significance. The stars depict statistical significance, (*) = p value < 0.05 and (**) = p value < 0.01.

A

		S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	-	-	-	-
	IG2			-	-	-	-	-	-	-
	IG3				-	-	-	-	-	-
	MG					-	-	-	-	-
	B						-	-	-	-
	B+3							*	-	-
	B+5								-	*
	B+10									-
	B+20									

B

		S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	-	-	-	-
	IG2			-	-	-	-	-	-	-
	IG3				-	-	-	-	-	-
	MG					-	-	-	-	-
	B						-	-	-	-
	B+3							-	-	*
	B+5								-	**
	B+10									-
	B+20									

Table 4-2 continued.

C

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	*	**	**	-	**	**
	IG2			-	*	*	**	-	**	**
	IG3				-	-	-	-	**	**
	MG					-	-	-	-	-
	B						-	-	-	-
	B+3							-	-	-
	B+5								*	-
	B+10									-
	B+20									

D

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	-	-	-	-
	IG2			-	-	-	-	-	-	*
	IG3				-	-	-	-	-	-
	MG					-	-	-	-	-
	B						-	-	-	-
	B+3							-	-	-
	B+5								-	-
	B+10									-
	B+20									

Table 4-2 continued.

E

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	-	-	-	-
	IG2			-	-	-	*	-	-	-
	IG3				-	-	-	-	-	-
	MG					-	-	-	-	-
	B						-	-	-	-
	B+3							-	-	-
	B+5								-	-
	B+10									-
	B+20									

F

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	-	-	-	-
	IG2			-	-	-	-	-	-	-
	IG3				-	-	-	-	-	-
	MG					-	-	-	-	-
	B						-	-	*	-
	B+3							-	-	-
	B+5								**	-
	B+10									-
	B+20									

Table 4-2 continued.

G

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	**	**	**	**
	IG2			-	-	-	-	-	-	-
	IG3				-	-	-	-	-	-
	MG					-	-	-	-	-
	B						-	-	-	-
	B+3							-	-	-
	B+5								-	-
	B+10									-
	B+20									

H

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	-	-	-	*
	IG2			-	-	-	-	-	-	-
	IG3				-	-	-	-	-	-
	MG					-	-	-	-	-
	B						-	-	-	-
	B+3							-	-	-
	B+5								-	-
	B+10									-
	B+20									

Table 4-2 continued.

I

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	**	-	**	**
	IG2			-	-	*	**	-	**	**
	IG3				-	-	**	-	*	**
	MG					-	**	-	**	**
	B						-	-	-	-
	B+3							-	-	-
	B+5								-	*
	B+10									-
	B+20									

J

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		**	-	-	**	*	-	**	**
	IG2			-	-	-	-	-	-	-
	IG3				-	*	-	-	*	**
	MG					-	-	-	-	*
	B						-	-	-	-
	B+3							-	-	-
	B+5								-	-
	B+10									-
	B+20									

Table 4-2 continued.

K

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	-	-	**	**
	IG2			-	-	-	-	-	**	**
	IG3				-	-	-	-	*	*
	MG					-	-	-	*	**
	B						-	-	-	-
	B+3							-	-	-
	B+5								-	-
	B+10									-
	B+20									

L

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		**	-	-	-	-	-	-	**
	IG2			-	-	-	-	-	-	**
	IG3				-	-	-	-	-	**
	MG					-	-	-	-	*
	B						-	-	-	-
	B+3							-	*	**
	B+5								-	-
	B+10									-
	B+20									

Table 4-2 continued.

M

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	**	-	-	*
	IG2			-	-	-	*	-	-	-
	IG3				-	-	-	-	-	-
	MG					-	-	-	-	-
	B						-	-	-	-
	B+3							-	-	-
	B+5								-	-
	B+10									-
	B+20									

N

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	-	-	-	-
	IG2			-	-	-	*	-	-	-
	IG3				-	-	-	-	-	-
	MG					-	-	-	-	-
	B						**	-	-	-
	B+3							-	-	**
	B+5								-	-
	B+10									-
	B+20									

Table 4-3 Table shows statistical significance of tocopherol contents of the *S.lycopersicum* fruit stages for the tissue of the epidermis, calculated using Tukey tests. The letters represent the different forms of tocopherol; A – alpha tocopherol, B – beta tocopherol, C – gamma tocopherol, D – delta tocopherol, E – total tocopherol. The stars indicate statistical significance (*) = p value < 0.05, (**) = p value < 0.01 and - = no significant difference.

A

		S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	*	**	**	*
	IG2			-	-	-	*	*	**	-
	IG3				-	-	*	*	**	-
	MG					-	*	**	**	-
	B						-	*	**	-
	B+3							-	-	-
	B+5								-	-
	B+10									-
	B+20									

B

		S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	-	-	-	-
	IG2			-	-	-	-	-	-	-
	IG3				-	-	-	-	-	-
	MG					-	-	-	-	-
	B						-	-	-	-
	B+3							-	-	-
	B+5								-	-
	B+10									-
	B+20									

Table 4-3 continued.

C

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	**	*	-	-
	IG2			-	-	-	*	*	-	-
	IG3				-	-	**	**	-	-
	MG					-	**	**	-	-
	B						-	*	*	-
	B+3							-	-	-
	B+5								-	-
	B+10									-
	B+20									

D

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	**	**	-	-
	IG2			-	-	-	**	**	-	-
	IG3				-	-	**	**	-	-
	MG						**	**	-	-
	B						**	**	-	-
	B+3							-	**	**
	B+5								**	*
	B+10									-
	B+20									

Table 4-3 continued.

E

	S.lycopersicum fruit stage								
S.lycopersicum fruit stage		IG1	IG2	IG3	MG	B	B+3	B+5	B+10	B+20
	IG1		-	-	-	-	**	**	**	-
	IG2			-	-	-	*	**	**	-
	IG3				-	-	*	**	**	-
	MG					-	*	**	**	-
	B						*	*	*	-
	B+3							-	-	-
	B+5								-	-
	B+10									-
	B+20									

Table 4-4 Table shows the Tukey test results for tocopherol contents of *S.pennellii* fruit stages. The tables are for the different forms of tocopherols; A – alpha tocopherol, B – beta tocopherol, C – gamma tocopherol and D – delta tocopherol. The dashed line (-) indicate no statistical significance (P>0.05), the stars (*) indicate statistical significance (P<0.05).

A

		S.pennellii fruit stage					
			2	3	4	5	6
S.pennellii fruit stage	1		-	-	-	-	-
	2			-	-	-	-
	3				-	-	-
	4					-	-
	5						-
	6						

B

		S.pennellii fruit stage					
		1	2	3	4	5	6
S.pennellii fruit stage	1		-	-	-	-	-
	2			-	-	-	-
	3				-	-	-
	4					-	-
	5						-
	6						

C

		S.pennellii fruit stage					
		1	2	3	4	5	6
S.pennellii fruit stage	1		-	-	-	-	*
	2			-	-	-	*
	3				-	-	-
	4					-	*
	5						*
	6						

D

		S.pennellii fruit stage					
		1	2	3	4	5	6
S.pennellii fruit stage	1		-	-	-	-	-
	2			-	-	-	-
	3				-	-	-
	4					-	-
	5						-
	6						

Chapter 5 appendix

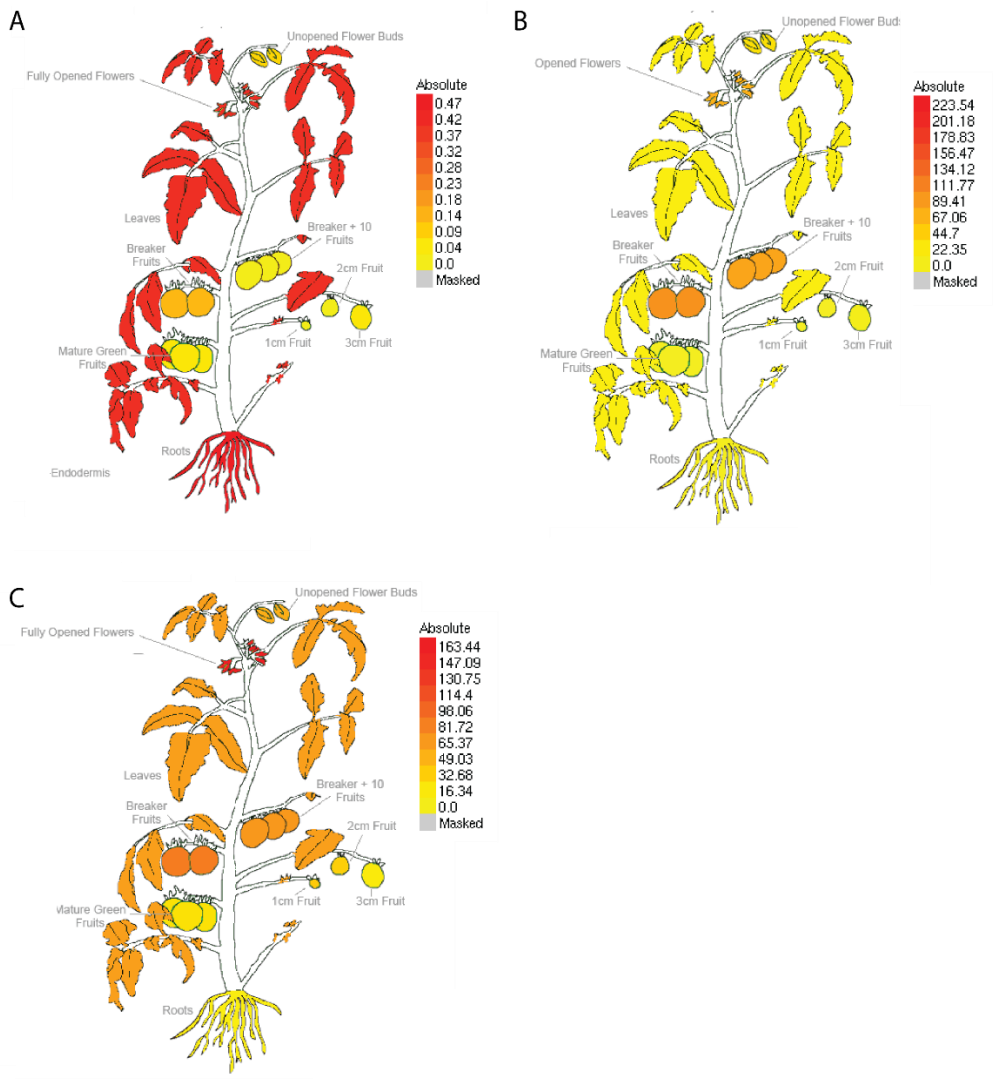


Figure 5-1 Expression of (A) *GGPS(1),* (B) *GGPS(2)* and (C) *GGPS(3)* in *S.lycopersicum* cv. Heinz. These pictures were adapted from the eFP browser (Patel, Koenig et al., 2013). The absolute (RPKM-normalised) expression values are shown on the scale bars.

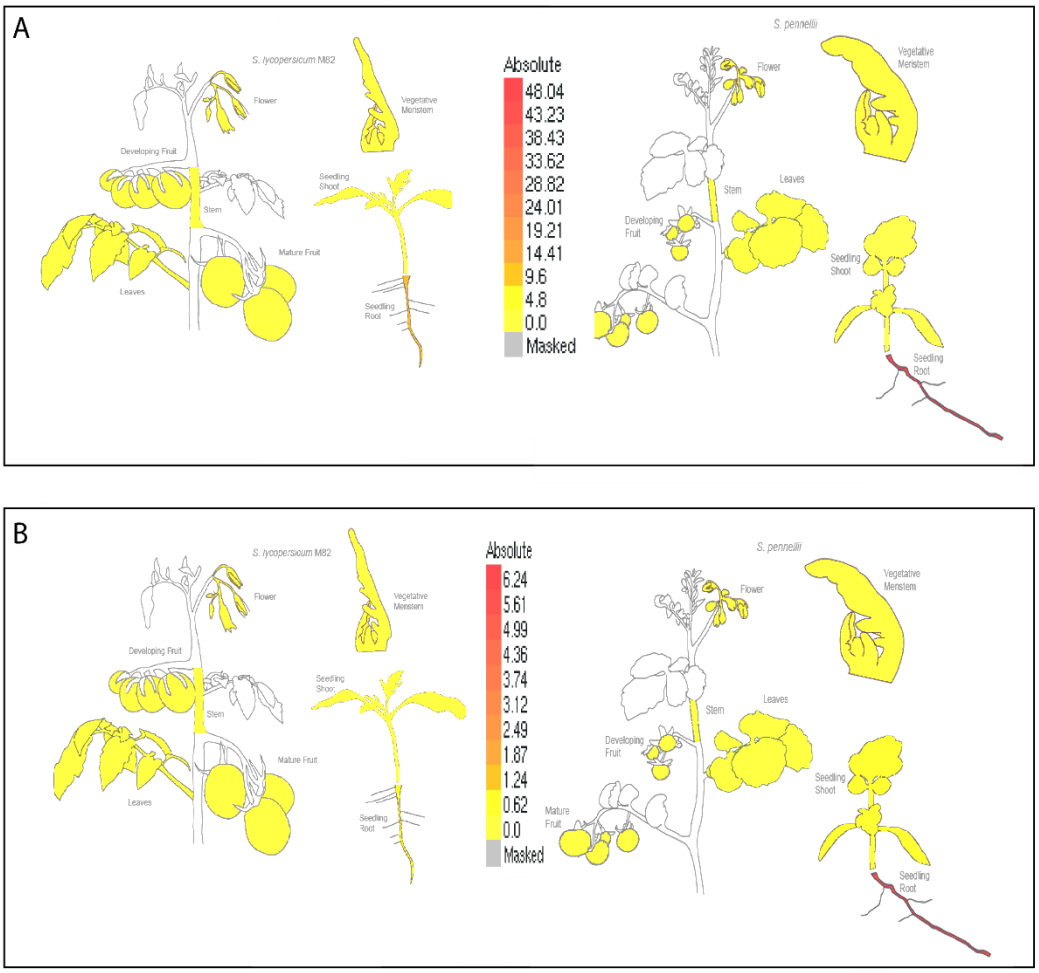


Figure 5-2 Expression of (A) SlMYB79 and its ortholog in *S.lycopersicum* and *S.pennellii,* (B) SlMYB71 and its ortholog in *S.lycopersicum* cv. Heinz and *S.pennellii.* These pictures were adapted from the eFP browser (Patel, Koenig et al., 2013). The absolute (RPKM-normalised) expression values are shown on the scale bars.

Chapter 7 appendix

List of *S.pennellii* x *S.lycopersicum* introgression lines that were not analysed in Chapter 7

IL2-3

IL2-5

IL3-2

IL3-3

IL4-2

IL4-4

IL6-2

IL6-2-2

IL7-5-5

IL8-1-5

IL8-3

IL8-3-1

IL9-1

IL9-1-2

IL9-3

IL11-2

IL11-3

IL12-1

References

ABUSHITA, A. A., HEBSHI, E. A., DAOOD, H. G. & BIACS, P. A. 1997. Determination of antioxidant vitamins in tomatoes. *Food Chemistry,* 60**,** 207-212.

AGGARWAL, B. B., SUNDARAM, C., PRASAD, S. & KANNAPPAN, R. 2010. Tocotrienols, the vitamin E of the 21st century: its potential against cancer and other chronic diseases. *Biochem Pharmacol,* 80**,** 1613-31.

AHSAN, H., AHAD, A. & SIDDIQUI, W. A. 2015. A review of characterization of tocotrienols from plant oils and foods. *J Chem Biol,* 8**,** 45-59.

ALBA, R., PAYTON, P., FEI, Z., MCQUINN, R., DEBBIE, P., MARTIN, G. B., TANKSLEY, S. D. & GIOVANNONI, J. J. 2005. Transcriptome and selected metabolite analyses reveal multiple points of ethylene control during tomato fruit development. *Plant Cell,* 17**,** 2954-65.

ALBERT, N. W., LEWIS, D. H., ZHANG, H., IRVING, L. J., JAMESON, P. E. & DAVIES, K. M. 2009. Light-induced vegetative anthocyanin pigmentation in *Petunia*. *J Exp Bot,* 60**,** 2191-202.

ALBERTS, R., TERPSTRA, P., LI, Y., BREITLING, R., NAP, J. P. & JANSEN, R. C. 2007. Sequence polymorphisms cause many false cis eQTLs. *PLoS One,* 2**,** e622.

ALMEIDA, J., QUADRANA, L., ASIS, R., SETTA, N., DE GODOY, F., BERMUDEZ, L., OTAIZA, S. N., CORREA DA SILVA, J. V., FERNIE, A. R., CARRARI, F. & ROSSI, M. 2011. Genetic dissection of vitamin E biosynthesis in tomato. *J Exp Bot,* 62**,** 3781-98.

ALMEIDA, J., ASIS, R., MOLINERI, V. N., SESTARI, I., LIRA, B. S., CARRARI, F., PERES, L. E. P. & ROSSI, M. 2015. Fruits from ripening impaired, chlorophyll degraded and jasmonate insensitive tomato mutants have altered tocopherol content and composition. *Phytochemistry,* 111**,** 72-83.

ALMEIDA, J., AZEVEDO, M. D., SPICHER, L., GLAUSER, G., VOM DORP, K., GUYER, L., CARRANZA, A. D., ASIS, R., DE SOUZA, A. P., BUCKERIDGE, M., DEMARCO, D., BRES, C., ROTHAN, C., PERES, L. E. P., HORTENSTEINER, S., KESSLER, F., DORMANN, P., CARRARI, F. & ROSSI, M. 2016. Down-regulation of tomato PHYTOL KINASE strongly impairs tocopherol biosynthesis and affects prenyllipid metabolism in an organ-specific manner. *J Exp Bot,* 67**,** 919-934.

ALSEEKH, S., TOHGE, T., WENDENBERG, R., SCOSSA, F., OMRANIAN, N., LI, J., KLEESSEN, S., GIAVALISCO, P., PLEBAN, T., MUELLER-ROEBER, B., ZAMIR, D., NIKOLOSKI, Z. & FERNIE, A. R. 2015. Identification and mode of inheritance of quantitative trait loci for secondary metabolite abundance in tomato. *Plant Cell,* 27**,** 485-512.

AUSTIN, J. R., FROST, E., VIDI, P. A., KESSLER, F. & STAEHELIN, L. A. 2006. Plastoglobules are lipoprotein subcompartments of the chloroplast that are permanently coupled to thylakoid membranes and contain biosynthetic enzymes. *Plant Cell,* 18**,** 1693-1703.

AZARI, R., TADMOR, Y., MEIR, A., REUVENI, M., EVENOR, D., NAHON, S., SHLOMO, H., CHEN, L. & LEVIN, I. 2010. Light signaling genes and their manipulation

towards modulation of phytonutrient content in tomato fruits. *Biotechnol Adv,* 28**,** 108-18.

BAILEY, P. C., DICKS, J., WANG, T. L. & MARTIN, C. 2008. IT3F: A web-based tool for functional analysis of transcription factors in plants. *Phytochemistry,* 69**,** 2417-2425.

BAILEY, P. C., DICKS, J., WANG, T. L. & MARTIN, C. 2012. IT3F: A web-based tool for functional analysis of transcription factors in plants *Phytochemistry,* 81**,** 176-176.

BARRANGOU, R., FREMAUX, C., DEVEAU, H., RICHARDS, M., BOYAVAL, P., MOINEAU, S., ROMERO, D. A. & HORVATH, P. 2007. CRISPR provides acquired resistance against viruses in Prokaryotes. *Science,* 315**,** 1709-1712.

BARRY, C. S., MCQUINN, R. P., CHUNG, M. Y., BESUDEN, A. & GIOVANNONI, J. J. 2008. Amino acid substitutions in homologs of the STAY-GREEN protein are responsible for the green-flesh and chlorophyll retainer mutations of tomato and pepper. *Plant Physiol,* 147**,** 179-87.

BASSEL, G. W., MULLEN, R. T. & BEWLEY, J. D. 2006. ABI3 expression ceases following, but not during, germination of tomato and Arabidopsis seeds. *J Exp Bot,* 57**,** 1291-7.

BAUD, S., MENDOZA, M. S., TO, A., HARSCOET, E., LEPINIEC, L. & DUBREUCQ, B. 2007. WRINKLED1 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during seed maturation in Arabidopsis. *Plant J,* 50**,** 825-38.

BAXTER, C. J., SABAR, M., QUICK, W. P. & SWEETLOVE, L. J. 2005. Comparison of changes in fruit gene expression in tomato introgression lines provides evidence of genome-wide transcriptional changes and reveals links to mapped QTLs and described traits. *J Exp Bot,* 56**,** 1591-1604.

BEHRINGER, C. & SCHWECHHEIMER, C. 2015. B-GATA transcription factors - insights into their structure, regulation, and role in plant development. *Front Plant Sci,* 6**,** 90.

BERGMULLER, E., PORFIROVA, S. & DORMANN, P. 2003. Characterization of an *Arabidopsis* mutant deficient in gamma-tocopherol methyltransferase. *Plant Mol Biol,* 52**,** 1181-1190.

BI, Y. M., ZHANG, Y., SIGNORELLI, T., ZHAO, R., ZHU, T. & ROTHSTEIN, S. 2005. Genetic analysis of *Arabidopsis* GATA transcription factor gene family reveals a nitrate-inducible member important for chlorophyll synthesis and glucose sensitivity. *Plant J,* 44**,** 680-92.

BINO, R. J., RIC DE VOS, C. H., LIEBERMAN, M., HALL, R. D., BOVY, A., JONKER, H. H., TIKUNOV, Y., LOMMEN, A., MOCO, S. & LEVIN, I. 2005. The light-hyperresponsive high pigment-2dg mutation of tomato: alterations in the fruit metabolome. *New Phytol,* 166**,** 427-38.

BIOCAT. 2018. *CRISPR/Cas9* [Online]. https://www.biocat.com/genome-engineering. [Accessed 12/09/18 2018].

BLUM, S., VARDI, M., LEVY, N. S., MILLER-LOTAN, R. & LEVY, A. P. 2010. The effect of vitamin E supplementation on cardiovascular risk in diabetic individuals with different haptoglobin phenotypes. *Atherosclerosis,* 211**,** 25-27.

BOAZ, M., SMETANA, S., WEINSTEIN, T., MATAS, Z., GAFTER, U., IAINA, A., KNECHT, A., WEISSGARTEN, Y., BRUNNER, D., FAINARU, M. & GREEN, M. S. 2000. Secondary prevention with antioxidants of cardiovascular disease in endstage renal disease (SPACE): randomised placebo-controlled trial. *Lancet,* 356**,** 1213-1218.

BOLGER, A., SCOSSA, F., BOLGER, M. E., LANZ, C., MAUMUS, F., TOHGE, T., QUESNEVILLE, H., ALSEEKH, S., SORENSEN, I., LICHTENSTEIN, G., FICH, E. A., CONTE, M., KELLER, H., SCHNEEBERGER, K., SCHWACKE, R., OFNER, I., VREBALOV, J., XU, Y. M., OSORIO, S., AFLITOS, S. A., SCHIJLEN, E., JIMENEZ-GOMEZ, J. M., RYNGAJLLO, M., KIMURA, S., KUMAR, R., KOENIG, D., HEADLAND, L. R., MALOOF, J. N., SINHA, N., VAN HAM, R. C. H. J., LANKHORST, R. K., MAO, L. Y., VOGEL, A., ARSOVA, B., PANSTRUGA, R., FEI, Z. J., ROSE, J. K. C., ZAMIR, D., CARRARI, F., GIOVANNONI, J. J., WEIGEL, D., USADEL, B. & FERNIE, A. R. 2014. The genome of the stress-tolerant wild tomato species *Solanum pennellii*. *Nat Genet,* 46**,** 1034-+.

BOLOTIN, A., QUINQUIS, B., SOROKIN, A. & EHRLICH, S. D. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology,* 151**,** 2551-61.

BORTESI, L. & FISCHER, R. 2015. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol Adv,* 33**,** 41-52.

BOTHA, C. E. J., CROSS, R. H. M., VAN BEL, A. J. E. & PETER, C. I. 2000. Phloem loading in the sucrose-export-defective (SXD-1) mutant maize is limited by callose deposition at plasmodesmata in bundle sheath-vascular parenchyma interface. *Protoplasma,* 214**,** 65-72.

BOU GHANEM, E. N., CLARK, S., DU, X., WU, D., CAMILLI, A., LEONG, J. M. & MEYDANI, S. N. 2015. The α-tocopherol form of vitamin E reverses age-associated susceptibility to *Streptococcus pneumoniae l*ung infection by modulating pulmonary neutrophil recruitment. *Journal Immunol* 194**,** 1090/1099.

BOWRY, V. W., INGOLD, K. U. & STOCKER, R. 1992. Vitamin-E in human low-density-lipoprotein - when and how this antioxidant becomes a prooxidant. *Biochem J,* 288**,** 341-344.

BRAMLEY, P. M. 2002. Regulation of carotenoid formation during tomato fruit ripening and development. *J Exp Bot,* 53**,** 2107-2113.

BROUNS, S. J. J., JORE, M. M., LUNDGREN, M., WESTRA, E. R., SLIJKHUIS, R. J. H., SNIJDERS, A. P. L., DICKMAN, M. J., MAKAROVA, K. S., KOONIN, E. V. & VAN DER OOST, J. 2008. Small CRISPR RNAs guide antiviral defense in Prokaryotes. *Science,* 321**,** 960-964.

BURTON, G. W., TRABER, M. G., ACUFF, R. V., WALTERS, D. N., KAYDEN, H., HUGHES, L. & INGOLD, K. U. 1998. Human plasma and tissue a-tocopherol

concentrations in response to supplementation with deuterated natural and synthetic vitamin E. *Am J Clin Nutr,* 67**,** 669-684.

BUTELLI, E., TITTA, L., GIORGIO, M., MOCK, H. P., MATROS, A., PETEREK, S., SCHIJLEN, E. G., HALL, R. D., BOVY, A. G., LUO, J. & MARTIN, C. 2008. Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat Biotechnol,* 26**,** 1301-8.

CAHOON, E. B., HALL, S. E., RIPP, K. G., GANZKE, T. S., HITZ, W. D. & COUGHLAN, S. J. 2003. Metabolic redesign of vitamin E biosynthesis in plants for tocotrienol production and increased antioxidant content. *Nat Biotechnol,* 21**,** 1082-7.

CANADY, M. A., MEGLIC, V. & CHETELAT, R. T. 2005. A library of *Solanum lycopersicoides* introgression lines in cultivated tomato. *Genome,* 48**,** 685-697.

CANENE-ADAMS, K., CAMPBELL, J. K., ZARIPHEH, S., JEFFERY, E. H. & ERDMAN JR, J. W. 2005. The tomato as a functional food. *J Nutr,* 135**,** 1226-1230.

CARRARI, F., BAXTER, C., USADEL, B., URBANCZYK-WOCHNIAK, E., ZANOR, M. I., NUNES-NESI, A., NIKIFOROVA, V., CENTERO, D., RATZKA, A., PAULY, M., SWEETLOVE, L. J. & FERNIE, A. R. 2006. Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. *Plant Physiol,* 142**,** 1380-1396.

CERNAC, A. & BENNING, C. 2004. WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in *Arabidopsis*. *Plant J,* 40**,** 575-85.

CHAPMAN, N. H., BONNET, J., GRIVET, L., LYNN, J., GRAHAM, N., SMITH, R., SUN, G., WALLEY, P. G., POOLE, M., CAUSSE, M., KING, G. J., BAXTER, C. & SEYMOUR, G. B. 2012. High-resolution mapping of a fruit firmness-related quantitative trait locus in tomato reveals epistatic interactions associated with a complex combinatorial locus. *Plant Physiol,* 159**,** 1644-57.

CHE, P., ZHAO, Z. Y., GLASSMAN, K., DOLDE, D., HU, T. X., JONES, T. J., OBUKOSIA, S., WAMBUGU, F. & ALBERTSEN, M. C. 2016. Elevated vitamin E content improves all-trans beta-carotene accumulation and stability in biofortified sorghum. *Proc Natl Acad Sci U S A,* 113**,** 11040-5.

CHEN, H., ZHANG, J., NEFF, M. M., HONG, S. W., ZHANG, H., DENG, X. W. & XIONG, L. 2008. Integration of light and abscisic acid signaling during seed germination and early seedling development. *Proc Natl Acad Sci U S A,* 105**,** 4495-500.

CHEN, R. H. & LIPSICK, J. S. 1993. Differential transcriptional activation by v-myb and c-myb in animal cells and *Saccharomyces cerevisiae*. *Mol Cell Biol,* 13**,** 4423-31.

CHENG, Z., SATTLER, S., MAEDA, H., SAKURAGI, Y., BRYANT, D. A. & DELLAPENNA, D. 2003. Highly divergent methyltransferases catalyze a conserved reaction in tocopherol and plastoquinone synthesis in cyanobacteria and photosynthetic eukaryotes. *Plant Cell,* 15**,** 2343-56.

CHIN, S. F., HAMID, N. A., LATIFF, A. A., ZAKARIA, Z., MAZLAN, M., YUSOF, Y. A., KARIM, A. A., IBAHIM, J., HAMID, Z. & NGAH, W. Z. 2008. Reduction of DNA damage in older healthy adults by Tri E tocotrienol supplementation. *Nutrition,* 24**,** 1-10.

CHIN, S. F., IBAHIM, J., MAKPOL, S., ABDUL HAMID, N. A., ABDUL LATIFF, A., ZAKARIA, Z., MAZLAN, M., MOHD YUSOF, Y. A., ABDUL KARIM, A. & WAN NGAH, W. Z. 2011. Tocotrienol rich fraction supplementation improved lipid profile and oxidative status in healthy older adults: A randomized controlled study. *Nutr Metab (Lond),* 8**,** 42.

CHITWOOD, D. H., KUMAR, R., HEADLAND, L. R., RANJAN, A., COVINGTON, M. F., ICHIHASHI, Y., FULOP, D., JIMENEZ-GOMEZ, J. M., PENG, J., MALOOF, J. N. & SINHA, N. R. 2013. A quantitative genetic basis for leaf morphology in a set of precisely defined tomato introgression lines. *Plant Cell,* 25**,** 2465-2481.

CHROST, B., FALK, J., KERNEBECK, B., MÖLLEKEN, H. & KRUPINSKA, K. 1999. *The chloroplast: from molecular biology to biotechnology*, Kluwer Academic Publishers.

CHUN, J., LEE, J., YE, L., EXLER, J. & EITENMILLER, R. R. 2006. Tocopherol and tocotrienol contents of raw and processed fruits and vegetables in the United States diet. *J Food Comp Anal,* 19**,** 196-204.

COLLAKOVA, E. & DELLAPENNA, D. 2001. Isolation and functional analysis of homogentisate phytyltransferase from *Synechocystis* sp PCC 6803 and arabidopsis. *Plant Physiol,* 127**,** 1113-1124.

COLLAKOVA, E. & DELLAPENNA, D. 2003. Homogentisate phytyltransferase activity is limiting for tocopherol biosynthesis in *Arabidopsis*. *Plant Physiol,* 131**,** 632-642.

CONG, L., RAN, F. A., COX, D., LIN, S., BARRETTO, R., HABIB, N., HSU, P. D., WU, X., JIANG, W., MARRAFFINI, L. A. & ZHANG, F. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science,* 339**,** 819-23.

CORDOBA, E., SALMI, M. & LEON, P. 2009. Unravelling the regulatory mechanisms that modulate the MEP pathway in higher plants. *J Exp Bot,* 60**,** 2933-2943.

CUBILLOS, F. A., COUSTHAM, V. & LOUDET, O. 2012. Lessons from eQTL mapping studies: non-coding regions and their role behind natural phenotypic variation in plants. *Curr Opin Plant Biol,* 15**,** 192-198.

DAL CIN, V., TIEMAN, D. M., TOHGE, T., MCQUINN, R., DE VOS, R. C., OSORIO, S., SCHMELZ, E. A., TAYLOR, M. G., SMITS-KROON, M. T., SCHUURINK, R. C., HARING, M. A., GIOVANNONI, J., FERNIE, A. R. & KLEE, H. J. 2011. Identification of genes in the phenylalanine metabolic pathway by ectopic expression of a MYB transcription factor in tomato fruit. *Plant Cell,* 23**,** 2738-53.

DAVULURI, G. R., VAN TUINEN, A., FRASER, P. D., MANFREDONIA, A., NEWMAN, R., BURGESS, D., BRUMMELL, D. A., KING, S. R., PALYS, J., UHLIG, J., BRAMLEY, P. M., PENNINGS, H. M. & BOWLER, C. 2005. Fruit-specific RNAi-

mediated suppression of DET1 enhances carotenoid and flavonoid content in tomatoes. *Nat Biotechnol,* 23**,** 890-5.

DELKER, C., SONNTAG, L., JAMES, G. V., JANITZA, P., IBANEZ, C., ZIERMANN, H., PETERSON, T., DENK, K., MULL, S., ZIEGLER, J., DAVIS, S. J., SCHNEEBERGER, K. & QUINT, M. 2014. The DET1-COP1-HY5 pathway constitutes a multipurpose signaling module regulating plant photomorphogenesis and thermomorphogenesis. *Cell Rep,* 9**,** 1983-1989.

DELLAPENNA, D. 2005. Progress in the dissection and manipulation of vitamin E synthesis. *Trends Plant Sci,* 10**,** 574-9.

DELMAS, F., SANKARANARAYANAN, S., DEB, S., WIDDUP, E., BOURNONVILLE, C., BOLLIER, N., NORTHEY, J. G. B., MCCOURT, P. & SAMUEL, M. A. 2013. ABI3 controls embryo degreening through Mendel's I locus. *Proc Natl Acad Sci USA,* 110**,** E3888-E3894.

DELTCHEVA, E., CHYLINSKI, K., SHARMA, C. M., GONZALES, K., CHAO, Y., PIRZADA, Z. A., ECKERT, M. R., VOGEL, J. & CHARPENTIER, E. 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature,* 471**,** 602-7.

DIETRICH, M., TRABER, M. G., JACQUES, P. F., CROSS, C. E., HU, Y. Q. & BLOCK, G. 2006. Does gamma-tocopherol play a role in the primary prevention of heart disease and cancer? A review. *J Am Coll Nutr,* 25**,** 292-299.

DROGE-LASER, W., SNOEK, B. L., SNEL, B. & WEISTE, C. 2018. The *Arabidopsis* bZIP transcription factor family-an update. *Curr Opin Plant Biol,* 45**,** 36-49.

DRUKA, A., POTOKINA, E., LUO, Z., BONAR, N., DRUKA, I., ZHANG, L., MARSHALL, D. F., STEFFENSON, B. J., CLOSE, T. J., WISE, R. P., KLEINHOFS, A., WILLIAMS, R. W., KEARSEY, M. J. & WAUGH, R. 2008. Exploiting regulatory variation to identify genes underlying quantitative resistance to the wheat stem rust pathogen *Puccinia graminis* f. sp. tritici in barley. *Theor Appl Genet,* 117**,** 261-72.

DUBOS, C., LE GOURRIEREC, J., BAUDRY, A., HUEP, G., LANET, E., DEBEAUJON, I., ROUTABOUL, J. M., ALBORESI, A., WEISSHAAR, B. & LEPINIEC, L. 2008. MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. *The Plant Journal,* 55**,** 940-953.

DUBOS, C., STRACKE, R., GROTEWOLD, E., WEISSHAAR, B., MARTIN, C. & LEPINIEC, L. 2010. MYB transcription factors in *Arabidopsis*. *Trends Plant Sci,* 15**,** 573-81.

DWIYANTI, M. S., YAMADA, T., SATO, M., ABE, J. & KITAMURA, K. 2011. Genetic variation of gamma-tocopherol methyltransferase gene contributes to elevated alpha-tocopherol content in soybean seeds. *BMC Plant Biol,* 11**,** 152.

ENFISSI, E. M., BARNECHE, F., AHMED, I., LICHTLE, C., GERRISH, C., MCQUINN, R. P., GIOVANNONI, J. J., LOPEZ-JUEZ, E., BOWLER, C., BRAMLEY, P. M. & FRASER, P. D. 2010. Integrative transcript and metabolite analysis of nutritionally enhanced DE-ETIOLATED1 downregulated tomato fruit. *Plant Cell,* 22**,** 1190-215.

ENFISSI, E. M. A., FRASER, P. D., LOIS, L. M., BORONAT, A., SCHUCH, W. & BRAMLEY, P. M. 2005. Metabolic engineering of the mevalonate and non-mevalonate isopentenyl diphosphate-forming pathways for the production of health-promoting isoprenoids in tomato. *Plant Biotech J,* 3**,** 17-27.

ESHED, Y. & ZAMIR, D. 1995. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield associated QTL. *Gen Soc Am,* 141**,** 1147-1 162.

EVANS, H. M. & BISHOP, K. S. 1922. Fetal Resorption: on the existence of a hitherto unrecognized dietary factor essential for reproduction. *Science*, 56, 650-651

EXPOSITO-RODRIGUEZ, M., BORGES, A. A., BORGES-PEREZ, A. & PEREZ, J. A. 2008. Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. *BMC Plant Biol,* 8**,** 131.

FALK, J., ANDERSEN, G., KERNEBECK, B. & KRUPINSKA, K. 2003. Constitutive overexpression of barley 4-hydroxyphenylpyruvate dioxygenase in tobacco results in elevation of the vitamin E content in seeds but not in leaves. *FEBS Letters,* 540**,** 35-40.

FAOSTAT. 2010. *Food Supply - Crops Primary Equivalent* [Online]. Available: http://faostat3.fao.org/download/FB/CC/E [Accessed 10/03 2015].

FERNANDEZ-POZO, N., MENDA, N., EDWARDS, J. D., SAHA, S., TECLE, I. Y., STRICKLER, S. R., BOMBARELY, A., FISHER-YORK, T., PUJAR, A., FOERSTER, H., YAN, A. & MUELLER, L. A. 2015a. The Sol Genomics Network (SGN)--from genotype to phenotype to breeding. *Nucleic Acids Res,* 43**,** D1036-41.

FERNANDEZ-POZO, N., ROSLI, H. G., MARTIN, G. B. & MUELLER, L. A. 2015b. The SGN VIGS Tool: user-friendly software to design Virus-Induced Gene Silencing (VIGS) constructs for functional genomics. *Mol Plant,* 8**,** 486-488.

FERNIE, A. R., TADMOR, Y. & ZAMIR, D. 2006. Natural genetic variation for improving crop quality. *Curr Opin Plant Biol,* 9**,** 196-202.

FERRARI, S., GALLETTI, R., DENOUX, C., DE LORENZO, G., AUSUBEL, F. M. & DEWDNEY, J. 2007. Resistance to *Botrytis cinerea* induced in *Arabidopsis* by elicitors is independent of salicylic acid, ethylene, or jasmonate signaling but requires PHYTOALEXIN DEFICIENT3. *Plant Physiol,* 144**,** 367-379.

FINN, R. D., ATTWOOD, T. K., BABBITT, P. C., BATEMAN, A., BORK, P., BRIDGE, A. J., CHANG, H. Y., DOSZTANYI, Z., EL-GEBALI, S., FRASER, M., GOUGH, J., HAFT, D., HOLLIDAY, G. L., HUANG, H. Z., HUANG, X. S., LETUNIC, I., LOPEZ, R., LU, S. N., MARCHLER-BAUER, A., MI, H. Y., MISTRY, J., NATALE, D. A., NECCI, M., NUKA, G., ORENGO, C. A., PARK, Y., PESSEAT, S., PIOVESAN, D., POTTER, S. C., RAWLINGS, N. D., REDASCHI, N., RICHARDSON, L., RIVOIRE, C., SANGRADOR-VEGAS, A., SIGRIST, C., SILLITOE, I., SMITHERS, B., SQUIZZATO, S., SUTTON, G., THANKI, N., THOMAS, P. D., TOSATTO, S. C. E., WU, C. H., XENARIOS, I., YEH, L. S., YOUNG, S. Y. & MITCHELL, A. L. 2017.

InterPro in 2017-beyond protein family and domain annotations. *Nuc Acid Res,* 45**,** D190-D199.

FITZPATRICK, T. B., BASSET, G. J. C., BOREL, P., CARRARI, F., DELLAPENNA, D., FRASER, P. D., HELLMANN, H., OSORIO, S., ROTHAN, C., VALPUESTA, V., CARIS-VEYRAT, C. & FERNIE, A. R. 2012. Vitamin deficiencies in humans: can plant science help? *Plant Cell,* 24**,** 395-414.

FONSECA, S., FERNANDEZ-CALVO, P., FERNANDEZ, G. M., DIEZ-DIAZ, M., GIMENEZ-IBANEZ, S., LOPEZ-VIDRIERO, I., GODOY, M., FERNANDEZ-BARBERO, G., VAN LEENE, J., DE JAEGER, G., FRANCO-ZORRILLA, J. M. & SOLANO, R. 2014. bHLH003, bHLH013 and bHLH017 are new targets of JAZ repressors negatively regulating JA responses. *Plos One,* 9.

FRANCO-ZORRILLA, J. M., LOPEZ-VIDRIERO, I., CARRASCO, J. L., GODOY, M., VERA, P. & SOLANO, R. 2014. DNA-binding specificities of plant transcription factors and their potential to define target genes. *Proc Natl Acad Sci USA,* 111**,** 2367-72.

FRASER, P. D., TRUESDALE, M. R., BIRD, C. R., SCHUCH, W. & BRAMLEY, P. M. 1994. Carotenoid biosynthesis during tomato fruit development (evidence for tissue-specific gene expression). *Plant Physiol,* 105**,** 405-413.

FRASER, P. D., ENFISSI, E. M., HALKET, J. M., TRUESDALE, M. R., YU, D., GERRISH, C. & BRAMLEY, P. M. 2007. Manipulation of phytoene levels in tomato fruit: effects on isoprenoids, plastids, and intermediary metabolism. *Plant Cell,* 19**,** 3194-211.

GALPAZ, N., GONDA, I., SHEM-TOV, D., BARAD, O., TZURI, G., LEV, S., FEI, Z. J., XU, Y. M., MAO, L. Y., JIAO, C., HAREL-BEJA, R., DORON-FAIGENBOIM, A., TZFADIA, O., BAR, E., MEIR, A., SA'AR, U., FAIT, A., HALPERIN, E., KENIGSWALD, M., FALLIK, E., LOMBARDI, N., KOL, G., RONEN, G., BURGER, Y., GUR, A., TADMOR, Y., PORTNOY, V., SCHAFFER, A. A., LEWINSOHN, E., GIOVANNONI, J. J. & KATZIR, N. 2018. Deciphering genetic factors that determine melon fruit-quality traits using RNA-Seq-based high-resolution QTL and eQTL mapping. *Plant J,* 94**,** 169-191.

GANGAPPA, S. N. & KUMAR, S. V. 2017. DET1 and HY5 control PIF4-mediated thermosensory elongation growth through distinct mechanisms. *Cell Rep,* 18**,** 344-351.

GAO, Z., DANEVA, A., SALANENKA, Y., VAN DURME, M., HUYSMANS, M., LIN, Z. C., DE WINTER, F., VANNESTE, S., KARIMI, M., VAN DE VELDE, J., VANDEPOELE, K., VAN DE WALLE, D., DEWETTINCK, K., LAMBRECHT, B. N. & NOWACK, M. K. 2018. KIRA1 and ORESARA1 terminate flower receptivity by promoting cell death in the stigma of *Arabidopsis*. *Nat Plants,* 4**,** 365-+.

GARNEAU, J. E., DUPUIS, M. E., VILLION, M., ROMERO, D. A., BARRANGOU, R., BOYAVAL, P., FREMAUX, C., HORVATH, P., MAGADAN, A. H. & MOINEAU, S. 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nat,* 468**,** 67-71.

GILAD, Y., RIFKIN, S. A. & PRITCHARD, J. K. 2008. Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends Genet,* 24**,** 408-15.

GILMARTIN, P. M., SAROKIN, L., MEMELINK, J. & CHUA, N.-H. 1990. Molecular Light Switches for Plant Genes. *Plant Cell,* 2**,** 369-378.

GONZALEZ, A., ZHAO, M., LEAVITT, J. M. & LLOYD, A. M. 2008. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant J,* 53**,** 814-27.

GOUJON, M., MCWILLIAM, H., LI, W., VALENTIN, F., SQUIZZATO, S., PAERN, J. & LOPEZ, R. 2010. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucl Acid Res,* 38**,** W695-9.

GUEVARA-GARCIA, A., SAN ROMAN, C., ARROYO, A., CORTES, M. E., GUTIERREZ-NAVA, M. D. & LEON, P. 2005. Characterization of the *Arabidopsis* clb6 mutant illustrates the importance of posttranscriptional regulation of the methyl-D-erythritol 4-phosphate pathway. *Plant Cell,* 17**,** 628-643.

GUPTA, S. K., SHARMA, S., SANTISREE, P., KILAMBI, H. V., APPENROTH, K., SREELAKSHMI, Y. & SHARMA, R. 2014. Complex and shifting interactions of phytochromes regulate fruit development in tomato. *Plant Cell Environ,* 37**,** 1688-702.

HANSEN, B. G., HALKIER, B. A. & KLIEBENSTEIN, D. J. 2008. Identifying the molecular basis of QTLs: eQTLs add a new dimension. *Trends Plant Sci,* 13**,** 72-77.

HARDTKE, C. S., GOHDA, K., OSTERLUND, M. T., OYAMA, T., OKADA, K. & DENG, X. W. 2000. HY5 stability and activity in *Arabidopsis* is regulated by phosphorylation in it COP1 binding domain. *EMBO J* 19**,** 4997-5006.

HARRIS, W. H. & SPURR, A. R. 1969. Chromoplasts of tomato fruits. II. The Red Tomato. *Amer J Bot,* 56**,** 380-389.

HAVAUX, M., EYMERY, F., PORFIROVA, S., REY, P. & DORMANN, P. 2005. Vitamin E protects against photoinhibition and photooxidative stress in Arabidopsis thaliana. *Plant Cell,* 17**,** 3451-69.

HELLENS, R. P., EDWARDS, E. A., LEYLAND, N. R., BEAN, S. & MULLINEAUX, P. M. 2000. pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol Biol,* 42**,** 819-32.

HERNANDEZ, C., MATHIS, A., BROWN, D. J. F. & BOL, J. F. 1995. Sequence of RNA-2 of a Nematode-Transmissible Isolate of Tobacco Rattle Virus. *J Gen Virol,* 76**,** 2847-2851.

HERRMANN, K. M. 1995. The shikimate pathway - early steps in the biosynthesis of aromatic compounds. *Plant Cell,* 7**,** 907-919.

HOFIUS, D., HAJIREZAEI, M. R., GEIGER, M., TSCHIERSCH, H., MELZER, M. & SONNEWALD, U. 2004. RNAi-mediated tocopherol deficiency impairs photoassimilate export in transgenic potato plants. *Plant Physiol,* 135**,** 1256-68.

HORTENSTEINER, S. & KRAUTLER, B. 2011. Chlorophyll breakdown in higher plants. *Biochim Biophys Acta,* 1807**,** 977-988.

HÖRTENSTEINER, S. 2013. *Plastid development in leaves during growth and senescence,* Springer Dordrecht Heidelberg New York London Springer.

Advances in photosynthesis and respiration including bioenergy and related processes. Editors: Govindjee, G. Sharkey, T.D. 363-392.
HOSOMI, A., ARITA, M., SATO, Y., KIYOSE, C., UEDA, T., IGARASHI, O., ARAI, H. & INOUE, K. 1997. Affinity for α-tocopherol transfer protein as a determinant of the biological activities of vitamin E analogs. *FEBS Letters,* 409**,** 105-108.
HUANG, H., GAO, H., LIU, B., FAN, M., WANG, J. J., WANG, C. L., TIAN, H. X., WANG, L. X., XIE, C. Y., WU, D. W., LIU, L. Y., YAN, J. B., QI, T. C. & SONG, S. S. 2018. bHLH13 regulates jasmonate-mediated defense responses and growth. *Evol Bioinf,* 14, 1-8.
HUNTER, S. C. & CAHOON, E. B. 2007. Enhancing vitamin E in oilseeds: unraveling tocopherol and tocotrienol biosynthesis. *Lipids,* 42**,** 97-108.
HUSAINEID, S. S. H., KOK, R. A., SCHREUDER, M. E. L., HANUMAPPA, M., CORDONNIER-PRATT, M. M., PRATT, L. H., VAN DER PLAS, L. H. W. & VAN DER KROL, A. R. 2007. Overexpression of homologous phytochrome genes in tomato: exploring the limits in photoperception. *J Exp Bot,* 58**,** 615-626.
JIANG, Q., CHRISTEN, S., SHIGENAGA, M. & AMES, B. N. 2001. γ-Tocopherol, the major form of vitamin E in the US diet, deserves more attention. *Amer J Clin Nutr,* 74**,** 714-722.
JIANG, Q. 2014. Natural forms of vitamin E: metabolism, antioxidant, and anti-inflammatory activities and their role in disease prevention and therapy. *Free Rad Biol Med,* 72**,** 76-90.
JIANG, W., BIKARD, D., COX, D., ZHANG, F. & MARRAFFINI, L. A. 2013. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol,* 31**,** 233-9.
JIANG, Y., CHEN, B., DUAN, C., SUN, B., YANG, J. & YANG, S. 2015. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl Environ Microbiol,* 81**,** 2506-14.
JIMENEZ-GOMEZ, J. M., WALLACE, A. D. & MALOOF, J. N. 2010. Network Analysis Identifies ELF3 as a QTL for the shade avoidance response in *Arabidopsis*. *PLoS Genet,* 6.
JIMENEZ, A., CREISSEN. G, KULAR, B., FIRMIN, J., ROBINSON, S., VERHOEYEN, M. & MULLINEAUX, P. 2002. Changes in oxidative processes and components of the antioxidant system during tomato fruit ripening. *Planta,* 214**,** 751-758.
JIN, H. L. & MARTIN, C. 1999. Multifunctionality and diversity within the plant MYB-gene family. *Plant Mol Biol,* 41**,** 577-585.
JOHNSEN, M. G., RASMUSSEN, O. F., ALBRECHTSEN, M. & BORKHARDT, B. 1991. In vivo expression of the 29000Mr protein from RNA-2 of pea early. *J Gen Virol,* 72**,** 1223-1227.
JU, J., PICINICH, S. C., YANG, Z. H., ZHAO, Y., SUH, N., KONG, A. N. & YANG, C. S. 2010. Cancer-preventive activities of tocopherols and tocotrienols. *Carcinogenesis,* 31**,** 533-542.
KANWISCHER, M., PORFIROVA, S., BERGMULLER, E. & DORMANN, P. 2005. Alterations in tocopherol cyclase activity in transgenic and mutant plants

of *Arabidopsis* affect tocopherol content, tocopherol composition, and oxidative stress. *Plant Physiol,* 137**,** 713-723.

KARUNANANDAA, B., QI, Q. G., HAO, M., BASZIS, S. R., JENSEN, P. K., WONG, Y. H. H., JIANG, J., VENKATRAMESH, M., GRUYS, K. J., MOSHIRI, F., POST-BEITTERMILLER, D., WEISS, J. D. & VALENTIN, H. E. 2005. Metabolically engineered oilseed crops with enhanced seed tocopherol. *Metab Eng,* 7**,** 384-400.

KHANNA, S., ROY, S., SLIVKA, A., CRAFT, T. K., CHAKI, S., RINK, C., NOTESTINE, M. A., DEVRIES, A. C., PARINANDI, N. L. & SEN, C. K. 2005. Neuroprotective properties of the natural vitamin E alpha-tocotrienol. *Stroke,* 36**,** 2258-64.

KIRSH, V. A., HAYES, R. B., MAYNE, S. T., CHATTERJEE, N., SUBAR, A. F., DIXON, L. B., ALBANES, D., ANDRIOLE, G. L., URBAN, D. A., PETERS, U. & TRIAL, P. 2006. Supplemental and dietary vitamin E, beta-carotene, and vitamin C intakes and prostate cancer risk. *J Nat Cancer Inst,* 98**,** 245-254.

KOENIG, D., JIMENEZ-GOMEZ, J. M., KIMURA, S., FULOP, D., CHITWOOD, D. H., HEADLAND, L. R., KUMAR, R., COVINGTON, M. F., DEVISETTY, U. K., TAT, A. V., TOHGE, T., BOLGER, A., SCHNEEBERGER, K., OSSOWSKI, S., LANZ, C., XIONG, G., TAYLOR-TEEPLES, M., BRADY, S. M., PAULY, M., WEIGEL, D., USADEL, B., FERNIE, A. R., PENG, J., SINHA, N. R. & MALOOF, J. N. 2013. Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. *Proc Natl Acad Sci U S A,* 110**,** E2655-62.

KOKSAL, M., HU, H. Y., COATES, R. M., PETERS, R. J. & CHRISTIANSON, D. W. 2011. Structure and mechanism of the diterpene cyclase ent-copalyl diphosphate synthase. *Nat Chem Biol,* 7**,** 431-433.

KRANZ, H. D., DENEKAMP, M., GRECO, R., JIN, H., LEYVA, A., MEISSNER, R. C., PETRONI, K., URZAINQUI, A., BEVAN, M., MARTIN, C., SMEEKENS, S., TONELLI, C., PAZ-ARES, J. & WEISSHAAR, B. 1998. Towards functional characterisation of the members of the MYB family in *Arabidopsis thaliana*. *Plant J,* 16**,** 263-276.

KRIEGER-LISZKAY, A. & TREBST, A. 2006. Tocopherol is the scavenger of singlet oxygen produced by the triplet states of chlorophyll in the PSII reaction centre. *J Exp Bot,* 57**,** 1677-84.

KRIS-ETHERTON, E. M., HECKER, K. D., BONANOME, A., COVAL, S. M., BINKOSKI, A. E., HILPERT, K. F., GRIEL, A. E. & ETHERTON, T. D. 2002. Bioactive compounds in foods their role in the prevention of cardiovascular disease and cancer. *Amer J Med* 113**,** 71-88.

KRUK, J., HOLLANDER-CZYTKO, H., OETTMEIER, W. & TREBST, A. 2005. Tocopherol as singlet oxygen scavenger in photosystem II. *J Plant Physiol,* 162**,** 749-57.

KUROHA, T., NAGAI, K., KUROKAWA, Y., NAGAMURA, Y., KUSANO, M., YASUI, H., ASHIKARI, M. & FUKUSHIMA, A. 2017. eQTLs regulating transcript variations associated with rapid internode elongation in deepwater rice. *Front Plant Sci,* 8, 1-16.

LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R.,

THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics,* 23**,** 2947-8.

LEE, J. M. & GIOVANNONI, J. Transcriptome profiling of ripe fruit from a *S. lycopersicum* (M82) parent and a set of lines with distinct introgressed *S. pennellii* segments using Illumina RNA-seq analysis. Available: http://ted.bti.cornell.edu/cgi-bin/TFGD/digital/home.cgi [Accessed: 14/10/2014].

LEE, R. H., WANG, C. H., HUANG, L. T. & CHEN, S. C. 2001. Leaf senescence in rice plants: cloning and characterization of senescence up-regulated genes. *J Exp Bot,* 52**,** 1117-21.

LEVY, A. P., GERSTEIN, H. C., MILLER-LOTAN, R., RATHER, R., MCQUEEN, M., LONN, E. & POGUE, J. 2004. The effect of vitamin E supplementation on cardiovascular risk in diabetic individuals with different haptoglobin phenotypes. *Diabetes Care,* 27**,** 2767-2767.

LI, J. 2018. *Dissecting regulatory eQTLs of the carotenoid biosynthetic pathway in tomato fruit using S.lycopersicum x S.pennellii introgression lines.* PhD Thesis Department of Metabolic Biology, John Innes Centre, University of East Anglia.

LI, J. F., AACH, J., NORVILLE, J. E., MCCORMACK, M., ZHANG, D., BUSH, J., CHURCH, G. M. & SHEEN, J. 2013. Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol,* 31**,** 686-8.

LI, Y., ZHOU, Y., WANG, Z., SUN, X. & TANG, K. 2010. Engineering tocopherol biosynthetic pathway in *Arabidopsis* leaves and its effect on antioxidant metabolism. *Plant Sci,* 178**,** 312-320.

LIM, Y. & TRABER, M. G. 2007. Alpha-tocopherol transfer protein (α-TTP) insights from alpha-tocopherol. *Nutr Res Prac,* 1**,** 247-253.

LINKIES, A. & LEUBNER-METZGER, G. 2012. Beyond gibberellins and abscisic acid: how ethylene and jasmonates control seed germination. *Plant Cell Rep,* 31**,** 253-70.

LIPPMAN, Z. B., SEMEL, Y. & ZAMIR, D. 2007. An integrated view of quantitative trait variation using tomato interspecific introgression lines. *Curr Opin Genet Dev,* 17**,** 545-52.

LIRA, B. S., ROSADO, D., ALMEIDA, J., DE SOUZA, A. P., BUCKERIDGE, M. S., PURGATTO, E., GUYER, L., HORTENSTEINER, S., FRESCHI, L. & ROSSI, M. 2016. Pheophytinase knockdown impacts carbon metabolism and nutraceutical content under normal growth conditions in tomato. *Plant Cell Physiol,* 57**,** 642-53.

LIRA, B. S., GRAMEGNA, G., TRENCH, B. A., ALVES, F. R. R., SILVA, E. M., SILVA, G. F. F., THIRUMALAIKUMAR, V. P., LUPI, A. C. D., DEMARCO, D., PURGATTO, E., NOGUEIRA, F. T. S., BALAZADEH, S., FRESCHI, L. & ROSSI, M. 2017. Manipulation of a senescence-associated gene improves fleshy fruit yield. *Plant Physiol,* 175**,** 77-91.

LIU, Y. S., ROOF, S., YE, Z. B., BARRY, C., VAN TUINEN, A., VREBALOV, J., BOWLER, C. & GIOVANNONI, J. 2004. Manipulation of light signal transduction as a

means of modifying fruit nutritional quality in tomato. *Proc Nat Acad Sci USA,* 101**,** 9897-9902.

LOFFREDO, L., PERRI, L., DI CASTELNUOVO, A., IACOVIELLO, L., DE GAETANO, G. & VIOLI, F. 2015. Supplementation with vitamin E alone is associated with reduced myocardial infarction: a meta-analysis. *Nutr Metab Cardiovasc Dis,* 25**,** 354-63.

LU, Y., RIJZAANI, H., KARCHER, D., RUF, S. & BOCK, R. 2013. Efficient metabolic pathway engineering in transgenic tobacco and tomato plastids with synthetic multigene operons. *Proc Natl Acad Sci U S A,* 110**,** E623-32.

LUO, J., BUTELLI, E., HILL, L., PARR, A., NIGGEWEG, R., BAILEY, P., WEISSHAAR, B. & MARTIN, C. 2008. AtMYB12 regulates caffeoyl quinic acid and flavonol synthesis in tomato: expression in fruit results in very high levels of both types of polyphenol. *Plant J,* 56**,** 316-26.

MACH, J. 2015. Phytol from degradation of chlorophyll feeds biosynthesis of tocopherols. *Plant Cell,* 27**,** 2676-2676.

MAEDA, H., SONG, W., SAGE, T. L. & DELLAPENNA, D. 2006. Tocopherols play a crucial role in low-temperature adaptation and phloem loading in *Arabidopsis*. *Plant Cell,* 18**,** 2710-2732.

MAEDA, H. & DELLAPENNA, D. 2007. Tocopherol functions in photosynthetic organisms. *Curr Opin Plant Biol,* 10**,** 260-265.

MAEDA, H., SONG, W., SAGE, T. & DELLAPENNA, D. 2014. Role of callose synthases in transfer cell wall development in tocopherol deficient *Arabidopsis* mutants. *Front Plant Sci,* 5.

MAES, M., AMIT, E., DANIELI, T., LEBENDIKER, M., LOYTER, A. & FRIEDLER, A. 2014. The disordered region of *Arabidopsis* VIP1 binds the *Agrobacterium* VirE2 protein outside its DNA-binding site. *Prot Eng Design Select,* 27**,** 439-446.

MAJEWSKI, J. & PASTINEN, T. 2011. The study of eQTL variations by RNA-seq: from SNPs to phenotypes. *Trends Genet,* 27**,** 72-9.

MAO, Y., ZHANG, H., XU, N., ZHANG, B., GOU, F. & ZHU, J. K. 2013. Application of the CRISPR-Cas system for efficient genome engineering in plants. *Mol Plant,* 6**,** 2008-11.

MARAS, J. E., BERMUDEZ, O. I., QIAO, N., BAKUN, P. J., BOODY-ALTER, E. L. & TUCKER, K. L. 2004. Intake of alpha-tocopherol is limited among US adults. *J Amer Dietetic Assoc,* 104**,** 567-575.

MARIN-RODRIGUEZ, M. C., ORCHARD, J. & SEYMOUR, G. B. 2002. Pectate lyases, cell wall degradation and fruit softening. *J Exp Bot,* 53**,** 2115-2119.

MARRAFFINI, L. A. & SONTHEIMER, E. J. 2008. CRISPR interference limits horizontal gene transfer in *Staphylococci* by targeting DNA. *Science,* 332**,** 1843-1845.

MARTIN, C. & PAZ-AREZ, J. 1997. MYB transcription in plants. *Trends in Genetics,* 13**,** 67-73.

MATAS, A. J., YEATS, T. H., BUDA, G. J., ZHENG, Y., CHATTERJEE, S., TOHGE, T., PONNALA, L., ADATO, A., AHARONI, A., STARK, R., FERNIE, A. R., FEI, Z. J., GIOVANNONI, J. J. & ROSE, J. K. C. 2011. Tissue- and cell-type specific transcriptome profiling of expanding tomato fruit provides insights into

metabolic and regulatory specialization and cuticle formation. *Plant Cell,* 23**,** 3893-3910.

MEHRTENS, F., KRANZ, H., BEDNAREK, P. & WEISSHAAR, B. 2005. The *Arabidopsis* transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiol,* 138**,** 1083-96.

MENE-SAFFRANE, L. & DELLAPENNA, D. 2010. Biosynthesis, regulation and functions of tocochromanols in plants. *Plant Physiol Biochem,* 48**,** 301-9.

MEYDANI, S. N., MEYDANI, M., BLUMBERG, J. B., LEKA, L. S., SIBER, G., LOSZEWSKI, R., THOMPSON, C., PEDROSA, M. C., DIAMOND, R. D. & STOLLAR, B. D. 1997. Vitamin E supplementation and in vivo immune response in healthy elderly subjects. A randomized controlled trial. *JAMA,* 277**,** 1380-6.

MEYDANI, S. N., MEYDANI, M., BLUMBERG, J. B., LEKA, L. S., PEDROSA, M., DIAMOND, R. & SCHAEFER, E. J. 1998. Assessment of the safety of supplementation with different amounts of vitamin E in healthy older adults. *Amer J Clin Nutr,* 68**,** 311-318.

MILMAN, U., BLUM, S., SHAPIRA, C., ARONSON, D., MILLER-LOTAN, R., ANBINDER, Y., ALSHIEK, J., BENNETT, L., KOSTENKO, M., LANDAU, M., KEIDAR, S., LEVY, Y., KHEMLIN, A., RADAN, A. & LEVY, A. P. 2008. Vitamin E supplementation reduces cardiovascular events in a subgroup of middle-aged individuals with both type 2 diabetes mellitus and the haptoglobin 2-2 genotype: a prospective double-blinded clinical trial. *Arterioscler Thromb Vasc Biol,* 28**,** 341-7.

MOCCHEGIANI, E., COSTARELLI, L., GIACCONI, R., MALAVOLTA, M., BASSO, A., PIACENZA, F., OSTAN, R., CEVENINI, E., GONOS, E. S., FRANCESCHI, C. & MONTI, D. 2014. Vitamin E-gene interactions in aging and inflammatory age-related diseases: implications for treatment. A systematic review. *Ageing Res Rev,* 14**,** 81-101.

MOJICA, F. J., DIEZ-VILLASENOR, C., GARCIA-MARTINEZ, J. & SORIA, E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol,* 60**,** 174-82.

MORSE, A. M., WHETTEN, R. W., DUBOS, C. & CAMPBELL, M. M. 2009. Post-translational modification of an R2R3-MYB transcription factor by a MAP Kinase during xylem development. *New Phytol,* 183**,** 1001-13.

MUNNE-BOSCH, S. & ALEGRE, L. 2002. Interplay between ascorbic acid and lipophilic antioxidant defensin chloroplasts of water-stressed *Arabidopsis* plants. *FEBS Lett,* 524**,** 145-148.

MUNNÉ-BOSCH, S. & ALEGRE, L. 2002. The function of tocopherols and tocotrienols in plants. *Crit Rev Plant Sci,* 21**,** 31-57.

NAITO, Y., HINO, K., BONO, H. & UI-TEI, K. 2015. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics,* 31**,** 1120-3.

NAWKAR, G. M., KANG, C. H., MAIBAM, P., PARK, J. H., JUNG, Y. J., CHAE, H. B., CHI, Y. H., JUNG, I. J., KIM, W. Y., YUN, D. J. & LEE, S. Y. 2017. HY5, a positive

regulator of light signaling, negatively controls the unfolded protein response in *Arabidopsis*. *Proc Natl Acad Sci U S A,* 114**,** 2084-2089.

NGUYEN, N. H., JEONG, C. Y., KANG, G. H., YOO, S. D., HONG, S. W. & LEE, H. 2015. MYBD employed by HY5 increases anthocyanin accumulation via repression of MYBL2 in *Arabidopsis*. *Plant J,* 84**,** 1192-1205.

NIH. 2016. *Vitamin E: Fact Sheet for Health Professionals* [Online]. National Institutes for Health. Available: https://ods.od.nih.gov/factsheets/VitaminE-HealthProfessional/ [Accessed 14/09/18 2018].

OFNER, I., LASHBROOKE, J., PLEBAN, T., AHARONI, A. & ZAMIR, D. 2016. Solanum pennellii backcross inbred lines (BILs) link small genomic bins with tomato traits. *Plant J,* 87**,** 151-160.

ORZAEZ, D., MEDINA, A., TORRE, S., FERNANDEZ-MORENO, J. P., RAMBLA, J. L., FERNANDEZ-DEL-CARMEN, A., BUTELLI, E., MARTIN, C. & GRANELL, A. 2009. A visual reporter system for virus-induced gene silencing in tomato fruit based on anthocyanin accumulation. *Plant Physiol,* 150**,** 1122-34.

PACKER, L., WEBER, S. U. & RIMBACH, G. 2001. Molecular Aspects of a-Tocotrienol Antioxidant Action and Cell Signalling. *Amer Soc Nutr Sci,* 131**,** 369S-73S.

PARK, H. A., KUBICKI, N., GNYAWALI, S., CHAN, Y. C., ROY, S., KHANNA, S. & SEN, C. K. 2011. Natural vitamin E alpha-tocotrienol protects against ischemic stroke by induction of multidrug resistance-associated protein 1. *Stroke,* 42**,** 2308-14.

PATEL, R. *Tomato eFP browser* [Online]. [Accessed 12/06/2018 2018].

PELLAUD, S., BORY, A., CHABERT, V., ROMANENS, J., CHAISSE-LEAL, L., DOAN, A. V., FREY, L., GUST, A., FROMM, K. M. & MENE-SAFFRANE, L. 2018. WRINKLED1 and ACYL-COA:DIACYLGLYCEROL ACYLTRANSFERASE1 regulate tocochromanol metabolism in *Arabidopsis*. *New Phytol,* 217**,** 245-260.

PELLEGRINI, N., SERAFINI, M., COLOMBI, B., DEL RIO, D., SALVATORE, S., BIANCHI, M. & BRIGHENTI, F. 2003. Total antioxidant capacity of plant foods, beverages and oils consumed in Italy assessed by three different in vitro assays. *J Nutr,* 133**,** 2812-2819.

PEPPER, A., DELANEY, T., WASHBURN, T., POOLE, D. & CHORY, J. 1994. Det1, a negative regulator of light-mediated development and gene-expression in *Arabidopsis*, encodes a novel nuclear-localized protein. *Cell,* 78**,** 109-116.

PEREZ-FONS, L., WELLS, T., COROL, D. I., WARD, J. L., GERRISH, C., BEALE, M. H., SEYMOUR, G. B., BRAMLEY, P. M. & FRASER, P. D. 2014. A genome-wide metabolomic resource for tomato fruit from *Solanum pennellii*. *Sci Rep,* 4**,** 3859.

POLTURAK, G., GROSSMAN, N., VELA-CORCIA, D., DONG, Y., NUDEL, A., PLINER, M., LEVY, M., ROGACHEV, I. & AHARONI, A. 2017. Engineered gray mold resistance, antioxidant capacity, and pigmentation in betalain-producing crops and ornamentals. *Proc Natl Acad Sci USA,* 114**,** 9062-9067.

PORFIROVA, S., BERGMULLER, E., TROPF, S., LEMKE, R. & DORMANN, P. 2002. Isolation of an *Arabidopsis* mutant lacking vitamin E and identification of a cyclase essential for all tocopherol biosynthesis. *Proc Nat Acad Sci USA,* 99**,** 12495-12500.

POTOKINA, E., DRUKA, A., LUO, Z., WISE, R., WAUGH, R. & KEARSEY, M. 2008. Gene expression quantitative trait locus analysis of 16 000 barley genes reveals a complex pattern of genome-wide transcriptional regulation. *Plant J,* 53**,** 90-101.

POURCEL, C., SALVIGNOL, G. & VERGNAUD, G. 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology,* 151**,** 653-63.

PRABHAKAR, V., LOTTGERT, T., GIGOLASHVILI, T., BELL, K., FLUGGE, U. I. & HAUSLER, R. E. 2009. Molecular and functional characterization of the plastid-localized phosphoenolpyruvate enolase (ENO1) from *Arabidopsis thaliana*. *FEBS Lett,* 583**,** 983-91.

PUAH, C. W., CHOO, Y. M., MA, A. N. & CHUAH, C. H. 2007. The Effect of Physical Refining on Palm Vitamin E (Tocopherol, Tocotrienol and Tocomonoenol) *Amer J App Sci,* 4**,** 374-377.

PUCHTA, H. 2005. The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J Exp Bot,* 56**,** 1-14.

QUADRANA, L., ALMEIDA, J., OTAIZA, S. N., DUFFY, T., CORREA DA SILVA, J. V., DE GODOY, F., ASIS, R., BERMUDEZ, L., FERNIE, A. R., CARRARI, F. & ROSSI, M. 2013. Transcriptional regulation of tocopherol biosynthesis in tomato. *Plant Mol Biol,* 81**,** 309-25.

QUADRANA, L., ALMEIDA, J., ASIS, R., DUFFY, T., DOMINGUEZ, P. G., BERMUDEZ, L., CONTI, G., DA SILVA, J. V. C., PERALTA, I. E., COLOT, V., ASURMENDI, S., FERNIE, A. R., ROSSI, M. & CARRARI, F. 2014. Natural occurring epialleles determine vitamin E accumulation in tomato fruits. *Nat Comms,* 5.

QUIRINO, B. F., NOH, Y. S., HIMELBLAU, E. & AMASINO, R. M. 2000. Molecular aspects of leaf senescence. *Trends Plant Sci,* 5**,** 278-82.

RANJAN, A., BUDKE, J. M., ROWLAND, S. D., CHITWOOD, D. H., KUMAR, R., CARRIEDO, L., ICHIHASHI, Y., ZUMSTEIN, K., MALOOF, J. N. & SINHA, N. R. 2016. eQTL regulating transcript levels associated with diverse biological processes in tomato. *Plant Physiol,* 172**,** 328-40.

REYES, J. C., MURO-PASTOR, M. I. & FLORENCIO, F. J. 2004. The GATA family of transcription factors in *Arabidopsis* and rice. *Plant Physiol,* 134**,** 1718-32.

RICHTER, R., BEHRINGER, C., MULLER, I. K. & SCHWECHHEIMER, C. 2010. The GATA-type transcription factors GNC and GNL/CGA1 repress gibberellin signaling downstream from DELLA proteins and PHYTOCHROME-INTERACTING FACTORS. *Genes Dev,* 24**,** 2093-104.

RIEWE, D., KOOHI, M., LISEC, J., PFEIFFER, M., LIPPMANN, R., SCHMEICHEL, J., WILLMITZER, L. & ALTMANN, T. 2012. A tyrosine aminotransferase involved in tocopherol synthesis in *Arabidopsis*. *Plant J,* 71**,** 850-9.

RIPPERT, P., SCIMEMI, C., DUBALD, M. & MATRINGE, M. 2004. Engineering plant shikimate pathway for production of tocotrienol and improving herbicide resistance. *Plant Physiol,* 134**,** 92-100.

RONEN, G., CARMEL-GOREN, L., ZAMIR, D. & HIRSCHBERG, J. 2000. An alternative pathway to beta -carotene formation in plant chromoplasts discovered by map-based cloning of beta and old-gold color mutations in tomato. *Proc Natl Acad Sci USA,* 97**,** 11102-7.

SAKURABA, Y., PARK, S. Y., KIM, Y. S., WANG, S. H., YOO, S. C., HORTENSTEINER, S. & PAEK, N. C. 2014. *Arabidopsis* STAY-GREEN2 is a negative regulator of chlorophyll degradation during leaf senescence. *Molecular Plant,* 7**,** 1288-1302.

SALGHETTI, S. E., MURATANI, M., WIJNEN, H., FUTCHER, B. & TANSEY, W. P. 2000. Functional overlap of sequences that activate transcription and signal ubiquitin-mediated proteolysis. *Proc Natl Acad Sci USA,* 97**,** 3118–3123.

SALGHETTI, S. E., CAUDY, A. A., CHENOWETH, J. G. & TANSEY, W. P. 2001. Regulation of transcriptional activation domain function by ubiquitin. *Science,* 293**,** 1651-3.

SAPRANAUSKAS, R., GASIUNAS, G., FREMAUX, C., BARRANGOU, R., HORVATH, P. & SIKSNYS, V. 2011. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nuc Acid Res,* 39**,** 9275-82.

SATTLER, S. E. 2003. Characterization of tocopherol cyclases from higher plants and Cyanobacteria. Evolutionary implications for tocopherol synthesis and function. *Plant Physiol,* 132**,** 2184-2195.

SATTLER, S. E., GILLILAND, L. U., MAGALLANES-LUNDBACK, M., POLLARD, M. & DELLAPENNA, D. 2004. Vitamin E is essential for seed longevity and for preventing lipid peroxidation during germination. *Plant Cell,* 16**,** 1419-32.

SATTLER, S. E., MENE-SAFFRANE, L., FARMER, E. E., KRISCHKE, M., MUELLER, M. J. & DELLAPENNA, D. 2006. Nonenzymatic lipid peroxidation reprograms gene expression and activates defense markers in *Arabidopsis* tocopherol-deficient mutants. *Plant Cell,* 18**,** 3706-3720.

SAVIDGE, B., WEISS, J. D., WONG, Y. H. H., LASSNER, M. W., MITSKY, T. A., SHEWMAKER, C. K., POST-BEITTENMILLER, D. & VALENTIN, H. E. 2002. Isolation and characterization of homogentisate phytyltransferase genes from *Synechocystis* sp PCC 6803 and *Arabidopsis*. *Plant Physiology,* 129**,** 321-332.

SCHAUER, N., SEMEL, Y., ROESSNER, U., GUR, A., BALBO, I., CARRARI, F., PLEBAN, T., PEREZ-MELIS, A., BRUEDIGAM, C., KOPKA, J., WILLMITZER, L., ZAMIR, D. & FERNIE, A. R. 2006. Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat Biotech,* 24**,** 447-454.

SCHAUER, N., SEMEL, Y., BALBO, I., STEINFATH, M., REPSILBER, D., SELBIG, J., PLEBAN, T., ZAMIR, D. & FERNIE, A. R. 2008. Mode of inheritance of primary metabolic traits in tomato. *Plant Cell,* 20**,** 509-523.

SCHLEDZ, M., SEIDLER, A., BEYER, P. & NEUHAUS, G. 2001. A novel phytyltransferase from *Synechocystis* sp PCC 6803 involved in tocopherol biosynthesis. *Febs Letters,* 499**,** 15-20.

SCHNEIDER, C. 2005. Chemistry and biology of vitamin E. *Mol Nutr Food Res,* 49**,** 7-30.

SEO, Y. S., KIM, S. J., HARN, C. H. & KIM, W. T. 2011. Ectopic expression of apple fruit homogentisate phytyltransferase gene (MdHPT1) increases tocopherol in transgenic tomato (*Solanum lycopersicum* cv. Micro-Tom) leaves and fruits. *Phytochemistry,* 72**,** 321-9.

SERBINOVA, E., KAGAN, V., HAN, D. & PACKER, L. 1991. Free radical recycling and intramembrane mobility in the antioxidant properties of alpha-tocopherol and alpha-tocotrienol. *Free Radic Biol Med,* 10**,** 263-75.

SHALATA, A. & TAL, M. 1998. The effect of salt stress on lipid peroxidation and antioxidants in the leaf of the cultivated tomato and its wild salt-tolerant relative *Lycopersicon pennellii*. *Physiol Plant,* 104**,** 169-174.

SHAO, Y., ZHU, H. L., TIAN, H. Q., WANG, X. G., LIN, X. J., ZHU, B. Z., XIE, Y. H. & LUO, Y. B. 2008. Virus-induced gene silencing in plant species. *Russian J Plant Physiol,* 55**,** 168-174.

SHIN, J., PARK, E. & CHOI, G. 2007. PIF3 regulates anthocyanin biosynthesis in an HY5-dependent manner with both factors directly binding anthocyanin biosynthetic gene promoters in *Arabidopsis*. *Plant J,* 49**,** 981-94.

SHINOZAKI, Y., NICOLAS, P., FERNANDEZ-POZO, N., MA, Q., EVANICH, D. J., SHI, Y., XU, Y., ZHENG, Y., SNYDER, S. I., MARTIN, L. B. B., RUIZ-MAY, E., THANNHAUSER, T. W., CHEN, K., DOMOZYCH, D. S., CATALA, C., FEI, Z., MUELLER, L. A., GIOVANNONI, J. J. & ROSE, J. K. C. 2018. High-resolution spatiotemporal transcriptome mapping of tomato fruit development and ripening. *Nat Commun,* 9**,** 364.

SHINTANI, D. & DELLAPENNA, D. 1998. Elevating the vitamin E content of plants through metabolic engineering. *Science,* 282**,** 2098-2100.

SHINTANI, D. K., CHENG, Z. G. & DELLAPENNA, D. 2002. The role of 2-methyl-6-phytylbenzoquinone methyltransferase in determining tocopherol composition in *Synechocystis* sp PCC6803. *FEBS Letters,* 511**,** 1-5.

SIEVERS, F., WILM, A., DINEEN, D., GIBSON, T. J., KARPLUS, K., LI, W., LOPEZ, R., MCWILLIAM, H., REMMERT, M., SODING, J., THOMPSON, J. D. & HIGGINS, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol,* 7**,** 539.

SINGH, R. K., ALI, S. A., NATH, P. & SANE, V. A. 2011. Activation of ethylene-responsive p-hydroxyphenylpyruvate dioxygenase leads to increased tocopherol levels during ripening in mango. *J Exp Bot,* 62**,** 3375-85.

SMIRNOFF, N. 2000. Ascorbic acid: metabolism and functions of a multi-facetted molecule. *Curr Opin Plant Biol,* 3**,** 229-235.

SMITH, A. M., COUPLAND, G., DOLAN, L., HARBERD, N., JONES, J., MARTIN, C., SABLOWSKI, R. & AMEY, A. 2009. *Plant Biology*, Garland Science.

SPICHER, L., ALMEIDA, J., GUTBROD, K., PIPITONE, R., DÖRMANN, P., GLAUSER, G., ROSSI, M. & KESSLER, F. 2017. Essential role for phytol kinase and tocopherol in tolerance to combined light and temperature stress in tomato. *J Exp Bot*, 68, 5845-5856.

SPITZER-RIMON, B., MARHEVKA, E., BARKAI, O., MARTON, I., EDELBAUM, O., MASCI, T., PRATHAPANI, N. K., SHKLARMAN, E., OVADIS, M. & VAINSTEIN, A. 2010. EOBII, a gene encoding a flower-specific regulator of phenylpropanoid volatiles' biosynthesis in petunia. *Plant Cell,* 22**,** 1961-76.

SPITZER-RIMON, B., FARHI, M., ALBO, B., CNA'ANI, A., BEN ZVI, M. M., MASCI, T., EDELBAUM, O., YU, Y., SHKLARMAN, E., OVADIS, M. & VAINSTEIN, A. 2012. The R2R3-MYB-like regulatory factor EOBI, acting downstream of EOBII, regulates scent production by activating ODO1 and structural scent-related genes in petunia. *Plant Cell,* 24**,** 5089-105.

STACEY, M. G., CAHOON, R. E., NGUYEN, H. T., CUI, Y., SATO, S., NGUYEN, C. T., PHOKA, N., CLARK, K. M., LIANG, Y., FORRESTER, J., BATEK, J., DO, P. T., SLEPER, D. A., CLEMENTE, T. E., CAHOON, E. B. & STACEY, G. 2016. Identification of homogentisate dioxygenase as a target for vitamin E biofortification in oilseeds. *Plant Physiol,* 172**,** 1506-1518.

STEINHAUSER, M. C., STEINHAUSER, D., KOEHL, K., CARRARI, F., GIBON, Y., FERNIE, A. R. & STITT, M. 2010. Enzyme activity profiles during fruit development in tomato cultivars and *Solanum pennellii*. *Plant Physiol,* 153**,** 80-98.

STRACKE, R., WERBER, M. & WEISSHAAR, B. 2001. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr Opin Plant Biol,* 4**,** 447-456.

SUN, Y., ZHANG, X., WU, C., HE, Y., MA, Y., HOU, H., GUO, X., DU, W., ZHAO, Y. & XIA, L. 2016. Engineering herbicide-resistant rice plants through CRISPR/Cas9-mediated homologous recombination of acetolactate synthase. *Mol Plant,* 9**,** 628-31.

SUZUKI, Y. J., TSUCHIYA, M., WASSAIL, S. R., CHOO, Y. M., GOVIL, G., KAGAN, V. E. & PACKER, L. 1993. Structural and dynamic membrane properties of a-tocopherol and a-tocotrienol: implication to the molecular mechanism of their antioxidant potency. *Biochemistry,* 32**,** 10692-10699.

SVITASHEV, S., SCHWARTZ, C., LENDERTS, B., YOUNG, J. K. & MARK CIGAN, A. 2016. Genome editing in maize directed by CRISPR-Cas9 ribonucleoprotein complexes. *Nat Commun,* 7**,** 13274.

SWANSON-WAGNER, R. A., DECOOK, R., JIA, Y., BANCROFT, T., JI, T. M., ZHAO, X. F., NETTLETON, D. & SCHNABLE, P. S. 2009. Paternal dominance of trans-eQTL influences gene expression patterns in maize hybrids. *Science,* 326**,** 1118-1120.

TAKATSUJI, H., MORI, M., BENFEY, P. N., REN, L. & CHUA, N. H. 1992. Characterization of a zinc finger DNA-binding protein expressed specifically in petunia petals and seedlings. *EMBO J,* 11**,** 241-9.

TOLEDO-ORTIZ, G., HUQ, E. & RODRIGUEZ-CONCEPCION, M. 2010. Direct regulation of phytoene synthase gene expression and carotenoid

biosynthesis by phytochrome-interacting factors. *Proc Nat Acad Sci USA,* 107**,** 11626-11631.

TOLEDO-ORTIZ, G., JOHANSSON, H., LEE, K. P., BOU-TORRENT, J., STEWART, K., STEEL, G., RODRIGUEZ-CONCEPCION, M. & HALLIDAY, K. J. 2014. The HY5-PIF regulatory module coordinates light and temperature control of photosynthetic gene transcription. *PLoS Genet,* 10**,** e1004416.

TRABER, M. G., ELSNER, A. & BRIGELIUS-FLOHE, R. 1998. Synthetic as compared with natural vitamin E is preferentially excreted as K-CEHC in human urine studies using deuterated K-tocopheryl acetates. *FEBS Letters,* 437**,** 145-148.

TRABER, M. G. & ATKINSON, J. 2007. Vitamin E, antioxidant and nothing more. *Free Radic Biol Med,* 43**,** 4-15.

TREBST, A., DEPKA, B. & HOLLANDER-CZYTKO, H. 2002. A specific role for tocopherol and of chemical singlet oxygen quenchers in the maintenance of photosystem II structure and function in *Chlamydomonas reinhardtii*. *FEBS Lett,* 516**,** 156-60.

TREBST, A. 2003. Function of beta-carotene and tocopherol in photosystem II. *Z Naturforsch C,* 58**,** 609-20.

TSEGAYE, Y., SHINTANI, D. & DELLAPENNA, D. 2002. Overexpression of the enzyme p-hydroxyphenolpyruvate dioxygenase in Arabidopsis and its relation to tocopherol biosynthesis. *Plant Physiol Biochem,* 40**,** 913-920.

TSUGAMA, D., LIU, S. K. & TAKANO, T. 2016. The bZIP protein VIP1 is involved in touch responses in *Arabidopsis* roots. *Plant Physiol,* 171**,** 1355-1365.

TZIN, V., MALITSKY, S., AHARONI, A. & GALILI, G. 2009. Expression of a bacterial bi-functional chorismate mutase/prephenate dehydratase modulates primary and secondary metabolism associated with aromatic amino acids in *Arabidopsis*. *Plant J,* 60**,** 156-67.

TZIN, V. & GALILI, G. 2010a. The biosynthetic pathways for shikimate and aromatic amino acids in *Arabidopsis thaliana*. *Arabidopsis Book,* 8**,** e0132.

TZIN, V. & GALILI, G. 2010b. New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. *Mol Plant,* 3**,** 956-72.

TZIN, V., MALITSKY, S., BEN ZVI, M. M., BEDAIR, M., SUMNER, L., AHARONI, A. & GALILI, G. 2012. Expression of a bacterial feedback-insensitive 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase of the shikimate pathway in *Arabidopsis* elucidates potential metabolic bottlenecks between primary and secondary metabolism. *New Phytol,* 194**,** 430-9.

U.S. DEPARTMENT OF AGRICULTURE, A. R. S. 2018. USDA national nutrient database for standard reference, release. *version 3.9.5.* Nutrient Data Laboratory Home Page. Available: http://www.ars.usda.gov/nutrientdata [Accessed: 03/09/18]

VALENTIN, H. E., LINCOLN, K., MOSHIRI, F., JENSEN, P. K., QI, Q., VENKATESH, T. V., KARUNANANDAA, B., BASZIS, S. R., NORRIS, S. R., SAVIDGE, B., GRUYS, K. J. & LAST, R. L. 2006. The *Arabidopsis* vitamin E pathway gene5-1 mutant

reveals a critical role for phytol kinase in seed tocopherol biosynthesis. *Plant Cell,* 18**,** 212-24.

VAN EENENNAAM, A. L., LINCOLN, K., DURRETT, T. P., VALENTIN, H. E., SHEWMAKER, C. K., THORNE, G. M., JIANG, J., BASZIS, S. R., LEVERING, C. K., AASEN, E. D., HAO, M., STEIN, J. C., NORRIS, S. R. & LAST, R. L. 2003. Engineering vitamin E content: From *Arabidopsis* mutant to soy oil. *Plant Cell,* 15**,** 3007-3019.

VAN MOERKERCKE, A., HARING, M. A. & SCHUURINK, R. C. 2011. The transcription factor EMISSION OF BENZENOIDS II activates the MYB ODORANT1 promoter at a MYB binding site specific for fragrant petunias. *Plant J,* 67**,** 917-28.

VAN SANDT, V. S., SUSLOV, D., VERBELEN, J. P. & VISSENBERG, K. 2007. Xyloglucan endotransglucosylase activity loosens a plant cell wall. *Ann Bot,* 100**,** 1467-73.

VAUCHERET, H., BECLIN, C. & FAGARD, M. 2001. Post-transcriptional gene silencing in plants. *J Cell Sci,* 114**,** 3083-3091.

VENKATESWARAN, V., FLESHNER, N. E., SUGAR, L. M. & KLOTZ, L. H. 2004. Antioxidants block prostate cancer in lady transgenic mice. *Cancer Research,* 64**,** 5591-5596.

VERDONK, J. C., HARING, M. A., VAN TUNEN, A. J. & SCHUURINK, R. C. 2005. ODORANT1 regulates fragrance biosynthesis in petunia flowers. *Plant Cell,* 17**,** 1612-1624.

VOM DORP, K., HOLZL, G., PLOHMANN, C., EISENHUT, M., ABRAHAM, M., WEBER, A. P. M., HANSON, A. D. & DORMANN, P. 2015. Remobilization of phytol from chlorophyll degradation is essential for tocopherol synthesis and growth of *Arabidopsis*. *Plant Cell,* 27**,** 2846-2859.

WAN, C., LI, C., MA, X., WANG, Y., SUN, C., HUANG, R., ZHONG, P., GAO, Z., CHEN, D., XU, Z., ZHU, J., GAO, X., WANG, P. & DENG, X. 2015. GRY79 encoding a putative metallo-beta-lactamase-trihelix chimera is involved in chloroplast development at early seedling stage of rice. *Plant Cell Rep,* 34**,** 1353-63.

WANG, H., HONG, J. & YANG, C. S. 2015. delta-Tocopherol inhibits receptor tyrosine kinase-induced AKT activation in prostate cancer cells. *Mol Carcinog,* 55, 1728-1738.

WANG, L., LI, Q., ZHANG, A., ZHOU, W., JIANG, R., YANG, Z., YANG, H., QIN, X., DING, S., LU, Q., WEN, X. & LU, C. 2017. The phytol phosphorylation pathway is essential for the biosynthesis of phylloquinone, which is required for Photosystem I stability in *Arabidopsis*. *Mol Plant,* 10**,** 183-196.

WANG, X., OUYANG, Y., LIU, J., ZHU, M., ZHAO, G., BAO, W. & HU, F. B. 2014. Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of prospective cohort studies. *BMJ,* 349**,** g4490.

WANG, X. F., CHEN, Q. Y., WU, Y. Y., LEMMON, Z. H., XU, G. H., HUANG, C., LIANG, Y. M., XU, D. Y., LI, D., DOEBLEY, J. F. & TIAN, F. 2018. Genome-wide analysis

of transcriptional variability in a large maize-teosinte population. *Mol Plant,* 11**,** 443-459.

WATERHOUSE, P. M., WANG, M. B. & LOUGH, T. 2001. Gene silencing as an adaptive defence against viruses. *Nature,* 411**,** 834-842.

WATERS, M. T., MOYLAN, E. C. & LANGDALE, J. A. 2008. GLK transcription factors regulate chloroplast development in a cell-autonomous manner. *Plant J,* 56**,** 432-44.

WATERS, M. T., WANG, P., KORKARIC, M., CAPPER, R. G., SAUNDERS, N. J. & LANGDALE, J. A. 2009. GLK transcription factors coordinate expression of the photosynthetic apparatus in *Arabidopsis*. *Plant Cell,* 21**,** 1109-28.

WENTZELL, A. M., ROWE, H. C., HANSEN, B. G., TICCONI, C., HALKIER, B. A. & KLIEBENSTEIN, D. J. 2007. Linking metabolic QTLs with network and cis-eQTLs controlling biosynthetic pathways. *PLoS,* 3**,** 1687-1701.

WEST, M. A. L., KIM, K., KLIEBENSTEIN, D. J., VAN LEEUWEN, H., MICHELMORE, R. W., DOERGE, R. W. & CLAIR, D. A. S. 2007. Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in *Arabidopsis*. *Genetics,* 175**,** 1441-1450.

WILLMANN, M. R., MEHALICK, A. J., PACKER, R. L. & JENIK, P. D. 2011. MicroRNAs regulate the timing of embryo maturation in *Arabidopsis*. *Plant Physiol,* 155**,** 1871-1884.

XIE, K. & YANG, Y. 2013. RNA-guided genome editing in plants using a CRISPR-Cas system. *Mol Plant,* 6**,** 1975-83.

XU, D. B., CHEN, M., MA, Y. N., XU, Z. S., LI, L. C., CHEN, Y. F. & MA, Y. Z. 2015. A G-protein beta subunit, AGB1, negatively regulates the ABA response and drought tolerance by down-regulating AtMPK6-related pathway in *Arabidopsis*. *PLoS One,* 10**,** e0116385.

YANAGISAWA, S. 1998. Transcription factors in plants: Physiological functions and regulation of expression. *J Plant Res,* 111**,** 363-371.

YANG, W., CAHOON, R. E., HUNTER, S. C., ZHANG, C., HAN, J., BORGSCHULTE, T. & CAHOON, E. B. 2011. Vitamin E biosynthesis: functional characterization of the monocot homogentisate geranylgeranyl transferase. *Plant J,* 65**,** 206-17.

YAO, C., JOEHANES, R., JOHNSON, A. D., HUAN, T. X., LIU, C. Y., FREEDMAN, J. E., MUNSON, P. J., HILL, D. E., VIDAL, M. & LEVY, D. 2017. Dynamic role of trans regulation of gene expression in relation to complex traits. *Amer J Hum Genet,* 100**,** 571-580.

YASUMURA, Y., MOYLAN, E. C. & LANGDALE, J. A. 2005. A conserved transcription factor mediates nuclear control of organelle biogenesis in anciently diverged land plants. *Plant Cell,* 17**,** 1894-907.

YE, J., COULOURIS, G., ZARETSKAYA, I., CUTCUTACHE, I., ROZEN, S. & MADDEN, T. L. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinf,* 13**,** 134.

YE, J., HU, T., YANG, C., LI, H., YANG, M., IJAZ, R., YE, Z. & ZHANG, Y. 2015. Transcriptome profiling of tomato fruit development reveals transcription

factors associated with ascorbic acid, carotenoid and flavonoid biosynthesis. *PLoS One,* 10**,** e0130885.

ZHANG, C., CAHOON, R. E., HUNTER, S. C., CHEN, M., HAN, J. & CAHOON, E. B. 2013. Genetic and biochemical basis for alternative routes of tocotrienol biosynthesis for enhanced vitamin E antioxidant production. *Plant J,* 73**,** 628-39.

ZHANG, Y., BUTELLI, E., ALSEEKH, S., TOHGE, T., RALLAPALLI, G., LUO, J., KAWAR, P. G., HILL, L., SANTINO, A., FERNIE, A. R. & MARTIN, C. 2015. Multi-level engineering facilitates the production of phenylpropanoid compounds in tomato. *Nat Commun,* 6**,** 8635.

ZHAO, L., LU, L., ZHANG, L., WANG, A., WANG, N. & TANG, K. X. 2009. Molecular evolution of the E8 promoter in tomato and some of its relative wild species. *J Biosci,* 34**,** 71-83.

ZUBIETA, C., KOTA, P., FERRER, J. L., DIXON, R. A. & NOEL, J. P. 2002. Structural basis for the modulation of lignin monomer methylation by caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase. *Plant Cell,* 14**,** 1265-77.

ZULUAGA, D. L., GONZALI, S., LORETI, E., PUCCIARIELLO, C., DEGL'INNOCENTI, E., GUIDI, L., ALPI, A. & PERATA, P. 2008. *Arabidopsis thaliana* MYB75/PAP1 transcription factor induces anthocyanin production in transgenic tomato plants. *FunctPlant Biol,* 35**,** 606-618.