Silencing Cell Wall Remodeling Enzymes in Tomato and Analysis of the Effects on Fruit Softening

Nurul Hidayah binti Samsulrizal, MSc

Thesis submitted to the University of Nottingham in fulfilment of the requirements for the degree of Doctor of Philosophy

Division of Plants and Crop Sciences School of Biosciences, University of Nottingham Sutton Bonington Campus, Loughborough, LE12 5R

October 2016

ABSTRACT

Fruits are the seed dispersal units of flowering plants and fleshy fruits form an important part of the human diet. Ripening is a complex developmental process and involves many events such as textural and constitutional changes. The texture of fleshy fruits is one of the major criteria for consumer choice. However, the molecular determinants of ripening- associated changes in texture or "softening" are relatively poorly understood and seem to involve a large number of cell wall remodelling factors. The recent completion of the tomato genome sequence has revealed more than 50 cell wall structure-related genes that are expressed during fruit development and ripening and may impact texture changes in this fruit (Tomato Genome Consortium, 2012). The aim of the project was to compare, on a genome-wide scale, ripening-related gene expression in a range of fleshy fruits and especially those linked with cell wall remodelling. Then by identifying orthologous genes in different fruit species, to make predictions about those genes important for the softening process in all fleshy fruits. We hypothesise that there are a core set of genes that is always associated with fruit softening. Once this core set was identified we planned to use transgenic approaches to test the role of these cell wall genes in texture changes using tomato (*Solanum lycopersicum*) as a model system. Comparative genomics analysis of tomato (*S. lycopersicum*), banana (*Musa acuminata*), melon (*Cucumis melo*) and grape (*Vitis vinifera*), was undertaken using Inparanoid software. This analysis showed that a total of 8,982 (25.86%) gene models could be identified in common between all four genomes based on comparison of amino acid sequences. However, comparison of the expression patterns of cell wall remodelling genes across the various species revealed that most were expressed in tissues other than ripening fruits and of the fruit expressed genes only a small number were common between different species. We selected a *CELLULOSE SYNTHASE* (*CESA*) gene that was up-regulated in all fruit species for further analysis. The *PHYTOENE SYNTHASE* 1 (*PSY1*) was selected as a control, and *PECTATE LYASE* (*PL*) and *POLYGALACTURONASE* (*PG*) as genes that were highly expressed in many fruits including tomato. We then used the recently developed CRISPR/ Cas9 DNA editing technology to generate mutations in the target *PSY1*, *CESA*, *PL* and *PG* genes. A range of either single base insertions or deletions of more than ten bases were identified at Cas9 cleavage site in T_0 plants in all cases except for the *CESA* lines where no mutant plants were recovered; suggesting that knocking out this gene gives a lethal phenotype. The efficacy of CRISPR/Cas9 constructs in tomato fruits was readily apparent when observing the phenotype in *PSY1* line where yellow or orange fruits occurred in T_0 individuals. In the *PL* lines, substantially firmer tomato fruits were produced at the red ripe stage. There were some much smaller impacts on texture in the *PG* CRISPR lines. The CRISPR approaches described in this thesis provide methods to enhance texture and extend shelf life in tomato without compromising fruit quality.

ACKNOWLEDGEMENTS

First of all I would like to express my special thanks to my supervisor Professor Graham Seymour for the guidance and invaluable help throughout my Ph.D. He has set an example of excellence as a researcher, mentor, instructor, and role model. I really appreciate all his contributions of time, ideas and patience. I also very grateful to have Dr Rupert Fray as my co-supervisor who help me a lot during generate the CRISPR construct and advice in all techniques.

I would like to give thanks to all of my colleagues in Plant Science especially my colleagues in lab A49 past and present, especially Cheng Yan who help me a lot during vector construction for *PSYI* and *PG*. You all have given me some great moments during my study. I also would like to acknowledge Jonathan Jones, Vladimir Nekrasov and Sophien Kamoun, TSL and The Gatsby Charitable Foundation for provision of the CRISPR/Cas 9 vectors and The Ministry of Higher Education including International Islamic University of Malaysia for providing funding for my PhD.

My deepest gratitude to my family, especially to my beloved husband Haneef Zulkifle who has been my physical and emotional support during this periods, my dearest sons Afeef and Irfan, my parents Samsulrizal, Siti Aishah, my in-laws Zulkifle and Ziayah. Also, thanks to all my lovely Malay friends Maizatul, Shima, Fahmi, Nizaha, Zarani, Asilah, Engku Elini, Aini, Aishah, Syamila, Izzi and many more. Without their encouragement, understanding and love, it would be impossible to finish my study.

-Nurul Hidayah, Nottingham 2016-

TABLE OF CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

ABA	Abscisic acid
ACC	1-aminocyclopropane-1-carboxylic acid
ACS	1-aminocyclopropane-1-carboxylate synthase
Br+	breaker plus
Ca^{2+}	calcium ions
cDNA	complementary DNA
CTR	constitutive triple response
CWM	cell wall material
dH_2O	dionised water
ELF	elongation factor
ER	endoplasmic reticulum
ERF	ethylene-response factor
EST	expressed sequences tag
HEMA	2-hydroxyethyl methacrylate
HG	homoglacturonan
IAA	Indole-3-acetic acid
LR	London resin
MG	mature green
M-ML V	Moloney murine leukemia virus
NIL	near isogenic lines
PCR	polymerase chain reaction
PE	pectin esterase
PG	polygaclaturonase
PL	pectate lyase
PME	pectin methylesterase
PSY	phytoene synthase
RNAi	RNA interference
TBG	beta-galactosidase
QTL	quantitative trait loci

Rheological terms

G^*	complex modulus
η^*	complex viscosity
G''	loss modulus
γ	shear strain
τ	shear stress
η	shear viscosity
G'	storage modulus
$\gamma_{(t)}$	time dependent shear strain
$\tau_{(t)}$	time dependent shear stress

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction and Literature Review

CHAPTER 1: INTRODUCTION

1.1 Fleshy fruits

The structure known as a "fruit" is found only in the members of the Angiosperms. The botanical definition of a "fruit" is a ripened ovary, an organ that contains seeds, protecting these as they develop and often aiding in their dispersal (Van der Pijl, 1982). Fruits can be classified into two major groups which are (1) fleshy, in which the pericarp and accessory parts develop into succulent tissues and (2) dry, in which the entire pericarp becomes dry at maturity. Fleshy fruits include the berries such as tomato and banana where the seeds are surrounded by a fleshy mass of pericarp, drupes such as peach and apricot which have a stony endocarp and pomes such as apple where accessory tissue is the edible material. Fleshy fruits may also be aggregate multiple fruits, such as raspberries. Dry fruits can be indehiscent, e.g. nuts or dehiscent e.g capsules which at maturity open by pores or split to shed the contained seeds (Spujt, 1994). Some of the most important crop products are dry fruits, e.g. the caryopsis of cereals.

Fleshy fruits are an important part of a healthy diet. They provide essential vitamins and minerals, Ca, Fe and K, and their consumption is associated with a reduced risk of chronic diseases (Seymour et al, 2013). The awareness of the importance of fruits in our diet has driven the increase in demand for these crop products, but they are highly perishable

and have a short shelf-life. Many fleshy fruits are major food crops that have great economic importance. In 2013, the Food and Agriculture Organization of the United Nations reported the top five traded fruit products which are banana, tomato, apple, grape and orange (Table 1-1).

Table 1-1 The global production of five major (million tonnes) fruit commodities in 2012-2013. Data from FAOSTAT, Food and Agriculture Organization of the United Nations (2013)

Commodity	Production	Five biggest producers
Tomato	163.4	China, India, United States, Turkey, Egypt
Banana	139.2	India, Chad, Uganda, Philippines, Ecuador
Apple	80.8	China, United States, Turkey, Poland, India
Grape	67.1	China, United States, Italy, France, Spain
Orange	96.1	Brazil, United States, China, India, Mexico

1.1.1 Ripening Behaviour

Physiologically, fleshy fruit can be divided into two categories which are climacteric and non-climacteric. Climacteric fruit are those that show a burst of respiration at the onset of ripening and where ripening can be induced by ethylene. Well known examples of climacteric fruits include bananas, mangoes, papayas, and tomatoes. Non-climacteric fruit fail to show a burst of respiration or ethylene biosynthesis and include grapes, citrus and strawberries (Rhodes, 1971; Giovannoni, 2004; Barry and Giovannoni, 2007). This review will focus on climacteric fruit and especially tomato which is the model for studying ripening in fleshy fruits, specifically action of ethylene and the regulatory networks linked to ripening changes particularly texture.

1.2 Fruit development and ripening

1.2.1 Ethylene biosynthesis

Ethylene is a small hydrocarbon gas and associated with numerous plant responses and developmental processes in plants, including fruit ripening, root initiation, floral development, seed germination, senescence, and responses to numerous biotic and abiotic stresses (Gapper and Giovannoni, 2013). The regulation of ethylene biosynthesis consists of three steps. Firstly, L-methionine is converted to S-adenosyl methionine (SAM) by SAM synthetase. Then, ACC synthase (ACS) catalyses the

conversion of SAM to 1-aminocyclopropane-1-carboxylic acid (ACC). Finally, ACC is converted to ethylene by ACC oxidase (ACO) (Figure 1- 1) (Hamilton et al., 1990).

There are two systems that control the production of ethylene in plants, System 1 functions during vegetative and fruit development, whereas System 2 occurs during fruit ripening and floral senescence Figure 1-1B. In tomato, *LeACO* is encoded by multiple genes. Four of them have been well characterised (Tomato Genome Consortium 2012). During ethylene biosynthesis prior to ripening (System 1) the expression levels of the *LeACO* 1, 3 and 4 genes are low in mature green fruits and remain relatively constant. At the onset of ripening (System 2) the expression of *LeACO* 1 and 4 increases and is maintained, whereas *LeACO* 3 expression level increases transiently. The System 2 ethylene biosynthesis is autocatalytic and *LeACO* 1 and 4 are positively regulated by ethylene (Barry and Giovannoni, 2007). In 1990, Hamilton et al. reported that by silencing the expression of *LeACO* 1 normal ripening in tomato can be inhibited and this is also found to be the case when orthologous genes were silenced in melon (Ayub et al. 1996), apple (Tacken et al., 2010), papaya (Atkinson et al., 2011) and kiwi fruit (Che-Radziah et al., 2011).

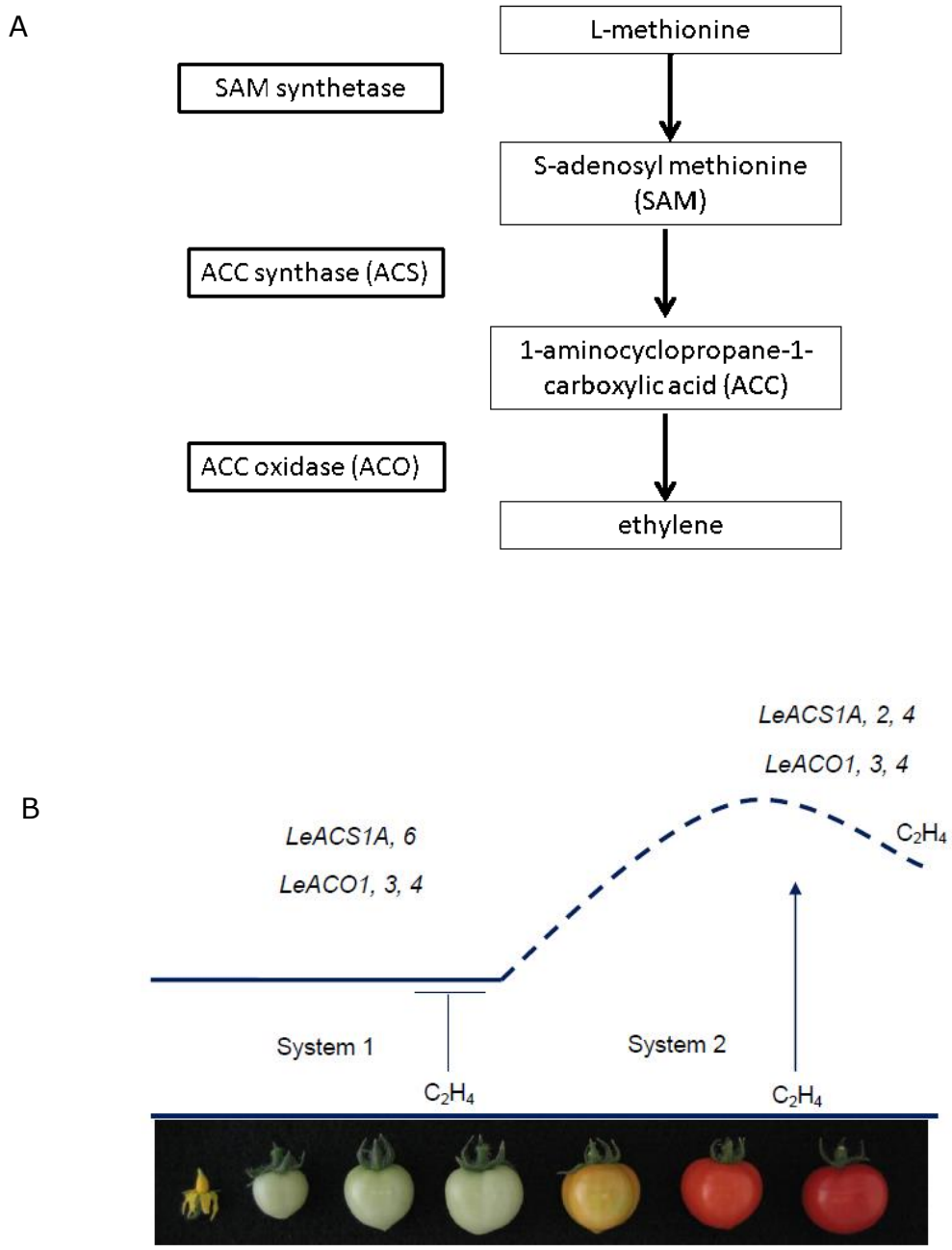


Figure 1-1 (A) The pathway of ethylene biosynthesis (redrawn from Gapper et al., 2013). (B) The differences of gene expression of *ACS* and *ACO* in System 1 and System 2 ethylene biosynthesis during fruit development and ripening in tomato. System 1 is auto-inhibitory (T) cannot initiate ripening and the System 2 is an autocatalytic process (↑) that will induce the ripening reaction. Image taken from Barry and Giovannoni, 2007.

In mature green fruit the expression level of *LeACS* 6 increases, but during fruit ripening it then decrease dramatically and this pattern of gene expression is linked to the regulation in system 1 ethylene synthesis in tomato fruit (Barry et al., 2000). *LeACS* 6 then has modest expression during the ripening phase (Gapper et al., 2013). However, *LeACS* 1A is expressed in mature fruits, but the expression level increases transiently at the onset of fruit ripening. Therefore, *LeACS* 1A seems to be involved in the regulation of ethylene synthesis during the transition from System 1 to System 2. In contrast, LeACS 2 and *LeACS* 4 expression levels are low in mature green fruit. However, the expression levels of *LeACS* 2 and *LeACS* 4 then increase dramatically during ripening and this indicates that autocatalytic System 2 ethylene synthesis is maintained by the combined expression of *LeACS* 2 and *LeACS* 4 (Barry and Giovannoni, 2007).

1.2.2 Ethylene receptors

Ethylene action is associated with ethylene receptors located in the endoplasmic reticulum (ER) membranes. The receptors are linked to a signal cascade that results in altered transcription of ethylene-mediated genes. In tomato, there are seven ethylene receptor genes (*LeETR1, LeETR2, LeETR3/Nr* (Never ripe), *LeETR4, LeETR5*, *LeETR6* and *LeETR7*) (Klee and Giovannoni, 2011) and five in *Arabidopsis* (Klee 2004). However, all these have similar ethylene binding affinities (O'Malley et al., 2005). Generally ethylene receptors contain three domains (1) a sensor domain that contains three or four hydrophobic regions, (2) a signalling domain also known as histidine (His) kinase and (3) a response regulator domain (Tieman and Klee, 1999).

ETRs can be divided into two subfamilies based on the protein structure. Subfamily-1 (LeETR1, LeETR2 and LeETR3/Nr) share the three N-terminal membrane-spanning domains and a conserved carboxy terminus His kinase domain, whereas Subfamily-2 (LeETR4, LeETR5 and LeETR6) contain a small number of a complete His-kinase domain at the carboxy terminus and possess an additional transmembrane spanning domain at the N terminus (Barry and Giovannoni, 2007). The transcript levels of *LeETR1*, *LeETR2* and *LeETR5* change little following exogenous application of ethylene to fruit in tomato, whereas LeETR3/Nr, LeETR4 and LeETR6 mRNA levels increased during ripening (Kevany et al., 2007).

Silencing of either *LeETR4* or *LeETR6* with a fruit specific promoter caused a slow ripening phenotype (Kevany et al., 2008).

The ethylene receptors act as homo and/or heterodimers which are linked by a disulfide bond at the amino terminal region (Figure 1-2). For binding ethylene, copper is used as a cofactor. The receptors interact with other proteins that are part of the ethylene signalling pathway. For example, CONSTITUTIVE TRIPLE RESPONSE 1 (CTR1) acts as negative regulator in the pathway of ethylene transduction. CTR1 has homology to mitogen-activated protein kinase kinase kinase (MAPKKK). The ethylene receptor is thought to activate CTR1 leading to inhibition of the ethylene response genes in the absence of ethylene. Binding of ethylene to the receptors deactivates the receptor and results in the loss of CTR1 activity. The inactivation of CTR1 then leads to induction of ETHYLENE-INSENSITIVE 2 (*EIN2*) and the signal is transduced to EIN-3-LIKE TRANSCRIPTION FACTORS (*EILs*). However, the mechanism of *EIN2* action still remains unclear. The *EIL* proteins bind to the promoters of ethylene response genes including *ETHYLENE RESPONSE FACTORS* (*ERFs*) to activate their transcription (Alexander and Grierson, 2002, Vandenbussche *et al*., 2007; Barry and Giovannoni, 2007; Klee and Giovannoni, 2011).

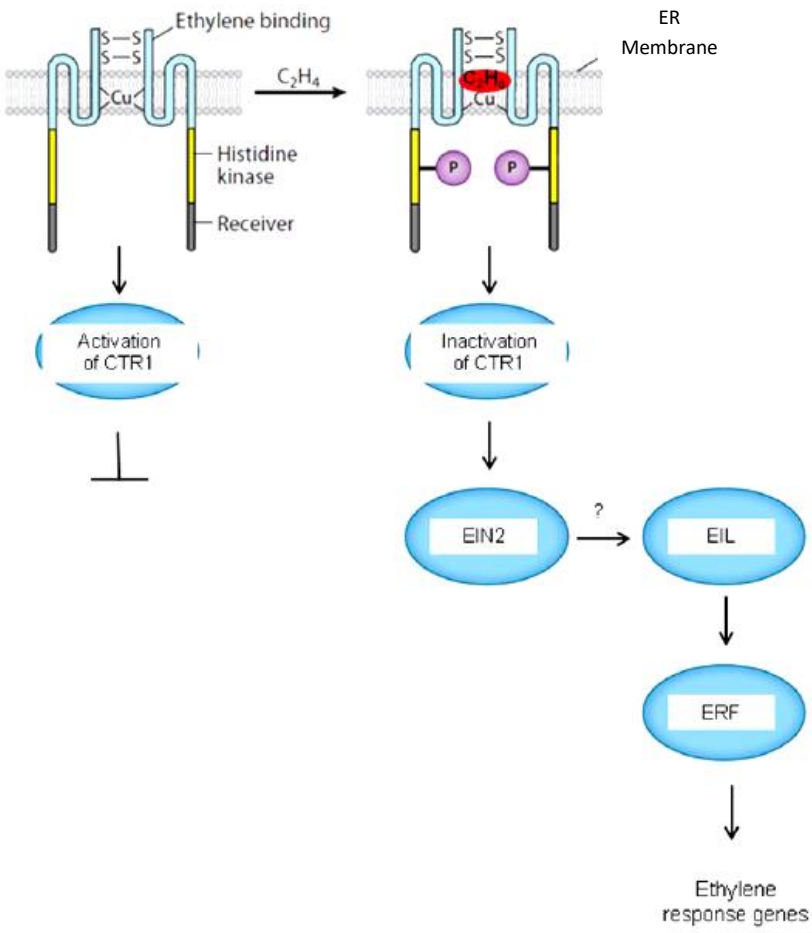


Figure 1-2 Model of ethylene signalling pathway. The ethylene receptor binds the endoplasmic reticulum (ER) membrane. The ethylene binding domain of ETR1 is shown as a homodimer stabilized with two disulfide bonds at the N-terminus. The copper acts as a cofactor for binding ethylene. The binding of ethylene to the ethylene receptor results in ethylene signal transduction which leads to diminished activity of the negative regulator, CTR1, causing release of inhibition to downstream signaling ETHYLENE-INSENSITIVE 2 (EIN2), EIN-3-LIKE TRANSCRIPTION FACTORS (EILs) and ETHYLENE RESPONSE FACTORS (ERFs) (from Klee and Giovannoni, 2011).

1.2.3 Transcriptional Regulation of Ripening

The ethylene signal transduction pathway works in concert with additional transcriptional networks which are linked to developmental cues. The precise molecular framework is yet to be elucidated although it almost certainly involves other hormones such as auxin and ABA driving the expression of developmental regulators. The hormone signalling network is intimately linked to the control of transcription.

In tomato (*Solanum lycopersicum*), there are a small number of rare single gene mutations that abolish normal ripening and these include *ripening inhibitor* (*rin*), *non-ripening* (*nor*) and *Colourless non-ripening* (*Cnr*). In the past 20 years, the genes underlying these mutations have all been cloned. *RIPENING INHIBITOR* (RIN; also called MADS-RIN), *NON-RIPENING* (NOR; also called NAC-NOR), and *COLORLESS NON-RIPENING* (CNR; also called SPL-CNR) have been identified as key transcription factor genes for the regulation of fruit ripening (Vrebalov et al., 2002; Manning et al., 2006; Giovannoni, 2007).

Ripening-Inhibitor (RIN)

RIN belongs to the *SEPALLATA* subfamily of *MADS* box TFs and is strongly expressed in a ripening-specific manner (Vrebalov et al., 2002). The *rin* mutation was identified as a deletion stretching over a part of the RIN gene and the intergenic region between RIN and the adjacent *MADS* box

gene *MACROCALYX* (also called *MADS-MC*) (Vrebalov et al., 2002). The *rin* mutant produces non-ripening fruit where the expression of numerous ripening-related genes is inhibited (Osorio et al., 2011). In normal ripening in tomato the RIN protein interacts directly with the promoters of many other regulatory genes and downstream effectors such as NOR, HB1, CNR and TDR4 and the ethylene biosynthetic genes *ACS2* and *ACS4*, ethylene receptors (*LeETR3*), ethylene response genes (*E4* and *E8*), cell wall modified genes (*PG* and *EXP1*) and carotenoid biosynthetic genes such as *PSY1* (Martel et al., 2011).

The *rin* mutation confers a slow ripening phenotype in the heterozygous condition and is commonly used in commercial lines to achieve extended shelf-life. However, the fruits bearing only one working copy of *rin* often fail to develop their full flavour or colour (Fujisawa et al., 2014). In strawberry, MADS-box genes (*FaMADS9*) were identified homolog with *RIN-MADS* gene in tomato (Vrebalov et al., 2002) and silencing of *FaMADS9* will inhibit the normal development and ripening in petal achene and receptacle tissue (Seymour et al., 2011). Recently, Elitzur et al. (2016) has also reported the importance of this MADS box genes in banana. There are two types of MADS box genes which are homologous with RIN-MADS in tomato, MaMADS1 and MaMADS2 and repressing this gene will delay the ripening process as well extended the colour development and softening event (Elitzur et al. 2016). These indicate that transcription regulators are conserved in both climacteric and non-climacteric species.

Non-ripening (NOR)

Mutations in *NOR* can inhibit the production of ethylene and fruit fail to ripen normally (Giovannoni, 2004 and Barry and Giovannoni, 2007). The gene at the *NOR* locus has been cloned and shown to be a *NAC* transcription factor, *LeNACNOR*. Transcript levels of a range of ripening genes are absent in mature *nor* fruits and their expression cannot be restored by exogenous ethylene (Giovannoni, 2007).

Colourless non-ripening (CNR)

Colourless non-ripening (*Cnr*) is a pleiotropic dominant mutation of tomato *(Solanum lycopersicum*) that results in fruits with a white pericarp displaying much reduced cell-to-cell adhesion (Thompson et al., 1999; Fraser et al., 2001). The gene at the *Cnr* locus has been cloned and shown to encode a *SQUAMOSA PROMOTER BINDING PROTEIN* or *SBP*-box gene (SBP/SPL, where SPL refers to SBP-like) (Manning et al., 2006). The protein is a transcription factor that is thought to target the promoters of *SQUAMOSA MADS*-box genes. In tomato, several MADS domain proteins, in addition to RIN, are present in ripening fruits including TOMATO AGAMOUS-LIKE1 (TAGL1), TAGL11, FUL1 (previously TDR4), and FUL2 (previously MBP7) (Hileman et al., 2006;Leseberg et al., 2008; Martel et al., 2011). FUL is involved in controlling ripening-related metabolic changes in tomato (Litt and Irish, 2003). An orthologue of *FUL* in bilberry is involved in the regulation of anthocyanin biosynthesis and transcription of *CHS* and *MYB2* (Jaakola *et al*., 2010). Silencing the expression of *FUL1*

and *FUL2* in tomato indicated that they influence ripening in an ethylene-independent manner and are also involved in cell wall modification (Bemer et al., 2012).

1.3 Cell Wall Modification and Fruit Softening

1.3.1 Plant Cell Wall

During ripening, the changes in texture involve remodeling of cell walls of the fruits including tomato and also alterations in tissue water-relations caused by modifications in the cuticle (Vicente et al., 2007). To better understand the relationship between cell wall remodeling and fruit softening an understanding of cell wall structure is necessary. Plant cell walls are complex composites which consist of the building blocks of cellulose, hemicellulose, pectin and protein (Keestra et al., 1973; Abersheim, 1976; Carpita and McCann, 2000).

Cellulose

Cellulose is composed of β-1-4-D-gulcan chains that associate through hydrogen bonds which exhibit a paracrystalline structure (Carpita and McCann, 2000). The hydrogen-bonding of many cellulose molecules will form long microfibrils, which are the major component of plant cell walls and approximately a third of dry mass of primary wall consist of cellulose (Ruiz-May and Rose, 2013).

Hemicelluloses

Hemicellulose consists of polysaccharides made up of a variety of sugars including xylans, mannans and glucomannans (mixed linkage glucans, β-(1→3, 1→4)-glucans or MLGs) and xyloglucans (XyGs), (Scheller and Ulvskov, 2010). Hemicelluloses that are primarily composed of xylosyl or arabinosyl residues are referred to as xyloglucans or arabinoglucans, respectively. Hemicelluloses commonly consist of xyloglucans and glucuronoarabinoxylans where the xyloglucans arrange to form linear chain of β-1,4-D- glucan having an α-D-xylose unit on the 6 position of each glucose unit (Carpita and McCann, 2000). They differ from glucuronoarabinoxylans in that they have a xylan backbone with arabinosyl and glucosyl units in the side chains. The basic structures are shown in Figure 1-3.

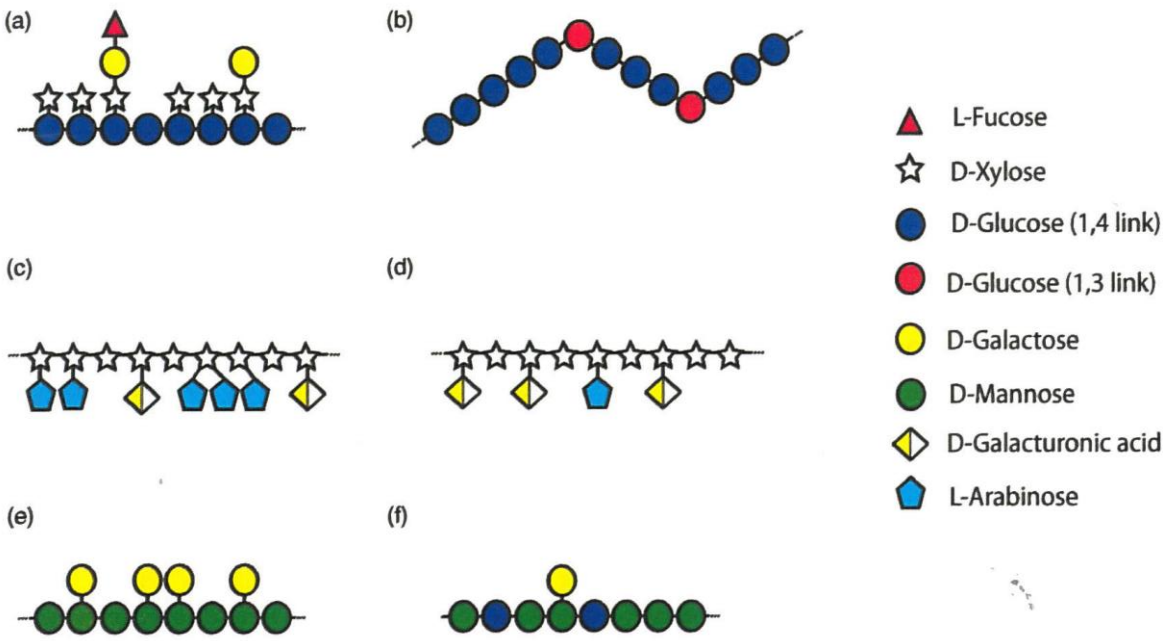


Figure 1-3 The illustration of hemicellulose classes found in plant cell walls (taken from Scheller and Ulvskov, 2010). The hemicellulose structure varies in different plants and tissues. (a) The xyloglucan with side chain that come from pea and Arabidopsis. (b) The linkage β-glucan which can be seen in Poales. (c) Glucuronoarabinoxylan (GAX) that command in monocot. (d) The GAX structure which typical from dicot. (e) Galactomannan seen in Fabaceae seeds. (f) Galactoglucomannan, which is characteristic of conifer wood (taken from Ruiz-May and Rose, 2013).

Pectins

Pectins are the most complex of the cell wall polysaccharides. They are composed of backbones of galacturonosyl uronic acid (GalA) residues and are present in primary cell walls (Cosgrove, 2005). They can be divided into three major classes (Figure 1-4). The first class is the simplest in structure which consists of homogalacturonan (HG), a linear polymer of GalA residues. The second class is rhamnogalacturonan I (RGI) that has alternating GalA and rhamnosyl (Rha) residues as a backbone (Mohnen, 2008). The third class is the polymer rhamnogalacturonan II (RGII) that has an HG backbone with structurally complex side chains and more than 20 different glycosyl linkages (O'Neil et al., 2004). RGII were found conserved across the plant kingdom (Matsunaga et al., 2004). However, the functional significance of the structural heterogeneity or conservation is still yet known (Ruiz-May and Rose, 2013).

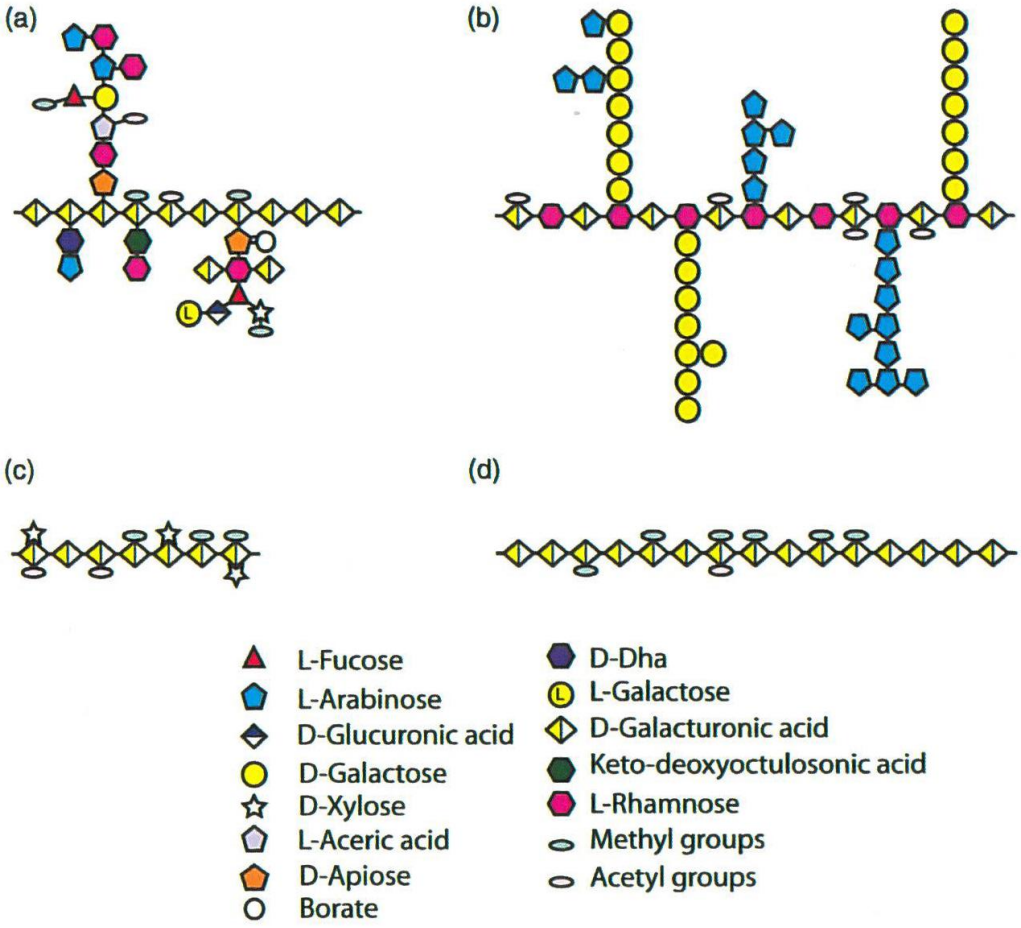


Figure 1-4 Schematic structure of pectin (from Ruiz-May and Rose 2013). (a) Rhamnogalacturonan II (RGII) conserved across plant taxa. (b) Rhamnogalacturonan I (RGI) (c) Xylogalacturonan (XGA) (d) homogalacturonan (HG) (from Ruiz-May and Rose, 2013).

Cell wall proteins

Apart from complex polysaccharides, plant primary cell walls also contain a range of structural proteins including hydroxyl-proline-rich glycoproteins (HPRG), proline-rich-proteins (PRPs), glycine-rich protein (GRPs) and arabinogalactan proteins (AGPs) (Carpita and McCann, 2000). They have a wide variety of roles such as defense against pathogens, cell wall reinforcement and signal transduction (Ringli et al. 2001; Seifert and Roberts, 2007; Deepak et al., 2010). In primary cell wall architecture, the network of cellulose microfibrils is thought to be cross linked by hydrogen bonding to hemicelluloses that are then embedded in pectin polysaccharides, which form a gel-like matrix. In dicot plants, the amount of cellulose and hemicellulose-xyloglucan are approximately equal and they are known as Type I cell walls. These contrasts with Type II, which generally have much lower pectin contents and are mostly found in monocots (Carpita and McCann, 2000). Tomato cell walls are rich in pectins and their composition has been investigated by the Seymour lab (Seymour et al, 1990) (Figure 1-5). It has long been known that the structure of tomato primary cell wall changes during ripening and this is most likely to be the principal cause of the loss of fruit firmness and tissue cohesion (Hobson and Davies, 1971, Brownleader et al., 1999 and Smith, 2001). There are more than 50 cell wall structure-related genes expressed in ripening tomato fruits (Tomato Genome Consortium, 2012). Of these a range of genes encoding hydrolytic enzymes are up-regulated during ripening.

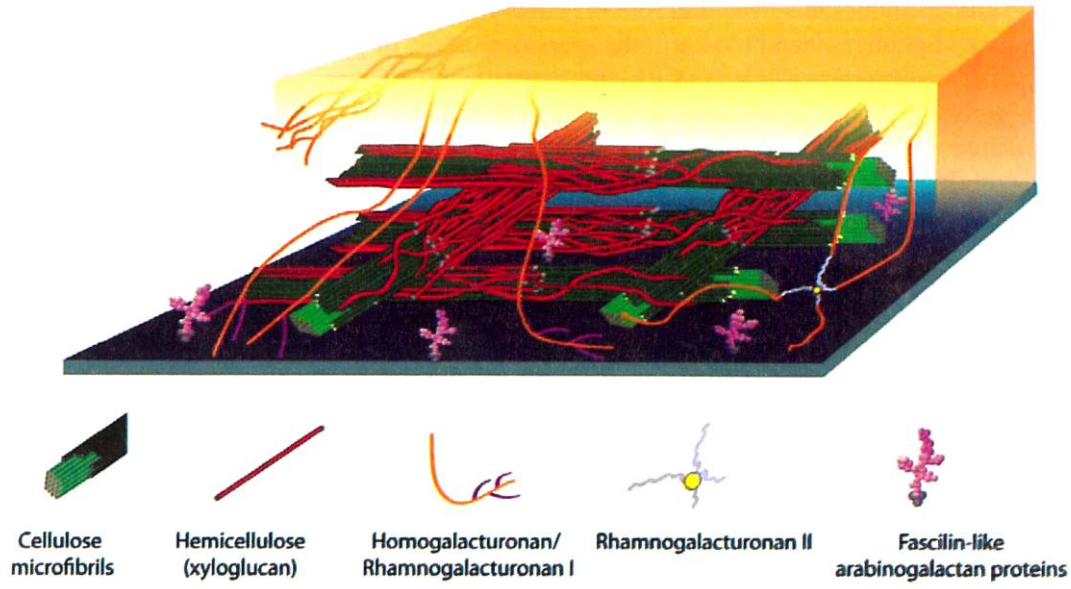


Figure 1-5 The primary cell wall structure model in plants showing major polysaccharide domains (from Ruiz-May and Rose, 2013).

1.3.2 Cell wall hydrolases

Polygalacturonase

Polygalacturonase (*PG*) was the first cell wall hydrolyzing enzyme in tomato that was thoroughly investigated for the role that its plays in fruit softening. Endo-PG functions to catalyze the hydrolysis of the internal 1,4-β-D-galacturonide linkages in the homogalacturonan pectins. *PG* activity is not detected in unripe fruit, but increases dramatically during fruit ripening (Hobson, 1962). Huber and O'Donoghue, (1993) reported that, in tomato and avocado *PG* shows very high activity compared to melon, apple and strawberry (Hadfield and Bennett, 1998). *PG* activity is correlated with pectin solubilization and depolymerization in the late stage of ripening (Hadfield and Bennett 1998) of tomato (Langley et al., 1994), melon (Hadfield et al., 1998) and peach (Lester et al., 1996).

The real test was to determine the effect of down-regulating *PG* on tomato fruit softening using genetic modification. Generating transgenic plants which showed reduced *PG* activity by expression of a *PG* antisense construct demonstrated that although pectin depolymerisation was reduced during the ripening period, fruit softening was unaffected (Smith et al., 1988; Smith et al., 1990). However, reduction of *PG* activity did not affect fruit color or other aspects of ripening (Schuch et al., 1991). Reduced PG tomato fruits were marketed as a GM food for a short while in the 1990s. In UK this GM product involved in tomato puree with improved viscosity. In the USA fresh fruits were sold. Over expression of *PG* in

transgenic *rin* mutant plants failed to induce softening in these mutant fruits (Giovannoni et al., 1989). These results demonstrated that *PG* is important in pectin degradation, but is not sufficient alone to induce pectin solubilization or fruit softening.

In contrast to the results in tomato, silencing of *PG* expression in apple and strawberry has increased the firmness of the ripe fruit and resulted in pectins with a lower degree of de-polymerization. Microscopic analysis of transgenic fruits also has shown smaller intra-cellular spaces and more cellular adhesion (Atkinson et al., 2012; Pose et al., 2013 and Quesada et al., 2009). In apple, the *MdPG1* gene shows novel developmental phenotypes including premature leaf shedding, malformed leaves and malfunctioning stomata.

Pectate lyase

PG is the best characterised cell wall remodelling enzyme in tomato, but it is not the only pectin degrading activity. Interestingly changes in tomato cell walls during fruit ripening indicate a role of several other genes products such as pectate lyase and α-L-arabinofuranosidase. Pectate lyase (EC4.2.2.2) catalyses the cleavage of unesterified of α-(1,4)-galacturosyl linkages giving the 4,5-unsaturated oligogalacturonates by β-elimination (Yoder et al., 1993). The expression of *PL* can be seen in plant organs which include pollen, anthers and pistils, but *PL* gene expression is also high in ripening fruits such as tomato, banana, strawberry and grape

(Marín-Rodríguez et al., 2002). There are 10 genes *PL*-like genes in the tomato genome (Tomato Genome Consortium, 2012), five of which show fruit-related expression.

The role of *PL* in tomato fruit softening has never been properly investigated. However, PLs are highly expressed in strawberry (Jiménez-Bermúdez et al., 2002) and banana fruit (Marín-Rodríguez et al., 2003). The firmness of strawberry fruit is increased when the *PL* gene is suppressed (Jiménez-Bermúdez et al., 2002). In two varieties of apple with different firmness, the softer variety was found to have the higher levels of *PL* enzyme (Harb et al., 2012). In banana, *PL* activity was not detected in preclimacteric fruits but the *PL* activity increased gradually from early climacteric till post climacteric (Tadakittisarn et al., 2009). These findings suggest an important role for *PL* in tomato fruit softening.

Pectin methylesterase

Pectin methylesterases (*PME*) are secreted into the cell wall where they act to remove the methyl groups from esterified pectins. This removal of methyl groups by *PME*, results in the pectins becoming negatively charged and also able to associate with other pectic polysaccharides through calcium bridges (Figure 1.6) (Ochoa-Villarreal et al., 2012). The degree and pattern of methyl esterification is also important as it impacts the rate at which other enzymes like *PG* can cleave the pectin (Pressey and

Avants, 1982). Lower methoxy pectins are the preferred substrate for endo-*PG*.

There are numerous isoforms of PME found in plant tissues (Pressey and Avants, 1972; Tucker et al., 1982). PMEs are present during early fruit development and are generally high in developing fruits, while their levels decrease during the ripening phase (Harriman et al., 1991; Tieman et al., 1992). That PME activity deceases during the later stages of ripening has reported in many fruits, such as strawberry (Barnes and Patchett, 1976), banana (Kanellis et al., 1989), melon (Fils-Lycaon and Buret, 1991) and peach (Glover and Brady, 1994).

There are many different PME isoforms that are active in developing tomato fruit. Suppression of *PMEU1* using an antisense construct resulted in fruit that showed faster softening (Phan et al., 2007). This suggests that PME may be involved in maintaining texture properties perhaps through enhancing the formation of calcium-bound gels. Suppression of PME leads to a higher degree of methyl esterification in pectin during the ripening stage than in WT fruits, but the fruit still appear to ripen normally, although in the later stages of ripening the fruits show very rapid deterioration (Tieman et al., 1992).

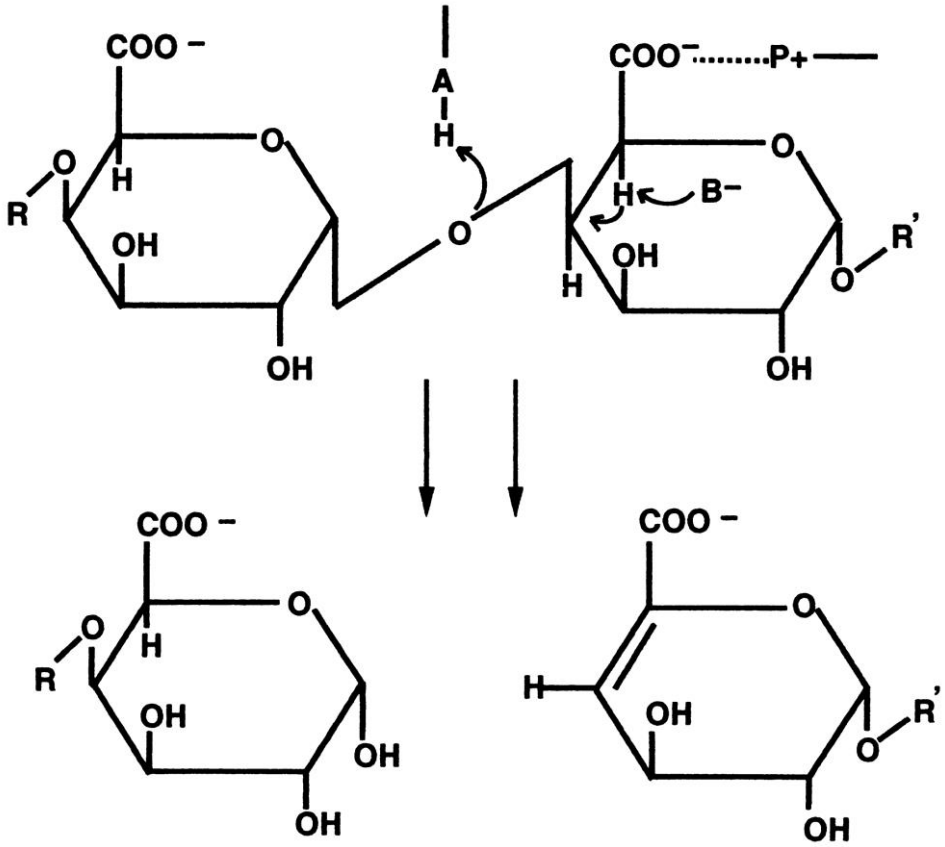


Figure 1-6 Schematic diagram of α-1,4-polygalacturonic acid cleavage by a β-elimination mechanism. The pectate lyases are expected to contribute a minimum of three groups to the catalytic mechanism: P+, which neutralizes the charge on the carboxylic acid group; B, a general base that abstracts the proton from C-5; and A, which is involved in the transfer of the final proton to the glycosidic oxygen, leaving a double bond between C-4 and C-5. (Herron et al., 2000)

β-galactosidase (*TBG*) and arabinofuranosidase

There are both exo- and endo- galactanases active in plants. They act on galactan or arabinose side chains and are mainly found in rhamnogalacturonan I (RGI) (Carpita and Gibeaut, 1993). As tomato fruits start to ripen, this galactosidase activity has been reported to increase and it correlates with the free galactosyl content. The activity continues to increase through the ripening process, with three isoforms of β-galactosidase (I, II and III) purified from ripening tomato (Pressey 1983), pear (Tateishi et al., 2001a), avocado (Tateishi et al., 2001b) and persimmon (Nakamura et al., 2003). Only β-galactosidase II activity significantly increases during tomato ripening and is closely related to softening (Pressey, 1983; Carey et al., 1995; Carrington and Presse, 1996).

β-galactosidase activity is encoded by a family of genes. In tomato there are seven different genes that have been identified *TBG1-TBG7* (Smith and Gross, 2000). However, the genome sequencing programm indicates only two of them are expressed during ripening, *TBG4* and *TBG6* (Tomato Genome Consortium, 2012). The β-galactosidase II isoform has been purified (Carey et al., 1995) and the *TBG4* gene matches this sequence (Smith et al, 2002). When *TBG4* was silenced in transgenic tomato plants the fruit showed a 40% reduction in softening (Smith et al., 2002) at the red ripe stage. The increase in firmness was correlated to a higher content of galactosyl and lower level of *TBG4* mRNA and exo-galactanase activity in the early stage of ripening (Smith et al., 2002). In

contrast, the suppression of the *TBG6* gene in tomato led to a reduction in fruit firmness and gave fruit that exhibited cracking and presence of double cuticle in the early stage of fruit development (Moctezuma et al., 2003).

α-L-arabinofuranosidase exhibits a similar pattern of activity to β-galactosidase. This enzyme is an exo-hydrolase that removes terminal and non-reducing arabinosyl residues from pectin, hemicellulosic and other glycoconjugates (Sozzi et al., 2002). α-L-arabinofuranosidase can been found during the ripening in both the control and the antisense ACC synthase tomato fruit (Sozzi et al., 2002). The activity is observed in ripening tomato, but its role in fruit softening remains unknown (Schroder et al., 2004).

Xyloglucan Endo-trans-glycosylase/ hydrolase (XTH)

Xyloglucan endotransglycosylase (*XET*) is now known as xyloglucan endotransglycosylase/hydrolase (*XTH*) (Rose et al., 2002). *XTH* enzymes can have one of two activities. Most *XTHs* act as *XETs* and catalyse the cleavage of the donor xyloglucan polymer and transfer the newly created reducing end to the non-reducing terminus of an acceptor xyloglucan polymer. The other role is to act preferentially as xyloglucan endohydrolase (*XEHs*), resulting in depolymerisation of xyloglucan (Figure 1-7) (Saladie et al., 2006). *XTH* activity has been detected in many fruit, such as tomato (Maclachlan and Brady, 1994), apple and kiwifruit (Redgwell and Fry, 1993; Percy et al., 1996). The activity of *XTH* is very high in young expanding tomato fruits and it declines slowly during fruit

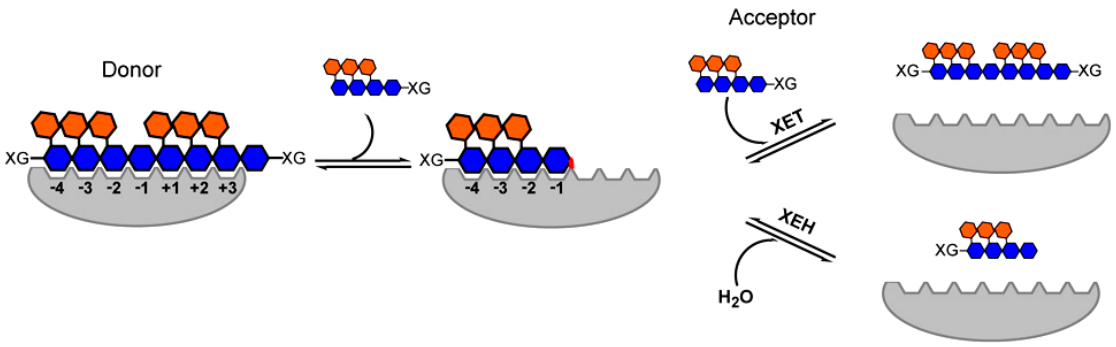


Figure 1-7 The model of depolymerisation in *XETs* and *XEHs*. *XTH* will cut xyloglucan chain and start again at the end of another xyloglucan molecule (xyloglucan endotransglycosylase, *XET* activity) or to water molecule (xyloglucan endotranshydrolase, *XEH* activity). The glucosyl of xyloglucan unit was colour in blue and xylosyl unit in orange. The image was adapted from Eklöf and Brumer, 2010.

maturation and slightly rises during ripening (Maclachlan and Brady, 1992; 1994). The role of these enzymes in fruit softening is unknown.

Endo-(1,4)-β-D-glucanase (EGases)

This class of enzymes are also known as cellulases and are able to hydrolase the internal linkages of 1,4-β-D-glucans (Brummell and Harpster, 2001). These enzymes have been implicated as being involved in a diverse range of processes including fruit softening, cell elongation and abscission (Brummell and Harpster, 2001). Plant *EGases* are encoded by multi gene families in tomato and there are at least two *EGases*, *LeCel1* and *LeCel2* that are highly expressed during fruit ripening (Lashbrook et al., 1994). However, later in ripening the expression of *LeCel1* slowly declines whereas the *LeCel2* is progressively expressed from breaker stage onward (Lashbrook et al., 1994). Suppression of *LeCel2* does not give any changes in firmness of tomato fruits (Brummell et al. 1999). Over-expression of pepper *CaCel1* gene also failed to impact changes in fruit softening (Harpster et al., 2002).

1.3.3 Non-enzymatic mechanisms

Expansins (Exp)

Expansins are cell wall modifying proteins that are present with no hydrolysis or transferase activity (McQueen-Mason et al., 1992). Expansin encoding genes are expressed in numerous plant tissues including hypocotyls (Mc-Queen-Mason et al., 1992), coleoptiles (Li et al., 1993),

leaves (Keller and Cosgrove, 1995), internodes (Cho and Kende, 1997), and fruit (Rose et al., 2000). In tomato there are seven expansin isoforms that have been identified. *LeExp1* acts as the main ripening related-expansin in tomato (Brummell et al., 1999; The Tomato Genome Consortium, 2012). Three expansin precursors (*EXPA3*, *EXPA4* and *EXPA5*) were present during fruit development prior to the onset of ripening (The Tomato Genome Consortium, 2012). Suppression of *LeExp1* increased fruit firmness by about 10%, while its over-expression led to an increase in hemicellulose depolymerization (Brummell et al., 1999). In combination with reduced PG activity, Powell et al., (2003) reported that *EXP1* and *PG* silenced tomato lines had significantly increased firmness and longer-shelf life, but the effects were quite modest. Additionally, the lines also had increased juice viscosity (Powell et al., 2003).

In strawberry, expansin (*FeExp2*) expression was detected during ripening (Civello et al., 1999). In peach several expansin genes are expressed but only *PpExp3* mRNA was suggested too play the major role in fruit softening (Hayama et al., 2003)

1.4 Changes in colour during ripening

1.4.1 Carotenoids

Carotenoids are isoprenoid molecules that exist in tissues that undergo photosynthetic processes. Biochemically, they are also classified as hydrophobic antioxidants. Tomato is fruit that contains very high levels of carotenoids including lycopene (Seymour et al., 2013). Lycopene acts as an antioxidant that helps to protect from prostate cancer and cardiovascular disease (Stacewicz-Sapuntzakis and Bowen, 2005). When tomato reaches the mature green stage, its plastids are chloroplasts and these contain chlorophylls and are able to photosynthesise. Then during breaker stage, the chloroplasts are converted into chromoplasts. Carotenoid compounds accumulate as the chloroplasts are converted to chromoplasts and the most significant increase occurs in lycopene, a carotenoid which gives tomato fruit their red colouration. The carotenoids accumulate in bodies called plastoglobuli that result from the breakdown of the thylakoids (Egea et al., 2010). In the process of ripening, the accumulated lycopene resulted in the increase of total levels of carotenoids by 10-14 times (Bramley, 2002). Figure 1-8 describes the biosynthetic steps involved in the generation of carotenoid pigments in plants. In 1993, Fray and Grierson discovered that *yellow flesh* (R) locus was responsible for the inhibition of carotenoid formation in tomato and resulted from a mutation in the phytoene synthase 1 gene (*PSY1*). This disrupted the conversion of geranylgeranyl diphosphate into phytoene and resulted in low levels of carotenoids such as lycopene.

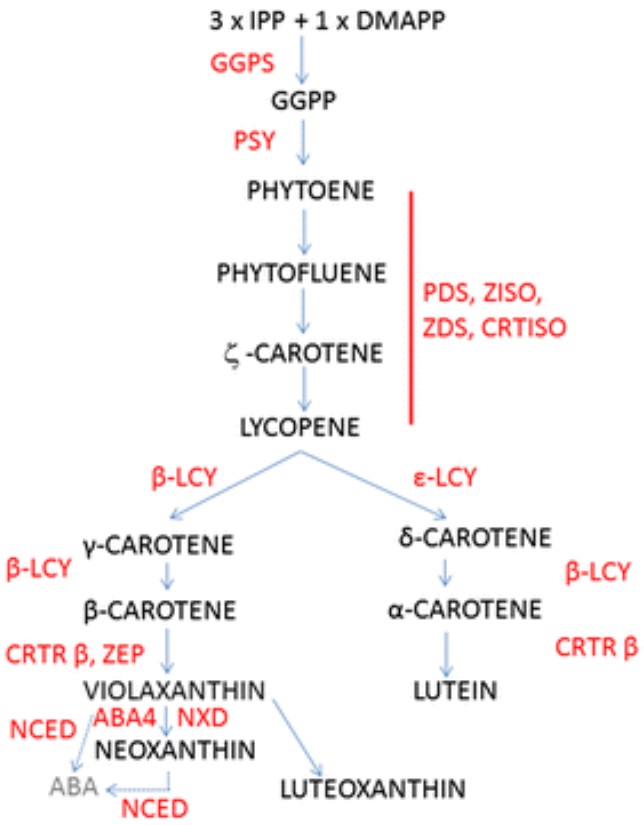


Figure 1-8 The carotenoid pigment biosynthesis pathway. The enzyme names are represent in red colour while black colour for intermediate compound. IPP = isopentenyl diphosphate, GGPS = GGPP synthase, GGPP = geranyl-geranyl pyrophosphate, PSY = phytoene synthase, PDS = phytoene desaturase, ZISO = zeta-carotene isomerase, ZDS = zeta-carotene desaturase, CRTISO = carotenoid isomerase, ε-LCY = lycopene ε-cyclase, β-LCY = lycopene β-cyclase, CRTR-β = β-carotene hydroxylase, ZEP = zeaxanthin epoxydase, NXD = neoxanthin synthase, CHL = chlorophyllases, ABA = abscisic acid (from Su et al. 2015).

1.4.2 Flavonoids

Flavonoids come from a group of phenolic secondary metabolites. They are plant pigments that can be found in a variety of fruits, vegetables and nuts. They produce yellow or red/blue coloured pigments in plants that draw the attention of pollinator animals and insects. They have various health benefits which include anti-inflammatory and anti-cancer properties (Chahar et al., 2011). Due to the health stimulating action of flavonoids, a major target is to increase their levels in crop plants. One of the ways is by gene modification, which can produce 78 times the level of flavanols through ectopic expression of a single biosynthetic enzyme, chalcone isomerase (Lee et al., 2007). The flavonoid synthesis pathway, consists of several distinct steps (Figure 1-9) with enzymes such as *CHS* and *DFR* which are involved in the early and later steps of the pathway. Mathews et al., (2003) discovered that by over expression of *ANT1,* a gene encoding a MYB transcription factor, they could enhance the levels of purple pigment in tomato fruits. More recently, similar genes from snapdragons, induce the production of anthocyanins in tomato and result in purple fruits (Butelli et al., 2008)

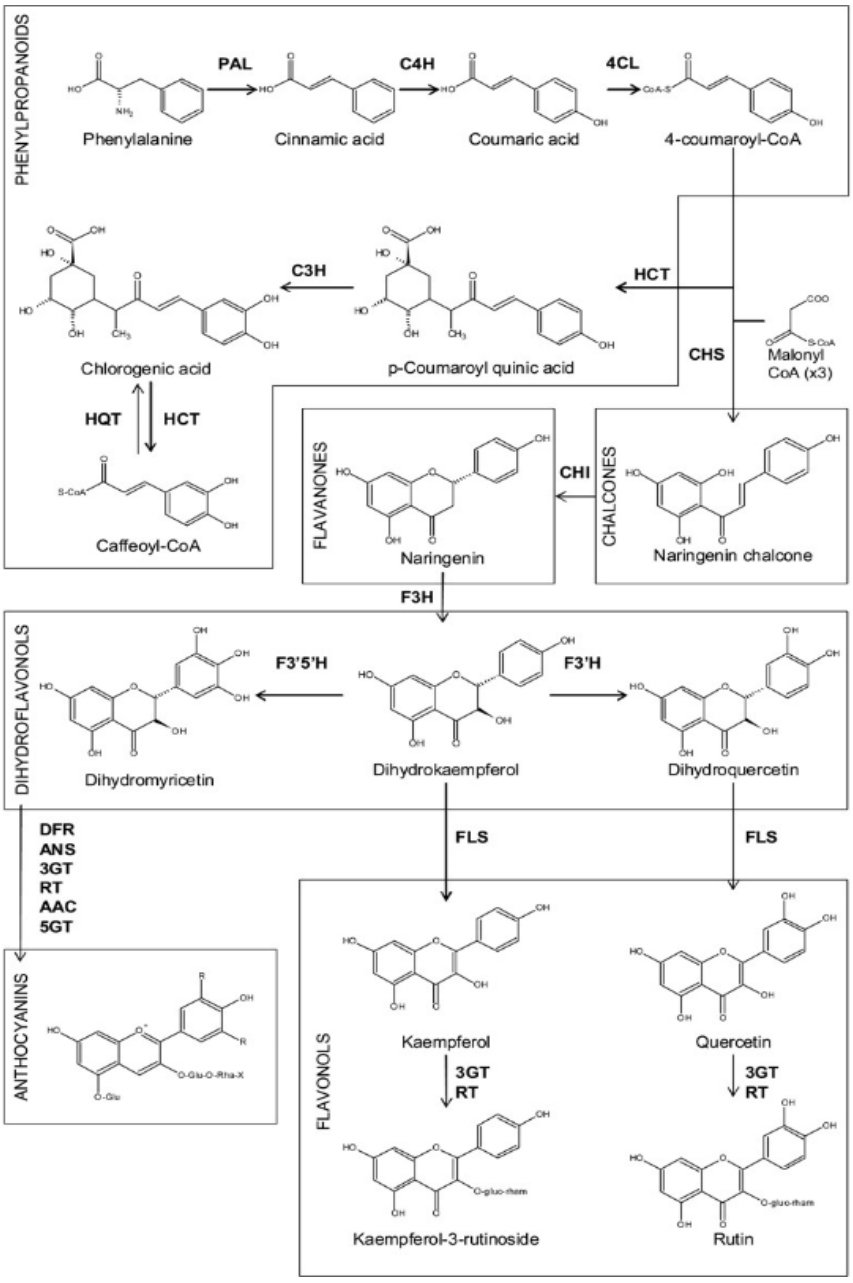


Figure 1-9 The schematic overview of flavonoid pathway. PAL, Phe ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; HCT, cinnamoyl-CoA shikimate/quinate transferase; C3H, p-coumaroyl ester 3-hydroxylase; HQT, hydroxycinnamoyl-CoA quinate transferase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone-3-hydroxylase; F3'H, flavonoid-3'-hydroxylase; F3'5'H, flavonoid-3'5'-hydroxylase; FLS, flavonol synthase; DFR, dihydroflavonol reductase; ANS, anthocyanidin synthase; 3GT, flavonoid-3-O-glucosyltransferase; RT, flavonoid 3-O-glucoside-rhamnosyltransferase; AAC, anthocyanin acyltransferase; 5GT, flavonoid-5-glucosyltransferase. The image from Ballester et al., 2010.

1.5 Modification of fruit flavour during ripening

The concentration of sugars and organic acids has a major effect on the taste of fruits. In early stages in fruit development (0-40 DPA), sugar will be transported from leaves (source) to developing fruit (sink). Sucrose is transported from the leaves and provides the carbon source for cell division and expansion in the fruits. In ripe tomato fruits, glucose and fructose form the major proportion of soluble sugars, with a small amount of sucrose (Carrari and Fernie, 2006). In tomato, two enzymes are involved in sucrose degradation. The first enzyme is sucrose synthase (SuSy) where this enzyme cleaves sucrose into UDP-glucose and fructose in the cytosol. The second enzyme is invertase which cleaves sucrose into glucose and fructose in vacuole (Wang et al., 1993; Nguyen-Quoc and Foyer, 2001). Some of the sugar also comes from starch that accumulates in the fruits and this reserve of carbohydrate can have an important impact on taste (Gillaspy et al., 1993; Klee and Giovannoni, 2011).

Amino acids such as glutamate and volatile compounds also play an important role in enhancing the flavour of tomatoes and other fruits (Tieman et al., 2006). In tomato, 400 volatile compounds were identified as being produced by the ripe fruit, but only 15-20 seem to be critical for flavour. Carotenoids, amino acids and fatty acids are precursors of volatile compounds. Volatile compounds can also act as dietary anticarcinogenics and antimicrobial agents (Goff and Klee, 2006; Klee, 2010; Osorio et al., 2010).

Table 1- 2 Volatile compounds that contribute to tomato flavour (Taken from Klee, 2010).

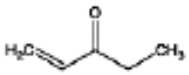
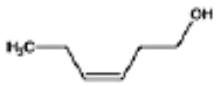
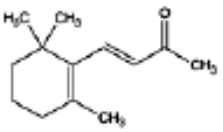
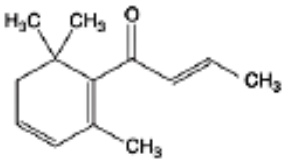
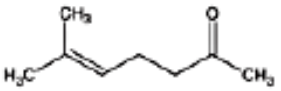
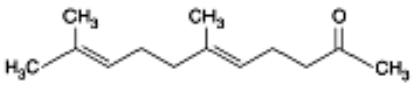
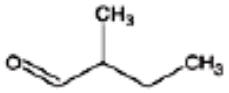
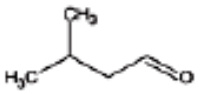
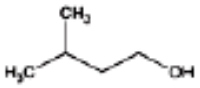
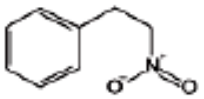
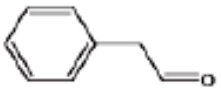
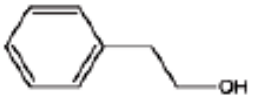
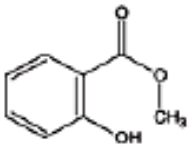
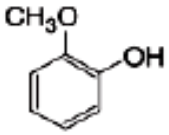
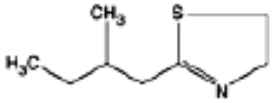
Volatile	Structure	Precursor
cis-3-hexenal		Fatty acid
Hexanal		Fatty acid
1-Penten-3-one		Fatty acid
trans-2-hexenal		Fatty acid
trans-2-heptenal		Fatty acid
cis-3-hexenol		Fatty acid
β-Ionone		Carotenoid
β-Damascenone		Carotenoid
6-Methyl-5-hepten-2-one		Carotenoid
Geranylacetone		Carotenoid

Table 1- 2 (Continued) Volatile compounds that contribute to tomato flavour.

Volatile	Structure	Precursor
2-Methylbutanal		Isoleucine
3-Methylbutanal		Leucine
3-Methylbutanol		Leucine
1-Nitro-2-phenylethane		Phenylalanine
Phenylacetaldehyde		Phenylalanine
2-Phenylethanol		Phenylalanine
Methyl salicylate		Phenylalanine and/or chorismate
Guaiacol (2-methoxyphenc		Phenylalanine?
Isobutylthiazole		Unknown

1.6 Sequenced Fruit Genomes and identification of genes underlying quality traits

There are eighteen fruit genome sequences that have been published to date (Table 1-3). Although most of the fruit genomes have fairly complete sequences there is a lot of work to do to finish the genome assemblies including physical and functional annotation. The tomato genome assembly is the most complete of all of the fruit species so far, although the functional annotation is still incomplete (Gapper et al., 2014). In some genome projects additional complexities have arisen e.g. apple with gene duplications and heterozygosity making assembly especially difficult (Velasco et al., 2010).

The sequenced genomes provide the means to identify the types and number of genes expressed during specific ripening processes. They also transform the ability to identify genes underlying quantitative trait loci associated with fruit quality characteristics. The enormous increase in data related to ripening-related genes provides a rich new source of information to explore the mechanistic basis of changes in fruit colour, texture and flavour (Figure 1-10).

Table 1-3 Fruit genomes that have been sequenced.

Common Name	Family	Scientific name	References
Grape	*Vitaceae*	*Vitis vinifera*	Jaillon et al., 2007; Velasco et al., 2007
Papaya	*Caricaceae*	*Carica papaya*	Ming et al., 2008
Cucumber	*Cucurbitaceae*	*Cucumis sativus*	Huang et al., 2009
Apple	*Rosaceae*	*Malus domestica*	Velasco et al., 2010
Strawberry	*Rosaceae*	*Fragaria vesca*	Shulaev et al., 2011
Date palm	*Arecaceae*	*Phoenix dactylifera*	Al-Dous et al., 2011
Tomato	*Solanaceae*	*Solanum lycopersicum*	Tomato Genome Consortium, 2012
Banana	*Musaceae*	*Musa acuminata*	D'Hont et al., 2012
Melon	*Cucurbitaceae*	*Cucumis melo*	Garcia-Mas et al., 2012
Chinese plum	*Rosaceae*	*Prunus mume*	Zhang et al., 2012
Pear	*Rosaceae*	*Pyrus bretschneideri*	Wu et al., 2013
Watermelon	*Cucurbitaceae*	*Citrullus lanatus*	Guo et al., 2013
Peach	*Rosaceae*	*Prunus persica*	Verde et al., 2013
Kiwifruit	*Actinidiaceae*	*Actinidia chinensis*	Huang et al., 2013
Pepper	*Solanaceae*	*Capsicum annuum*	Qin et al., 2014

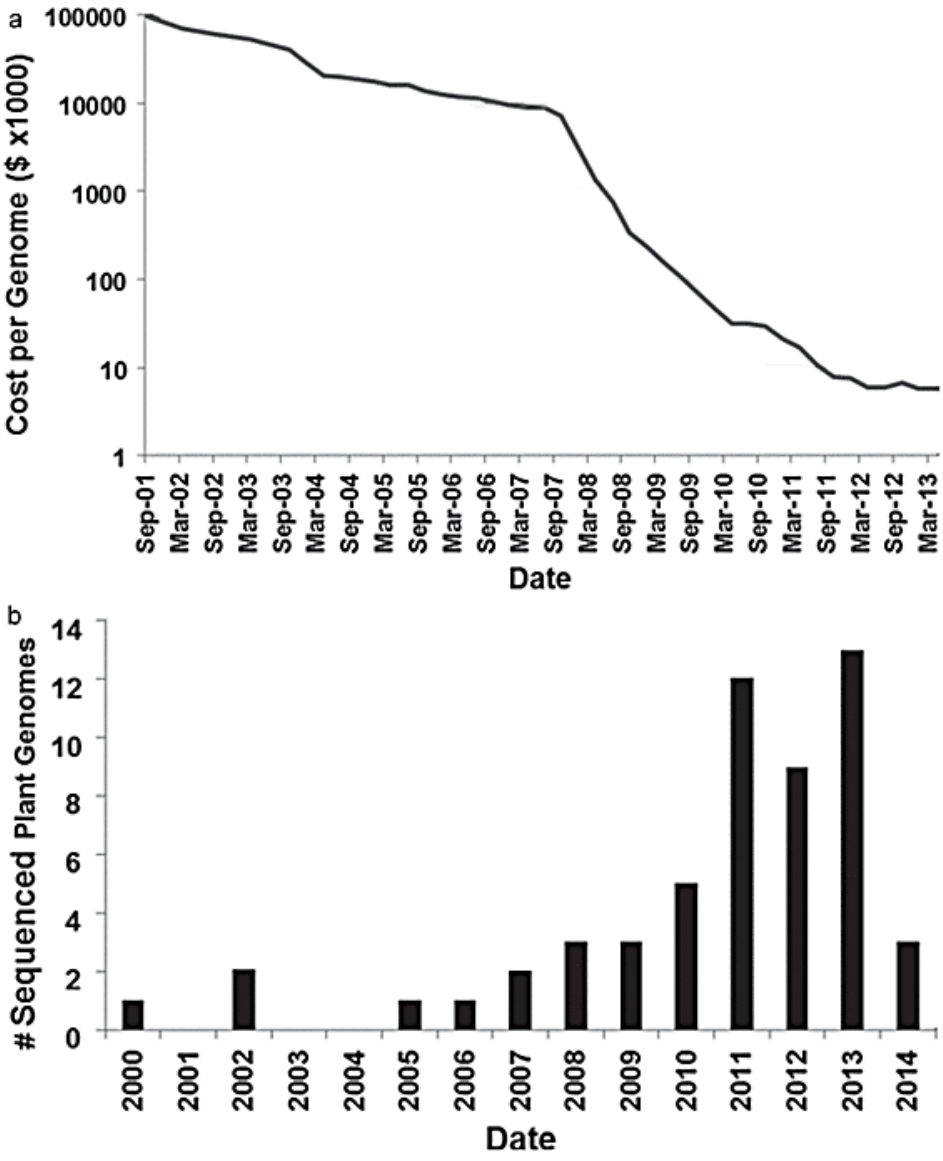


Figure 1-10 The cost and plant genomes this decade. (a) The cost of sequencing a human genome during the twenty-first century, data uploaded from the National Human Genome Research Initiative (NIH, http://genome.gov/sequencingcosts). (b) Increase in the number of sequenced plant genomes. (Taken from Gapper et al., 2014)

1.7 Aims and Objectives of This Project

The aim of the project was to investigate the mechanism of tomato fruit softening drawing on genome-wide information from a range of fruit species and using the new technology of genome editing. The specific objectives were:

1. Make a comparative genomics study of ripening-related gene expression in a range of fleshy fruits especially that linked with cell wall remodelling in order to identify common and unique subsets of genes encoding cell wall enzymes across fruit bearing species.

2. Make predictions about common genes likely to be responsible for the softening events in all fruit and select several candidate genes to develop constructs that can alter texture using CRISPR-Cas9 technology specifically in tomato.

3. Analyse and investigate the function of the candidate gene(s) in the transgenic tomato lines.

4. Investigate the role of additional specific cell wall genes in fruit softening during ripening where appropriate.

Chapter 2

General Materials and Methods

CHAPTER 2: MATERIALS AND METHODS

2.1 Plant materials

Tomato (*Solanum lycopersicum*) plants cv. Ailsa Craig, wild type and those containing CRISPR/Cas9 constructs, were grown under standard glasshouse conditions with 16 hours of day length and temperatures of, 20°C during the day time and 16°C at night. Supplemental lighting was supplied when required. Plants were grown in Pro C2 coarse potting compost (Levington) in 7.5 liter pots with irrigation that contain organic nutrients, seaweed and humates (Vitax organic). In order to select fruits at the required developmental stages the flowers were tagged at anthesis and then fruits were harvested at mature green which was taken as 40 days post anthesis (dpa), breaker (B), when the first sign of a yellow or orange colour becomes noticeable, and then at selected days post breaker, B+4 (orange ripe) and B+7 (red ripe).

2.2 Genomic DNA Extraction

Tomato genomic DNA was extracted according to the protocol in the DNeasy Plant Mini Kit (QIAGEN Ltd, Manchester, UK). Fresh leaves were collected and kept in frozen tubes at -80°C in the freezer. The leaves were then vigorously ground with a plastic pestle within an eppendorf tube. Then, 400 µL of AP1 buffer was added to 100 mg of frozen leaf tissue

along with 4 µL RNase. The solutions were then mixed using a votex mixer. The tubes were then placed in an incubator (Techne DI-block DB-2A) for 10 minutes at 65^{o}C and the contents were mixed 3-4 times during the incubation period by inverting the tube. Then, 130 µL of AP2 buffer was added to each tube and the mixture incubated on ice for 5 minutes to precipitate proteins and polysaccharides. The samples were then centrifuged for 5 minutes at 14,000 rpm. The flow-through, which was approximately 450 µL, was transferred into a new 1.5 mL tube and mixed with 675 µL of AP3/E buffer. After that, 650 µL of the mixture was pipetted into a DNeasy Mini Spin Column and centrifuged for one minute at 8,000 rpm. The flow-through was discarded and the collection tubes were reused with the remaining sample and this step was repeated twice. Then after discarding the flow-through and the collection tube, the DNeasy Spin Column was placed into a new 2 mL collection tube. 500 µL of AW buffer was added to the column and centrifuged for one minute at 8,000 rpm. The flow-through was discarded and this step was repeated by centrifuging at 14,000 rpm for two minutes. Finally, the DNeasy Spin Column was transferred to a clean 2 mL micro-centrifuge tube and 50 µL of AE buffer was pipetted onto the DNeasy membrane. The tube was incubated for five minutes at room temperature and centrifuged for one minute at 8,000 rpm. The DNeasy Mini Spin Column was then placed into a new 2 mL centrifuge tube and last step was repeated and a second sample collected.

2.3 Quick DNA Extraction

In some cases a rapid DNA extraction method was used due to high number of plants samples that needed to be analyzed. Fresh leaves (~50-100 mg) were collected in 96 well plates and stored on ice during collection. 3mm stainless steel ball was placed into each well. Then 600 µL of speed DNA extraction buffer (Appendix 9.1) was added at room temperature and immediately shaken for 5 minutes in a shaker (Mixer mill, 8000). The plate was centrifuged at 5,600 rpm for 20 minutes (Allegra 25R centrifuge). Then, 300 µL of supernatant was transferred with long narrow pipette tips to a fresh plate (Thermofisher storage plate) containing 300 µL of 100% isopropanol. The solutions were mixed and incubated for 5 minutes at room temperature. The plate was centrifuged again at 5,600 rpm for 20 minutes. The supernatant was removed by aspiration and the pellet was dried under desiccation for 15 minutes. Finally, the DNA was re-suspended in 0.1 X TE buffer (50 µL) for 3 hours at 4°C.

2.4 RNA Extraction

RNA extraction was carried out according to PureLink® RNA Mini Kit (Life Technologies, Dorset UK). Tomato fruit at various stages of ripening (MG, B, B+4 and B+7) were harvested and the pericarp frozen in liquid nitrogen. The frozen tissue was then ground to be fine powder in a coffee grinder. A 100 mg of sample was transferred to a fresh 2 mL tube and 700 µl Lysis buffer (Appendix 9.2) and 7 µL 2-mercaptoethanol were

added before mixing by vortexing (until cell pellet is completely dispersed) and centrifuging for 2 minutes at 12,000 x g. The supernatant was then transferred to a new 2 mL fresh tube and 70% ethanol added (1:1 volume) and mixed by vortexing. Then 700 µL of the sample was transferred to a Spin Cartridge before centrifuging at 12,000 × g for 15 seconds at room temperature. The flow through was discarded and the same step repeated with the remaining sample. 700 µL of wash buffer I (Appendix 9.3) was added to the Spin Cartridge and centrifuged at 12,000 × g for 30 seconds at room temperature. The wash step was repeated by using 500 µL of wash buffer II (Appendix 9.3) and centrifuging at 12,000 × g for 30 seconds at room temperature and the flow through discarded. The empty Spin Cartridge was centrifuged again at 12,000 × g for 1-2 minutes to dry the membrane with attached the RNA. The column was then placed in a fresh 1.5 mL tube and 30-100 µL of RNase–Free Water added to the centre of the Spin Cartridge. The wetted cartridge was then incubated at room temperature for 1-2 minutes before centrifuging again for 2 minutes at ≥12,000 × g. The sample was collected in a clean tube and purified RNA stored at -80°C.

DNA contamination in the RNA samples was removed using the Ambion DNA-*free* Kit (Life Technologies, Dorset, UK). Each sample from total RNA (1.5 µg) was treated with 10xDNase1 buffer and 1 µL of DNase 1. The solution was mixed gently before incubating at room temperature for 25 minutes. DNase inactivation reagent was then added into the reaction tubes and the solution mixed well by flicking or vortexing. The tubes were incubated at room temperature for 2 minutes with agitation

every 40 seconds before centrifuging at 10,000 rpm for 90 seconds. The supernatant was carefully transferred into a fresh tube and store at -80°C until required.

2.5 Determination of DNA and RNA concentrations

A NanoDrop ™ spectrophotometer (Thermo Scientific, Loughborough, UK) was used to measure the concentration of DNA in each sample. Before the concentration of sample was determined, the NanoDrop TM spectrophotometer was calibrated by using 1 µL of distilled water. The distilled water was wiped off with a tissue and 1 µL of DNA sample was placed on the NanoDrop ™. The ng μL^{-1}, A260, A280 260/280 and 260/230 values were recorded. The DNA concentration (ng μL^{-1}) was used to calculate the DNA dilutions.

2.6 Primer design

General Primers

Oligonucleotide primers were designed using the Primer-BLAST programme at the NCBI [http://www.ncbi.nlm.nih.gov/tools/primer-blast/] website and then synthesised by Eurofins MWG Operon (Anzingerstr. 7a, 85560 Ebersberg, Germany). Primers were diluted according to manufacturer's instructions and stored at -20°C. Stock concentrations for primers were diluted by 10X TE buffer and then 20 µL

of stock primer solution and 180 µL of water were mixed to give the working primer solutions.

2.7 Polymerase Chain Reaction (PCR)

PCR reactions were performed using Techne TC-412 PCR machine. Each reaction tube contained a total reaction mixture of 20 µL composed of 8 µL of water, 10 µL of PCR buffer (HotStarTaq Master Mix), 1 µL of genomic DNA, 0.5 µL of forward and reverse primer (MWG). Genomic DNA from tomato was used unless otherwise stated. The general PCR programme involved a denaturing step of 95°C for 3 minutes, 35 cycles of 94°C for 30 seconds, 59°C for 45 seconds, and 72°C for 1 minute and also 10 minutes for final extension at 72°C. All PCR products were kept in 4°C.

2.8 Quantitative RT-PCR (QPCR)

2.8.1 cDNA synthesis

The DNA free RNA sample was diluted to give a concentration of 100ng/µL using RNase free water (Promega). For annealing the oligo primer to RNA, 5 µL of total RNA, 1 µL of random primers (0.5 µg μL^{-1}, C1181, Promega) and 9 µL of RNase free water (Promega) were mixed together before being incubated at 70°C for 5 minutes using Techne TC-412 PCR machine. The samples were then transferred directly onto ice. For Reverse Transcriptase reaction 5 µL of 5x MMLV Reverse Transcriptase reaction

buffer, 1.25 µL of 1 mM dNTPs mix, 0.5 µL of RNase inhibitor (N2511, Promega, Southampton UK), 1 µL of MMLV Rev Trans (M170A, Promega) and 2.25 µL of Nuclease-free water were added to the annealed RNA before incubating for 10 minutes at room temperature and then 42°C for 1 hour (Technie TC-41 PCR machine). Finally, 75 µL dH_2O was added to the first strand cDNA to obtain total volume 100 µL before keeping at -20°C as a stock solution.

2.8.2 Quantitative RT-PCR (QPCR)

The SYBR Green1 qPCR was run on a LightCycler480 System (Roche Applied Science, West Sussex, UK). The PCR reaction contained a 5 µL cDNA pool, 7.5 µL of 2X SYBR Green PCR Master Mix (Life Technologies), 0.5 µL of 10 mM forward primer, 0.5 µL of 10 mM reverse primer and 1.5 µL of water that give in a final volume of 15 µL. Three biological replicates per cDNA pool were used for our candidate genes and an internal standard, elongation factor (EF). Standard curves for each gene were run concurrently.

The PCR conditions consisted of an initial denaturation step at 95°C for 10 min, followed by 50 cycles of 95°C for 10 seconds, 60°C for 50 seconds, and 72°C for 1 seconds, and a final cooling step of 40°C for 10 minutes. Standard curves were used to calculate relative mRNA concentrations from Crossing Point (the point at which the fluorescent signal is statistically significantly above background to the specific fluorescence threshold) values using absolute quantification with

LightCycler480 software release 1.5 (Roche Applied Science) and normalized to the reference gene EF (primer sequences in Appendix 9.8). Housekeeping genes that have stable transcript expression in our treatments were used to normalize with the target gene and the value was calculated using the equation 2.1 below (German et al., 2003).

$$\text{Relative gene expression} = \frac{2^{Cp\ of\ Referece}}{2^{Cp\ of\ Target}}$$

$$= 2^{-(Cp\ of\ Target\ -\ Cp\ of\ reference)}$$

$$= 2^{-\Delta Cp} \qquad (2.1)$$

2.9 CRISPR/Cas9 Vector construction

2.9.1 Primer Design for Target Genes

The CRISPR vectors and construct design information were obtained originally from Jonathan Jones, Vladimir Nekrasov and Sophien Kamoun, TSL and The Gatsby Charitable Foundation. In this study, we designed the sgRNA by selecting a 20 bp target region within the gene of interest. The information from TSL indicated that the target sequence should contain 5' GNNNN NNNNN NNNNN NNNNN NGG 3' (Mali et al., 2013) for purposes of detecting mutations introduced by Cas9 using the restriction enzyme site loss method (Voytas and Gao, 2014). Restriction site at the 3' half of the target sequence as Cas9 cuts towards the 3' end of the target were also included (Jinek et al., 2012). Forward (F) and reverse (R) primers for PCR were designed manually. During the primer design, we ensured that our 20 bp target regions only located in the exon of the gene with a minimum T_m of 58°C. The full sequence of candidate gene was obtained from the tomato genome browser [http://solgenomics.net/], sequence version 2.40 (Appendix 9-9b).

2.9.2 CRISPR Level 0 vector construction

In order to find the best annealing temperature a gradient PCR reaction was performed (Eppendorf Mastercycler) with temperatures ranging from 57 to 61°C. Each reaction tube contained 25 µL composed of 9 µL of

water, 12 µL of PCR buffer (2X Q5 Master Mix), 1 µL of pICH86966::AtU6p::sgRNA with concentration 100 ng μL^{-1}, 1.5 µL of forward and 1.5 µL reverse primer (MWG). The conditions for the PCR were, denaturing step of 95°C for 3 minutes, 35 cycles of 94°C for 30 seconds, 59°C for 45 seconds, and 72°C for 1 minute, and a final extension step of 10 minutes at 72°C.

The level 0 vector that contained the sgRNA using pICH86966::AtU6p::sgRNA construct (Addgene Plasmid 46966) as a template was purified for the next cut-ligation step in level 1. PCR purification was undertaken according to manufacturer's instructions (Sigma-Aldrich Company Ltd, Dorset, UK). The GenElute plasmid mini spin columns were placed in a collection tube provided and 0.5 mL of the column preparation solution was added to each mini spin column and centrifuged at 12,000 rpm for 30 seconds. Then, 5 volumes of binding solution were added to 1 volume of PCR reaction and mixed. The solution was transferred into the binding column and the column centrifuged at maximum speed for 1 minute. The flow through was discarded and 0.5 mL of wash solution was added to the column. The column was then centrifuged again for a minute at 13,000 rpm, the flow through was again was discarded and the column centrifuged for a further 2 minutes. Then, the column was placed in a new 1.5 mL microcentrifuge tube. The DNA was eluted by the addition of 30 µL of elution solution or water into the column. The column was left for around 5 minutes and then centrifuged. The 1.5 mL microcentrifuge tube was then placed in a freezer at -20°C.

2.9.3 Ligation CRISPR Level 1

The cut-ligation is performed as described in Weber et al. (2011). This, cut-ligation reaction using *BsaI* with a Level 0 construct pICSL01009::AtU6p (SpecR) and a Level 1 destination vector pICH47751 (*CarbR*) are illustrated in Figure 2-1.

The cut-ligation reaction was in a total volume of 10 µL in a 0.5 mL Eppendorf tubes composed of 1 µL of 10X CutSmart™ Buffer (NEB, UK), 1 µL of 10 mM Adenosine 5'-Triphosphate (ATP), 1 µL of BsaI (20,000 units ml^{-1}) (NEB, UK), 1 µL of T7 DNA Ligase (3,000,000 units ml^{-1}) (NEB, UK), 1 µL of sgRNA PCR product (100 ng μL^{-1}), 1 µL of pICSL01009::AtU6p (100 ng μL^{-1}), 1 µL of pICH47751 (100 ng μL^{-1}) and 2 µL of sterilised water. The PCR temperature conditions were 20 cycles of 37°C for 5 minutes, 20°C for 5 minutes and extension step of 20 minutes at 70°C. The reaction was incubated at room temperature (22 to 25°C) overnight before transforming into competent cells.

2.9.4 Transformation of Entry Clone into *E. coli* Competent Cells

Two microliters of each cloning reaction or 100 ng of DNA vectors were added into a vial of 40 µL of chemically competent cells of *E. coli* (DH5α chemical treated) and mixed gently. The vial was incubated on ice for 20 minutes. The cells in the vial were then heat-shocked for 45 seconds at 42°C without shaking. The vial was then placed on immediately on ice and left to stand for 5 minutes, and 250 µL of LB medium (Appendix 9.4) was added into the vial. The cells were incubated at 37°C for 2 hours with shaking (200 rpm). The cells from each transformation were then spread 20, 40 and 100 µL on a pre-warmed LB plates containing 50 µL mL^{-1} carbenicillin and 120 µL mL^{-1} Xgal. Then, the culture was incubated at 37°C for overnight.

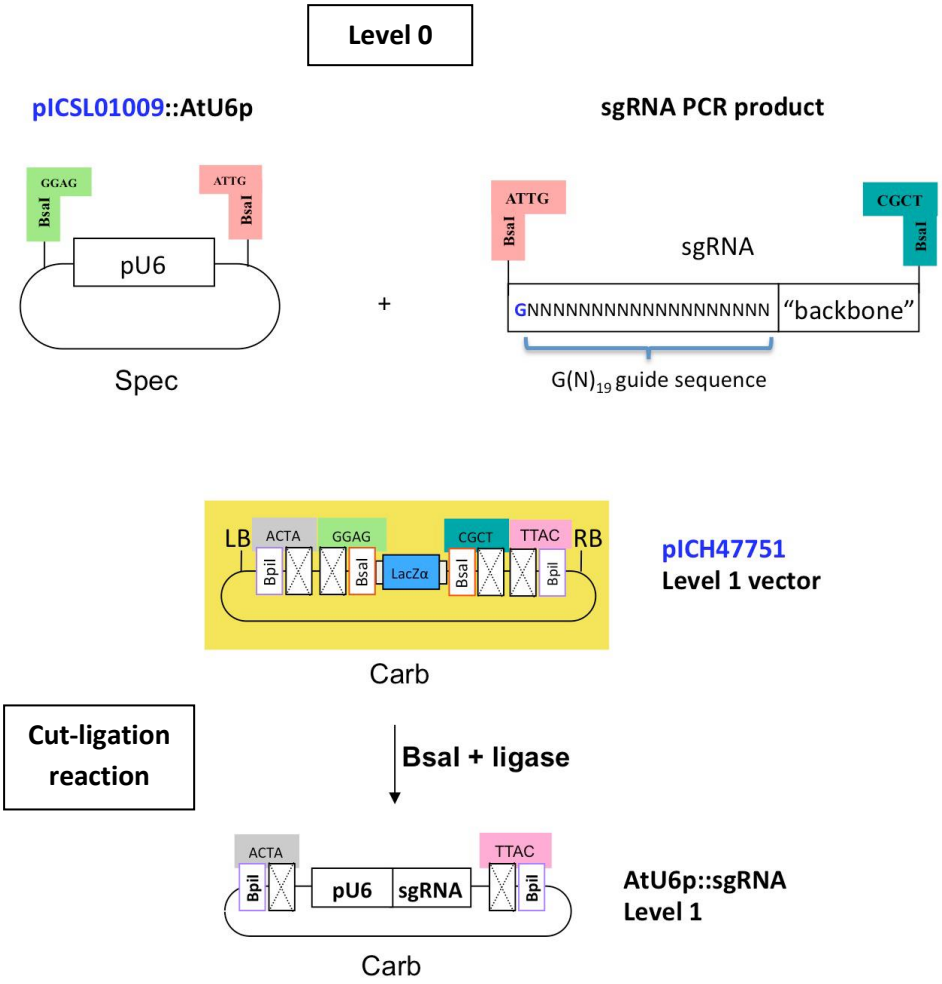


Figure 2-1 Cas9/sgRNA expressing constructs. The target site was introduced into sgRNA by PCR and then combined with the U6 promoter (Level 0). Then, cut-ligation reactions were done to generate suitable Level 1 vector and transformed to Level 2 destination vector using Golden-gate cloning method. Finally, the transformation of the Level 2 construct into *Agrobacterium* was undertaken and the next step was to proceed with plant transformation.

2.9.5 Plasmid purification and sequencing

Bacterial colonies were picked and cultured in 10 mL of LB medium containing a selective antibiotic at 37°C for overnight for *E. coli*. The plasmid DNA was then isolated from the transformed cells using the PureLink Quick Plasmid Miniprep Kit (Invitrogen, Dorset, UK). The transformed cells, which were in 10 mL of LB medium containing selective antibiotic, were centrifuged at 3,200 g for 6 minutes at 25°C. The LB medium was discarded then the pellet was resuspended in 250 µL of the Resuspension solution (Appendix 9-3). The cell suspension was transferred to a microcentrifuge tube. The bacteria were resuspended completely by pipetting up and down until no cell clumps remained. Next, 250 µL of the Lysis solution (Appendix 9-2) was added and mixed thoroughly by inverting the capped tube until the mixture was homogeneous and then the tube was incubated at room temperature for 5 minutes. After that, 350 µL of the Precipitation Buffer (Appendix 9-3) was added and mixed immediately by inverting the tube 4 to 6 times. The cell suspension was then centrifuged at 12,000 rpm for 10 minutes.

The supernatant was transferred to the supplied spin column by pipetting, before centrifuging 13,200 rpm for 1 minute. The flow through solution was discarded then the column was placed back into the same collection tube. 700 µL of the wash solution (Appendix 9.3) was added to the spin column which was then centrifuged for 60 seconds at 13,200 rpm and the flow-through discarded. The column was then placed back into the same collection tube and the previous step repeated. The column was

then centrifuged 13,200 rpm for 1 minute to remove residual wash solution. The spin column was transferred into a fresh 1.5 mL microcentrifuge tube. 75 µL of Elution buffer (Appendix 9-3) was added to the center of spin column membrane and incubated for 1 min at room temperature before centrifuging for 2 minutes. The 1.5 mL microcentrifuge tube was then placed in a freezer at -20°C.

2.9.6 Ligation CRISPR Level 2

This reaction was set up using the reagents in the order shown to transform chemically competent *E. coli*. The cut-ligation reaction was in a total volume of 12 µL in a 0.5 mL Eppendorf tubes as detailed in Table 2-1. The reaction was incubated at room temperature (22 to 25°C) overnight before transforming into competent cells. The transformed cells were then plated on LB plate (50 µL ml^{-1} kanamycin) and incubated at 37°C overnight. The next day, a single colony was picked and plasmid purification undertaken for sequencing. The isolated plasmid DNA was analysed by PCR to confirm the presence and correct orientation of the insert. The PCR products were separated by electrophoresis in 2% (w/v) agarose gel in 0.5X TAE at 100 V for 30 min. The PCR reaction was in a total volume of 10 µL in a 0.5 mL Eppendorf tubes composed of 2 µL of water, 5 µL of PCR buffer (2X Q5 Master Mix, NEB), 1 µL of forward primer and 1 µL reverse primer (100 ng μL^{-1}). The conditions for the PCR were, denaturing step of 98°C for 5 minutes, 30 cycles of 98°C for 10 seconds, 60°C for 30 seconds, and 72°C for 1 minute, and a final extension step of 10 minutes at 72°C.

Table 2-1 Cut-ligation reaction mixture and reaction conditions

10X CutSmart™ Buffer (NEB)	1 µL
10 mM Adenosine 5'-Triphosphate (ATP) (NEB)	1 µL
20 unit ml^{-1} BpiI (BbsI) (NEB)	1 µL
3000 unit ml^{-1} T7 DNA Ligase (NEB)	1 µL
10X NEBuffer 2 (NEB)	1 µL
pICH47732::NOSp::NPTII-OCST	1 µL
pICH47742::35Sp::Cas9-NOST	1 µL
pICH47751::AtU6p::sgRNA	1 µL
pICH41766	1 µL
pAGM4723	1 µL
Sterilised water	2 µL

Step	Temperature	Time	Cycle
Cycle 1	37°C	5 minutes	20
Cycle 2	20°C	5 minutes	20
Extension	70°C	20 minutes	1
Hold	22°C	~	1

2.10 Transformation of recombinant destination vectors into electrocompetent *Agrobacterium*

Fifty ng of plasmid DNA in 2 µL was added into 50 µL of *Agrobacterium tumefaciens* C58 competent cells, which were then incubated for a few min on ice. *Agrobacterium* cells were transferred to a chilled electroporation cuvette. The cuvette was placed in the electroporation apparatus and subjected to a single pulse of 2.2 kV for 5 ms. 300 µL of LB medium was then added to the cuvette and pipetted up and down. The cell suspension was transferred to 1.5 mL microcentrifuge tube before incubating at 28°C for 2 hours with shaking. Next, the incubated cells from each transformation were spread as 1 (adding 10 µL of LB medium), 20 and 100 µL aliquots on pre-warmed LB plates containing selective antibiotic and 50 ng mL^{-1} of rifampicin and incubated at 28°C for 2 days.

Bacterial colonies were picked and cultured in 8 mL of LB medium containing selective antibiotic and 50 ng mL^{-1} of rifampicin. The cells were incubated at 28°C for 2 days. After that, the plasmid DNA was isolated from the transformed cells. The isolated plasmid DNA was analysed by PCR to confirm the presence and correct orientation of the insert. The PCR products were separated by electrophoresis in 1% (w/v) agarose gel in 0.5X TAE at 100 V for 30 minutes.

2.11 Plant Transformation

To generate explant material for transformation, tomato seeds were sterilized for 10 minutes with 50% bleach solution (Sainsbury's Supermarkets Ltd, UK). The seed were then repeatedly washed with sterile water and then placed on KCMS medium (Murashige and Skoog, 1962) (Appendix 9.4). The Petri dishes were kept at 4°C in refrigerator for 1 day before being transferred to the tissue culture room. Seedlings were grown at 21°C for 9-10 days under fluorescent light. 10 mL of LB medium including 50 µg mL^{-1} kanamycin was then inoculated with *Agrobacterium*, containing the plasmid of interest and incubated at 28°C for 2 days with shaking (250 rpm). On the same day, the cotyledons were cut at both extremities and put into liquid MS (Appendix 9.4). Then the cotyledons were laid upside down in Petri dishes containing solid 2Z media (Appendix 9.4). The Petri dishes were placed in the dark at $22^{\pm}1$°C in a growth chamber for 24 hours. The cultured-*Agrobacterium* was centrifuged at 3000 rpm for 10 min. It was then resuspended in liquid MS to an optical density (OD_{600}) close to 0.2-0.3. The cotyledons were collected from the 2Z Petri dishes, and soaked in the bacterial suspension for 15 min with shaking. The cotyledons were then dried on a sterile filter paper and laid again on the Petri dishes with solid 2Z media. Petri dishes were incubated in the dark for 48 hours in a growth chamber. The cotyledons were placed upside down on Petri dishes of 2Z medium with 100 mg L^{-1} kanamycin and 80 mg L^{-1} carbenicillin (Appendix 9.4) for 15 days for regeneration of plantlets. The Petri dishes were sealed with tape and transferred into a

growth chamber. The cotyledons were transferred to new medium at least every 15 days. Callus and shoots normally were seen within 3 weeks after transformation.

For further shoot organogenesis, the callus and shoots were excised away from dying cotyledons and placed on Petri dishes of 2Z medium as mentioned above. After further 3 weeks, shoots with meristems were cut from the callus and transferred to antibiotic containing rooting media, Appendix 9.4. Shoots were found to root within 7-10 days. They were then transferred into wet soil in 1.5 L pots and the agar were rinsed from the roots with sterile water. The plantlets were then grown in a 60 x40 cm of a plastic box with a lid for 7 days in the growth room at 22°C during day time with 16 hours light and 20°C at night. The lid of the plastic box was then removed and when the plants were of a reasonable size, they were transferred into bigger pots and placed in the glasshouse.

2.12 Colour Measurements

Fruit colour was measured using a Minolta colorimeter CR400 (Minolta, Japan). Readings were taken based on the L*, a* and b* Hunter colour scale (Figure 2.2). A tomato colour index (TCI) value was calculated using equation 2.2 (Richardson and Hobson, 1987).

$$TCI = \frac{2000a^*}{L^*(a^{*2} + b^{*2})^{0.5}} \qquad (2.2)$$

2.13 Texture Measurement

For fruit texture measurements a 6-mm transverse section was cut from each fruit, and the maximum load was measured using a Lloyd Instrument LF plus machine equipped with a 10-N load cell and 1.6-mm flat-head cylindrical probe. Maximum load was the force required to penetrate the pericarp tissue at 10 mm min^{-1} to a 4mm depth. Measurements were taken separately from the outer and inner pericarp and in duplicate. For the purposes of this study, outer pericarp was defined as below the skin but before the vascular boundary. Inner pericarp was defined as the cells between the vascular boundary and the endodermis. This approach followed the protocol described in Chapman et al. (2012).

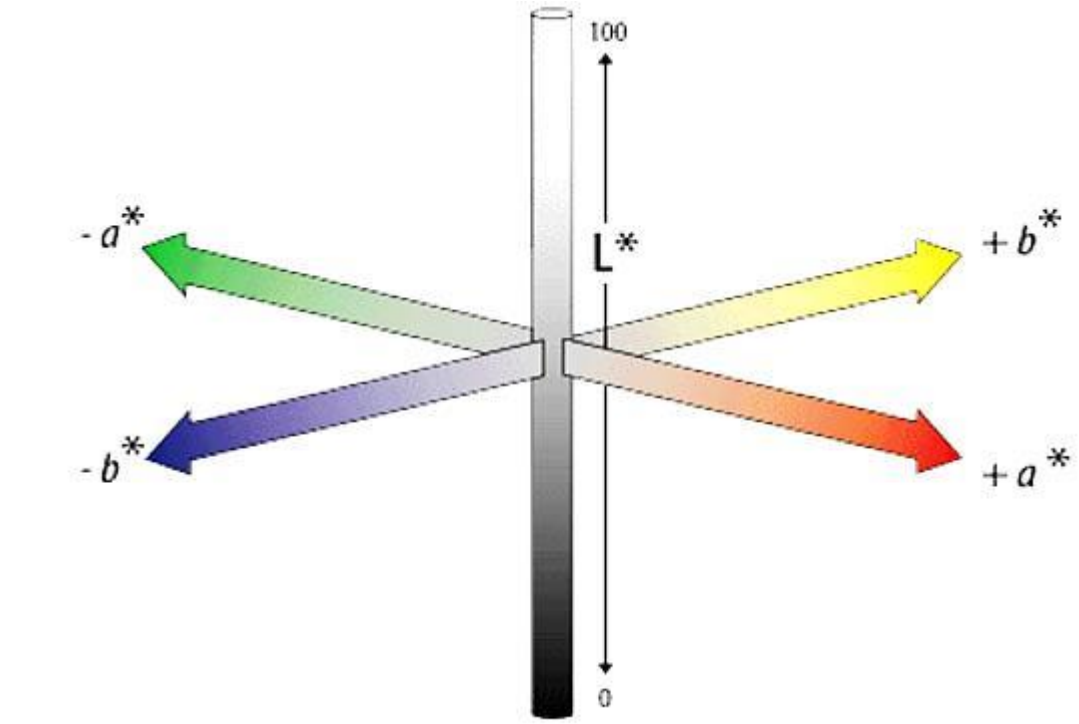


Figure 2-2 The basis of colour perception using Hunter L*, a*, b* colour space. The L* represent two different value which are white (maximum) 100 and black (minimum) 0.The a* and b* have no specific limit. Positive a* is red, negative a* is green. Positive b* is yellow and negative b* is blue. From http://www.mtk.nyme.hu/.

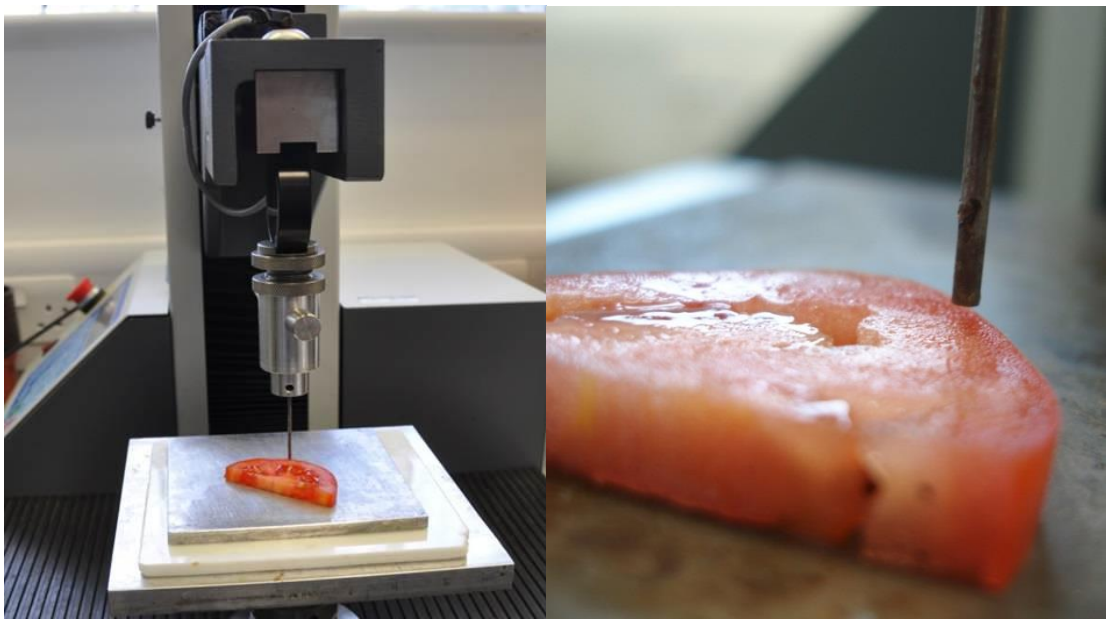


Figure 2-3 The TAPlus (Lloyd Instruments) performance food texture analyser. A 6-mm transverse section was cut from each fruit, and the maximum load reguired to penetrate the sample to a depth of 4 mm was measured with a calibrated texture tester run by NEXYGENPlus software.

Chapter 3

Comparative Genomics Analysis

CHAPTER 3: COMPARATIVE GENOMICS ANALYSIS

3.1 Introduction

The revolution in DNA sequencing technology in the last 10 years has enabled the sequencing and assembly of hundreds of genomes from organisms across the tree of life. These genomes include those from a wide range of fruit species (see Chapter 1). The first objective of this project was to undertake a comparative genomics study to reveal similarities and differences in the types of cell wall-structure genes expressed in a range of fleshy fruit species. The fruit species chosen were tomato (*Solanum lycopersicum*), melon (*Cucumis melo*), grape (*Vitis vinifera*) and banana (*Musa acuminata*). Tomato is the model for investigations on the mechanistic basis of ripening and will be our experimental platform and the genome assemblies for melon, grape and banana were of sufficient quality, annotated and with expression data to allow the bioinformatics analysis. The approach that was undertaken was to identify related gene families and identify orthologous genes across these species.

Gene families can be defined as a set of genes inherited from a common ancestor that maintain their sequence and functional similarities (Doolittle, 1981). This concept includes the gene paralogs (pairs of genes with similarities in sequence in the same species) and orthologs (genes that share similar sequences and have the same function in different

species) (Figure 3-1). Genes in the same family are expected to maintain their molecular structure and biochemical functions in different organisms (Chothia and Lesk, 1986) and sequence clustering helps to identify the gene family.

Identification of gene family members involves a search for all pairs of sequences that have similarities (Liu and Rost, 2003). Genes in the same group can be assumed to have an evolutionary relationship and similarities in sequence, biochemical function, and protein structure (Enright et al., 2002). The identification of a gene family can help the functional annotation of gene sequences where no experimental information exists. Annotation of genes based on the information in a gene family is also more accurate than annotations based on the similarity of individual genes (Devos and Valencia, 2000). Comparative genomics provides new and powerful tools for comparison between gene families at a large-scale (Lander et al., 2001; Lynch and Conery, 2003; Snel et al., 2002; Tatusov et al., 1997).

Studying relationships between genes within a family can provide important structure-function information and provides evidence for ancient genome duplications and neofunctionalisation as is apparent in tomato (Tomato Genome Consortium, 2012). The changes in number of genes that are involved in certain biological processes could occur in several scenarios, for example, gene duplication that leads to gene families with multiple copies of genes which encode the same or related

functions. The classic model in molecular biology assumes that duplication of a gene will generate several new genes that are free to be mutated, as long as one of the original genes retains its function (Ohno, 1970). The most likely outcome for these paralogues is that they degenerate into pseudogenes that are not transcribed (*nonfunctionalization*) (Nei and Roychoudhury, 1973; Petrov and Hartl, 2000). Alternatively paralogs may still be functional and through mutation diverge from the original function (*neofunctionalization*) (Ohno 1970). The tomato genome exhibits characteristics that are consistent with an ancient triplication event, gene loss and neofunctionalisation related to genes controlling ripening (Tomato Genome Consortium, 2012). The same finding also has been reported in banana genome where there have identified three round of duplication in the *Musa* lineage (D'Hont et al., 2012).

3.1.1 Comparative Genomic Software

There are many software programmes that are available for comparative genomics analysis and these include Inparanoid (O'Brien et al., 2005), OrthoMCL (Li et al., 2003) and VISTA (Frazer et al., 2004). In this study, Inparanoid Version 4.1 in combination with Multiparanoid script was used to identify all orthologous groups of genes among fleshy fruit (tomato, melon, grape and banana). The Inparanoid algorithm was used to search orthologous sequences that have paralogs within and between genome sequences (Berglund et al., 2008).

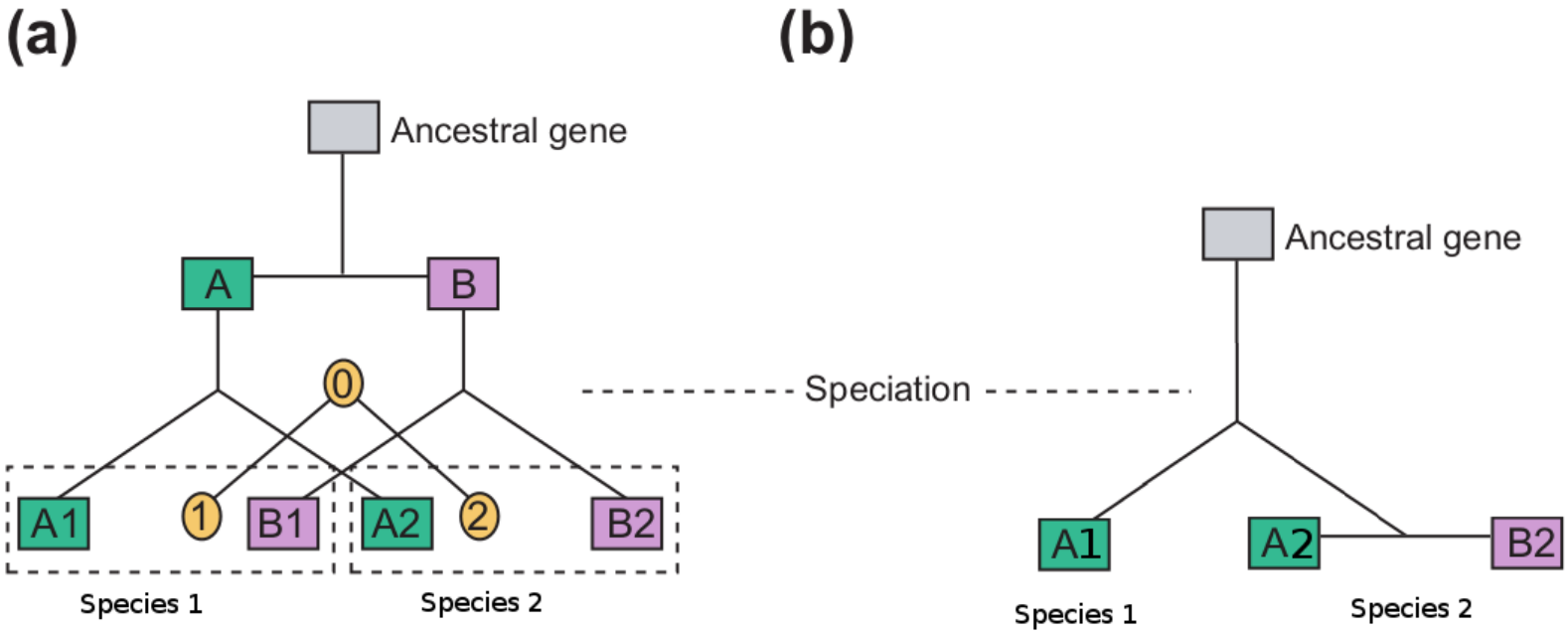


Figure 3-1 Evolutionary descent of an ancestral gene to paralogs and orthologs following gene duplication in species 0, and then speciation to yield species 1 and 2. (a) The resulting relationship between paralogs and orthologs. The ancestral genome had two copies of this gene (A and B) which were paralogs. At some point, the ancestral species split into two daughter species, each of whose genomes contained copies of the ancestral duplicated gene (A1, A2 and B1, B2). (b) This diagram indicates duplication occurs after a speciation. Therefore, A2 and B2 are orthologs of A1 and B1; A2 and B2 are paralogs of each other. The image edited from Jensen (2001).

The approach used by this software is like reciprocal–blast. Figure 3.2 shows the principle of the algorithms used here. The protein sequences from the genomes of two species are included as inputs in Fasta format. Blastp search all-against-all sequence pairs were carried out and the sequences with the best hits were identified. Finally, the various sequences are grouped and reveal the orthologous pairs. The orthologous pairs that have been identified between the genomes are based on the best match (contains the lower E-values). The matching is then used to identify protein inparalogs and outparalogs (Remm et al., 2001). Inparalogs are defined as paralogous genes which are also known as co-orthologs and outparalogs are defined as paralogs that were duplicated before the speciation and are not necessarily all present in the same species (Figure 3.2).

The cell wall genes in melon, grape and banana were compared with those in tomato. The hypothesis to be tested was that cell wall structure-related genes that are common to all fleshy fruits and expressed during softening are likely to be key targets to allow the manipulation of fruit firmness. These genes will be the targets of our CRISPR / Cas9 experiments in transgenic tomato.

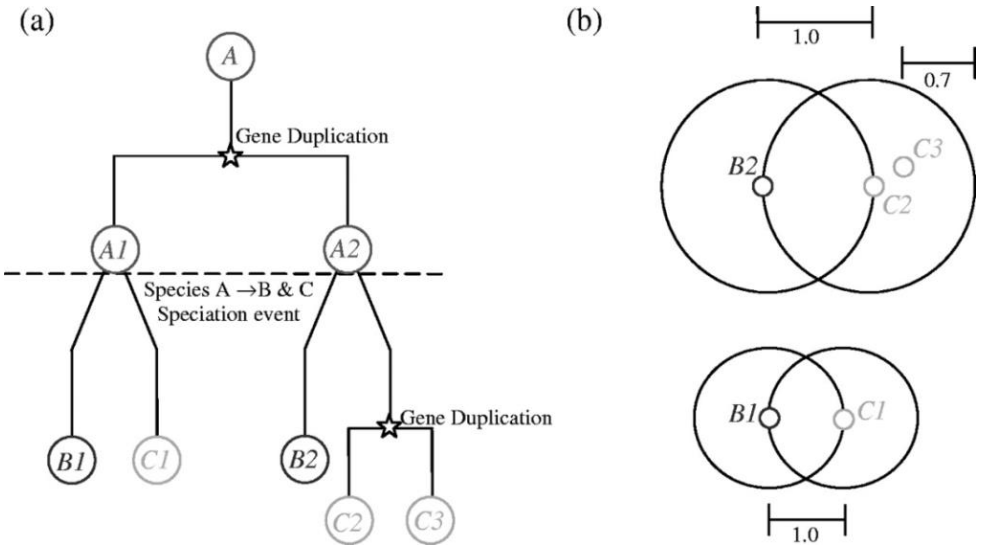


Figure 3-2 A hypothetical gene tree and the resulting Inparanoid clusters are shown to illustrate inparalog (and thus co-ortholog) and outparalog assignments.

(a) Protein A in an ancestral species 'A' undergoes gene duplication. A speciation event occurs which gives rise to the two lineages leading to species 'B' and 'C'. In the C genome the genes C2 and C3 are inparalogs since their gene duplication occurred after speciation; they are co-orthologous to the B2 gene (one common ancestral protein upon speciation). B1 is an outparalog of the C2 and C3 genes, as are B1 of B2 (duplication and divergence prior to speciation).

(b) B2 and C2 are the original seed-orthologous pair (all inparalogs are clustered around this pair), thus both receiving an inparalog score of 1.0. Other inparalogs (in this case C3) are scored according to their relative similarity to the seed-inparalog (here C2). Inparalog score of C3 = (Blast [C2:C3]−Blast[C2:B2])/(Blast[C2:C2]−Blast[C2:B2]) where Blast[X:Y] is the averaged blast score between X and Y in bits. In this case C2 is relatively more similar to B2 than C3 is, and thus C3 receives a lower inparalog score (0.7). C1 and B1 are orthologous to each other but are outparalogs of the other cluster and thus form a cluster of their own.

Source: Adapted from O'Brien et al., 2005.

3.2 Material and Methods

3.2.1 Similarity Search using BLAST

BLAST (Basic Local Alignment Search Tool) is a software tool that enables the comparison of DNA sequences or protein sequences that have been registered in different databases (Altschul et al., 1990). BLAST can be used to predict the functional and evolutionary relationships between sequences as well as help in predicting the sequences of the gene family of interest. This software produces a couple of segments with high scoring segment pairs (HSPs) which represent local alignments between the sequences of interest and those sequences in the database. BLAST results are the result of the alignment with percent identities, score, E-value (Expect) and P-value (Probability) (Altschul et al., 1990).

BLAST programs used in this study were blastp (Protein-protein BLAST), blastn (Nucleotide-nucleotide BLAST) and tblastx (Nucleotide 6-frame translation-nucleotide 6-frame translation). The BlastP program compares a protein query with protein sequences that are available in the database (Altschul et al., 1990). The parameter used is the default parameter which is BLOSUM (BLOcks SUbstitution Matrix) BLOSUM62 with opening value space is 11 and the extension value space is 1. The BLAST

program that is used in this study was downloaded from NCBI database and installed into a LINUX system on an external server. The similarity search analyses were then undertaken by batch collection (*batch blast*).

3.2.2 BLAST Database Format

In this study, formatDB was used for making a custom BLAST database. FormatDB is a tool that was used in molecular bioinformatics to format protein or nucleotide databases for BLAST. The first step was assembling all of the sequences in one file and then they are saved in fasta format. In this study, we selected all of the putative cell wall structure-related genes in each genome. This file was then used to construct the index for the BLAST database using a program from NCBI using a "formatdb" command. FormatDB was the used to construct a BLAST database called "customBLASTdb".

3.2.3 Comparative Analysis using InParanoid and MultiParanoid

Inparanoid (O'Brien et al., 2005) is one of the software tools that can be used in the comparative analysis of genomes. This software is freely available and can be downloaded from [http://inparanoid.sbc.su.se/cgi-bin/index.cgi]. In this study, the genome sequences from several species of fleshy fruit, banana (*Musa acuminata*), melon (*Cucumis melo*) and grape (*Vitis vinifera*) were compared with that of tomato (*S.*

lycopersicum). Only genomes that were assembled, annotated and had associated fruit-related expression data were used. The comparative genomic analysis used the stand alone Inparanoid program software version 4.1. This allowed the generation of groups of pairwise orthologous clusters of proteomes. All the pairwise proteome data from the Inparanoid approach were then brought together by using Multiparanoid script (Appendix 9.5).

The Multiparanoid method is a combination of the Inparanoid algorithm (Östlund, et al., 2010; Remm, et al., 2001) and also the Multiparanoid algorithm (Alexeyenko et al., 2006). This method begins with finding orthologs between each pair of species and paralogs for each of the organisms that can be acceptable based on being homologous by BLAST search in their own genome. MultiParanoid is a powerful approach for searching for gene clusters among multiple strains so the pairwise orthologous clusters that had been generated from Inparanoid were directly transferred to it. The local matched region was no less than 25% of the longer gene protein sequence and the global matched region was no less than 50% of the longer gene protein sequence for each pair genes of the same cluster and the minimum score value and E-value in BLAST were 50 and $1e^{-8}$.

3.2.4 Gene Classification by Gene Ontology (GO)

Gene Ontology (GO) is a software tool that also acts as a database to aid gene classification. The main purpose of this software is to provide a comparison of the classification of genes among genomes being studied (The Gene Ontology Consortium, 2000). It overcomes issues in linking genes that have been annotated by different researchers when comparison of genes between species is required (Stein, 2001). There are three major ontology groups in the GO which are Molecular Function, Biological Process and Cellular Component (The Gene Ontology Consortium, 2000).

From Multiparanoid analysis in this study, more than 8000 genes were designated orthologous in all fruit species studied, BLAST2GO software was then used in group classification. By using this software, GO analysis could be undertaken on a large number of sequences and the results easily exported to Excel format. BLAST2GO software was downloaded from [https://www.blast2go.com/] and installed into the Windows/Linux or Mac platform using JAVA applications (Conesa et al., 2005). Analysis using BLAST2GO started with similarity searches between the query and other sequences in the NCBI database by using the Blast application. Then the GO classification process was undertaken by mapping the sequences to GO database. All of the processes were

performed using the annotation rules concept in which each annotation case is examined at certain reliability levels (Stein, 2001).

3.2.5 Manual Curation with Transcriptome Database

The results from GO classification were used to identify all genes that had a putative cell wall base on cellular component distribution. Then, by using tomato genome as our reference genome, the cell wall remodelling genes were selected out of all the proteins that were identified as orthologs in all four genomes. The curations also have been done by using all transcriptome data which focusing on those genes expressed during ripening. Figure 3.3 shows the outline of the method. All of these data were then exported into Excel format and then, sorted manually by using Vertical Lookup function (VLOOKUP).

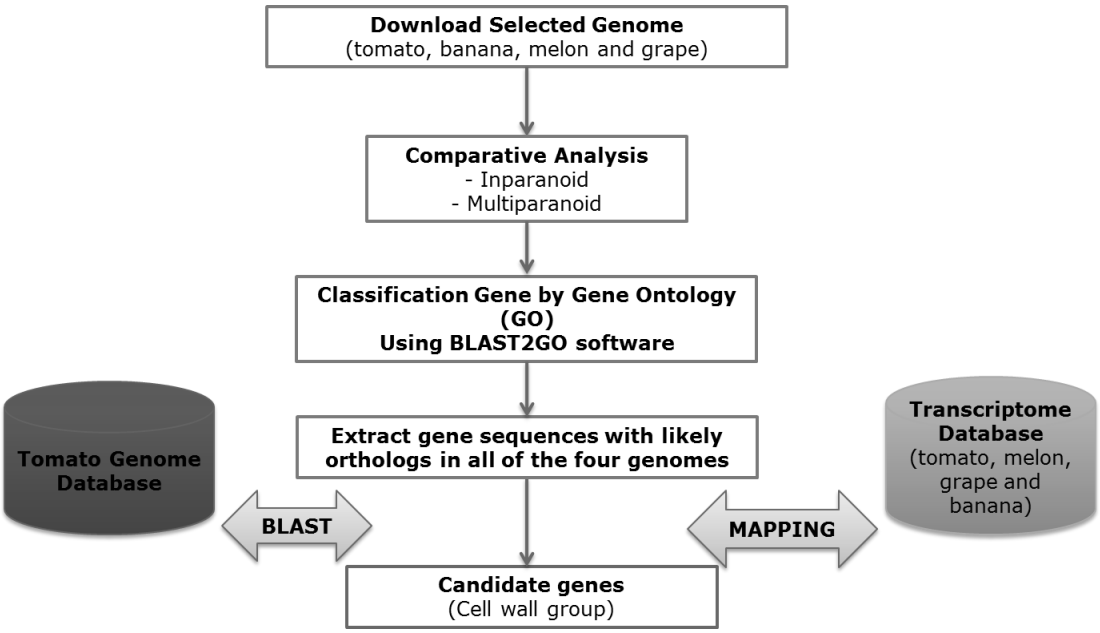


Figure 3-3 Outline of candidate gene identification approach. The combinations of inparanoid and multiparanoid were used to identify the orthologous groups in all of the genomes. The candidate genes from cell wall remodelling group were manually selected from transcriptomic data, GO classification and also annotation from tomato genome.

3.3 Results and Discussion

3.3.1 Ortholog Analysis

In this research, comparative analysis was undertaken with the detection of orthologous groups using the Inparanoid program which applies *all-versus-all* sequence comparisons of two genomes with the special rules of cluster analysis. Although phylogenetic tree is a well-established method to distinguish orthologs and has been used to study the evolution of organisms, it is time-consuming and prone to errors (Remm et al., 2001). Thus, the Inparanoid program was used as an alternative to the phylogenetic method.

The *S. lycopersicum* gene models were compared to the list of genes from *C. melo, V. vinifera* and *M. acuminata*. The orthologs for each genome that were generated from Inparanoid program were sorted and viewed using Microsoft Excel program. The numbers of genes shared among all four species were obtained and calculated. Then, the cell wall remodelling genes in this group were identified. *V. vinifera* and *C. melo* were compared with *S. lycopersicum* because all of them were dicot genomes and the genomes have been completely sequenced (Garcia-Mas et al., 2012; Jaillon et al., 2007; Velasco et al., 2007; Tomato Genome Consortium, 2012). *Musa acuminata* was used in a comparative analysis

because it is a representative for monocot fleshy fruit bearing species (D'Hont et al., 2012).

From this analysis (Figure 3.4), a total of 3,013 of the predicted *S. lycopersicum* genes have orthologs in *C. melo*, whereas 2,763 of the gene models are shared in *V. vinifera* sequences and 1,607 of the gene models are represented by orthologous sequences in *M. acuminata*. The different trends of orthologous relationships between *S. lycopersicum* with the two dicots, *C. melo* and *V. vinifera*, and with the monocot *M. acuminata* likely reflect evolutionary processes that occurred in the ancestral genomes of each group. Figure 3.4 shows that all the dicot species (*S. lycopersicum*, *C. melo*, and *V. vinifera*) share a total of 1,340 orthologous sequences while 8,982 orthologous genes were found to be common to all four species. This latter set may reflect a basic gene tool kit for adaptation in the environment (Banks et al., 2011). These sets of genes also provide the information for studies relating to the evolution of fleshy fruits species and are the foundation set for selecting the common cell wall remodelling group.

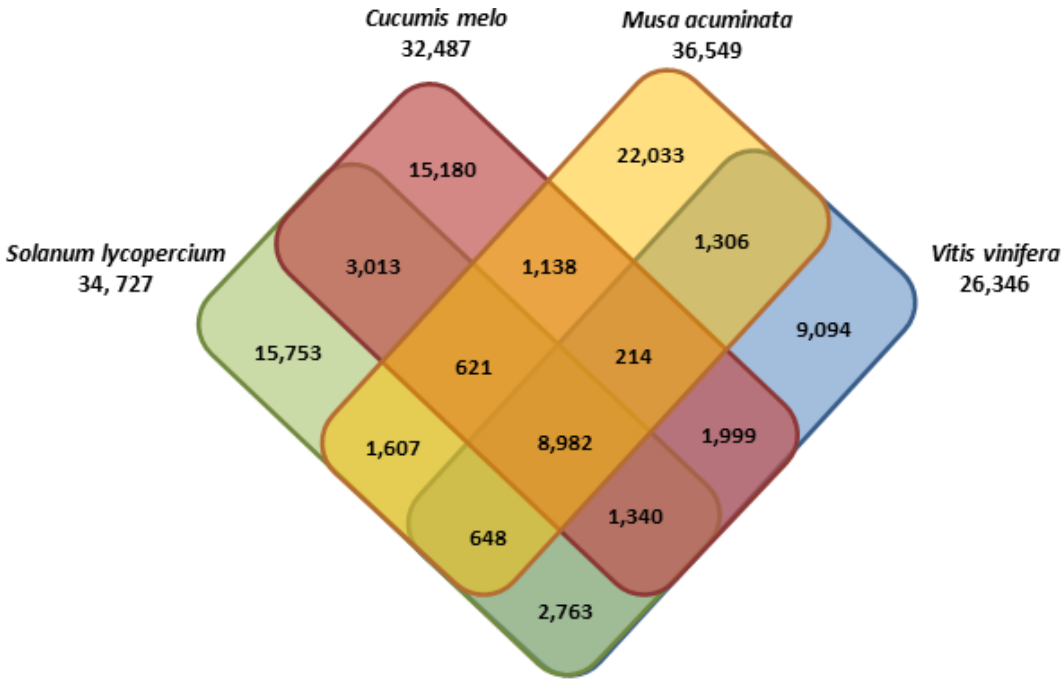


Figure 3-4 Venn diagram of the orthologous genes between *S. lycopersicum, M. acuminata, C. melo* and *V. vinifera*. Numbers in the area of overlap indicate the number of orthologs predicted by reciprocal Inparanoid v 2.0 analysis (threshold E-value = $1x10^{-5}$).

3.3.2 Gene Classification from GO

The Gene Ontology (GO) [http://www.geneontology.org] approach is probably the most widespread and the most extensive annotation scheme for the functional description of gene products (Ashburner et al., 2000). Other systems such as enzyme codes (Schomburg et al., 2004), KEGG pathways (Ogata et al., 1999), FunCat (Ruepp et al., 2004), or COG (Tatusov et al., 2003) are also widely used within molecular databases for annotation of gene sequences. In this research, we analysed the GO terms of the orthologous fruit gene set, which is all the genes that had orthologs across the four genomes studied. The functional classification of cell wall remodelling genes was carried out based on Gene Ontology classification. A straight forward mapping of gene sequences was made using homology searches (blast hit) (Conesa et al., 2005) to retrieve the GO terms associated with the hit obtained from the Blast results.

The 8,982 sequences that had significant blast hits were loaded for GO mapping in Blast2GO suite (Table 3-1). GO terms of a total of 7,791 sequences were successfully assigned based on Gene Ontology Consortium in term of biological process, molecular function and cellular components. These three terms are structured controlled vocabularies (ontologies) that describe gene products in GO (Seki and Mostafa, 2008). Biological process describes the operations and set of molecular events which are accomplished by one or more ordered assemblies of molecular functions. Meanwhile, molecular function represents the elemental activities of a gene product at the molecular level, for instance, catalytic or binding activities, and does not specify where, when or in what context the action takes place. As for cellular components this refers to parts of a cell or its extracellular environment (Gene Ontology Consortium, 2003). A total of 7,791 GO terms were loaded through mapping to Gene Ontology database. From these, all of the sequences were assigned to biological processes (Figure 3.5), molecular functions (Figure 3.6) and cellular components (Figure 3.7). GO terms could be assigned for more than one terms for one sequence and each category was divided into other subcategories.

Table 3-1 Gene Ontology classification analysis generated by Blast2GO suite.

	Quantity	**Percentage (%)**
Sequence has significant similarity (e-value ≤10^{-5})	8,982	100
Sequence has Gene Ontology assignment	7,791	86.74

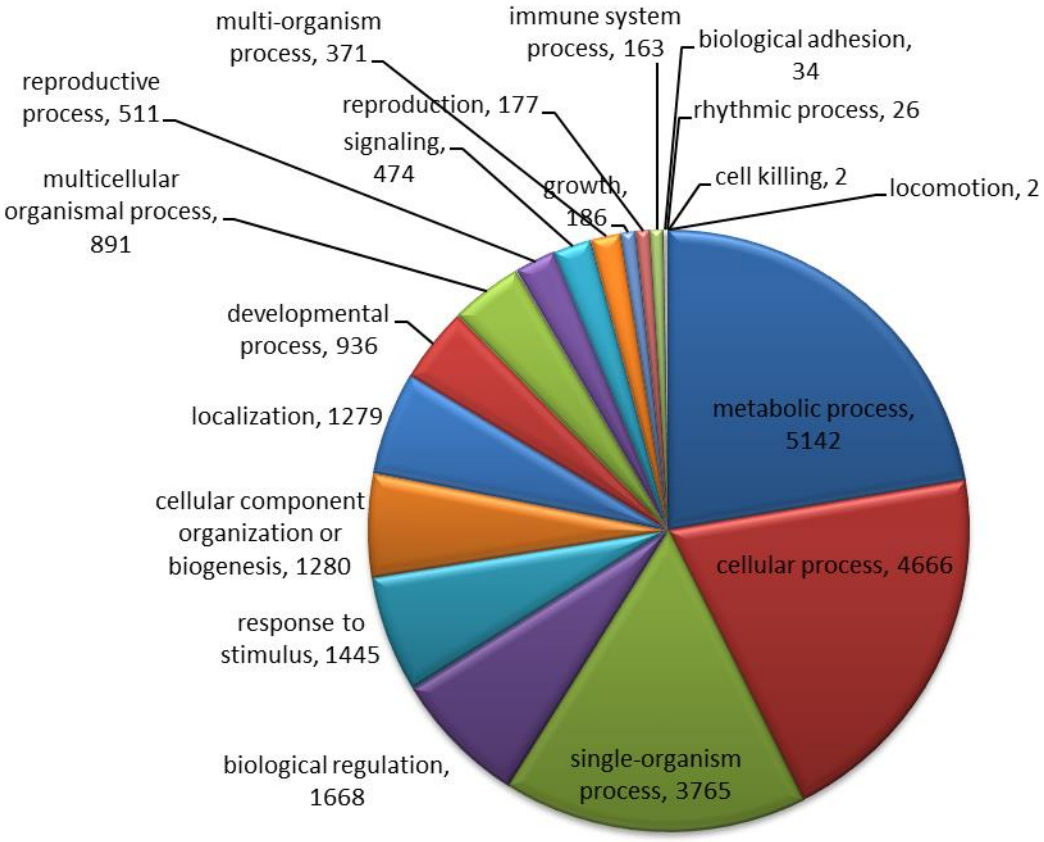


Figure 3-5 Biological process term at second level of GO classification that was generated using BLAST2GO software. There were 19 categories of second level biological process term. The metabolic process and cellular process were the most dominant second level , with 57% and 51% since these metabolic processes are crucial for plants with fleshy fruits.

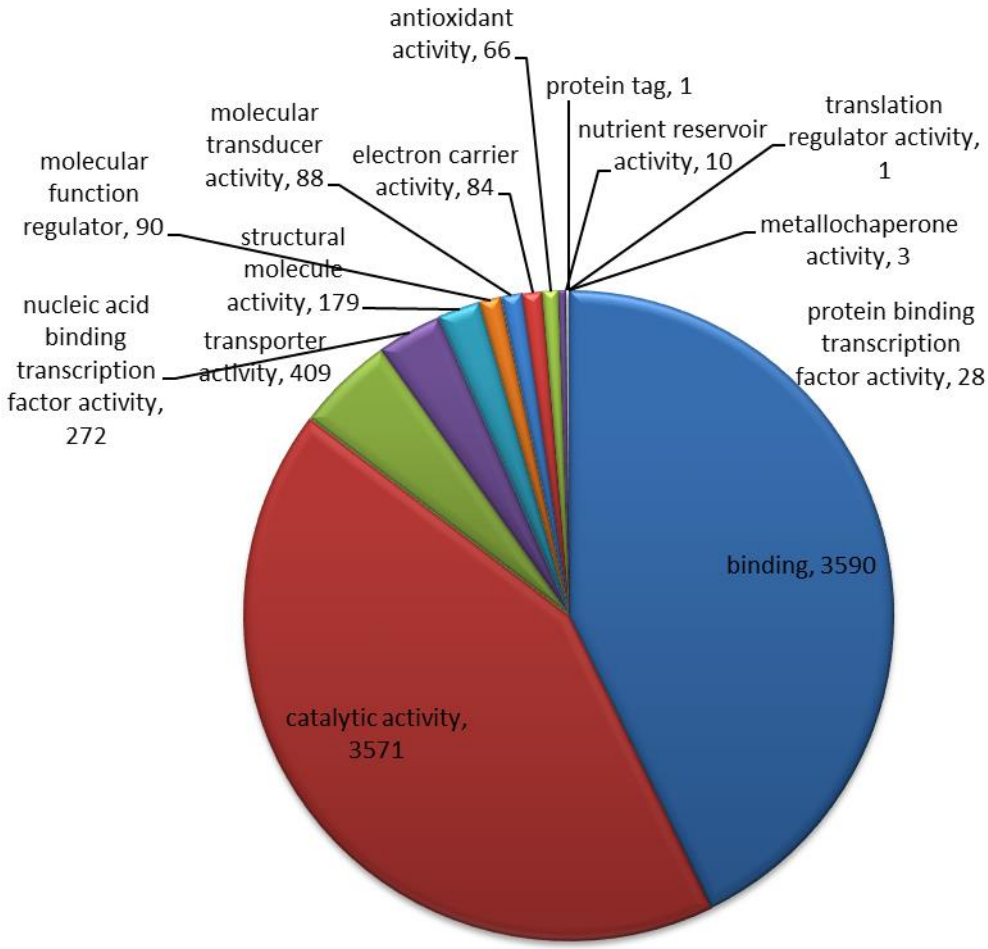


Figure 3-6 Molecular function at second level of GO classification that generated using BLAST2GO software. There were 14 categories of second level molecular function terms that have been successfully assigned for all sequences in database.

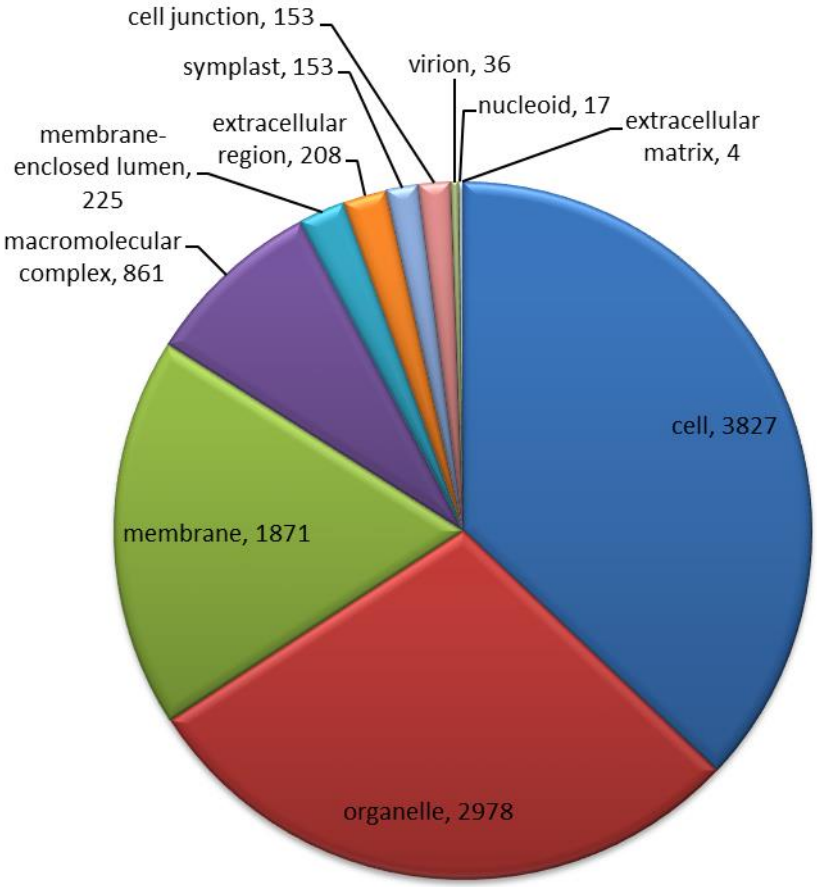


Figure 3-7 Cellular component at second level of GO classification that was generated using BLAST2GO software. There were 11 categories assigned cellular component term and most of the ortholog genes are located in cell, organelle, membrane, macromolecular complex, membrane-enclosed lumen, extracellular region, symplast, cell junction, virion, nucleoid and extracellular matrix.

3.3.3 Data Mining for Cell Wall Remodeling Genes

Fruit softening is a complex process characterized by sequential disassembly and degradation of cell wall components mediated cooperatively by cell wall modifying enzymes (Seymour et al., 2013). There is also evidence for renewed cell wall biosynthesis during this process (Vicente et al., 2007). Five GO IDs were used in the identification of cell wall genes common between all of the orthologous groups (Figure 3.8), these were GO:0005623, GO:0005618, GO:0044464, GO:0030312, and GO:0071944. Then, by using tomato as a model, comparison between the GO classification and tomato genes annotated as linked to cell wall structure, and remodelling (Tomato Genome Consortium, 2012) was undertaken manually using Microsoft Excel.

The Tomato Genome Consortium (2012) in Supplementary Table has reported there were 718 genes that were identified as cell wall structure-related genes in the tomato genome. This information was then mapped to the orthologous cell wall sequences. The mapping resulted revealed that there is 262, 261, 252 and 198 cell wall structure-related genes where sequence were highly related in tomato, melon, grape and banana. Figure 3.9, Figure 3.10, Figure 3.11 and Figure 3.12 showed the classification all of cell wall genes based on GO terms.

Although there are numerous cell wall-related genes, not all are involved in the fruit softening process. For example in tomato, out of more than 700 genes linked to cell wall metabolism, only just over 50 were expressed in developing and ripening fruits (Tomato Genome Consortium, 2012). Thus, in a study of peach and nectarine cultivars only 14 cell wall related genes changed in expression in all of the cultivars tested (Nilo et al. 2012). Cell wall classification results using GO terms were mapped to transcriptome data from each of the fruits to identify the genes expressed during fruit development and ripening.

In tomato, 52 cell wall-related genes have been identified as being expressed during fruit development and ripening. However, only 12 genes showed large changes in expression during the ripening process and these included PME, PL, PG and XET (Tomato Genome Consortium, 2012). In melon and grape, approximately 100-200 cell wall genes were identified as expressed during fruit development and ripening (Appendix 9.6) which included those that encoded polygalacturonase, pectate lyase, cellulose, and xyloglucan endotransglucosylase. In banana fruits, around 90 genes were expressed during the developing and ripening stage (D'Hont et al., 2012), with ripening-expressed genes being shown in Appendix 9.7. During ripening in bananas some of the most highly expressed cell wall genes were those encoding pectate lyase (2 genes), polygalacturonases (6 genes), pectinacetyl esterases (6 genes), xyloglucan endotransglucosylase/hydrolases (5 genes) and expansins (3 genes) (Mbéguié-A-Mbéguié et. al, 2009).

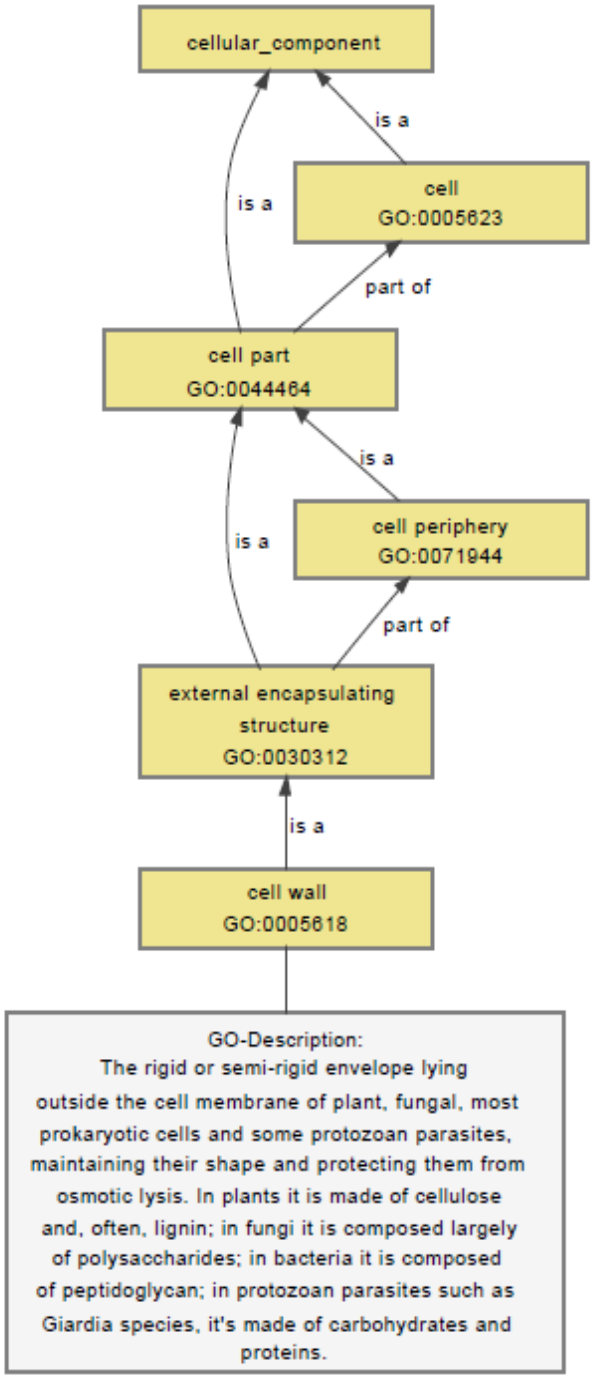


Figure 3-8. Cell wall related genes terms using Gene Ontology classification. The flow chart indicates how the hierarchical classification involved in the GO suite. There are five GO IDs classified in cell wall group which are GO:0005623, GO:0005618, GO:0044464, GO:0030312, and GO:0071944.

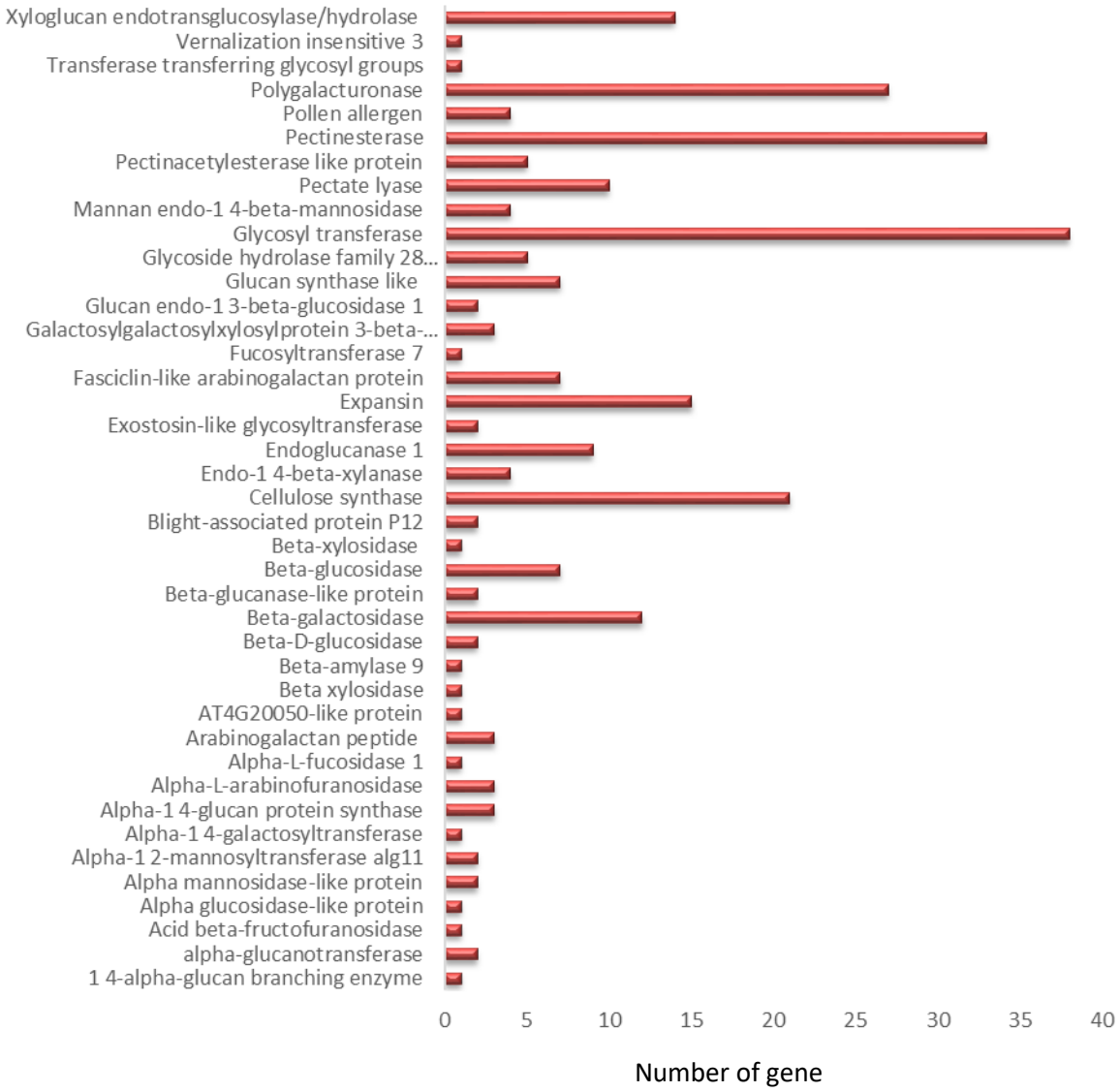


Figure 3-9 List of sequences based on cell wall function classified by Gene Ontology terms in tomato. The top five cell wall genes are glycosyl transferase (38), Pectinesterase (33), Polygalacturonase (27), Cellulose synthase (21), and Expansin (15).

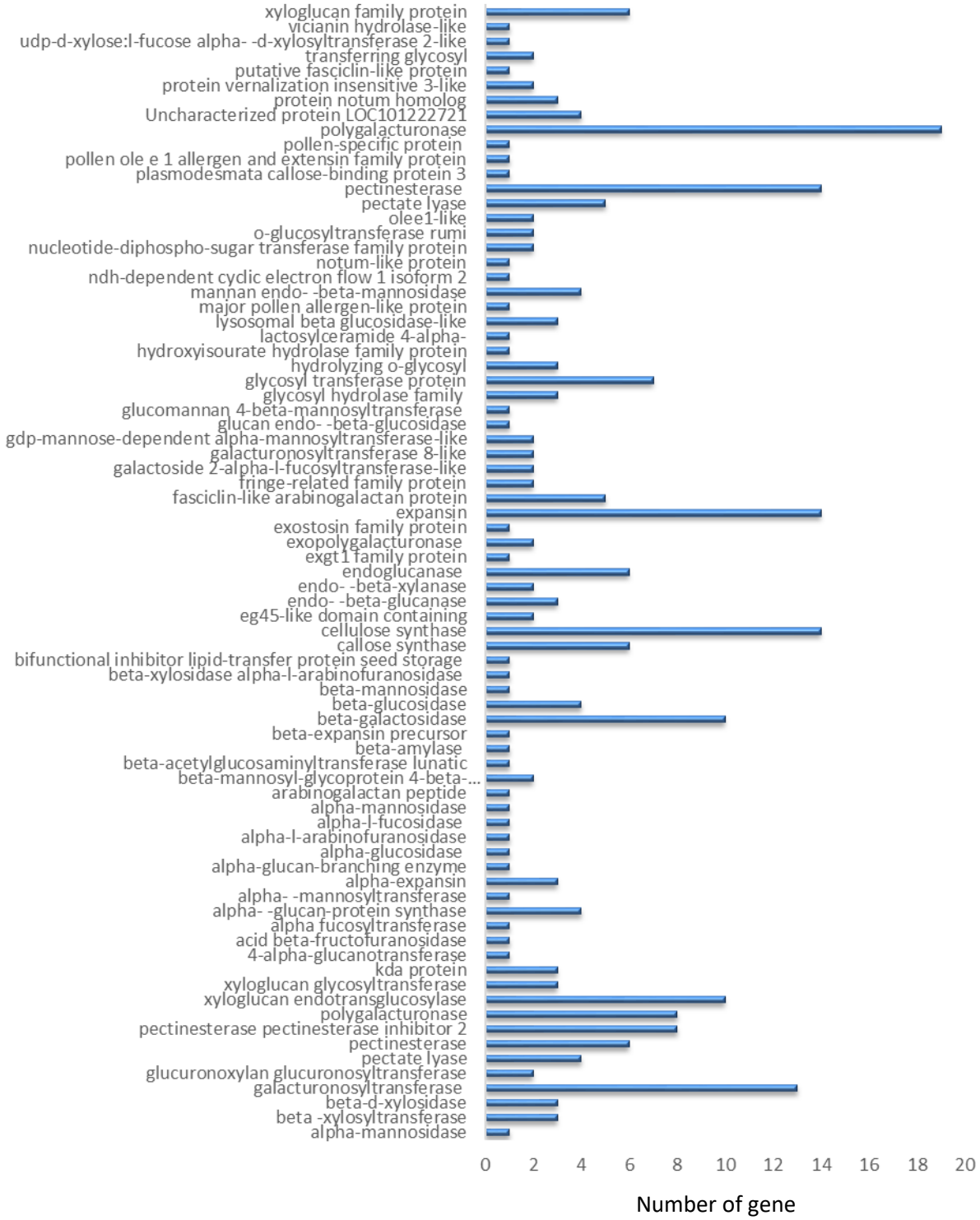


Figure 3-10 List of sequences based on cell wall function classified by Gene Ontology terms in melon. The top five cell wall genes are polygalacturonase (19), pectinesterase (14), expansin (14), cellulose synthase (14) and beta-galactosidase (10).

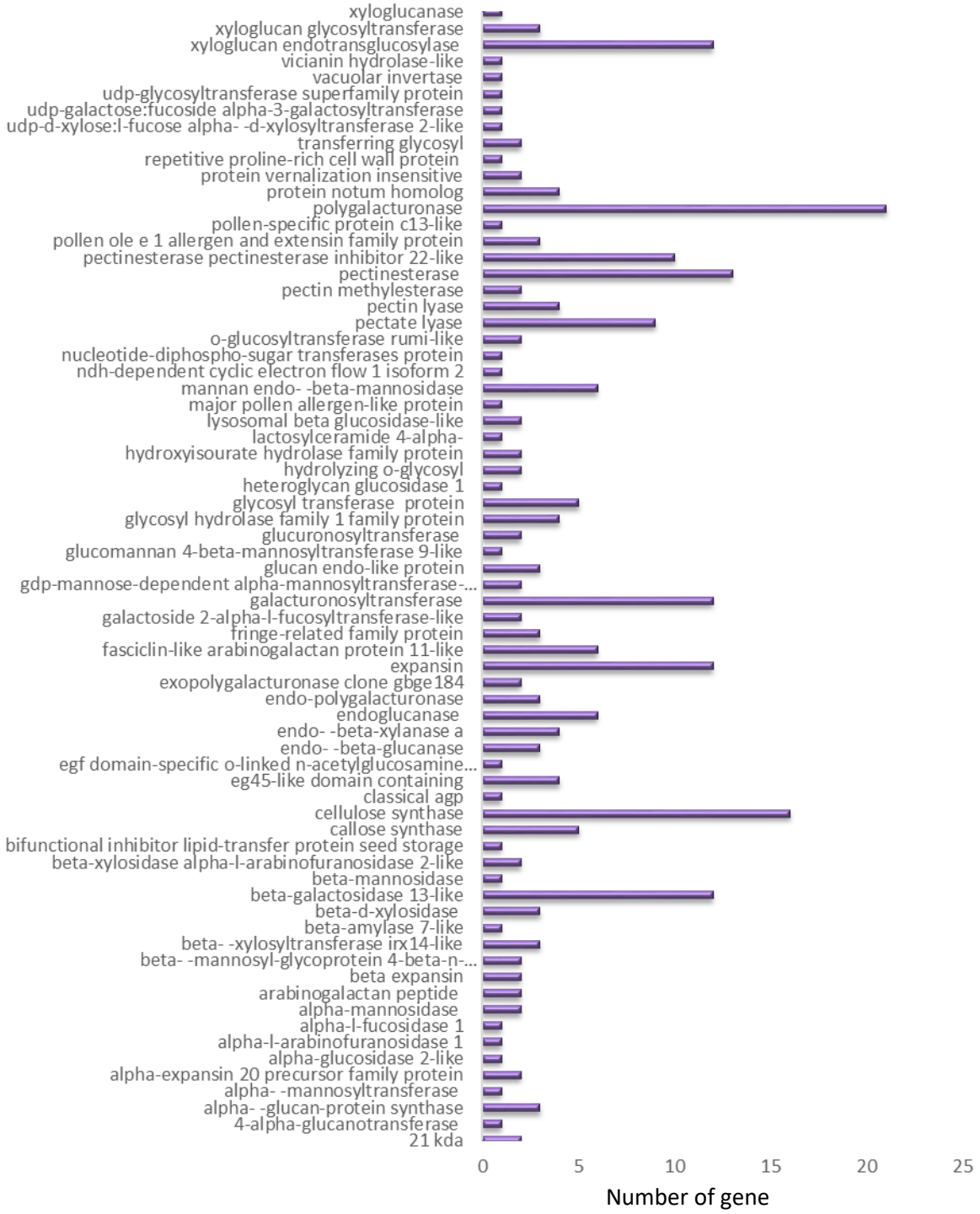


Figure 3-11 List of sequences based on cell wall function classified by Gene Ontology terms in grape. The top five cell wall genes are polygalacturonase (21), cellulose synthase (16), pectinesterase (13), expansin (12) and xyloglucan endotransglucosylase (12).

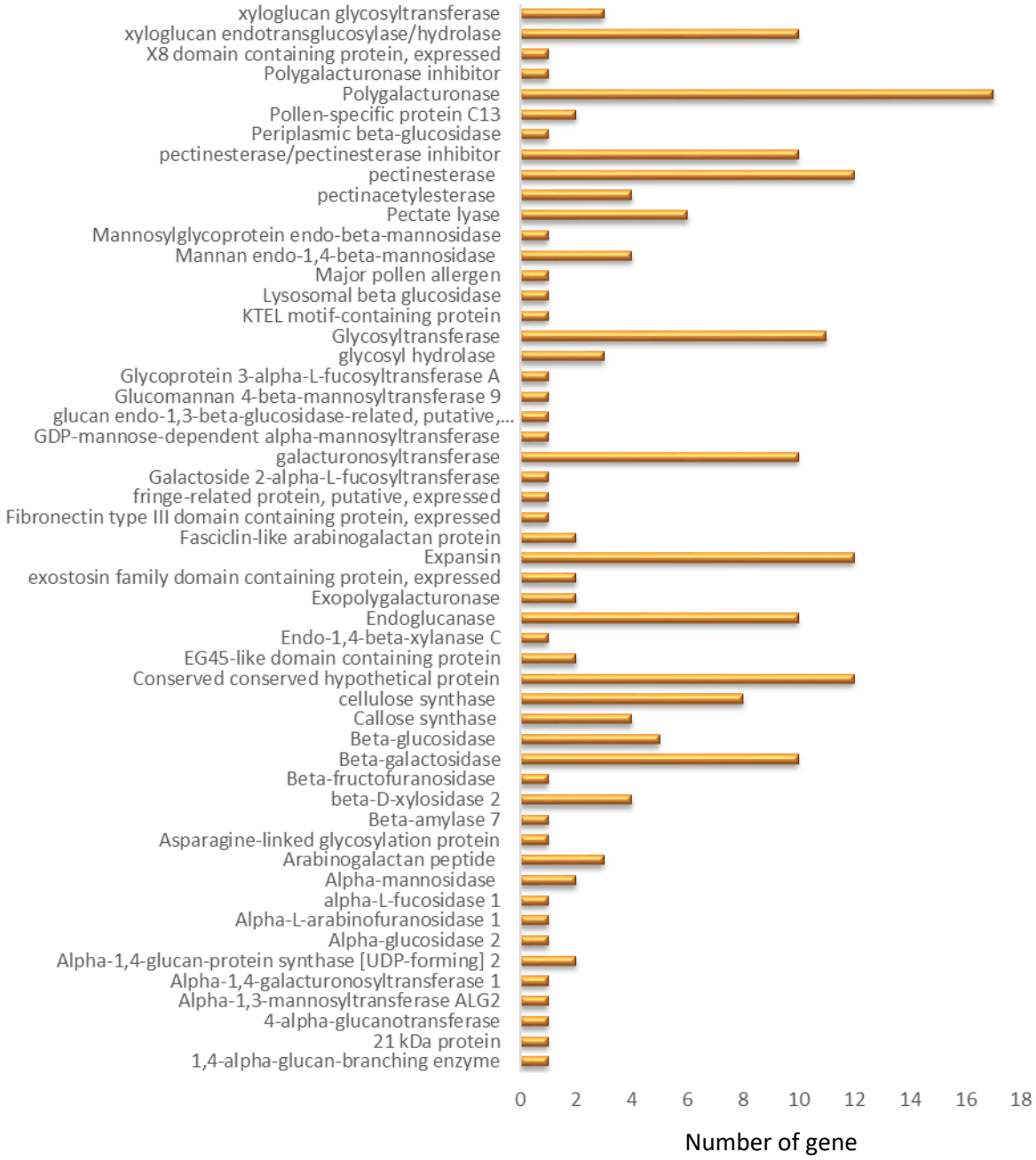


Figure 3-12 List of sequences based on cell wall function classified by Gene Ontology terms in banana. The top five cell wall genes are polygalacturonase (17), pectinesterase (12), Expansin (12), Beta-galactosidase (10) and xyloglucan endotransglucosylase/hydrolase (10).

Although the range of fleshy fruit species studied showed the expression of many of the same families of cell wall-related genes, the most surprising observation was that only a small number of truly orthologous genes were apparent that were expressed in even three of the four these species (Table 3-2). These include β-glucosidase, cellulose synthase, expansins, polygalactruonase and pectate lyase which were expressed in a wide range of fruits including others not studied in detail here such as apple (Atkinson et al. 2012) and strawberry (Jiménez-Bermúdez et al., 2002; Vladimir et al., 2011)

In tomato there are three *PL* genes that are expressed during fruit development and ripening (Solyc03g111690, Solyc05g014000 and Solyc06g083580) (Figure 3-13). The only one of these genes that has a close ortholog in melon, grape or banana was Solyc05g014000. However, this gene is expressed in tomato principally, during fruit development (Tomato Genome Consortium, 2012; D'Hont et al., 2012). Solyc06g083580, is expressed only in developing tomato fruits, but expression during ripening is very low (The Tomato Genome Consortium, 2012; Supplemental data) which fruit expressed orthologues in melon and grape (Garcia-Mas et al., 2012; Jaillon et al., 2007). Interestingly, Solyc03g111690 is highly expressed in tomato during fruit ripening, but it does not have a close ortholog in melon, grape or banana. Pectate lyase was shown to be important in fruit softening in banana (Marín-Rodríguez et al., 2003), strawberry (Shulaev et al., 2011), apple (Velasco et al.,

2010) and very recently work in the Seymour lab has shown it is very important in tomato (Uluisik et al., 2016).

Tomato *PG* is perhaps the best know pectin degrading enzyme in tomato and is encoded by the gene Solyc10g080210. Orthologues of this gene are present in melon, grape and banana and fruit related expression of these orthologs occurs in tomato, melon and grape. In apple the *PG* most highly expressed during ripening (Atkinson et al. 2012) was more closely related to the tomato gene Solyc05g049980 (Figure 3-14) which was lowly expressed in ripening tomato fruits. In addition, Solyc06g009200, which is not expressed in developing or ripening tomato fruits, is orthologous to a gene that modulates softening in strawberry (García-Gago et al., 2009). These data help highlight that species utilize a range of gene family members during cell wall disassembly.

Table 3-2 Cell wall structure-related genes expressed in tomato and likely orthologues in melon, grape and banana. These genes are expressed during fruit development and ripening.

Gene Description	Short name	Tomato Gene ID	Melon Gene ID	Grape Gene ID	Banana Gene ID
Xyloglucan endotransglucosylase/hydrolase 9	SlXTH3d	Solyc03g093080	MELO3C017478P1	GSVIVT01029162001	GSMUA_Achr10P30870_001
Xyloglucan endotransglucosylase/hydrolase 14	SlXTH5	Solyc01g081060	MELO3C003441P1	GSVIVT01013055001	GSMUA_Achr1P23980_001
Xyloglucan endotransglucosylase/hydrolase 5	SlXTH16	Solyc07g052980	MELO3C018785P2	GSVIVT01000416001	GSMUA_Achr9P29980_001
Xyloglucan endotransglucosylase/hydrolase 2	SlXTH26	Solyc05g005680	MELO3C012004	GSVIVT01020228001	GSMUA_Achr3P10480_001
Xyloglucan endotransglucosylase/hydrolase		Solyc03g031800	MELO3C002480P1	GSVIVT01012635001	GSMUA_Achr1P08500_001
Beta-galactosidase	TBG6	Solyc02g084720	MELO3C007872P1	GSVIVT01018853001	GSMUA_Achr9P21610_001
Polygalacturonase A	PG-2a	Solyc10g080210	MELO3C023556	GSVIVT01033303001, GSVIVT01033364001	GSMUA_Achr10P08930_001
Polygalacturonase		Solyc05g049980	MELO3C006092P1	GSVIVT01019405001	GSMUA_Achr11P02870_001
Pectinesterase inhibitor		Solyc06g009190	MELO3C015963P1	GSVIVT01028041001	GSMUA_Achr11P05430_001
Pectinesterase inhibitor		Solyc07g017600	MELO3C013699P1	GSVIVT01023135001	GSMUA_Achr7P15620_001
Pectate lyase 1-27		Solyc06g083580	MELO3C012791	GSVIVT01028548001	GSMUA_Achr7P04580_001
Pectate lyase		Solyc05g014000	MELO3C002319P1	GSVIVT01000592001, GSVIVT01007582001	GSMUA_Achr6P28260_001
Mannan endo-1 4-beta-mannosidase		Solyc02g084990	MELO3C007842	GSVIVT01009746001, GSVIVT01018923001	GSMUA_Achr4P02940_001
Glucan endo-1 3-beta-glucosidase 1		Solyc03g115200	MELO3C017220	GSVIVT01007873001	GSMUA_Achr4P32790_001

Fasciclin-like arabinogalactan protein 19		Solyc07g045440	MELO3C024192P1	GSVIVT01014684001	GSMUA_AchrUn_randomP25500_001
Fasciclin-like arabinogalactan protein 10		Solyc10g005960	MELO3C024938P1	GSVIVT01030085001	GSMUA_AchrUn_randomP25500_001
Expansin	LeEXP1	Solyc06g051800	MELO3C025907P1, MELO3C003134P1	GSVIVT01024946001	GSMUA_Achr2P16370_001
Expansin (EXPA3)		Solyc03g031840	MELO3C015695P2	GSVIVT01023857001	GSMUA_Achr1P19730_001
Endoglucanase 1	Cel8	Solyc08g082250	MELO3C016287P1	GSVIVT01019523001	GSMUA_Achr4P08520_001
Endoglucanase 1		Solyc04g081300	MELO3C003760P1	GSVIVT01009881001	GSMUA_Achr4P19910_001
Cellulose synthase-like		Solyc11g066820	MELO3C017935P1	GSVIVT01028071001	GSMUA_Achr3P24160_001
Cellulose synthase		Solyc01g087210	MELO3C023114	GSVIVT01033297001	GSMUA_AchrUn_randomP09460_001
Cellulose synthase		Solyc08g061100	MELO3C003689	GSVIVT01035830001	GSMUA_Achr5P06050_001

***Note**: - colour coding show genes expressed in fruit (green) or no fruit expression but other tissue (red)*
-The rows that highlighted with grey were our candidate genes that were used for CRISPR-Cas9 knockouts

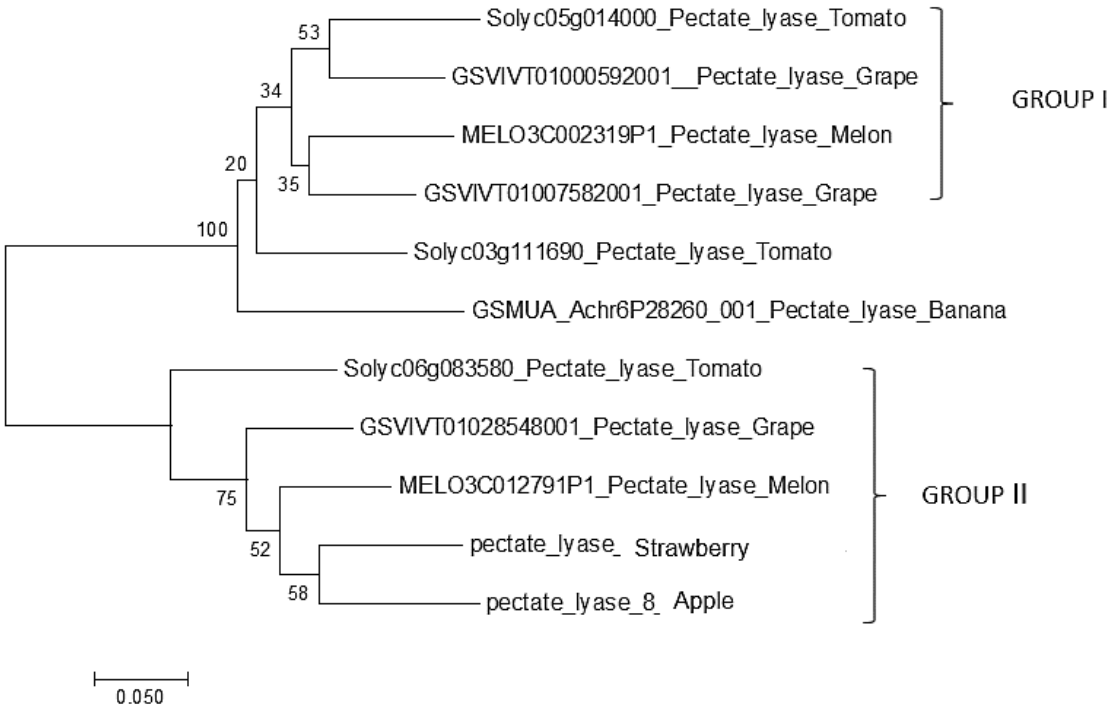


Figure 3-13. Molecular phylogenetic tree of pectate lyase (*PL*) amino acid sequences and the similar amino acid sequences in studied fruits. The dendrogram was generated by Mega 7.0 software using MUSCLE for the alignment and the maximum likelihood method for the construction of the phylogeny. Bootstrap tests were performed using 1,000 replicates and the percentage of the bootstrap value are shown in each branch where the value exceeds 50% is considered significant. The branch lengths are proportional to the phylogenetic distances.

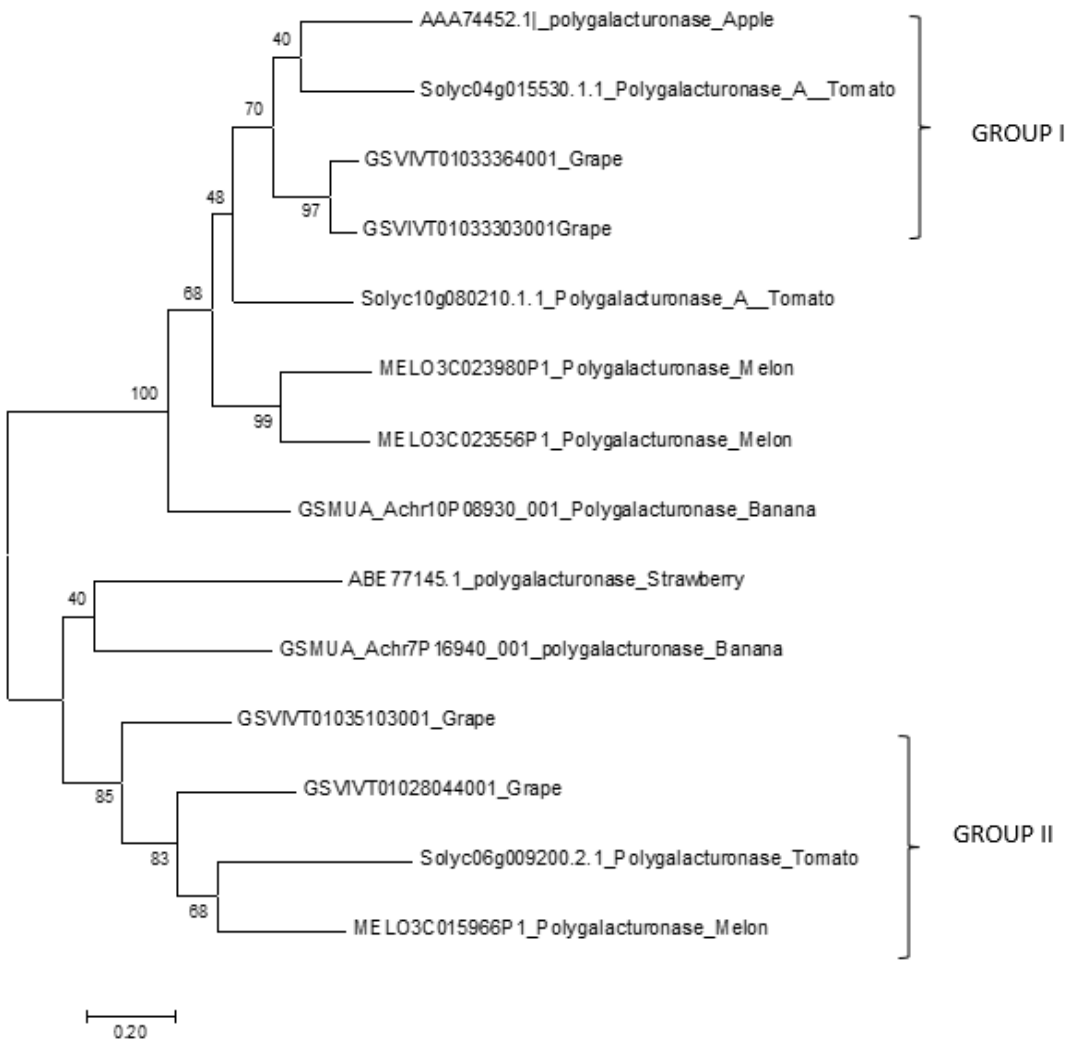


Figure 3-14 Molecular phylogenetic tree of polygalacturonase (*PG*) amino acid sequences and the similar amino acid sequences in studied fruits. The dendrogram was generated by Mega 7.0 software using MUSCLE for the alignment and the maximum likelihood method for the construction of the phylogeny. Bootstrap tests were performed using 1,000 replicates and the percentage of the bootstrap value are shown in each branch where the value exceeds 50% is considered significant. The branch lengths are proportional to the phylogenetic distances.

One of the only gene families where orthologues showed expression in all fruits were those encoding a *CesA*-like gene (Solyc08g061100) and also a glucan endo-beta glucosidase-like protein (Solyc03g115200) (Table 3-2). During fruit development cellulose synthases are highly expressed in tomato and then their levels decrease at the breaker stage (Tomato Genome Consortium 2012). They are likely to be involved in the biosynthesis of cellulose (Delmer and Haigler, 2002), but the function of the CesA-like gene product from Solyc08g061100 has not been investigated in fruits. With the role of the putative glucan endo-beta-glucosidase-like protein is even more obscure. In tomato its expression declines during fruit development and then increases during breaker stage (Figure 3-15). Glucan endo-beta-glucosidase has a role in callose decomposition and others have reported that they are expressed during ripening (Goulao and Oliveira, 2008).

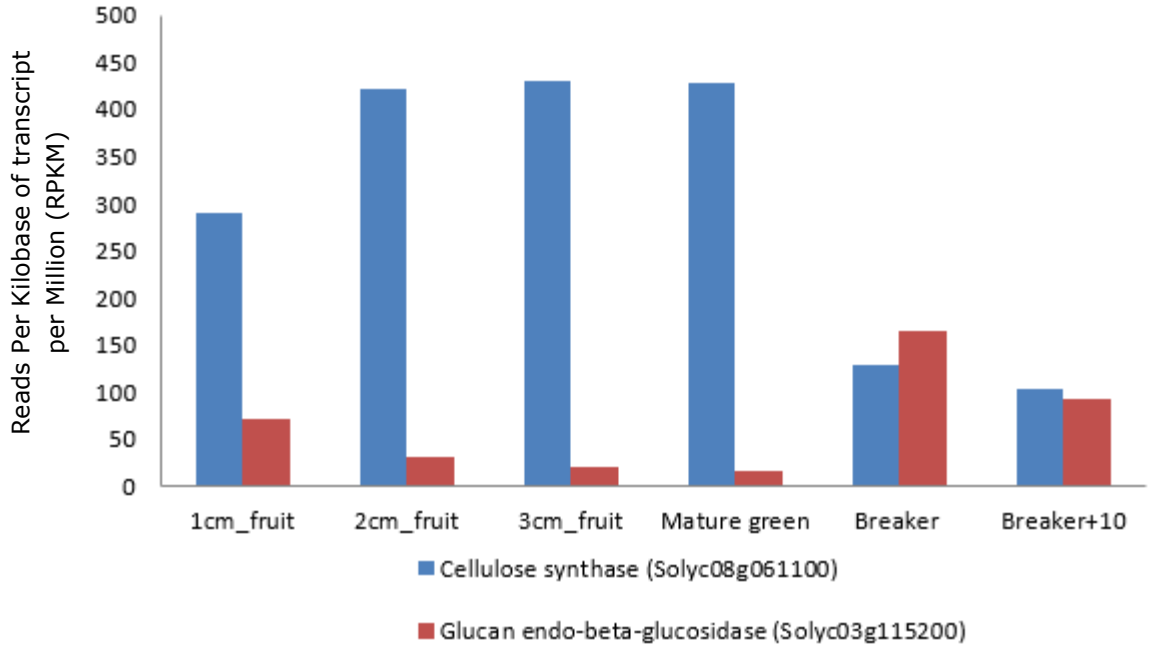


Figure 3-15 Gene expression patterns based on Reads Per Kilobase of transcript per Million mapped reads (RPKM) value of *CesA*-like gene and β-glucosidase in tomato during fruit development and ripening (Tomato Genome Consortium, 2012).

3.4 Conclusion

A comparison of the relationship between cell wall related genes in ripening tomato, melon, grape and banana revealed that there were only a small number of cell wall genes that were likely orthologues and expressed in all fruits that were surveyed. The aim of the project was to select genes that were common to all the fleshy fruit species examined with respect to expression during fruit ripening. A limited number of these would then be targeted for mutation by the CRISPR / Cas9 system (Bortesi and Fischer, 2014).

Comparison of the expression patterns of the cell wall structure-related genes revealed that most were expressed in tissues other than ripening fruits and of the fruit expressed genes, only a cellulose synthase catalytic subunit (*CesA*) and a glucan endo-beta-glucosidase had close orthologues that were fruit expressed in all four species. The substrate for the glucan endo-beta-glucosidases was not known so it was decided that attention be focused on the *CesA* gene and two other cell wall genes known to be important in tomato and other fruits and previously only silenced with RNAi.

In Chapter 4 CRISPR lines were generated for *CesA* and also *PHYTOENE SYNTHASE 1* (*PSY1*). *PSY1* was chosen as a control since it will give yellow fruit when knocked out. Additional DNA editing was performed

on, *POLYGALACTURONASE* Solyc10g080210 (Chapter 5) and *PECTATE LYASE* Solyc03g111690 (Chapter 6). *PG* has long been thought to be important for tomato fruit softening and although no effects were seen on fruit texture in RNAi lines when *PG* was suppressed (Smith et al, 1990) it is possible that a knockout line may have a more pronounced phenotype. The role of *PL* in tomato softening has never been explored in detail, but gene expression data suggests that it may have an important role (Marín-Rodríguez et al., 2002) and it has been shown to be important in texture changes in apple (Velasco et al., 2010) and strawberry (Jiménez-Bermúdez et al., 2002). Again CRISPR technology provides a precise method to test the role of a specific *PL* in tomato softening.

Chapter 4

Generation of Mutations in Tomato *PSY1* and *CesA-like* using the CRISPR-Cas9 Gene Editor

CHAPTER 4: GENERATION OF MUTATIONS IN TOMATO *PSY1* AND *CesA-LIKE* USING THE CRISPR-CAS9 GENE EDITOR

4.1 Introduction

The first gene silencing technologies to be used focused on antisense RNA and RNA interference (Saurabh et al., 2014). An advance on these technologies has been DNA editing with zinc finger nucleases (ZFN) (Beerli and Barbas, 2002) and TAL Effector Nucleases or TALENs (Li et al., 2011; Chen and Gao, 2013). More recently the use of Clustered Regulatory Interspaced Short Palindromic Repeats (CRISPR) /Cas9 vectors has provided a 'breakthrough' in the technology of specific genome editing to generate knockout mutants.

The type II prokaryotic adaptive immune system CRISPR systems are widespread in bacteria which are used as single endonuclease (Sorek et al., 2013). There are four genes that are involved in this mechanism including Cas9 nuclease, trans-activating crRNA (tracrRNA), crRNA and two non-coding RNA. Cas9 will form a complex with a synthetic single-guide RNA (sgRNA) that consists of ~20 bp guide sequence in crRNA and *trans*-activating crRNA (Deltcheva et al., 2011; Jinek et al., 2012). The sgRNA guides Cas9 which is complementary to gDNA followed by double-stranded DNA break (DSB) of target DNA. The Cas9 has a HNH nuclease domain and a RuvC-like domain

which each cleave one strand of double-stranded DNA (Figure 4-1) (Jinek et al., 2012). The DSB is then repaired by two types of repair mechanism (1) non-homologous end joining (NHEJ) which often leads to small insertions and deletions that interrupt gene function and (2) Homologous recombination (HDR) which introduce with sequence that homologous to the target site that allows precise mutations to the area (Figure 4-1).

According to Feng et al., (2013) the CRISPR/Cas9 system generates mutations by using Cas9 driven from the cauliflower mosaic virus constitutive 35S promoter (CaMV 35S promoter) and the synthetic sgRNA from the AtU6-26 promoter in *Arabidopsis* or OsU6-2 promoter in rice. A customized sgRNA encoded by a sequence of ~100 nt is required to target a specific sequence. Cas9 does not have to be reengineered for each new target site. Thus, the sgRNA: Cas9 system is therefore much more straightforward than RNA interference (RNAi). Demonstration of efficacy of genome editing using CRISPR-Cas9 system has been reported in *Arabidopsis thaliana* (Li et al., 2013; Feng et al., 2014), rice (Zhang et al., 2014), sorghum (Jiang et al., 2013), tomato (Brooks et al., 2014) and wheat (Wang et al., 2014).

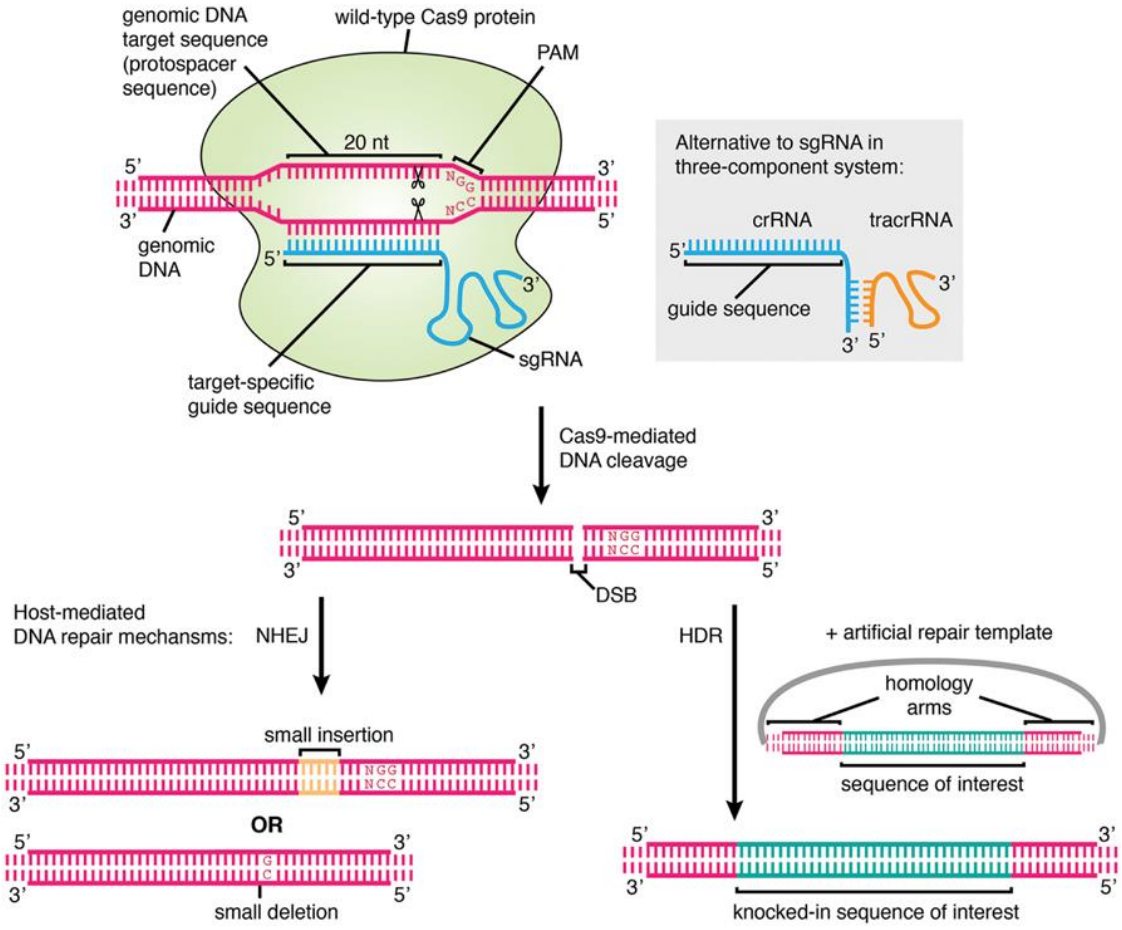


Figure 4-1 Schematic illustrating the engineered type II CRISPR-Cas9 system. The complex Cas9 (sgRNA, crRNA and tracrRNA) will complement with the DNA target upstream of the PAM site. The Cas9 endonuclease will cleave the DNA approximately 3 bases before the end of seed region. Cleavage results in a double-strand break (DSB) which is repaired by host-mediated two DNA repair mechanisms: 1. Non-homologous end joining (NHEJ) which creates random insertions or deletions (indels) at the targetd site or 2. Homologous recombination which creates precise changes based on template DNA (taken from Agrotis and Ketteler, 2015).

4.1.1 Targets for CRISPR/Cas9 DNA Editing

In Chapter 3 the bioinformatics analysis indicated a number of genes in tomato that were expressed during fruit development and ripening and have likely orthologues in at least three of the four fruit species. The time and resources required to generate CRISPR/Cas9 mutations in transgenic tomato mean that we have to limit our targets to no more than four genes. This is the first CRISPR transgenics made in the lab and therefore it was decided to include as one of those targets *PHYTOENE SYNTHASE 1* (*PSY1*) to provide a control. *PSY1* is the rate limiting step in the biosynthesis of carotenoids in tomato (Figure 4-2). The enzyme catalyses the formation of phytoene from geranyl geranyl pyrophosphate and knockouts mutations in *PSY1* has distinctive yellow fruits (Fray et al., 1993). The remaining three tomato genes selected for gene editing were (1) Solyc08g061100, encoding a putative cellulose synthase-like protein, (2) Solyc10g080210, which encodes polygalacturonase 2A (*PG2a*) and (3) Solyc03g111690, a gene encoding a pectate lyase (*PL*).

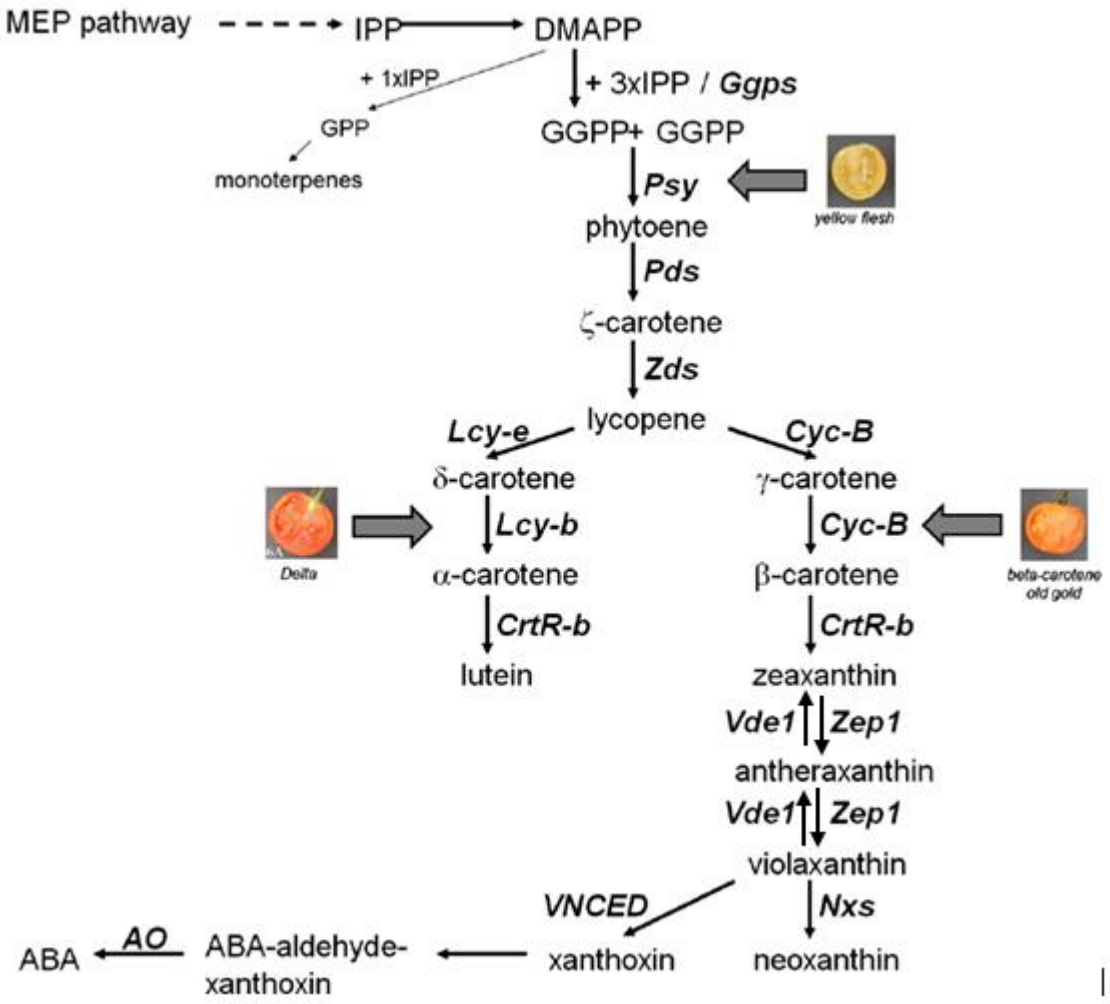


Figure 4-2 The pathway of carotenoids biosynthesis and the correlation of gene silencing in generated tomato mutants. Phytoene synthase (PSY), phytoene desaturase (PDS), ζ -carotene desaturase (ZDS), carotene isomerase (CRTISO), lycopene-ß-cyclase (LCY-B, CYC-B), lycopene-ε- cyclase (LCY-E), zeaxanthin epoxidase (ZEP) and violaxanthin de-epoxidase (VDE), neoxanthin synthase (NXS), ß- and ε-hydroxylases (CYP97A, CYP97C, CRT-B1 and B2) (from Bergougnoux, 2014).

The cellulose synthase or *CesA-like* gene has orthologues in all four fruit species and is expressed in tomato during both fruit development and ripening (Figure 4-3). *PG2a* was selected as it has likely orthologues in three of the four species and these are expressed during ripening. Also this gene and the encoded protein have been intensively studied in tomato, but a knockout line has never been generated. Previous transgenic tomato lines where *PG* was silenced used antisense RNA and a low level of *PG* gene expression and activity could still be detected in the fruits (Smith et al., 1988 and Smith et al., 1990). It has been postulated that this low level of *PG* activity could account for the normal softening in these tomato fruits and we wanted to test this hypothesis.

The *PL* gene does not appear in Table 3-3 because there are no direct orthologues in melon, grape or banana. However, *PL* is known to be important in strawberry fruit softening and a very recent experiment in the Seymour lab (Uluisik et al, 2016) has demonstrated that this gene when silenced by RNAi can significantly affect softening in tomato. *PLs* are also expressed in grape and banana (Table 3-3). Work in the Seymour lab on *PL* using RNAi (Uluisik et al, 2016) has provided the first report of a gene that when silenced significantly affects tomato fruit softening and therefore this gene was selected as the final target for this project. This chapter will present the data on the targeting of the *PSY1* and *CESA*-like genes by CRISPR/Cas9. Then, in Chapter5 and 6 it will be focusing on the experiments with *PG* and *PL*.

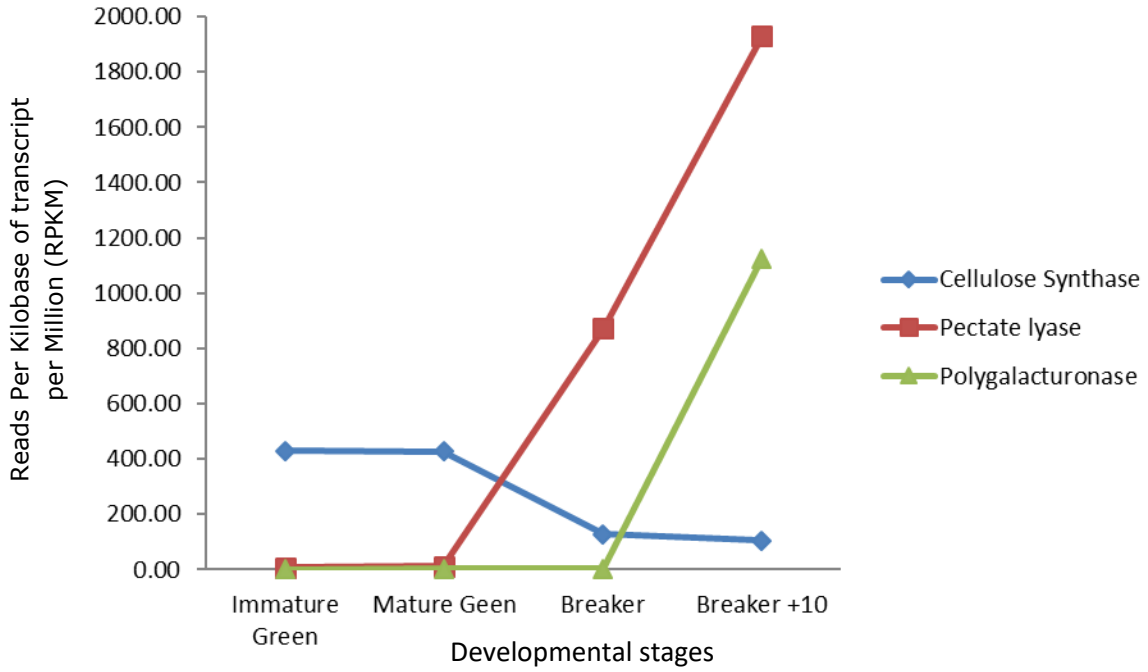


Figure 4-3 The expression level of selected genes (*PG, CesA-like* and *PL*) that have been identified as expressed during fruit development and ripening in tomato (Tomato Genome Consortium, 2012)

4.2 Material and Methods

4.2.1 Plant materials

Plant materials are described in detail in Section 2.1.

4.2.2 Vector construction for generating stable transgenic plants

Vector construction Level 0

Generation of the *PSY*::CRISPR lines was undertaken with Dr Cheng Yang a visitor to the Seymour lab during this project. The process for the construction of the CRISPR/Cas 9 vectors is summarized in Figure 4-4. The pICH86966::AtU6p::sgRNA construct (Addgene Plasmid 46966), pICSL01009::AtU6p (SpecR), pICH47751 (*CarbR*), pICH47732::NOSp::NPTII-OCST, pICH47742::35Sp::Cas9-NOST, pICH41766 and pAGM4723 were used to generate the CRISPR/Cas9 stable transgenic plants. The sequences for the cellulose synthase-like gene (*CesA*-like) (Solyc08g061100) and phytoene synthase (*PSY1*) gene (Solyc03g031860) were obtained from the newly released tomato genome assembly at [http://solgenomics.net/]. Forward (F) and reverse (R) primers for PCR are shown in Table 4-1. The sequences of Solyc08g061100 and Solyc03g031860 were amplified using gradient PCR by Q5 Master Mix (Finnzymes, UK) following the standard protocol (Chapter 2, section 2.8). The PCR conditions are shown in Table 4.2.

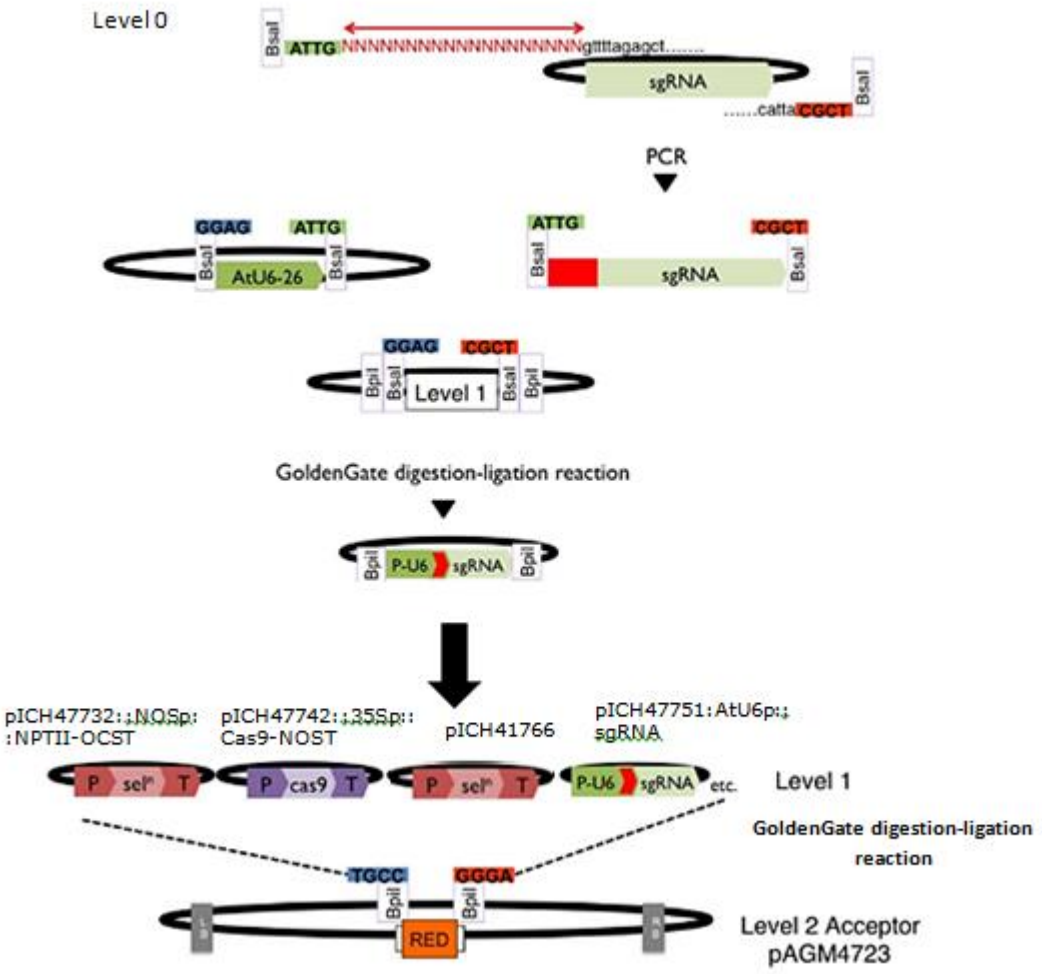


Figure 4-4 Flowchart describing steps involved in making CRISPR/Cas9 construct. PCR amplicon by using the sgRNA as a PCR template are made (Level 0). Then, the amplicon was used to combine with the small RNA promoter (AtU6p). The construct pICH47751:AtU6p::sgRNA level 1, pICH47732::NOSp::NPTII-OCST, pICH47742::35Sp::Cas9-NOST, pICH41766 were then assembled to pAGM4723 using Golden-Gate approach to generate Level 2 vector that was then transformed into *Agrobacterium*.

Table 4-1 Primer sequence of gene for *CesA*-like and *PSY1*

Primer ID	Primer sequence 5’ to 3’
8CeSa061100.F	TGTGGTCTCAATT**GTCTTTTTGGGGCACAGCGGGGG**TTTTAGAGCTAGAAATAGCAAG
*PSY1*001.F	TGTGGTCTCAATT**GGCAGGCAGCCTTGGTGAAG**GTTTTAGAGCTAGAAATAGCAAG
*PSY1*002.F	TGTGGTCTCAATT**GAAGCCGGATATACCTATTCG**TTTTAGAGCTAGAAATAGCAAG
*PSY1*003.F	TGTGGTCTCAATT**GAATGTCTGTTGCCTTGTTAG**TTTTAGAGCTAGAAATAGCAAG
*PSY1*004.F	TGTGGTCTCA**ATTGGTGGAAAGCAAACTAATAAG**TTTTAGAGCTAGAAATAGCAAG
RevCRISPR	TGTGGTCTCAAGCGTAATGCCAACTTTGTAC

Table 4-2 PCR analysis condition

Step	Temperature	Time	Cycle
Initial denaturation	98°C	5 minutes	1
Denaturation	98°C	10 seconds	30
Annealing	60°C	30 seconds	30
Elongation	72°C	30 seconds	30
Extention	72°C	2 minutes	1

Vector construction Level 1

To construct the level 1 vector (Figure 4-4), the Level 0 construct pICSL01009::AtU6p (SpecR) and a Level 1 destination vector pICH47751 (*CarbR*) were cut with *BsaI* (as detailed in 2.9.3) and a ligation reaction performed in a total volume of 10 µL in a 0.5 mL Eppendorf tubes. The reactions were incubated at room temperature (22 to 25°C) overnight before transforming into competent cells following protocols in Section 2.9.4. The cells from each transformation were then spread as aliquots of 20, 40 and 100 µL on a pre-warmed LB plates containing 100 µL mL^{-1} carbenicillin and 120 µL mL^{-1} X-gal.

The culture was incubated at 37°C overnight. The plasmid DNA was isolated from the transformed cells using the PureLink Quick Plasmid Miniprep Kit (Invitrogen, UK) as discussed in Section 2.9.5. The products were then sent to Eurofins- MWG for sequencing (Appendix 9-11). The resulting sequences were analyzed by the BLAST two alignment programs and compared with the original target sequence to confirm that the correct plasmid product had been isolated. Figure 4-4 shows the assembling of Level 1 Cas9/sgRNA expressing constructs with *CesA*-like and *PSY1* genes.

Vector construction Level 2

The cut-ligation reactions CRISPR/Cas9 for Level 2 (Appendix 9-12) involved a mixture of pICH47732::NOSp::NPTII-OCST, pICH47742::35Sp::Cas9-NOST, pICH47751::AtU6p::sgRNA, pICH41766, pAGM4723 with BpiI (BbsI) restriction enzyme. This reaction was in a total volume of 12 µL in a 0.5 mL Eppendorf tubes and the PCR analysis conditions were as described in Section 2.9.3. The reactions were then incubated at room temperature (22 to 25°C) overnight before transforming into competent cell. The transformation of entry clone to *E. coli* competent cell was carried out as described in Section 2.9.6. The transformed cells were then be plated on LB plate (50 µL mL^{-1} kanamycin) and incubated at 37°C overnight. The next day a single colony was picked up and was picked and inoculated into liquid LB before purified using PureLink Quick Plasmid Miniprep Kit (Invitrogen, UK) for sequencing.

The final Level 2 Cas9/sgRNA construct was transformed into *Agrobacterium* EHA105 by electroporation as described in Chapter 2.10. After 4 days incubation, the colonies on LB plates containing 50 µg mL^{-1} of rifampicin and kanamycin 100 µg mL^{-1} were then isolated and purified using PureLink Quick Plasmid Miniprep Kit (Invitrogen, UK). Then, the colonies were isolated, inoculated and the DNA purified. The plasmid DNA was analysed by PCR to confirm the presence and correct orientation of the insert. The PCR reaction was in a total volume of 10 µL in a 0.5 mL Eppendorf tubes as detailed in Table 2.1

4.2.3 Generation of stable transgenic plants

The methodology in plant transformation is described in detail in Chapter 2, Section 2.11.

4.2.4 Identification of the PG transgenic lines

To identify the presence of T-DNA in the transgenic plant lines a set of PCR reactions was performed using *CesA*-like and *PSY1* gene specific primer and the NPTII::Cas9 primer (Appendix 9.8). The DNA from wild-type Ailsa Craig was used as a positive control for the gene specific primer pair and as negative control for T-DNA mutation. The PCR condition for this method was followed the reaction in Table 4-3 and the sequence of transgene was confirmed by sequencing (Appendix 9.9)

4.2.5 Identification of single copy homozygous lines

The T_0 lines were grown to fruiting and T_1 seed were collected. Then 10-12 T_1 seed were then sown for each of the line. The DNA from each of the plants was extracted using DNeasy Plant Mini Kit (Qiagen) (Chapter 2, Section 2.2). PCR was used to establish the presence of the Cas9 gene and also to obtain the edited *CesA*-like and *PSY1* sequences. The PCR products were then sent for sequencing.

4.2.6 QPCR determination of gene expression

The RNA was extracted from *PSY1* fruits at stage MG, B, B+4 and B+7 (Chapter 2.4). DNA contamination was removed before generating the cDNA template (Section 2.8.1). Quantification of *PSY1* gene expression in the PSY knockout and wild-type Ailsa Craig (AC^{++}) line was undertaken using SYBR® Green quantitative RT-PCR. The serial dilutions, qPCR reactions and qPCR conditions followed the 2X SYBR® Green PCR master mix protocol (also see Chapter 2.8.1). An elongation factor was used as a housekeeping gene. The relative transcript abundance of the *PSY1* genes and elongation factor gene were calculated from the equation (2.1) in Chapter 2.8 (German *et al*., 2003).

4.2.7 Analysis of fruit characteristics

Measurement of fruit texture, colour and Brix

Methods followed those described in Chapter 2. The texture of transgenic fruits was determined by measuring the maximum load of pericarp tissues using the Lloyd Instrument LF plus machine. Peel colour was measured with a Minolta colorimeter CR400 and total soluble solids using a hand held refractometer.

Table 4- 3 PCR reactions and conditions for checking the endogenous genes.

PCR analysis reaction

Genomic DNA	3 µL
Forward primer (10 pM µL-1)	0.5 µL
Reverse primer (10 pM µL-1)	0.5 µL
Q5® High-Fidelity 2X Master Mix	5.0 µL
Sterilised water up to	12 µL

PCR analysis condition

Step	Temperature	Time	Cycle
Initial denaturation	98°C	5 minutes	1
Denaturation	98°C	10 seconds	28
Annealing	60°C	30 seconds	28
Elongation	72°C	30 seconds	28
Extention	72°C	10 minutes	1

4.3 Results and Discussion

4.3.1 Vector construction of *CESA-like* and *PSYI*

The PCR amplicons of the target genes, Solyc08g061100 and Solyc03g031860, with the length ~100 bp (Figure 4-5). According to the product sizes, the most convenient primer products were 8CeSa061100 and *PSY1*002. These products were then cloned into vector pICSL01009::AtU6p (SpecR) and shown to be correct by sequencing (Appendix 8-12). By using Golden Gate cloning approach Level 1 construct were assemble to Level 2 vector that involved (pICH47751::AtU6p::sgRNA, pICH47732::NOSp::NPTII-OCST, pIcH47742::35Sp::Cas9-NOST, Pich41766 Linker and pAGM4723. The presences of DNA for *CesA-like* and *PSY1* in Level 2 construct were confirmed using PCR (Figure 4.6). The destination vector Level 2 Cas9/sgRNA construct was then transformed into *Agrobacterium* EHA105 by electroporation as described in Chapter 2. Figure 4.7 (A) and (B) shows the PCR products that were amplified from the isolated plasmid DNA (Level 2) for both of samples *CesA-like* and *PSY1* using NPTII/Cas9 specific primer.

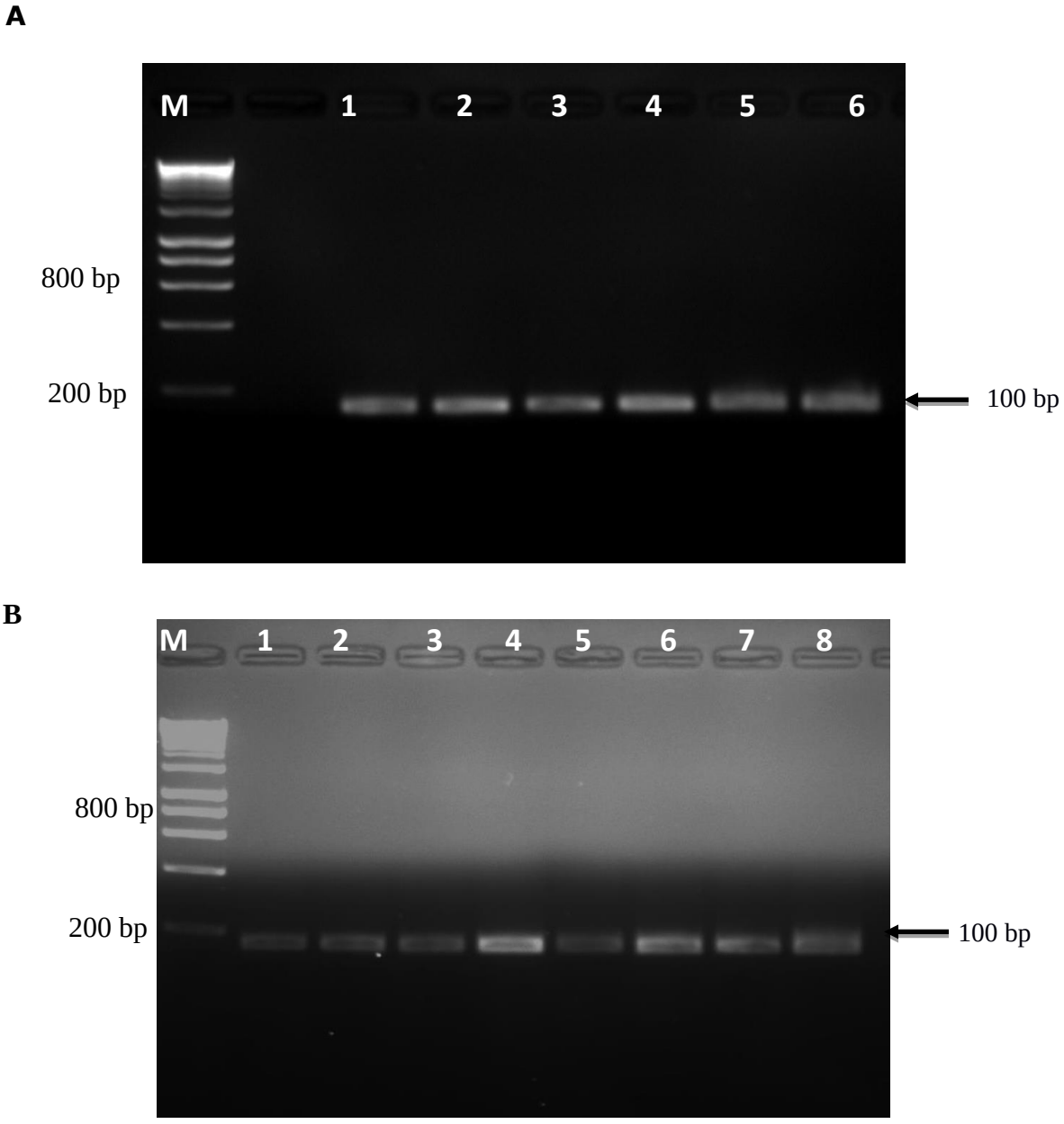
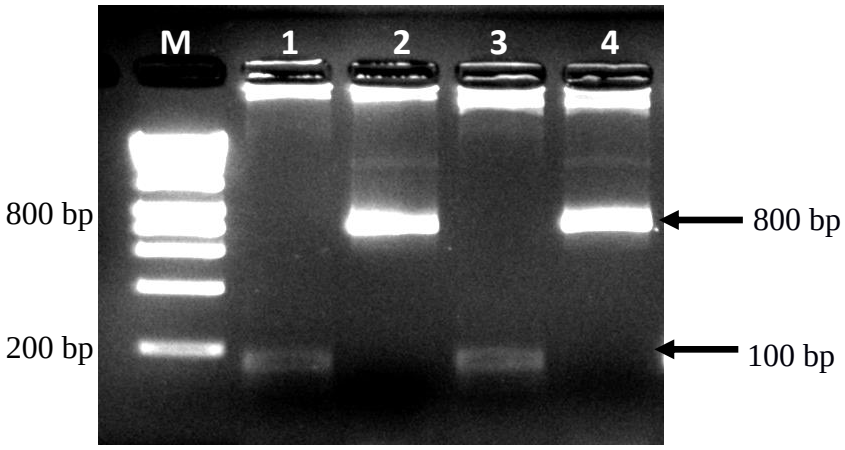


Figure 4-5 The gene specific *CesA-like* and *PSY1* PCR. The purified PCR products were separated by electrophoresis on a 2% (w/v) agarose gel in 0.5X TAE at 100 V for 30 minutes. A) Lane M: Hyperladder Marker. Lane 1-6 PCR product of sgRNA pICH86966::AtU6p::sgRNA_PDS construct that contained the CesA-like target sequence. B) Lane M: Hyperladder Marker. Lane 1-8 PCR product of sgRNA pICH86966::AtU6p::sgRNA_PDS construct that contain the *PSY1* target sequence.

A



B

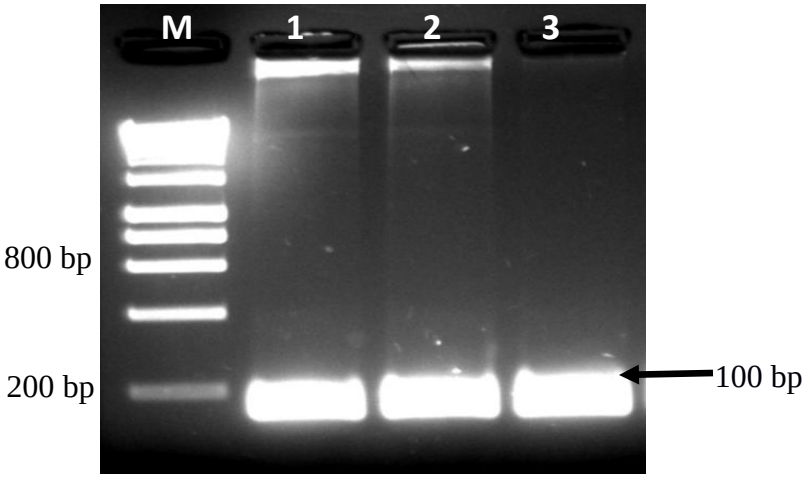
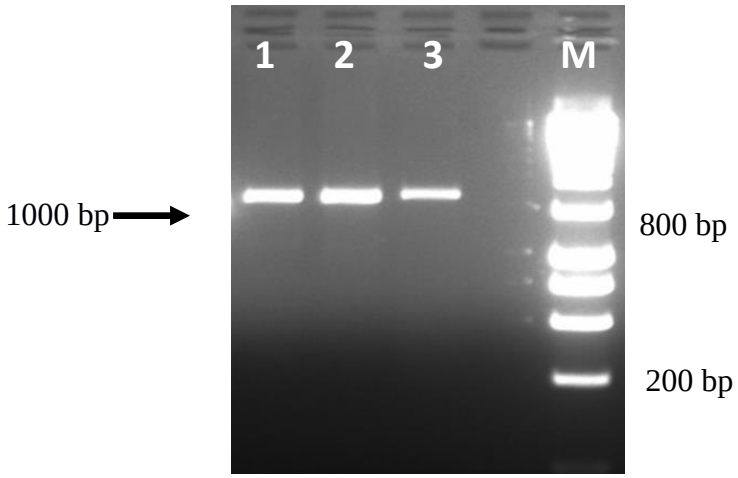


Figure 4-6 Cloning of *CESA-like* and *PSY1* gene fragment into the Level 2 vector The PCR products of Level 2 construct were amplified from the entry clone. A) Lane M: Hyperladder Marker, lane 1 and 4: the construct amplified using sgRNA 8CS061100 primer. Lane 2 and 4: the Level 2 construct amplified using Kanamycin primer. Bands of expected sizes 600 bp and 100 bp were seen B) Lane M: Hyperladder Marker, lane 1-3: the construct that amplify using sgRNA PSY1002 specific primer.

A)



B)

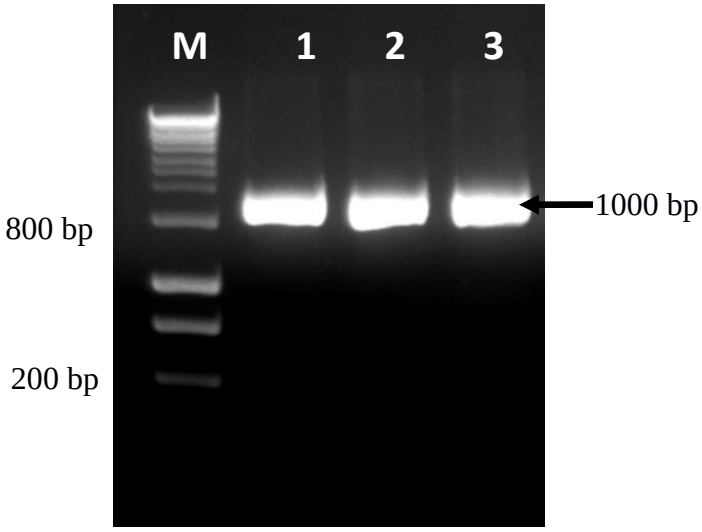


Figure 4-7 PCR products were amplified using the NPTII/Cas9 specific primers from *Agrobacterium* EHA105. The purified PCR products were separated by electrophoresis on a 2% (w/v) agarose gel in 0.5X TAE at 100 V for 30 minutes. Lane M: Hyperladder Marker, lanes 1, 2 and 3 PCR products of sgRNA constructs which were isolated from colonies of *Agrobacterium* EHA105.

4.3.2 Analysis of the *CESA-like* and *PSY1* transgenic lines

There were seven transgenic plant lines regenerated using the *CesA-like* construct and ten with the *PSY1* construct. These transgenic plants were then transferred to the glasshouse. To confirm the presence of the transgenes, DNA was extracted from leaves of the transformants. Primer NPTII/Cas9 was used to amplify the Cas9 region where the size of the product is 1,000 bp (Figure 4-8). The PCR generated an amplicon of the correct size of 1,000 bp in transformed plant lines No. 1, 2, 3, 4, 5, 6 and 7 for *CesA-like* and for *PSY1* lines No. 4, 9, 16, 17, 20, 22, 23, 26 37 and 38. In addition, no PCR amplification product was recovered in WT plants. The experiment demonstrated that these lines contained the Cas9 gene, but of course do not indicate if the endogenous gene has been mutated.

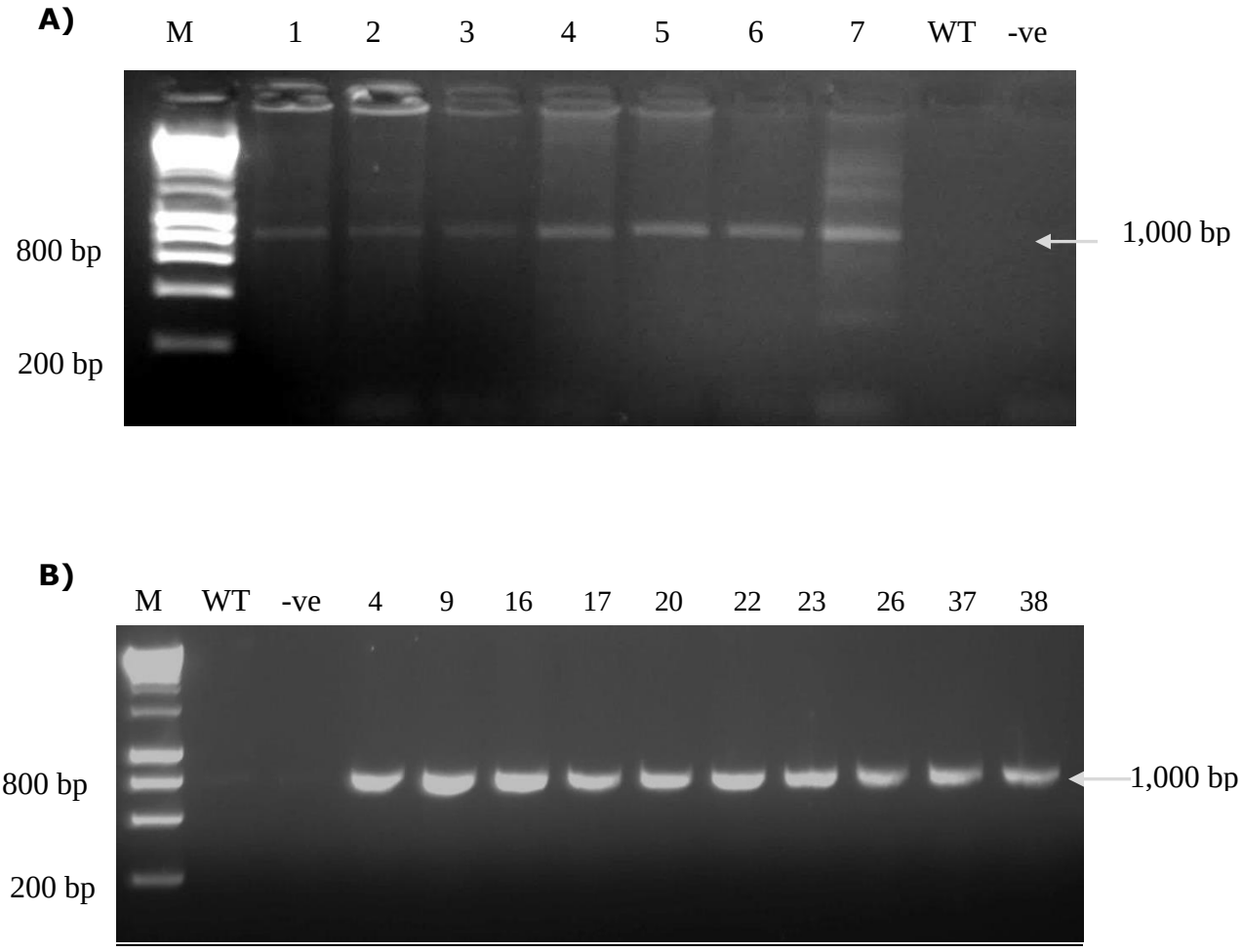


Figure 4-8 Amplification of transgenic lines (A) Lane M: Hyperladder Marker, Lane WT: wild-type, -ve: dH_2O; lanes 1-7 PCR products of CRISPR::*CesA-like* plants showing the transgenic lines that contain Cas9/sgRNA gene construct in T_0 plants. (B) Lane M: Hyperladder Marker, Lane WT: wild-type, -ve: dH_2O; lanes 4, 9, 16, 17, 20, 22, 23, 26, 37 and 38 PCR products of transgenic lines that contain the CRISPR::*PSY1* Cas9/sgRNA gene construct.

Validation of CRISPR/Cas9 driven mutations in the CesA-like and PSY1 genes

A PCR coupled with restriction digestion strategy as well as sequencing was used to validate the presence of mutations in the target genes in the T_0 lines. A PCR approach for the *CesA-like* gene was designed to give a single amplicon of 300 bp containing a BssSI site (Figure 4-9B). The guide RNA would be expected to target sequences in the region of this restriction site. This would result in a modification of the sequence at the site and prevent subsequent digestion by BssSI. However, the BssSI digestion pattern of PCR-amplified DNA from the 7 individual T_0 plants all show similar patterns to wild-type. All of the amplification products were digested by the enzyme. This suggests that the CRISPR/Cas9 targeting of the *CesA-like* gene has not been successful. Thus, to confirm this analysis the PCR products from the seven T_0 were purified and sent for sequencing. The sequencing results confirmed the absence of mutations in the target site of *CesA-like* gene (Figure 4-10). One explanation for these results is that the *CesA-like* gene is essential for proper regeneration of plants in tissue culture and knocking out this gene is lethal.

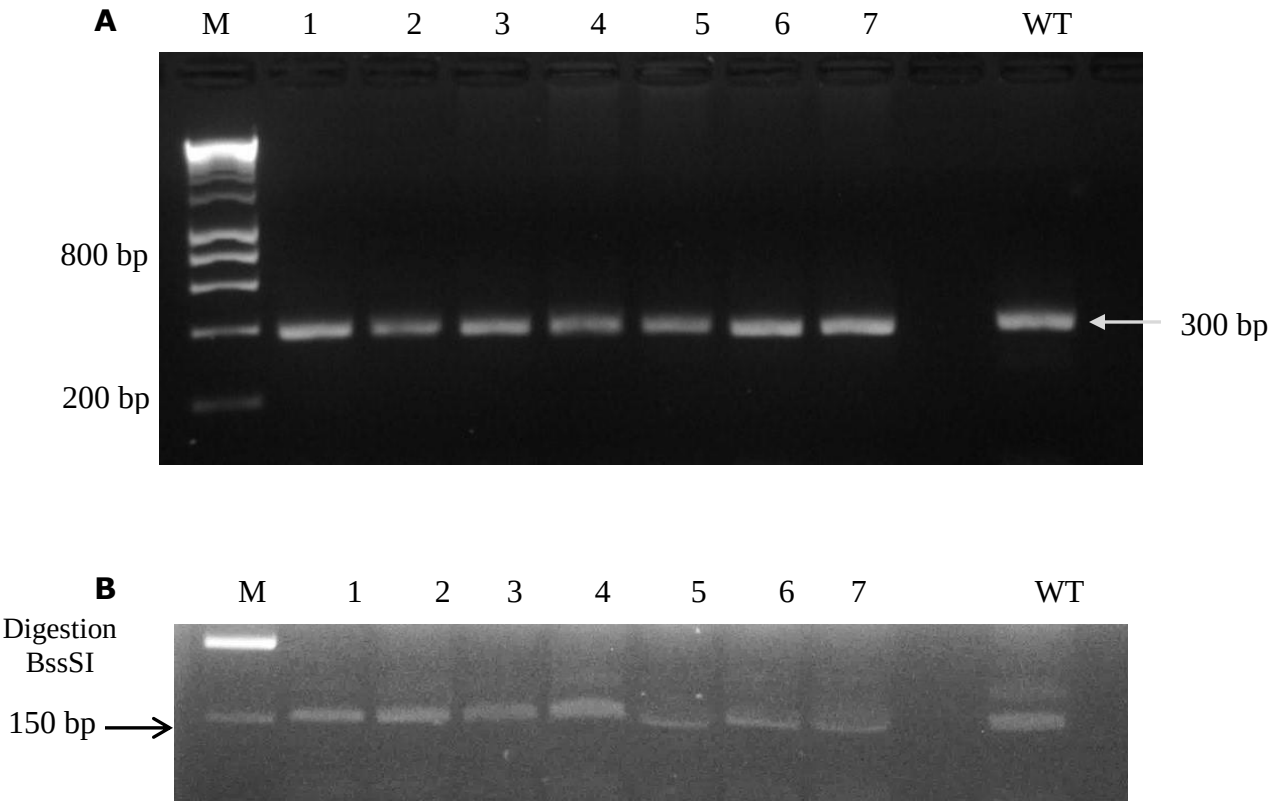


Figure 4-9 Efficiency of Cas9/sgRNA mutagenesis in *CesA-like.* (A) Lane M: Hyperladder Marker, Lane WT: wild-type; lanes 1-7 PCR amplification of the CRISPR::CesA lines to generate a 300 bp product that could then be challenged with BssS1. (B) Lane M: Hyperladder Marker, Lane WT: wild-type; lanes 1-7 PCR/Restriction digestion analysis demonstrates that all products can be cut with BssS1 indicating a lack of mutations in this sequence.

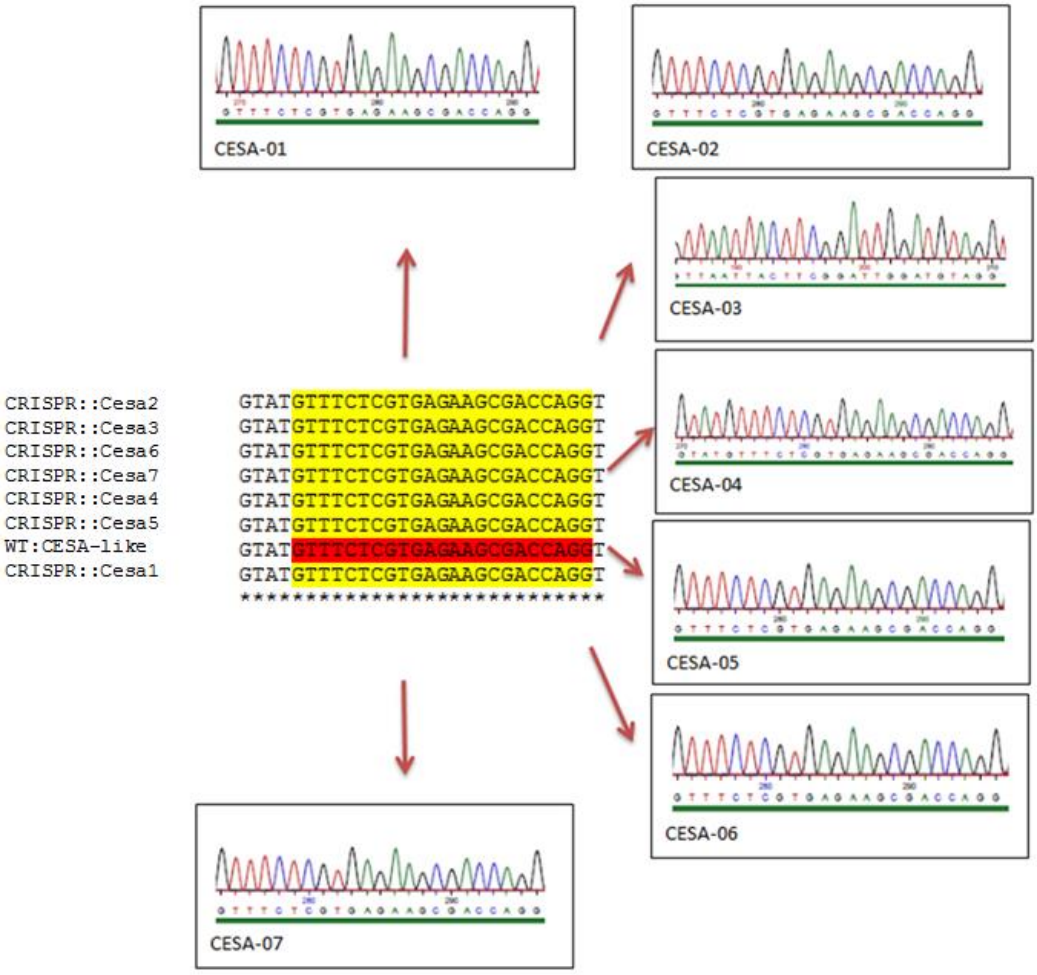


Figure 4-10 Confirmation of inheritance of a modified or non-modified *CesA-like* gene. Sequence alignment and DNA sequencing traces from sequencing of PCR amplification of CRISPR::CesA as target sites where there are no modification in the sequence level.

A similar approach was used to determine if mutations were apparent in the T_0 *PSY1* transgenic lines. In the normal or non-mutagenized plants (WT) the enzyme StyI will cleave the PCR product amplified from this gene (~350 bp) and generate two different fragments with sizes around (~200 bp and ~150 bp). The StyI digestion patterns of PCR-amplified DNA from 10 individual T_0 plants indicated a small number contained a StyI-resistant DNA product (Figure 4-11). To confirm Cas9/sgRNA-directed mutagenesis of the *PSY1* gene, the PCR products from the T_0 and wild type plants were purified and sequenced. The sequences of the target area in the *PSY1* gene showed mutations in the form of an insertion of a T nucleotide (line Psy-04) and deletions of several nucleotides of around 2 to 11 bp in lines *PSY1*6, Psy20, Psy23, Psy26, and Psy37 (Figure 4-12). Twelve seeds from CRISPR::*PSY1* T_0 lines 20 and 23 were grown to obtain homozygous lines in the T_1 generation.

4.3.3 Identification of Homozygous *PSY1* lines

Twelve seeds were grown from two T_0 transgenic lines (PSY-20 and PSY-23). When the seedlings were 2-3 weeks old, a small piece of leaf tissue (approximately 100 mg) was taken and used to extract the DNA by using

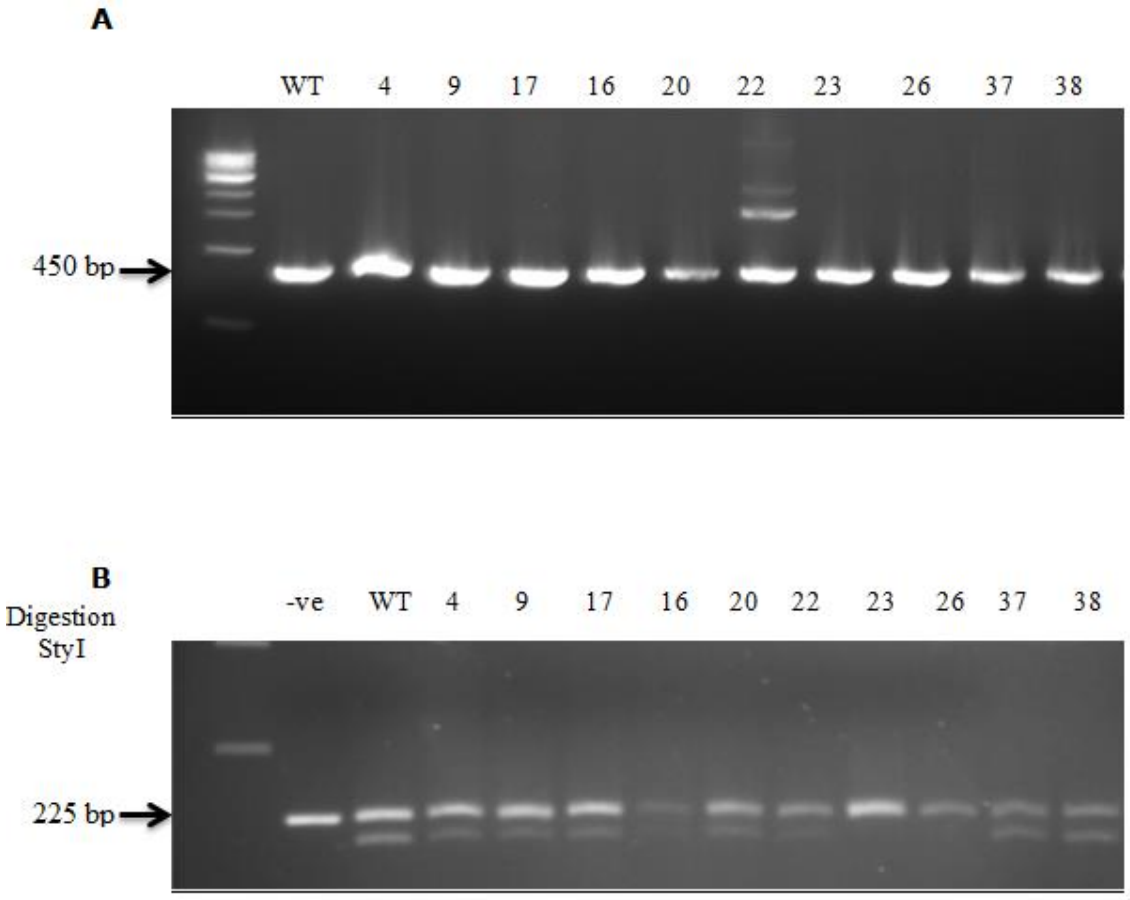


Figure 4-11 Efficiency of Cas9/sgRNA mutagenesis in *PSY1*. (A) Lane M: Hyperladder Marker, Lane WT: wild-type; lanes 4, 9, 17, 16, 20, 22, 23, 26, 37 and 38 PCR genotyping of ten representative CRISPR::*PSY1* plants using gene specific primer PSY1002. (B) Lane M: Hyperladder Marker, Lane WT: wild-type; lanes 4, 9, 17, 16, 20, 22, 23, 26, 37 and 38 PCR/Restriction analysis where wild type sequences are expected to give bands at ~200 bp and ~150 bp DNA resulting from StyI cleavage of the ~450bp PCR product.

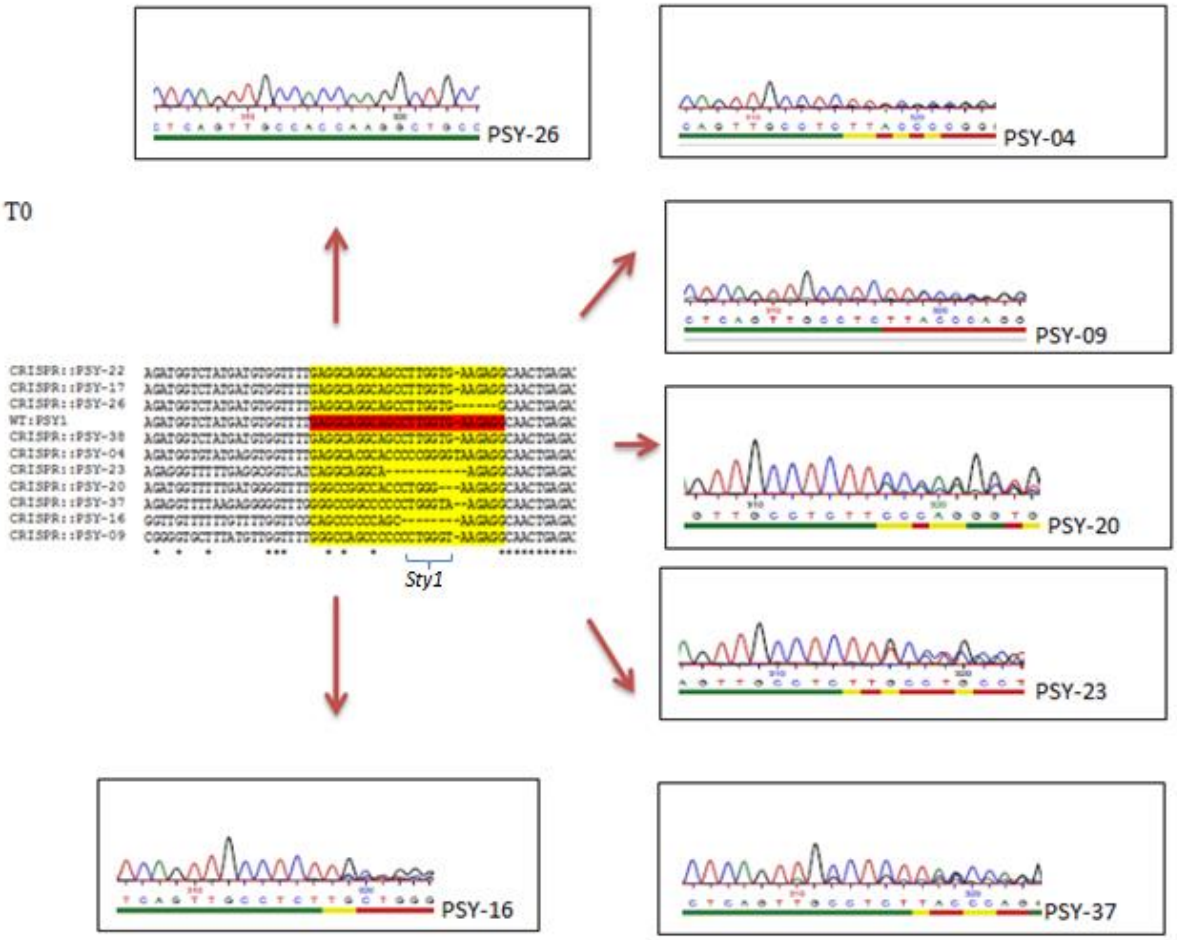


Figure 4-12 Confirmation of inheritance of a modified or non-modified *PSY1* gene. Sequence alignment and DNA sequencing traces from sequencing of PCR amplified CRISPR::*PSY1* as target sites in which there has been a Cas9/sgRNA-mediated gene modification. A range of insertion and deletions (indels) were identified from the sequencing results in T_0 plants which are line Psy04, *PSY1*6, Psy20, Psy23, Psy26 and Psy37. The PCR products were then cloned and resequenced to validate the indels and characterize additional DNA lesions.

DNeasy Plant Mini Kit (Qiagen) as described in Chapter 2, Section 2.2. Primers NPTII/Cas9 and *PSY1* (Appendix 9-8) were used to amplify the Cas9 gene and the *PSY1* target region to give amplicons of 800 bp and 350 bp (Figure 4-13) respectively. All of T_1 plants lines CRISPR::PSY20 and CRISPR::PSY23 contained the CRISPR/Cas9 transgene. The PCR samples from the *PSY1* target amplification were then purified for sequencing. The aims of this analysis was to identify lines harboring two identical copies of a CRISPR mutation in the *PSY1* target region and are therefore homozygous.

The sequence alignments for the progeny of the two T_0 lines are shown in Figure 4-14 and 4-15. For PSY-20, nine T_1 lines were recovered and these all contained mutations in the target region. The sequence alignment and DNA sequencing traces from sequencing result shows that most of the T_1 plants from lines CRISPR::PSY20 and CRISPR::PSY23 have mutation either deletion or insertion. In CRISPR::PSY20 two types of homozygous mutants were obtained from nine progenies which are single bp deletion (lines PSY-20-1, PSY-20-2, PSY-20-3 and PSY-20-7) and 2 bp deletion for lines PSY-20-4, PSY-20-8, PSY-20-10, PSY-20-5 and PSY-20-9. For ten progenies from CRISPR::PSY23 also have two type of mutation. PSY-23-2 and PSY-23-7 lines were homozygous for insertion of one base (A) and PSY-23-10, PSY-23-08, PSY-23-09, PSY-23-06, PSY-23-03, PSY-23-04, PSY-23-05 and PSY-23-06 were homozygous with deletions of eleven nucleotides identified in the target regions.

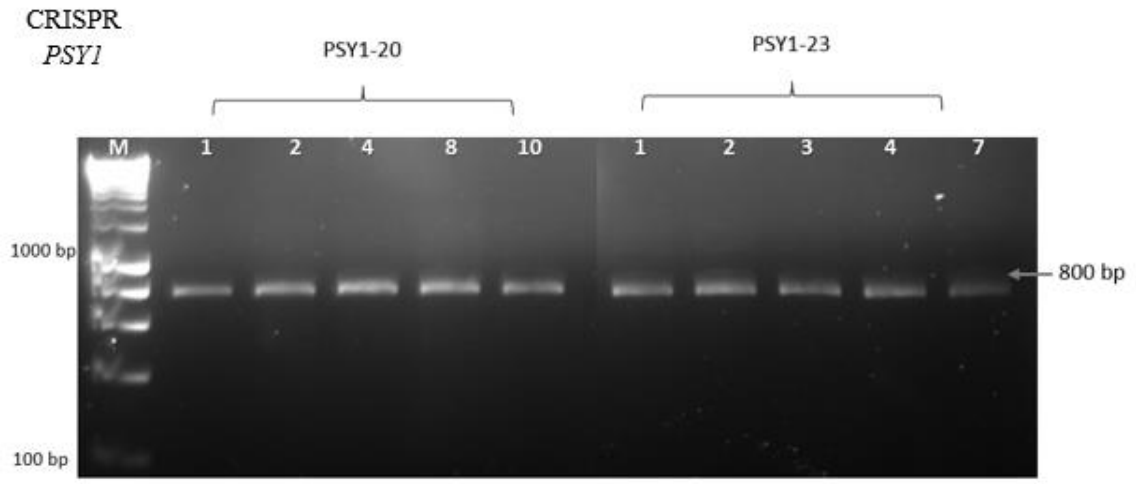


Figure 4-13 Confirmation of inheritance of modified *PSY1* gene in each of T_1 generation plants using NPTII/Cas9 primer. Out of the five progeny from plant line CRISPR::PSY20 and CRISPR::PSY23, all plants showed the presence of CRISPR/Cas9 transgene.

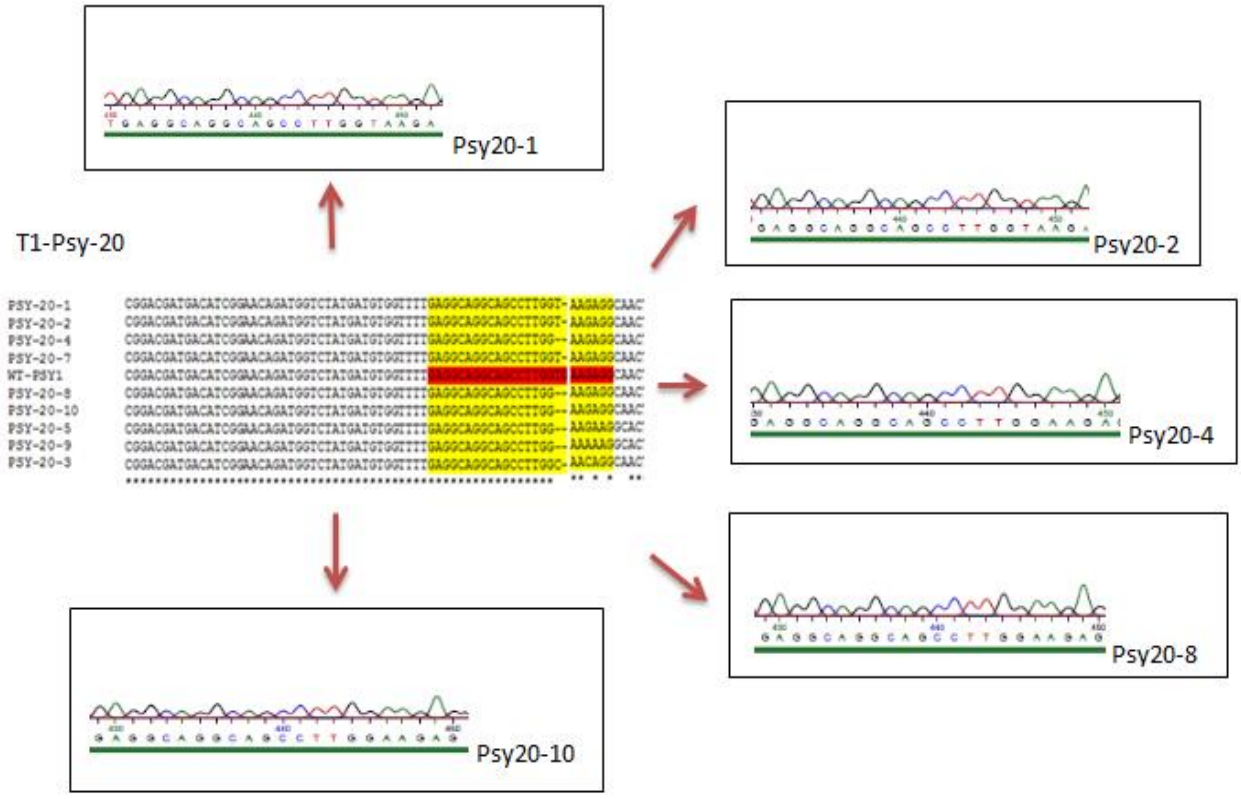


Figure 4-14 Inheritance of modified *PSY1* gene in each of T_1 progeny. (A) DNA sequence alignment and sequencing traces of CRISPR::PSY20. All on T_1 generation in CRISPR::PSY20 appeared homozygous based on lack of mixed sequence reads.

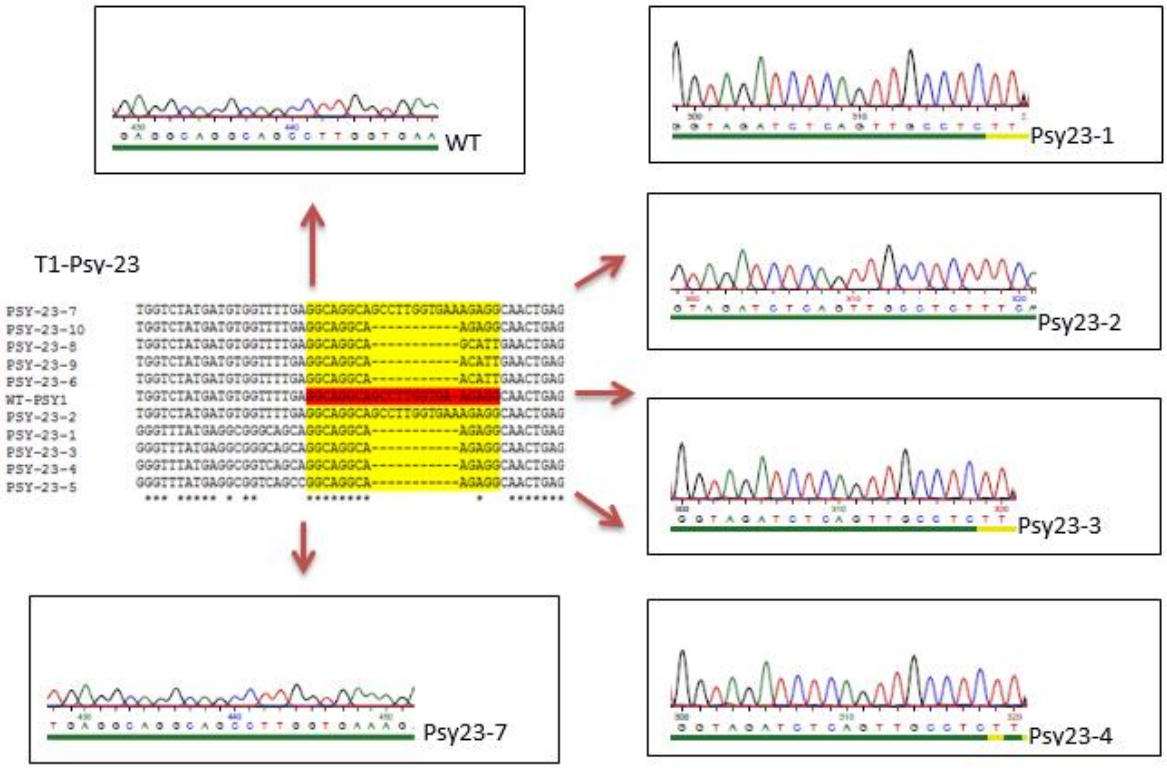


Figure 4-15 Inheritance of modified *PSY1* gene in each of T1 progeny**.** (A) DNA sequence alignment and sequencing traces of CRISPR::PSY23. From the lack of mixed sequence in T_1 generation in CRISPR::PSY23 progeny all appeared to be homozygous.

4.3.4 *PSY1* gene expression in the *PSY1*::CRISPR lines

PSY1 gene expression (Figure 4.16) was determined by QPCR in the pericarp of orange (breaker +4) and red ripe (breaker + 7) fruits. The expression of the *PSY1* was significantly reduced ($P<0.005$) in the CRISPR/Cas9 PSY-23. However, in normal tomato fruit, the *PSY1* gene (ID: Solyc03g031860) was expressed from breaker stage onward (The Tomato Genome Consortium 2012). The level of expression level in Ailsa Craig wild-type fruit was consistent with previously reported data (Figure 4.16). Other transgenic *PSY1* lines were examined by RT-PCR gave similar results with reduced expression in the CRISPR lines (Appendix 9-10). The indel mutations in the target site in *PSY1* leads to prematured stop codons and mRNAs that produced may have been degraded by the nonsense mediated mRNA decay pathway (Baker et al., 2004).

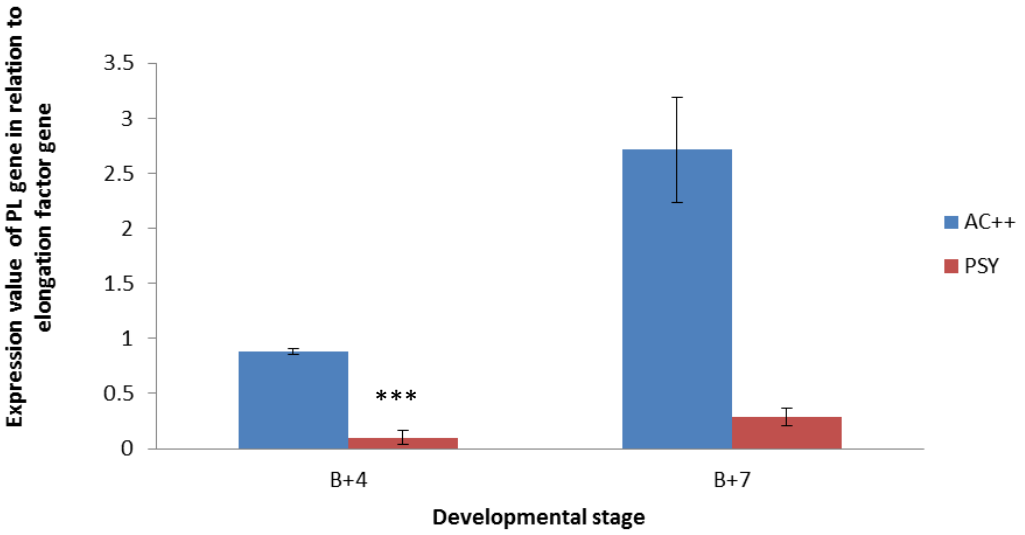


Figure 4.16 *PSY1* gene expression in wild type Ailsa Craig (AC++) and CRISPR::*PSY1* (*PSY*) tomato fruit pericarp at B+4 (breaker +4) and B+7 (breaker+7). The standard error of the mean (SEM) represented by vertical error bars with n=3. Significant differences to wild-type are indicated by *** (P<0.005) from T test analysis.

4.3.5 Phenotypic Analysis

Fruit phenotype

The phenotyping was undertaken on fruits from both PSY-20 and PSY-23 lines. CRISPR::*PSY1* transgenic fruits at mature green stage have a less intense green colour compared to WT. In addition, at Breaker+7 stages, normal tomato fruits (WT) produce red coloured fruits, but the CRISPR::*PSY1* fruits were all yellow in colour (Figure 4-17 and Figure 4-18). Interestingly, differences could also be seen in the colour of the flower petals, with *PSY1* flowers having a bleached/ pale yellow appearance (Figure 4-19). As reported by Fray and Grerson (1992) the yellow flesh *r*, mutation not only expresses in fruits but also shows in flower tissue. These results demonstrated that knocking out the *PSY1* gene leads to inhibition of lycopene biosynthesis in tomato and they are entirely consistent with previously published results using other methods (Fantini et al., 2013). In CRISPR::*PSY1*, the reduced *PSY1* gene expression where generated the yellow colour of the tomato fruits is due to the presence of naringenin chalcone and not carotenoids who biosynthesis is inhibited (Fray and Greson, 1992). The pale yellow colour of the flowers is likely to be due to the expression of the *PSY2* gene that is expressed in these tissues and not affected by our CRISPR mutation (Kachanovsky et al., 2012).

Figure 4-17 Comparison of fruit colour between WT and CRISPR::PSY1-23 fruits at B+7 in T_1 generations.

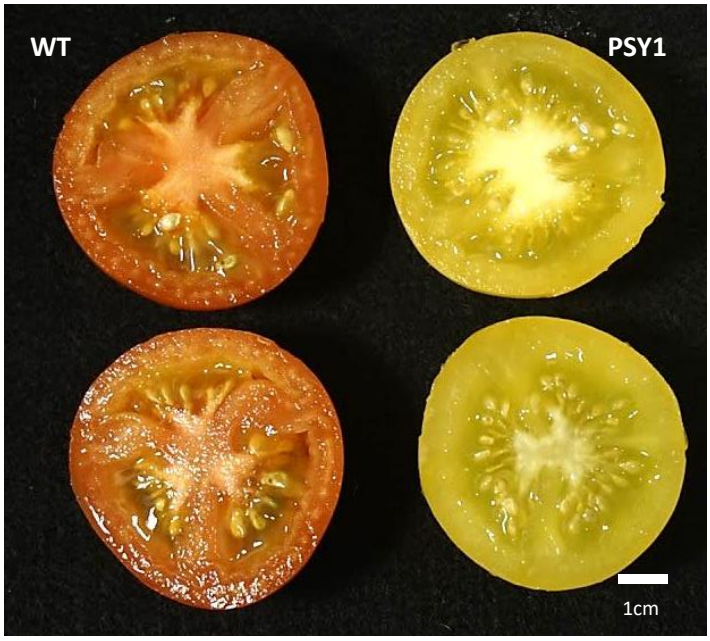


Figure 4-18 Transverse section of wild type and CRISPR::*PSY1*-23 fruits at red ripe (B+7) stage in T_1 generation.

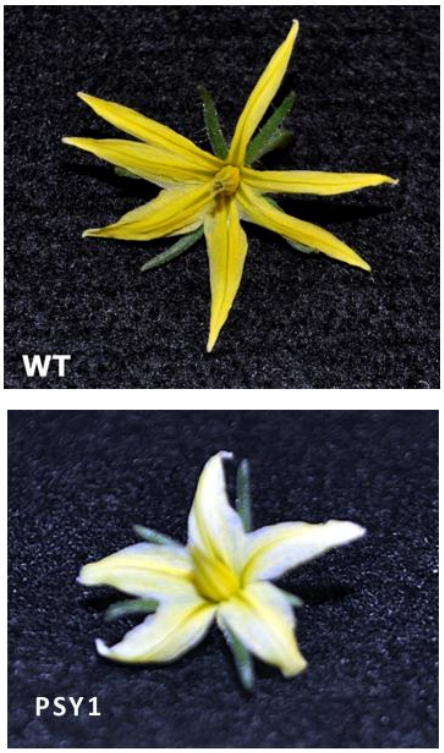


Figure 4-19 Comparison of flower colour between WT and CRISPR::*PSY1*-23 lines in T_1 generation.

4.3.4 Texture Measurement of fruit from homozygous *PSY*::CRISPR lines

Two independent homozygous transgenic lines (PSY-20 and PSY-23) were grown and the pericarp texture of fruits was tested to determine if there were any significant differences between transgenic T_1 and Ailsa Craig wild-type (WT). As shown in Figure 4.20 (A) and (B) the PSY-23 fruits have significantly ($P<0.05$) softer fresh texture compared to WT line especially at breaker +4 days (orange ripe) and breaker + 7 days (red ripe). This was the case for both the outer and inner pericarp tissue. These results are different from those reported by Fraser et al., (2007), which indicated that there were no significant differences in the fruit firmness between WT and lines with reduced *PSY1* gene expression. The reasons for these differences in fruit texture are not known.

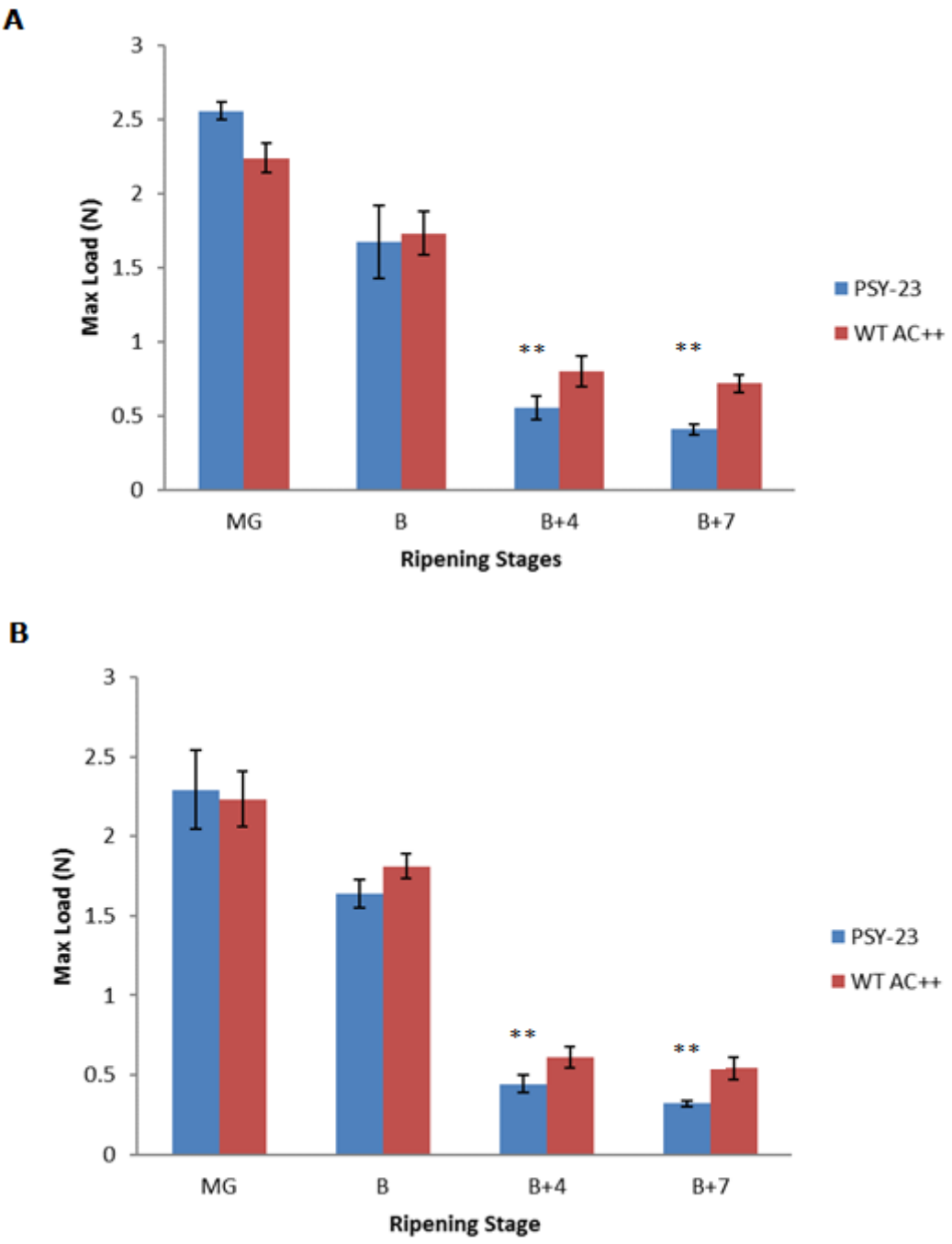


Figure 4.20 Pericarp fruit firmness of CRISPR::*PSY1* and Ailsa Craig WT fruits at various stages of development. The differences in maximum load in the (A) outer pericarp and (B) Inner pericarp at stages MG (Mature green), B (Breaker), B+4 (Orange ripe) and B+7 (Red ripe). The standard error of mean (SEM) is represented by vertical error bars, n=6. Significant differences to control line are indicated by ** ($P<0.05$) from T test analysis.

4.4 Conclusion

The failure to regenerate CRISPR / Cas9 driven mutations in the CESA-like gene was disappointing, but suggests the gene is necessary for plant development. It also suggests that in some instances CRISPR will not be the technology of choice for transgenic experiments where complete silencing of a target gene can be lethal. However, the data from analysis of the *PSY1* transgenic plants validated, with a strong visual phenotype, that our CRISPR/Cas9 approach could induce the mutations in specific genes efficiently and the mutant alleles were transmitted stably in T_1 progeny. Interestingly the CRISPR::*PSY1* fruits had altered texture but the reasons for this change are unclear and require further investigation. In Chapter 5 and 6 the production of CRISPR / Cas9 mutations in the pectin degrading enzymes polygalacturonase and pectate lyase are described and the effects of these mutations on fruit texture and cell wall disassembly are investigated.

Chapter 5

Generation of Mutations in Tomato *Polygalacturonase* using the CRISPR-Cas9 Gene Editor

CHAPTER 5: GENERATION OF MUTATIONS IN TOMATO *POLYGALACTURONASE* USING THE CRISPR-CAS9 GENE EDITOR

5.1 Introduction

Polygalacturonase (*PG*) (EC3.2.1.15) is a hydrolase that degrades pectic polysaccharides (Figure 5-1) and high levels of PG activity are present in many fleshy fruits during ripening including tomato. In tomato, the most abundant ripening-related PG consists of two isoforms *PG1* and *PG2*. Initial work on tomato PG was reported by Hobson (1964) who observed elevated levels of PG activity in ripening tomato associated with the softening process.

In 1981, Tucker et al. reported that there are two isoforms of PG in tomato, *PG1* forms heterodimer that contains of β-subunit glycoprotein. Meanwhile, *PG2* (a monomer) only contains a single catalytic polypeptide (Tucker et al., 1981). Due to its high level of extractable activity, tomato PG is one of the most widely researched cell wall degrading enzymes in fruits and *PG2* acts as an endo-acting enzyme able to hydrolyse the linear α-1,4-D-galacturonan present in cell walls. The biochemical evidence suggested a close relationship between PG activity and fruit softening in tomato. The advent of molecular techniques demonstrated that *PG2* mRNA was synthesised at the onset of ripening and transgenic plants were generated with an antisense RNA construct to test the role of *PG2* in ripening (Smith et al, 1988, Smith et al, 1990).

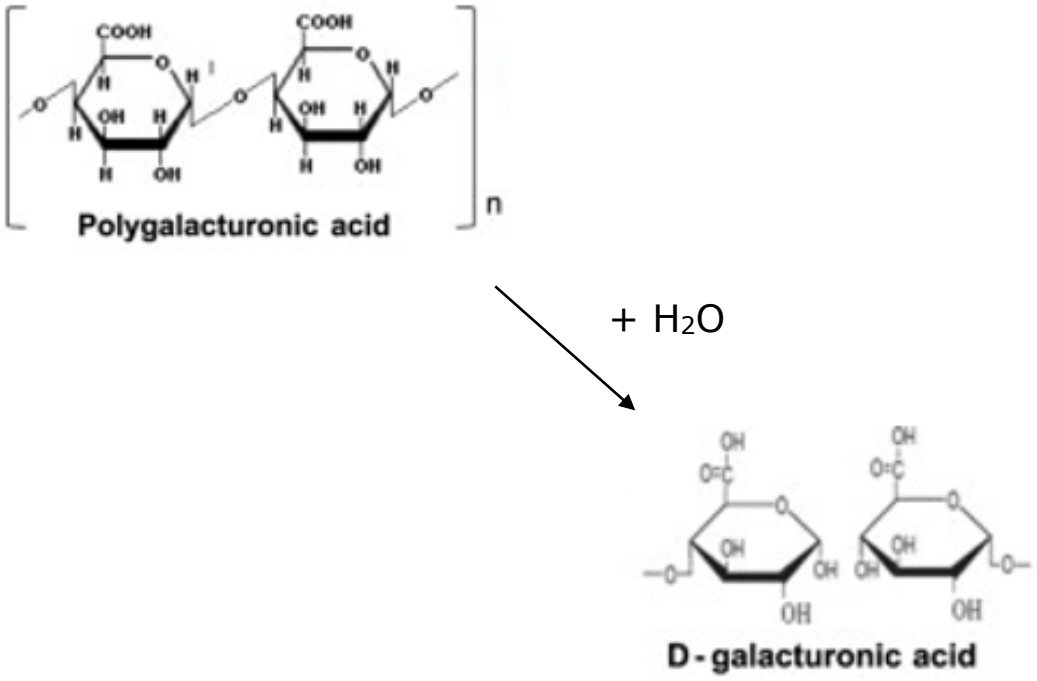


Figure 5-1 Polygalacturonase (PG) enzyme reaction. PG function is to hydrolyse the α-1-4 linkages galacturonosyl residues in polygalacturonic acid which produces D-galacturonic acid. PG has been implicated as one of the agents in the degradation of polyuronides during fruit ripening which contribute in altering fruit texture (taken from Ramírez-Tapias et al., 2014).

Unexpectedly transgenic tomato lines with PG levels reduced to 1% of normal showed no significant difference in the extent of fruit softening, but they did have lower levels of pectin depolymerisation during ripening (Smith et al., 1990). In a complementary experiment Giovannoni et al (1989) over expressed PG in the non-softening fruits of the tomato *ripening inhibitor* (*rin*) mutant. Although, *rin* showed 60% of the normal PG activity compared to the wild-type fruits, the experiment failed to induce softening in the mutant fruits. This indicated that the *PG* was important for pectin degradation, but was not the only factor involved in fruit softening. However, the antisense constructs used in the above experiments did not remove all the *PG* gene expression leaving open the possibility that low levels of this highly active enzyme were sufficient to cause the softening effects that were observed in the antisense lines. Additionally PG genes are expressed in a wide range of other fruits and in strawberry and apple have been shown to be responsible for a substantial part of fruit softening (García-Gago et al., 2009, Quesada et al, 2009, Atkinson et al, 2012).

The objective of experiments in this chapter was to generate a *PG* knockout using CRISPR/Cas9 DNA editing and examine the influence of this mutation on tomato fruit softening. The fruit from these lines can then be compared with those from other CRISPR lines, e.g. pectate lyase (Chapter 6).

5.2 Material and Methods

5.2.1 Plant materials

Plant materials are described in detail in Chapter 2.1. Fruit were harvested in the winter of 2015 and spring 2016. In the winter of 2015, fruit were harvested for texture test, colour total soluble solid and qPCR and in spring 2016 the experiment was repeated with fruit tissue used for texture testing, soluble solids, colour, enzyme assays, and viscosity analysis.

5.2.2 Vector construction for generating stable transgenic plants

Vector construction

Primers were designed (Table 5-1) using PG sequence (Solyc10g080210) from tomato genome browser at [www.solgenomics.net]. A specific sequence within the gene was then chosen to represent the guide RNA (Figure 5-2) and the vector for transformation was constructed following the protocol described in Chapter 4 that included the Cas9 nuclease. To generate the stable transgenic plants, constructs Level 1 (pICH47751::AtU6p::sgRNA, pICH47732::NOSp::NPTII-OCST, pIcH47742::35Sp::Cas9-NOST, Pich41766 Linker) were used and assembled to Level 2 (pAGM4723) via the Golden gate cloning method (Weber et al., 2011) (Chapter 4).

Table 5-1 Primer sequence of *PG* gene

Primer ID	Primer sequence 5’ to 3’
PG2A001.F	TGTGGTCTCAATT**GTGATTAATGTACTTAGCTT**GTTTTAG AGCTAGAAATAGCAAG
PG2A001.R	TGTGGTCTCAAGCGTAATGCCAACTTTGTAC

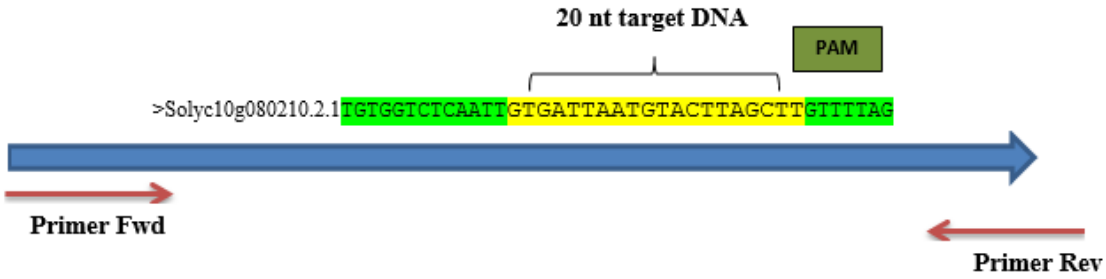


Figure 5-2 Cas9/sgRNA-mediated mutagenesis in *S. lycopersicum*. The 20 bp target region of Solyc10g080210 coding sequence that was selected from the sense DNA strand is highlighted in yellow.

5.2.3 Generation of stable transgenic plants

The methodology for plant transformation was described in detail in Chapter 2, Section 2.10.

5.2.4 Identification of the *PG* transgenic lines

The presence of T-DNA in the transgenic plant lines was determined by performing a set of PCR reactions using PG2A001 gene specific primer and the NPTII/Cas9 primer (Appendix 9.8). The DNA from wild-type AC^{++} was used as a positive control for the gene specific primer and as negative control for T-DNA mutation. The PCR conditions for this reaction are given in Table 4-5 and the presence of transgene was confirmed by sequencing (Appendix 9.9).

5.2.5 Identification of single copy homozygous lines

The T_0 tomato plants were grown to fruiting and twelve T_1 seed were sown for each T_0 line. The DNA from each of the plants was extracted using DNeasy Plant Mini Kit (Qiagen) (Chapter 2.2). The DNA samples were then used as templates for PCR amplification to check for the presence of the transgene using reaction conditions in Table 4-5. The PCR products were then sent for sequencing.

5.2.6 qPCR determination of gene expression

The methodology for qPCR was described in detail in Chapter 4, Section 2.6.

5.2.7 Determination of PG enzyme activity

Fruits were harvested at B+7 (breaker plus seven days) stages. The pericarp tissue of fruit was frozen at -80°C on the day of harvest and stored until required. Three fruit from each line were used for cell wall bound protein extraction and protein assays.

Extraction of cell wall bound protein

The protocol from Tucker et al., (1980), was used for protein cell wall extraction. About 10 g of frozen tomato pericarp tissue was taken from single fruit and ground using a coffee grinder. The homogenised tissue was then mixed with the extraction buffer that consists of 1 M sodium chloride (Fisher Scientific) and 0.05 M sodium acetate (Fisher Scientific) with the pH being adjusted to 6 with sodium hydroxide (Fisher Scientific). This mixture was then homogenized using Polytron PT Homogenizer for 10 minutes and then placed on the roller mixer for 3 hours at 4°C. The homogenate was then centrifuged for 30 minutes at 2000 g. The supernatant was then recovered and solid ammonium sulphate (Sigma-Aldrich) added to 80% saturation. This mixture was left overnight at 4°C before being centrifuged at 10,000 g for 20 minutes at 4°C. The

supernatant was then discarded and the pellet was dissolved in 3 mL of 0.15 M sodium chloride and 0.05 M sodium acetate and then stored in 2 mL aliquots at -20°C until required.

Protein quantification

The protein quantification protocol was adapted from Bradfod (1976) method by using standard concentration (0 – 20 µg) of bovine serum albumin (BSA) solution (Sigma-Aldrich). 20 µL of the BSA solutions were mixed well with 1000 µL dye reagent. Then, each of the standards was loaded into a 1 mL cuvette and the absorbance read at 595 nm using a Multiscan FC (Thermo Scientific, Loughborough, UK) microplate photometer. A standard curve was plotted and the protein content was calculated in µg mL^{-1} of assay volume.

Polygalacturonase assay

The level of PG activity was measured using 25 µL of protein sample diluted with 25 µL dH_2O. Then, 150 µL of an assay mixture was added containing 50 mM sodium acetate (Fisher Scientific) and 0.5% polygalacturonic acids (Fulka) at pH 4. This mixture was incubated at 37°C for 10 minutes. To terminate the enzyme reaction, 1 mL of 100mM sodium tetraborate (Sigma-Aldrich) and 200 µL of 1% cyanoacetamide (Sigma-Aldrich) was added to the sample and incubated at 100°C for 10 minutes. The absorbance of the resulting solution was read at 276 nm and

the PG activity expressed was calculated from standard curve produced from 0 – 10 µM galacturonic acid (Sigma-Aldrich).

5.2.8 Analysis of fruit characteristics

Analysis of fruit texture, colour and brix followed the same protocols as described in Chapter 4.2.7.

5.2.9 Determination of viscosity behaviour in *PG* tomato purees

Puree for viscosity analysis was obtained from cubed and peeled fresh tomato. The fruits were then blended in a coffee grinder for 1 minute until it gave a suitable texture. The measurement of viscosity for fresh puree was done using cone-on-plate rheometer (RHEOPLUS/32 V3.40) with a gap of 0.049mm at a constant temperature of 20°C. The viscosity and shear stress were measured within shear rates of 1-100 s^{-1} and 100-1 s^{-1}.

5.3 Results and Discussions

5.3.1 Vector construction PG transgenic lines

The fragment of sgRNA using the primer product from PG2A001 (Table 5-1) amplification from Level 0 was cut-ligated and cloned into vector Level 1 and assembled to Level 2. The resulting insert was shown to be correct by PCR (Figure 5-3 and 5-4) and sequencing (Appendix 9.8) and the transformations of CRISPR::Cas9 vector construct into *Agrobacterium* EHA105 was validated by PCR amplification using NPTII/Cas9 specific primer (Figure 5-5).

5.3.2 Analysis of the *PG* transgenic lines

Ten transgenic Ailsa Craig lines generated contained a CRISPR-Cas9 construct that was designed to down-regulate the ripening-related PG (Solyc10g080210). The presence of the T-DNA in all these plants was confirmed by PCR amplification (Figure 5-6) using NPTII/Cas9 primer with the expected size of the product as 1,000 bp. The PCR generated an amplicon of the correct size in transgenic plant lines No. 6, 10, 13, 14, 21, 24, 34, 36, 40, and 49 and there was no amplification product in WT plants as expected. This result demonstrated that these lines contained Cas9 gene.

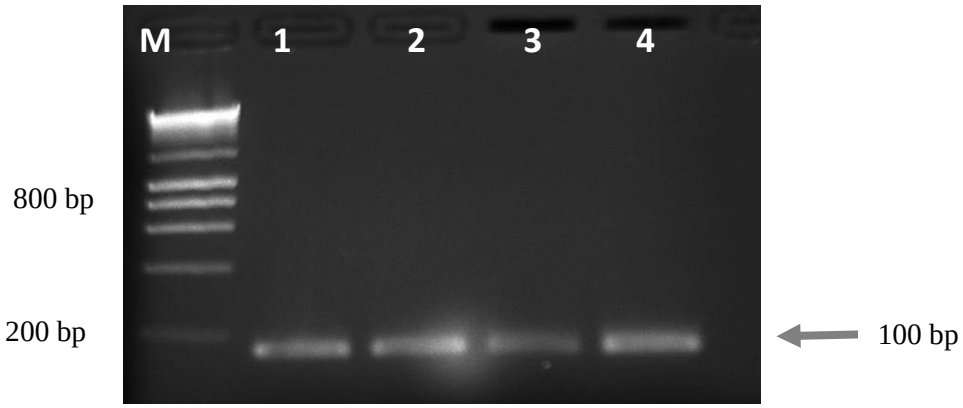


Figure 5-3 The gene specific PG PCR that was amplified using PG2A001 primer. The purified PCR products were separated by electrophoresis on a 2% (w/v) agarose gel in 0.5X TAE at 100 V for 30 minutes. A) Lane M: Hyperladder Marker. Lane 1-4 PCR product of sgRNA pICH86966::AtU6p::sgRNA_PDS construct that contained the PG target sequence.

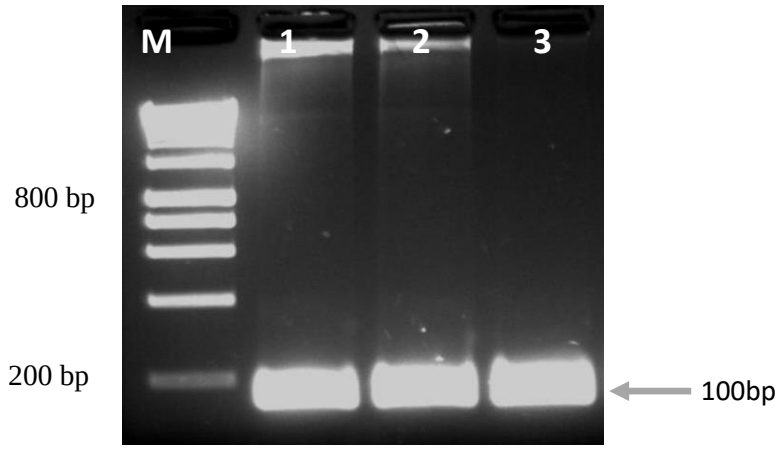


Figure 5-4 Cloning of *PG* gene fragment into the Level 2 vector. The PCR products of Level 2 construct were amplified by PG2A001 primer using the entry clone (pICH47732::NOSp::NPTII-OCST, pICH47742::35Sp::Cas9-NOST, pICH47751::AtU6p::sgRNA, pICH41766 and pAGM4723 vector recombination) to be a template. A) Lane M: Hyperladder Marker, lane 1-3: the construct CRISPR::PG contructs.

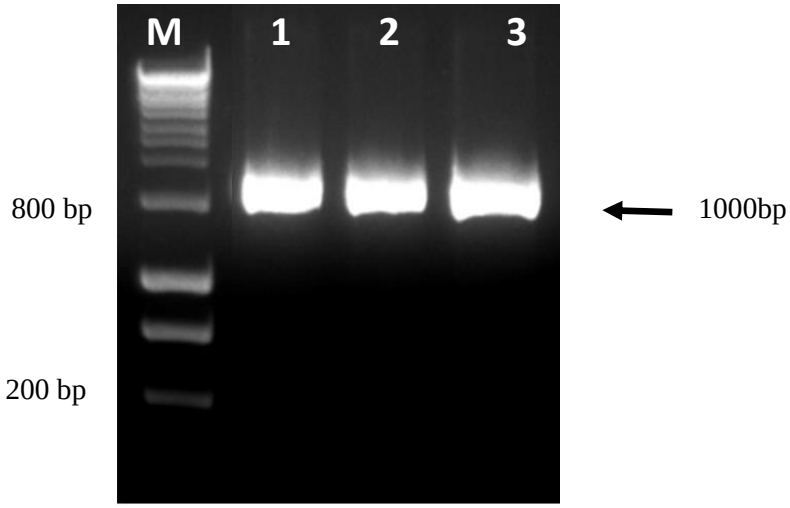


Figure 5-5 PCR products were amplified using the NPTII/Cas9 specific primers from *Agrobacterium* EHA105. The purified PCR products were separated by electrophoresis on a 2% (w/v) agarose gel in 0.5X TAE at 100 V for 30 minutes. Lane M: Hyperladder Marker, lanes 1-3 PCR products of CRISPR::PG constructs which were isolated from colonies of *Agrobacterium* EHA105.

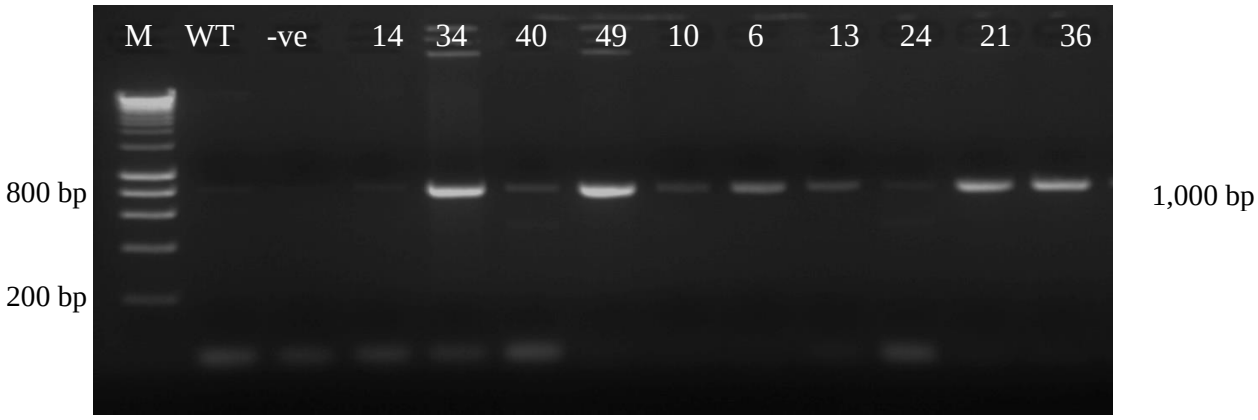


Figure 5-6 PCR genotyping of ten representative CRISPR::PG plants showing the transgenic lines that contain Cas9/sgRNA gene construct in T_0 plants. Lane M: Hyperladder Marker, lanes 14, 34, 40, 49, 10, 6, 13, 24, 21 and 36 PCR products that contain CRISPR::PG contructs.

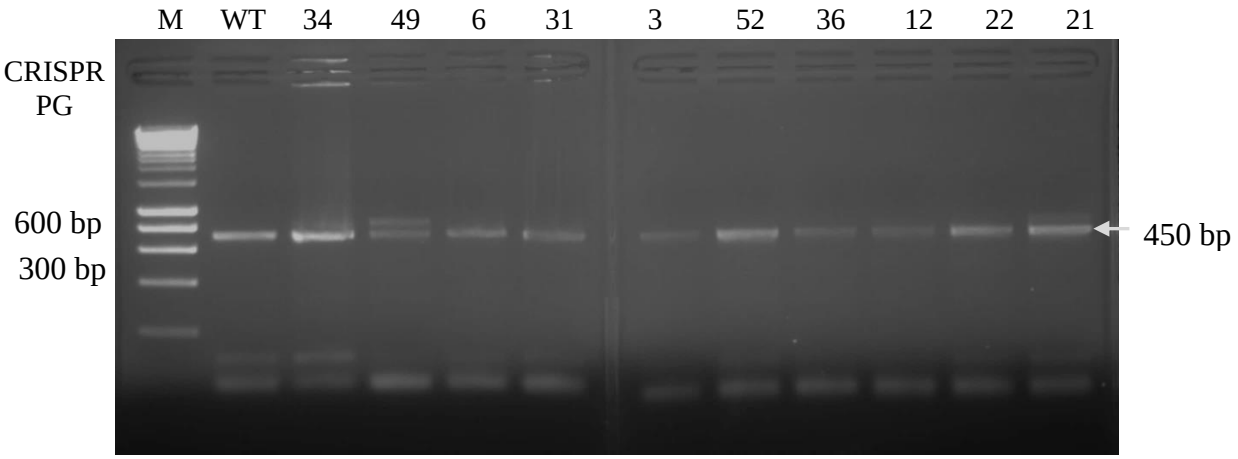
Validation of CRISPR / Cas9 driven mutations in the *PG* genes

Verification of mutagenesis in the endogenous gene was undertaken by a combination of a PCR and restriction digestion strategy, along with sequencing, in the T_0 lines (Figure 5-7). In the non-mutagenized plants (WT) the enzyme *RsaI* will cleave a 450 bp amplified PCR product into two fragments of 225 bp fragment size. Figure 5-7 shows *RsaI* digestion patterns of PCR-amplified DNA from 10 individual T_0. These results indicated that all of ten transgenic plants could be digested with *RsaI* activity and therefore there were no mutations at the restriction site. The PCR products from T_0 and wild-type plants were purified and sent for sequencing to confirm the site of mutagenesis in the *PG* gene. However, the sequences of target area in *PG* gene showed mutations in the form of an insertion of a C nucleotide in line PG-21 and deletion of single nucleotide in line PG-34. The remaining lines were similar to wild-type (Figure 5-8). Twelve seeds from CRISPR::PG T_0 lines 34 and 21 were grown to obtain homozygous lines in the T_1 generation.

5.3.3 Identification of Homozygous lines

To identify the homozygous line, the DNA was extracted from leaves of transgenic plants (PG-21 and PG-34) using DNeasy Plant Mini Kit (Qiagen) as described in Chapter 2, Section 2.2. Primers NPTII/Cas9 (Appendix 9-8) was used to amplify the Cas9 gene which gave amplicons of 800 bp (Figure 5-9).

A



B

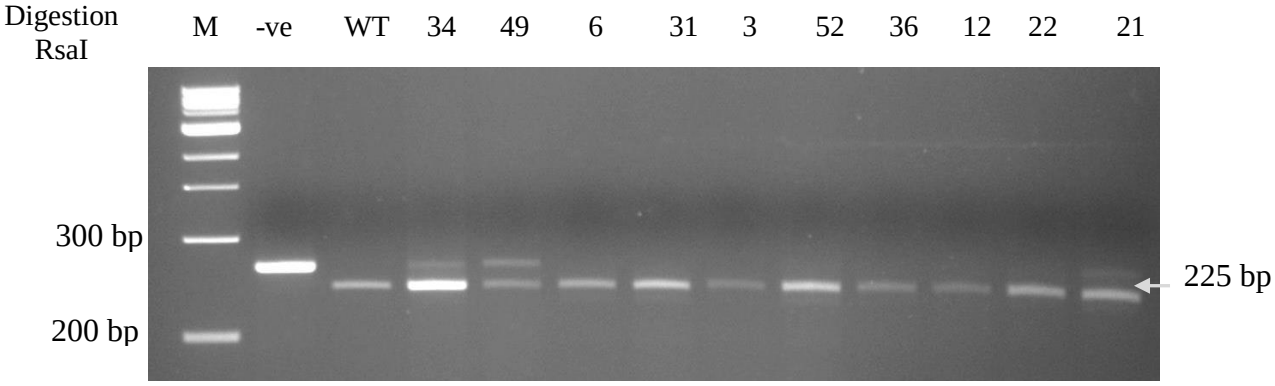


Figure 5-7 Efficiency of Cas9/sgRNA mutagenesis. (A) PCR genotyping of ten representative CRISPR::PG plants using gene specific primer (PG2A001). (B) PCR/Restriction analysis where wild type sequences are expected to give bands at ~225 bp DNA resulting from *RsaI* cleavage of the 450 bp PCR product. Lane M: Hyperladder Marker, lanes 34, 49, 6, 31, 3, 52, 36, 12, 22 and 21 PCR products of CRISPR::PG.

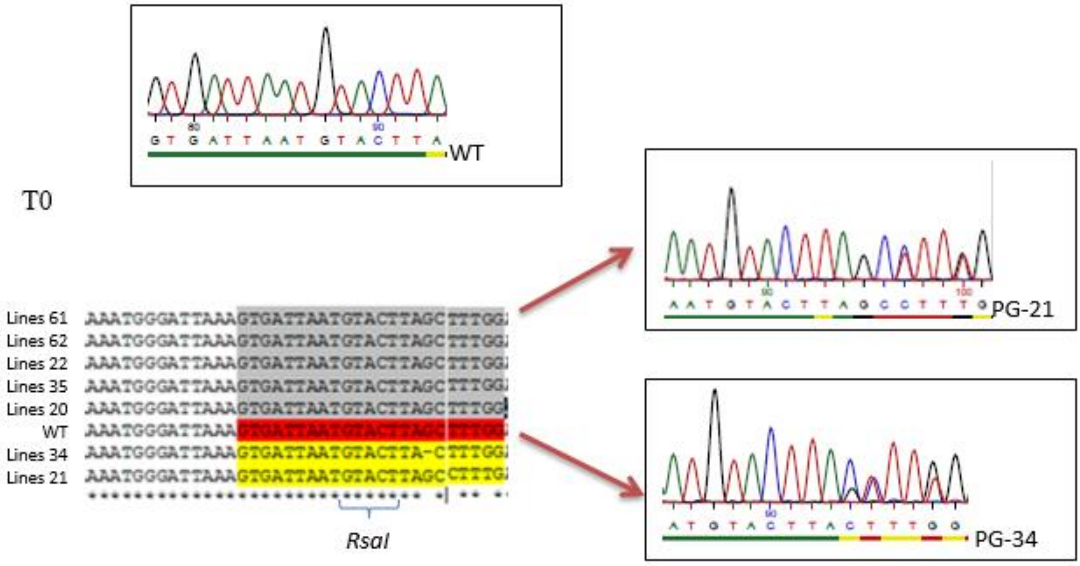


Figure 5-8 Confirmation of inheritance of a modified or non-modified *PG* gene. Sequence alignment and DNA sequencing traces from sequencing of PCR amplified CRISPR::PG as target sites in which there has been a Cas9/sgRNA-mediated gene modification. A range of insertion and deletions (indels) were identified from the sequencing results in T_0 plants which are lines PG-21 and PG-34. The PCR products were then cloned and re-sequenced to validate the indels and characterize additional DNA lesions.

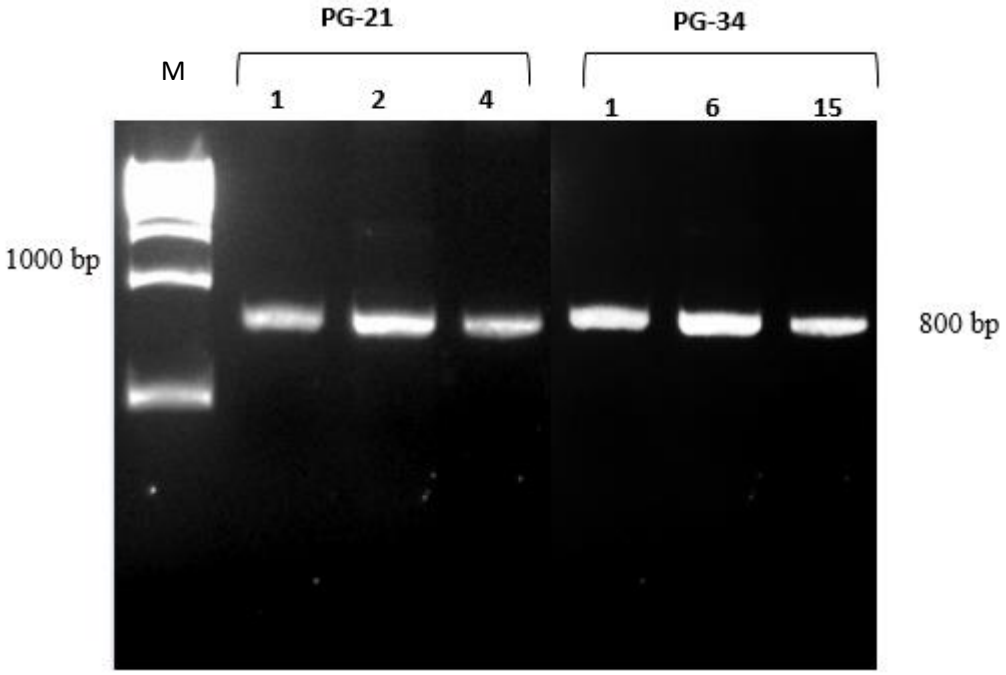
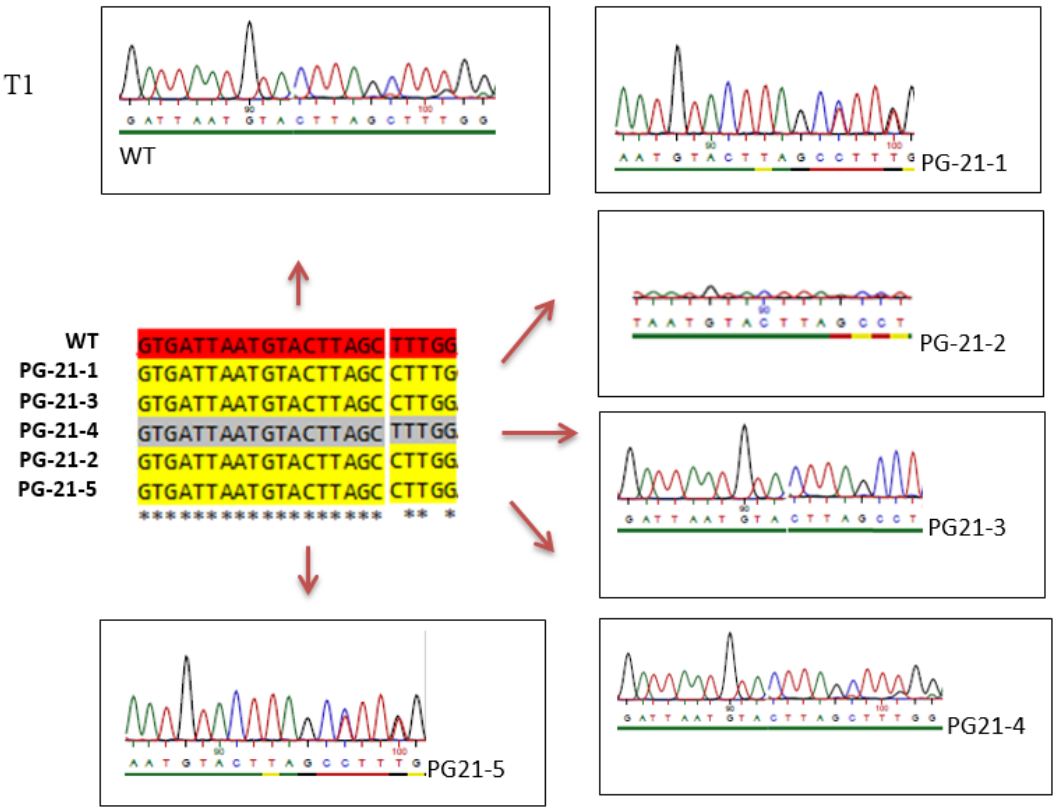


Figure 5-9 Confirmation of inheritance of modified *PG* gene in each of T_1 generation plants. Out of the three progeny from plant line CRISPR::PG21 and CRISPR::PG34, all plants showed the presence of CRISPR/Cas9 transgene. Lane M: 100 bp DNA ladder marker, lanes PG-21: 1, 2 and 4; PG-34, 1, 6 and 15 PCR products generated by amplification of NPTII/Cas9.

The PCR result indicated CRISPR/Cas9 transgene was present in all of the transgenic lines.

Putative homozygous lines in the T_1 generation were identified from the sequence analysis of the PCR products. In the case of the T_0 line 21 (Figure 5-10), T_1 progeny PG-21-2 and PG-21-3 were identified homozygous with a substitution at T to C, PG-21-1 and PG-21-5 were heterozygous and PG21-4 had a sequence identical to wild-type and was selected as an azygous control. Analysis of the T_1 progeny of the CRISPR::PG-34 line (Figure 5-11) identified PG-34-2, 3, 5 and 6 to be heterozygous and PG-34-4 to be an azygous line. The sequence analysis also showed two types of homozygous mutations in PG-34-1 and PG-34-15. In PG-34-1 four nucleotides were deleted in the target region and a single deletion of a G nucleotide was found in the PG-34-15 line (Figure 5-11). The base changes in the CRISPR PG lines resulted in stop codons in the coding sequence of *PG2a* (Figure 5-10B and 5-11B) and therefore we predicted that the mutant lines will have no PG enzyme activity.

(A)



(B)

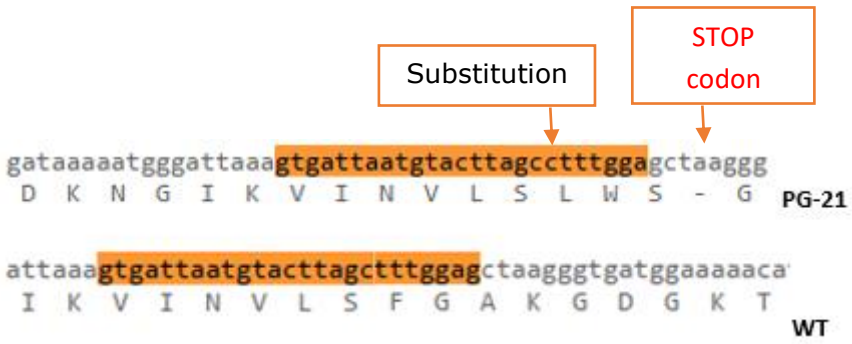
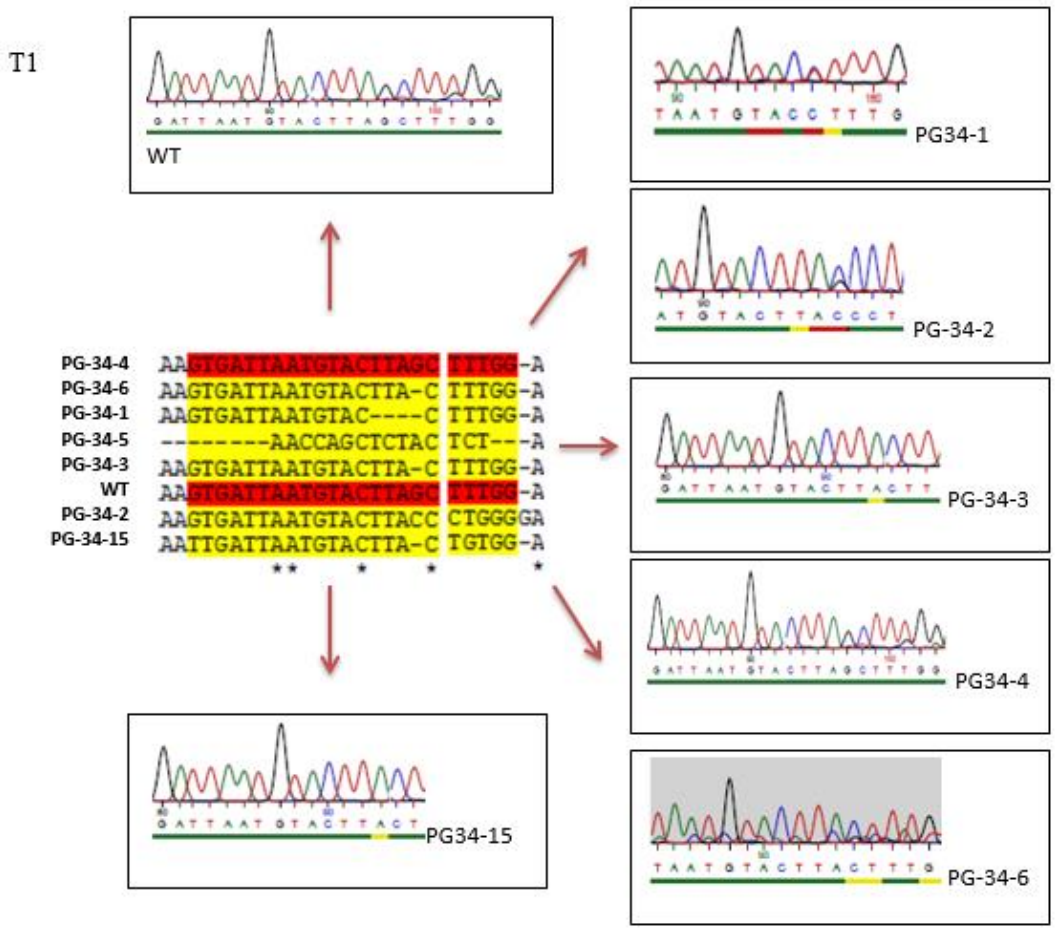


Figure 5-10 Inheritance of modified *PG* gene in each of T_1 progeny. (A) DNA sequence alignment and sequencing traces of CRISPR::*PG*-21 with varied of mutation rate at the target regions although all of them still inherited CRISPR/Cas9 from T_0. (B) Effect on protein sequence of transgenic lines i.e PG21-3.

(A)



(B)

Figure 5-11 Inheritance of modified *PG* gene in each of T_1 progeny. (A) DNA sequence alignment and sequencing traces of CRISPR::PG-34 with varied of mutation rate at the target regions although all of them still inherited CRISPR/Cas9 from T_0. (B) Effect on protein sequence of transgenic lines i.e PG34-1.

5.3.4 Gene expression of *PG*

PG gene expression was determined using QPCR at two stages of ripening, breaker+4 (orange) and breaker+7 (red ripe) fruits from CRISPR::PG-34 line. In tomato, *PG* (ID: Solyc10g080210) is expressed highly during ripening process (breaker stage onwards) as reported by The Tomato Genome Consortium (2012) (Figure 5-12). However, the expression level of *PG* in CRISPR::PG-34 was significantly reduced ($P<0.001$) at both ripening stages (Figure 5-12). This is supported by RT-PCR of the other PG transgenic lines which gave similar results (Appendix 9.10).

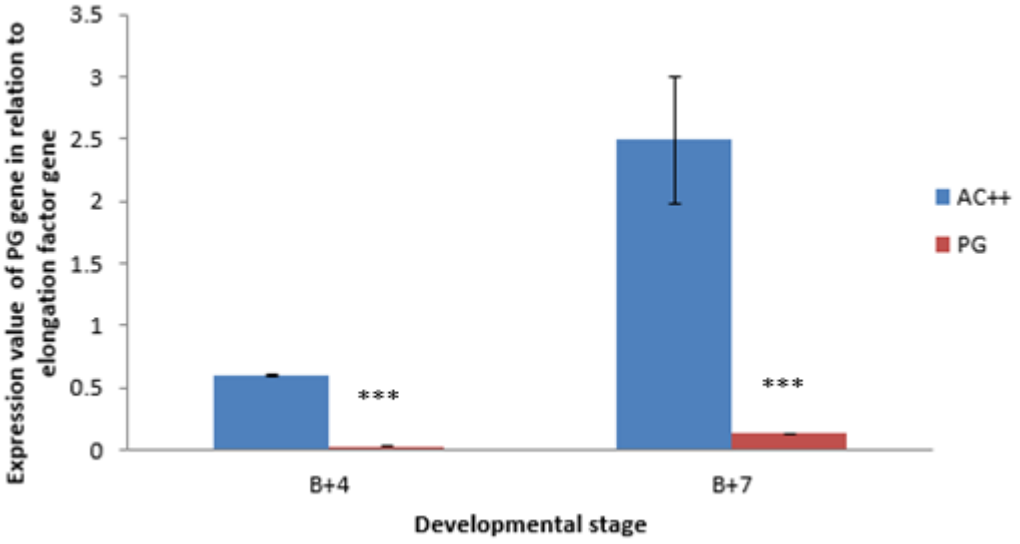


Figure 5.12 *PG* gene expression in wild-type Ailsa Craig and CRISPR::PG-34 (T_1 generation) at B+4 (breaker +4) and B+7 (breaker+7). The standard error of the mean (SEM) represented by vertical error bars with n=3. Significant differences to wild-type are indicated by *** ($P<0.001$) from T test analysis.

5.3.5 *PG* Enzyme Activity

PG enzyme assays were performed to determine if the CRISPR mutations had substantially reduced PG activity in ripe fruits. The enzyme activity was measured at red ripe (breaker+7) stage in fruits from line PG-34 (Figure 5-13). The PG activity was determined using the 2-cyanoacetamide method (Gross, 1982). The CRISPR::PG-34 activity was significantly reduced ($P<0.005$) compare to wild-type at the same stage of ripeness (Figure 5-13). The residual activity in the CRISPR line is probably accounted by the low levels of ripening-related expression of two other *PG* genes (Solyc05g049980 and Solyc09g075460) reported by the Tomato Genome Consortium (2012). Thus, the reduction of *PG* activity in CRISPR::PG line was entirely consistent with the generation of inactive PG protein.

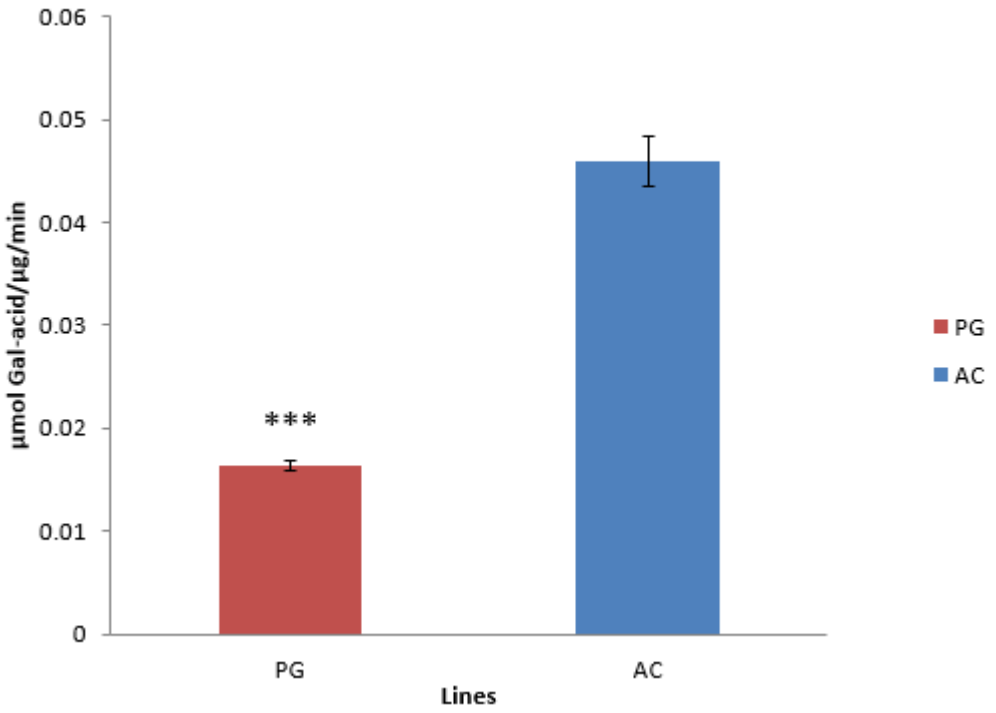


Figure 5-13 *PG* activity of CRISPR::PG-34 compare with wild-type during red ripe stages. The standard error mean of the (SEM) represented by vertical error bars with n=3. Significant differences to wild-type are indicated by *** (P<0.005) from T test analysis.

5.3.6 Phenotypic Analysis

Fruit phenotype

Fruits from CRISPR::*PG* transgenic lines and wild-type Ailsa Craig plants were collected at mature green and breaker +7 days stages from two seasons which is in winter 2015 and spring 2016. The phenotype analysis indicated that transgenic fruits at mature green had a slightly more intense green colour than those of wild-type (Figure 5-14). In the red ripe stage, the transgenic fruits also appeared a slightly more intense red. All the analysis was undertaken using the fruit harvested in winter 2015 season (Figure 5-15). Optical measurement of the colour of the fruits was undertaken using a Minolta Colormeter. The differences between fruits with respect to pericarp colour were not significant (Figure 5-16 A). Consistent with these observations, measurement of the total soluble solids also indicated that there were no significant differences between the fruits (Figure 5-16 B).

(A)

(B)

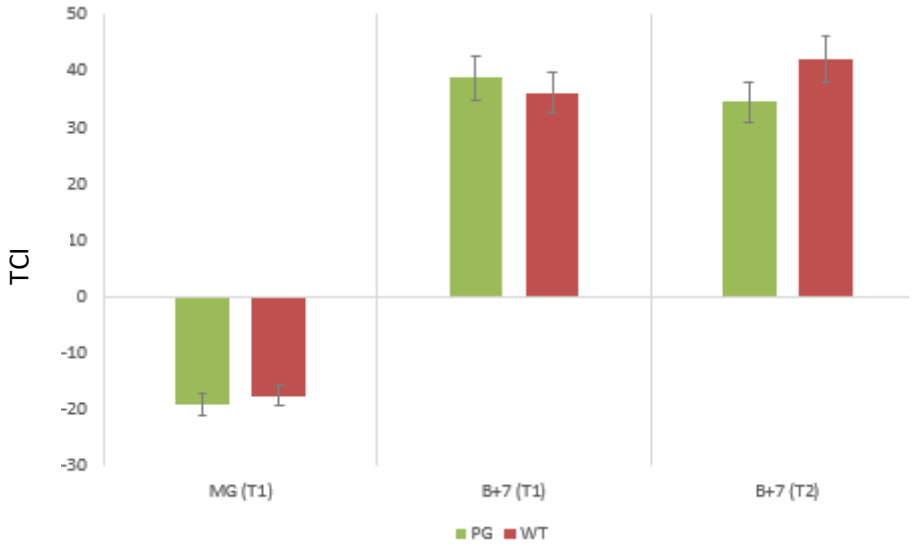
Figure 5-14 Comparison of fruit colour between WT and CRISPR::PG34 fruits at mature green stage from winter 2015 season. (A) Side view of WT and CRISPR::PG34. (B) Transverse section of wild type and CRISPR::PG34 fruits.

(A)

(B)

Figure 5-15 Comparison of fruit colour between WT and CRISPR::PG34 fruits at breaker+7 stages from winter 2015 season. (A) Side view of WT and CRISPR::PG34. (B) Transverse section of wild type and CRISPR::PG34 fruits.

(A)



(B)

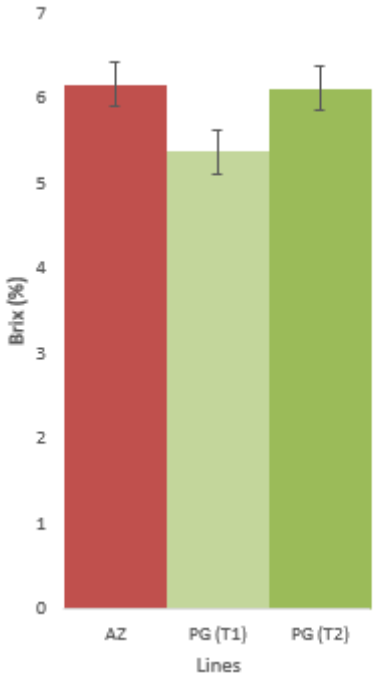


Figure 5-16 Fruit characteristics with respect to (A) Tomato colour index (TCI) values from wild type (WT AC++) and transgenic fruits (*PG*) containing CRISPR::*PG* transgene at MG and B+7 stages and (B) Brix measurement in red ripe fruits. The standard error of mean (SEM) is represented by vertical error bars and n=4 (biological replicates).

5.3.7 Texture Measurement on fruit from homozygous PG::CRISPR lines

Two independent lines (PG-21 and PG-34) were selected and grown for texture analysis. The results for line (PG-34) are shown in this Chapter and data for CRISPR::PG21 are shown in the Appendix 9.14. In this analysis, the maximum load of tomato pericarp tissue was measured to determine the firmness of the fruits. This involved measuring the force required to drive a 10 mm probe into a transverse slice of the fruit at a known speed and to a specific distance. These mechanical measures correlate with sensory measures of fruit texture (Chapman et al, 2012).

In the winter grown fruit (T_1), the firmness of CRISPR::PG fruit pericarp (Figure 5-17) was apparently enhanced to some degree at most stages of ripening as compared to the wild-type. However, based on T-test analysis the $P>0.05$ meaning that there are no significant differences in pericarp firmness between CRISPR::PG-34 line and control line in both inner and outer pericarp. A similar result can be seen in the T_2 generation (spring season), the firmness differences between the CRISPR::PG34 line and azygous controls were less pronounced and no significant ($P>0.05$) differences were apparent (Figure 5-17 B and Figure 5-18 B).

According to Giovannoni et al., (1989) and Smith et al., (1990) silencing of *PG* gene by antisense RNA had no effect on fruit softening. The small effect seen in the current study may reflect the complete knock down of the *PG* gene in comparison to silencing to about 1% of normal levels in the best antisense lines (Smith et al, 1990). However, the effects seen are small and the analysis is confined to single plants of two independent lines due to virus infection on other plants.

5.3.8 Viscosity behaviour in *PG* tomato purees

The puree from the lines of interest CRISPR::PG-34, PG-21 and azygous (as a control) were tested to investigate their material properties and gelling ability. The samples were placed in coaxial cylinders and then spinning probe will measured the viscosity of the sample. The Control Share Rate (CSR) is depending on time of shear stress and rate of the sample being measured. Figure 5-19 and Table 5-2 shows the comparison of the CSR between CRISPR::PG and azygous lines. In the original studies on *PG* the effect on paste viscosity indicated that *PG* action was responsible for depolymerisation of the tomato cell wall pectin (Smith et al., 1990). The higher molecular weight pectin in the antisense *PG* fruits resulted in increased tomato paste viscosity (Brummell and Labavitch 1997) and also see this in this CRISPR lines.

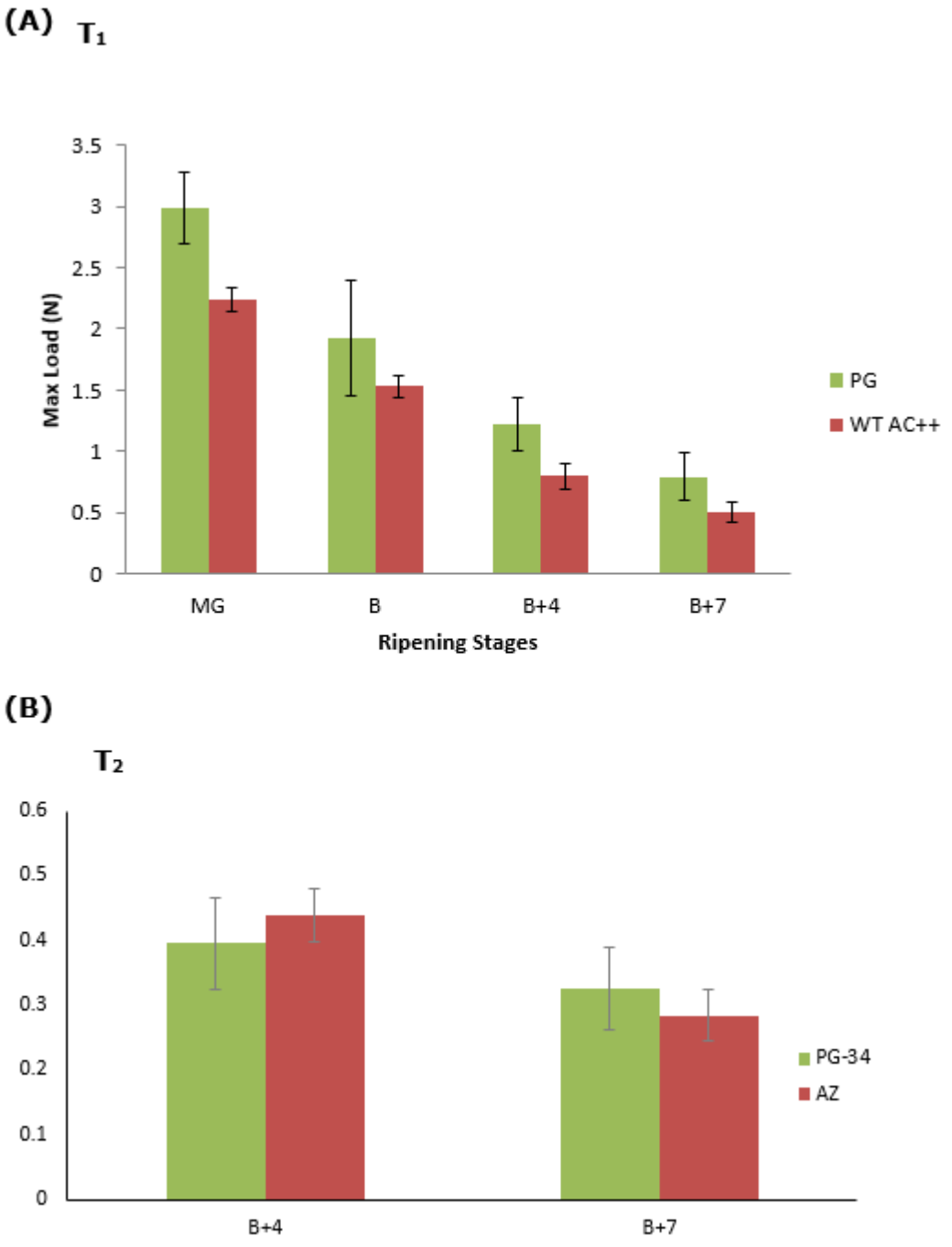


Figure 5.17 Pericarp firmness as result of CRISPR silencing of *PG*. (A) The differences in outer pericarp in T_1 generation for all development stages (B) Outer pericarp at stage B+4 (Orange ripe) and B+7 (Red ripe) in T_2 generation. The standard error of the mean (SEM) represented by vertical error bars with n=4. No significant differences to wild-type were indicated from T test analysis

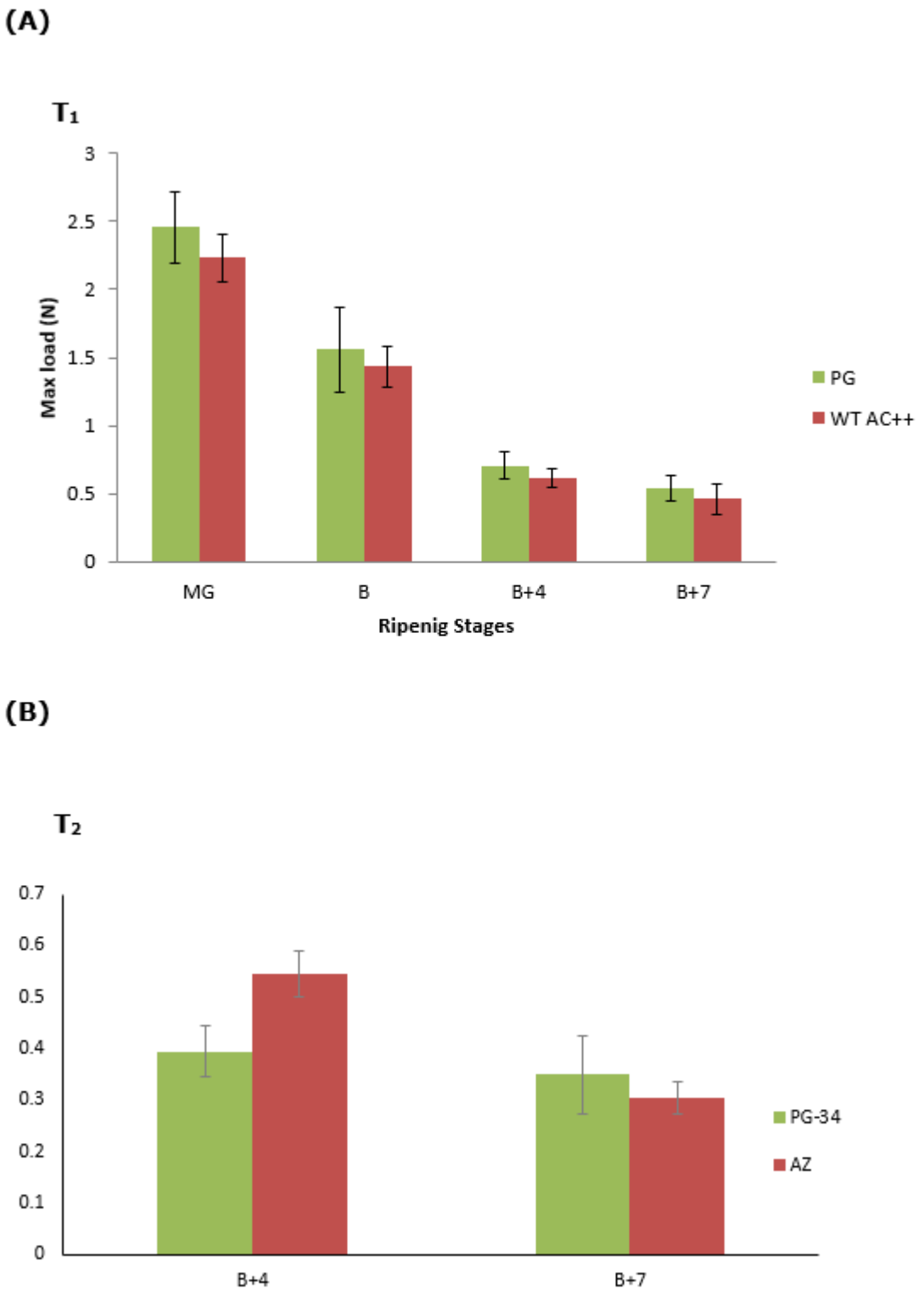


Figure 5.18 Inner pericarp firmness result of CRISPR silencing of *PG*. (A) The differences in inner pericarp in T_1 generation for all development stages (B) Inner pericarp at stage B+4 (Orange ripe) and B+7 (Red ripe) in T_2 generation. The standard error mean (SEM) represented by vertical error bars with n=4. No significant differences to wild-type are indicated from T test analysis

The viscosity, η of the puree samples at low (1 s^{-1}) and high (100 s^{-1}) shear rate is shown in Table 5-2. The η is significantly influenced by plant line at both 1 s^{-1} (P=0.02) and 100 s^{-1} (P≤0.05), with CRISPR::PG having a significantly higher η than azygous at 1s^{-1} and 100 s^{-1} (P<0.05). This indicated that *PG* puree was more elastic and had stronger gel like properties than poor structuring lines (azygous). According to Barnes et al. (1989), by measuring the yield point of the sample it also give information on the resistance flow. The higher the value of the yield point of the sample, the more elastic and gel-like is that material. Interestingly, several studies also indicated that the viscosity of juice and paste prepared from antisense *PG* lines also significantly increased due to low level of soluble pectin during processing (Schuch et al., 1991; Kramer et al., 1992; Brummell and Labavitch 1997; Kalamaki et al., 2003).

Table 5-2 Shear viscosity values at low (1 s^{-1}) and high (100 s^{-1}) shear rates for the CSR test of the tomato puree of CRISPR::PG and azygous line in Figure 5-19. Means are displayed with the error value as SD. Value significantly different to the control line are indicated by * (P<0.05). Significant values are displayed in the final row of the relevant column of each table.

Line	η at 1s^{-1}	η at 100s^{-1}
Azygous	3.26 (±0.434)	0.14 (±0.006)
CRISPR::PG	5.18 (±0.023)*	0.17 (±0.058)*
	P=0.02	P=0.05

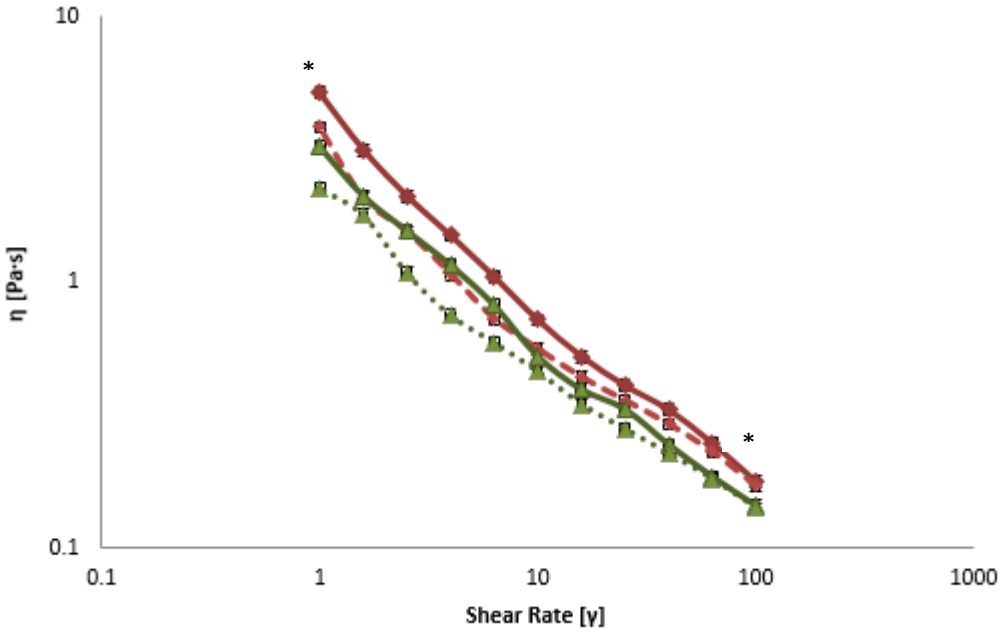


Figure 5-19 The CSR test data from CRISPR::PG and azygous line puree. The symbol ◆ represent CRISPR::PG and ▲ for azygous line. ——◆—— and ——▲—— are the viscosity values at 1-100 s^{-1} share rate. - -◆- - and ···▲··· are the viscosity values at 100-1 s^{-1} share rate. The standard error of mean (SEM) is represented by vertical error bars and n=4 (biological replicates). Value significantly different to the control line are indicated by ** ($P<0.05$).

5.4 Conclusion

The data from analysis of CRISPR::PG transgenic plants demonstrated that CRISPR/Cas9 can induce mutations in PG. The CRISPR::PG also shows a small effect on fruit texture and also impacts the characteristics of the fruit puree.

Chapter 6

Generation of Mutations in Tomato *Pectate Lyase* using the CRISPR-Cas9 Gene Editor

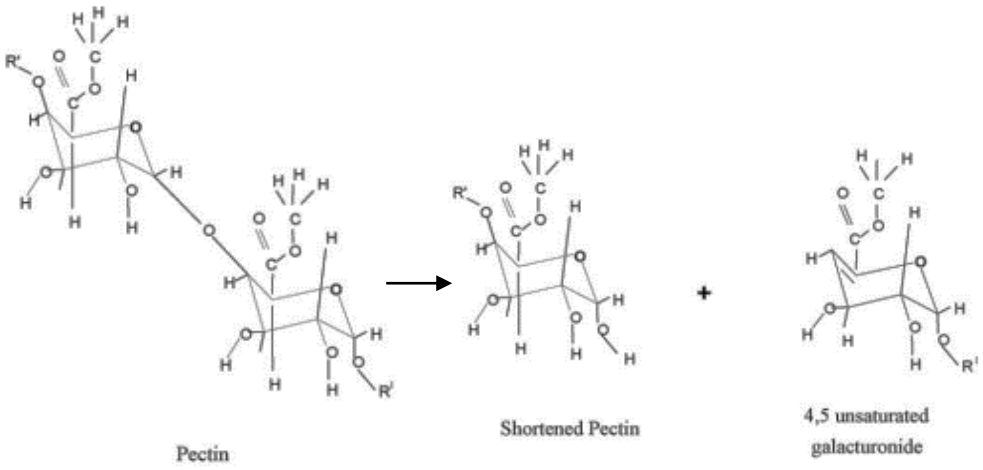
CHAPTER 6: GENERATION OF MUTATIONS IN TOMATO *PECTATE LYASE* USING THE CRISPR-CAS9 GENE EDITOR

6.1 Introduction

Pectate lyases (*PL*: EC 4.2.2.2), also known as pectate transeliminases, are pectin degrading enzymes which, like polygalacturonase (PG) described in Chapter 5, play a role in cell wall disassembly in plants. However, unlike PG these enzymes catalyze the cleavage of unesterified α-(1,4)-galacturosyl linkages by the mechanism of β-elimination (Figure 6-1A) rather than hydrolysis giving the 4,5-unsaturated oligogalacturonates (Yoder et al., 1993).

The activity of pectate lyases (PLs) was first investigated in detail in bacteria. *Erwinia carotovora* and *Bacillus* sp. were the first bacterial cultures shows to have *PL* activity. Three-dimensional structures of bacterial *PL* have been determined including PelC (Yoder et al., 1993 and Yoder and Jurnak, 1995) and PelE (Lietzke et al., 1994) from *Erwinia chrysanthemi* and BsPel from *Bacillus subtilis* (Pickersgill et al,., 1994) which has the unusual structure 'parallel β helix' where β-strands are folded into a large, superhelix (Figure 6.1B). The active sites of *PL* enzymes are highly conserved (Yadav et al., 2009).

(A)



(B)

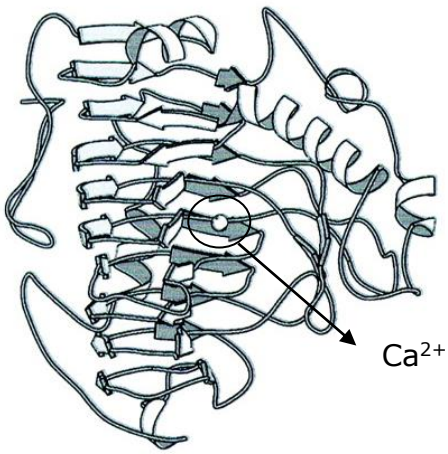


Figure 6-1 (A) Pectate lyase action that produces an unsaturated product (B) The 3D structure of *Bacillus subtilis* pectate lyase with calcium shown as a sphere. (Taken from Yadav et al., 2009 and Marín-Rodríguez et al., 2002).

The role of *PL* genes in plants is still not well understood. Sequences with homology to *PLs* in plants were first discovered in pollen, anthers and mature tomato flowers (Wing et al., 1989). *PL* activity was first measured in a recombinant enzyme derived from *Zinnia elegans* and this work indicated that this gene played an important role in the remodelling cell walls (Domingo et al., 1998). *PL* activity was also reported in banana (Pua et al., 2001; Marín-Rodríguez et al., 2002) and strawberry fruit (Medina-Escobar et al., 1997; Benítez-Burraco et al., 2003) during ripening.

A pioneering study on strawberry produced the first direct demonstration of the importance of PL in softening in that fruit (Marín-Rodríguez et al., 2002). The authors were able to demonstrate that firmness of strawberry fruit is increased when a ripening-expressed *PL* gene is suppressed by antisense RNA. More recently in two varieties of apple with different firmness characteristics, the softer variety was found to have high level of *PL* enzyme activity (Harb et al., 2012). In tomato, there are numbers of EST that have homolog to pectate lyases indicated from breaker and ripening fruit in cDNA libraries (Tomato Gene Index: www.tigr.org) (Marín-Rodríguez et al., 2002). There are five *PL* genes that have been identified as expressed in fruits (Tomato Genome Consortium, 2012). The most highly expressed gene in fruits is Solyc03g111690 (Figure 6-2).

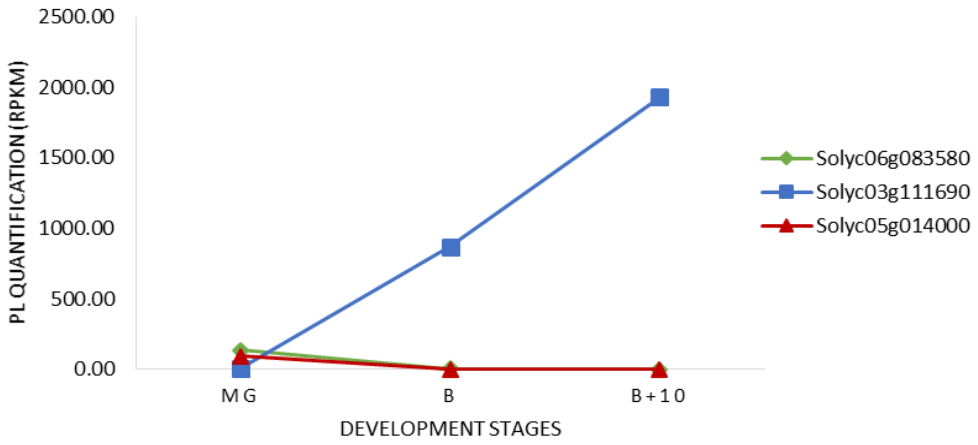


Figure 6-2 Pectate lyase gene expression patterns based on Reads Per Kilobase of transcript per Million mapped reads (RPKM) values in tomato during fruit development and ripening (Tomato Genome Consortium, 2012).

During the course of this project transgenic experiments, in the Seymour lab using an RNAi approach indicated that when this *PL* is silenced it leads to significantly firmer fruits (Uluisik et al., 2016). The aim of this study was to investigate the function of *PL* in tomato by using transgene constructs that result in a gene knockout using CRISPR-Cas9 system; in contrast to RNAi where only a degree of gene silencing is possible.

6.2 Material and Methods

6.2.1 Generation of mutatuions in tomato PL using CRISPR-Cas9

The sequences for pectate lyase (*PL*) Solyc03g111690, was obtained from the tomato genome assembly SL V2.40 at [http://solgenomics.net/]. The following protocols as described in Chapter 2.8.1 were used for designing the primers to generate the guide RNA (Table 6-1) (Figure 6-3). The protocols for vector construction followed those described in Chapter 4. The Solyc03g111690 gene fragment was amplified using gradient PCR with 2X Q5 High-Fidelity PCR Master Mix (NEB, UK) following the same protocol as described in Chapter 2.9.2.

6.2.2 Generation of stable transgenic plants

The methodology for plant transformation was described in detail in Chapter 2, Section 2.11.

Table 6-1. Primer sequence of gene for *pectate lyase*

Primer ID	Primer sequence 5’ to 3’
PL01.F	TGTGGTCTCAATT**GGGCGCTAGCGTACACATAGCG**TTTTAGAG CTAGAAATAGCAAG
PL01.R	TGTGGTCTCAAGCGTAATGCCAACTTTGTAC

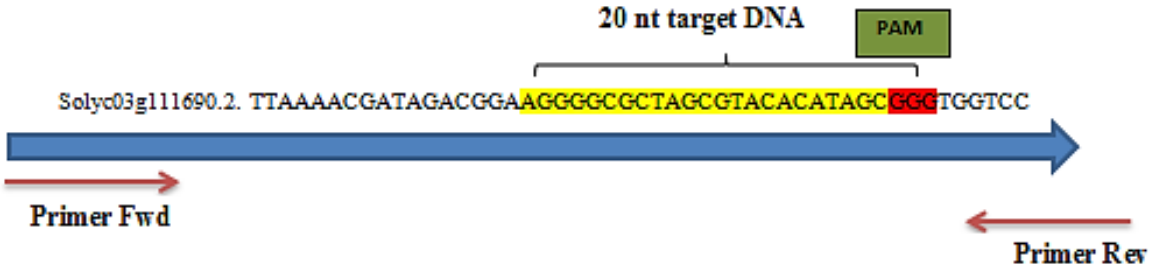


Figure 6-3 Cas9/sgRNA-mediated mutagenesis in *S. lycopersicum*. Scheme for Cas9/sgRNA-mediated mutagenesis of a non-functional (out-of-frame) mutant of Solyc03g111690 coding sequence.

6.2.3 Identification of the *PL* transgenic lines

A set of PCR reactions was undertaken to determine the presence of T-DNA insertion in the transgenic plants using the *PL* gene specific primer and NPTII/Cas9 primer (Appendix 9.8). The DNA from the wild-type Ailsa Craig was used as positive control for gene specific primer and as negative for T-DNA insertion. The presence of transgene was also confirmed by sequencing the PCR amplicons (Appendix 9.9).

6.2.4 Identification of single copy homozygous lines

The T_0 plants were grown to fruiting and T_1 seeds were collected. Then 12 T_1 seeds were sown from T_0 line for PL-05 and PL-11. The T_1 progeny were grown and the DNA from each of the plants was extracted followed the protocol in Chapter 2, Section 2.2. The DNA samples then were then used to check for the presence of the transgene by PCR and the purified PCR products were then sent for sequencing.

6.2.5 QPCR determination of *PL* gene expression

The RNA was extracted from *PL* fruits at the following stages of development, MG, B, B+4 and B+7. DNA contamination was removed before generating the cDNA template (Chapter 2.4). Quantification of *PL* gene expression was undertaken using SYBR® Green quantitative RT-

PCR. The serial dilutions, qPCR reactions and qPCR conditions followed the 2X SYBR® Green PCR master mix protocol (also see Chapter 2.8). An elongation factor was used as a housekeeping gene. The relative transcript abundance of the *PL* genes and elongation factor gene were calculated from the equation (2.1) in Chapter 2.8 (German et al., 2003).

6.2.6 Determination of *PL* enzyme activity

We followed the method described in Uluisik et al (2016). Acetone insoluble solids (AIS) were prepared for the measurement of PL enzyme activity. The skin was removed from about 10-20 gram of fresh breaker +7 days tomato pericarp tissue. The pericarp was then homogenised with 80% cold acetone and washed with 100% acetone until it turned to a pure white colour. The AIS powder was left overnight dry on the bench at room temperature. Then 80 mg of AIS was mixed in 10 mL of 50 mM Tris-HCL, pH 8 and stirred for 2 hours at room temperature. The samples were then centrifuged for 30 minutes at 9,300 x G at 25^{o}C. The absorbance of clear supernatant was measured at 235 nm. AIS prepared with a hot ethanol treatment to kill endogenous enzymes was used as control in this experiment.

6.2.7 Analysis of fruit characteristics

Analysis of fruit texture, colour and brix followed the same protocols as described in Chapter 4.27.

6.2.8 Determination of viscosity behaviour in *PL* tomato purees

The same protocol that was described in Chapter 5, Section 5.2.9 was used to determine the viscosity behaviour of the *PL* tomato puree.

6.3 Results and Discussion

6.3.1 Vector construction *PL* transgenic lines

Figure 6-4 shows the 100 bp of gene specific fragment that was amplified following the standard protocol in Chapter 2, Section 2.9. The sgRNA PL fragment was cloned into the Level 2 Golden Gate vector (pAGM4723). This was verified by PCR amplification of 100 bp sgRNA fragment and 600 bp kanamycin fragment (Figure 6-5). The destination vector was then transformed into *Agrobacterium* EHA105 by electroporation as described in Chapter 2.9. PCR and sequencing were used to confirm the presence and orientation of the insert (Figure 6-6).

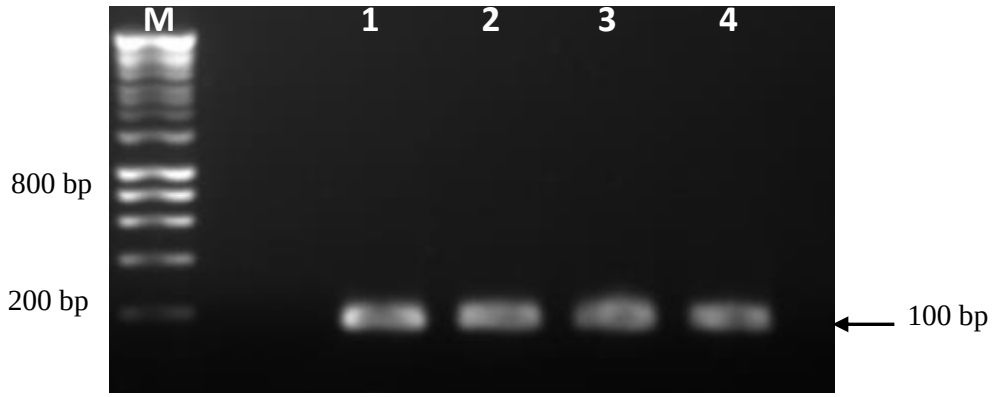


Figure 6-4 The gene specific *PL* PCR. The purified PCR products were separated by electrophoresis on a 2% (w/v) agarose gel in 0.5X TAE at 100 V for 30 minutes. A) Lane M: Hyperladder Marker. Lane 2-5 PCR product of sgRNA pICH86966::AtU6p::sgRNA_PDS construct that contained the *PL* target sequence.

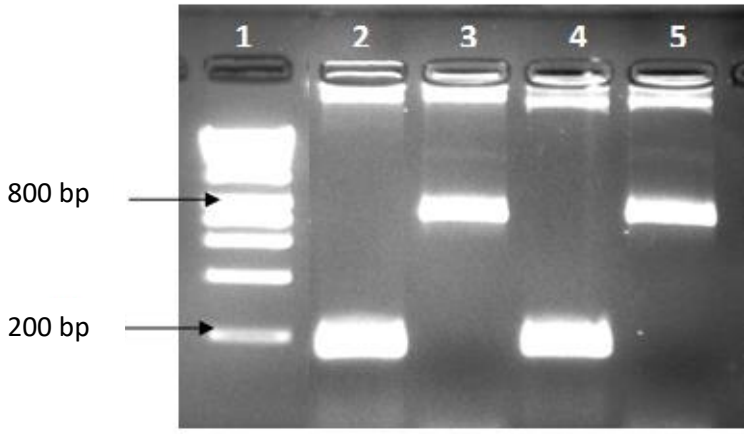


Figure 6-5 Cloning of *PL* gene fragment into the Level 2 vector. The PCR products of Level 2 construct were amplified by two different primers using the entry clone as a template. Lane 1: Hyperladder marker, lanes 2 and 4: fragments amplified by gene specific primers (PL01) and lane 3 and 5 fragments amplified using the NPTII/Cas9 primers.

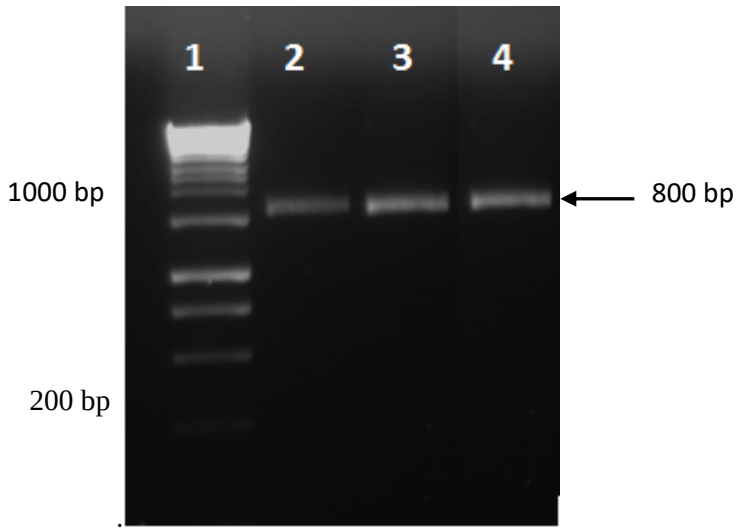


Figure 6-6 Transformation of CRISPR/Cas9 construct into *Agrobacterium.* PCR products were amplified using the NPTII/Cas9 specific primers from *Agrobacterium* EHA105. Lane 1: Hyperladder Marker, lanes 2,3 and 4 PCR products of sgRNA constructs which were isolated from colonies of *Agrobacterium* EHA105.

6.3.2 Analysis of the *PL* transgenic lines

There were thirteen transgenic Ailsa Craig T_0 lines generated which contained the sgRNA::CRISPR construct designed to mutate a ripening related *PL* (Solyc3g111690) gene. PCR amplification confirmed the presence of the transgene using NPTII/Cas9 primer (Appendix 9-8) and yielded an amplicon of the correct size of 800 bp (Figure 6-7) in transformed plant lines No. 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12 and 13. There was no amplification product generated in line 3 and the WT control.

Validation of CRISPR / Cas9 driven mutations in the *PL* genes

A combined PCR / restriction digestion strategy was used first to test for the presence of mutations in the target gene in the T_0 lines. A PCR approach was designed to give single amplicon of 300 bp covering the region of the *PL* gene containing the guide RNA sequence and containing a *BmtI* site. The fragmentation pattern from the restriction digest can be seen in Figure 6-8. Out of thirteen T_0 lines, all of samples showed similar pattern to WT with complete digestion of the PCR product which gave 2 x 150 bp fragments. This indicated that none of the CRISPR::PL lines were resistant to *BmtI* activity. To confirm the presence of Cas9/sgRNA-directed mutations in *PL* gene all of 300 bp PCR products were then purified and sequenced.

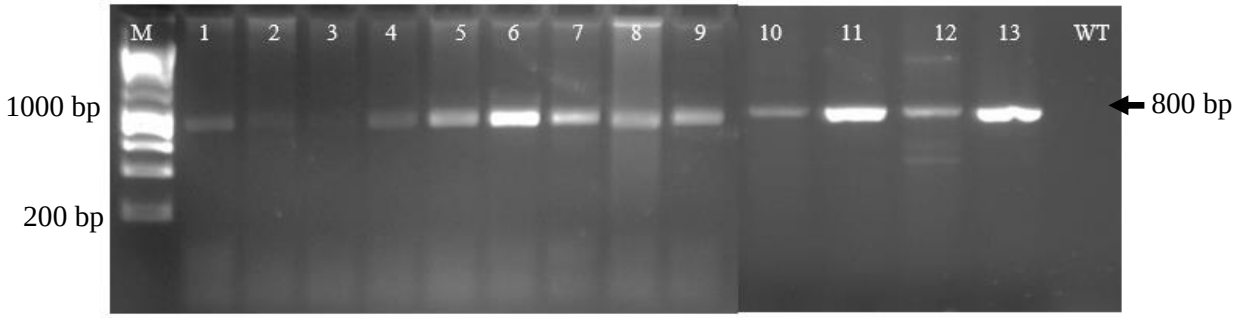
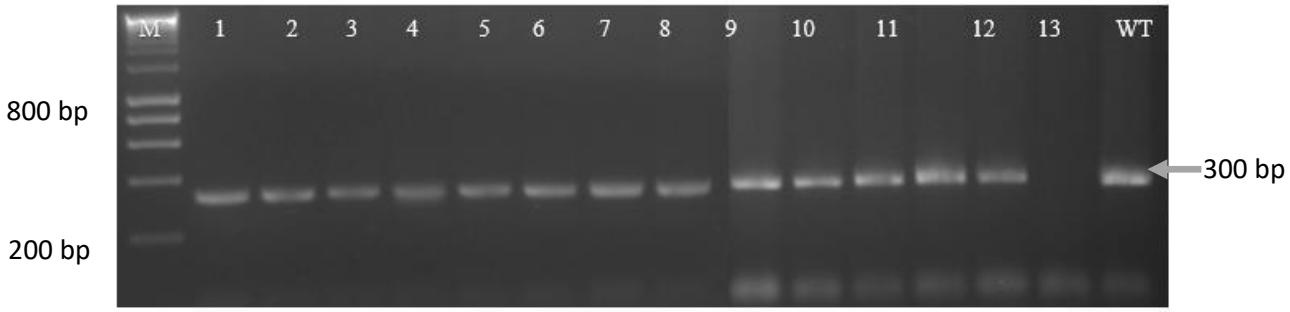


Figure 6-7 Amplification of transgenic lines. PCR genotyping of thirteen representative CRISPR::PL plants showing the transgenic lines that contain Cas9/sgRNA gene construct in T_0 plants. The PCR products were then cloned and sequenced to validate the insertion in plant and characterize additional DNA lesions.

A)



B)

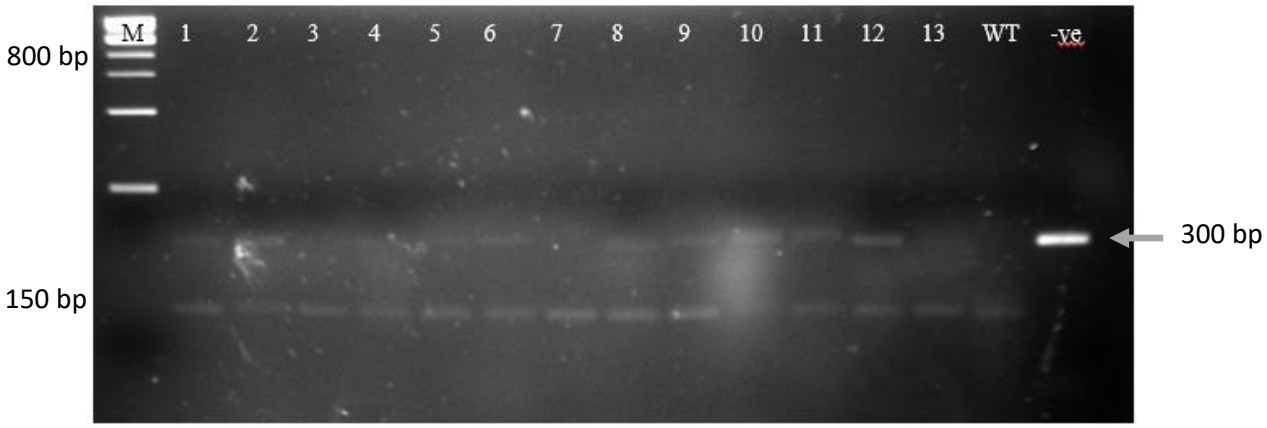


Figure 6-8 Efficiency of Cas9/sgRNA mutagenesis. (A) PCR amplicons generated from T_0 lines. Amplicons contained sequence of BmtI restriction enzyme in initial PL gene were amplified using PL01 gene specific primer. (B) PCR products from (A) were digested with BmtI. Wild type sequences will yield 2X ~150 bp fragments from cleavage of the ~300bp PCR product. The PCR from Level 2 plasmid was used as negative control. Lane M: Hyperladder Marker, lanes 1-13 PCR products of CRISPR::PL.

Sequence alignment and inspection of the DNA sequencing traces of the PCR amplified region of the *PL* gene targeted for CRISPR indicated various mutations at the target site including insertion of an A nucleotide (line PL-13), a T nucleotide in line (PL-05 and PL-11) and deletion of a G nucleotide in line PL-02 (Figure 6-9). The sequence of the remaining lines (PL-01, PL-07, PL-03, PL-06, PL-04, PL-09 and PL-10) was similar to wild type. Figure 6-9 also indicated that sequences alignment clearly show the mutation was at the end of 20 bp of sgRNA which is near to the PAM site. This was the reason why all of the CRISPR::PL lines still can be digested by *BmtI* enzyme as the *BmtI* site was located at the beginning of the 20 bp sequence.

The selected CRISPR::PL-05 and CRISPR::PL-11 to take through to the T_1 generation and obtain homozygous lines. CRISPR::PL-13 was not selected because this line did not produce fruit with viable seeds. In the initial screening, it was assumed that PL02 was same as the wild type. However, after further analysis it became apparent that there was a deletion of a G nucleotide at the target site. However due to time constraint no further analysis was undertaken.

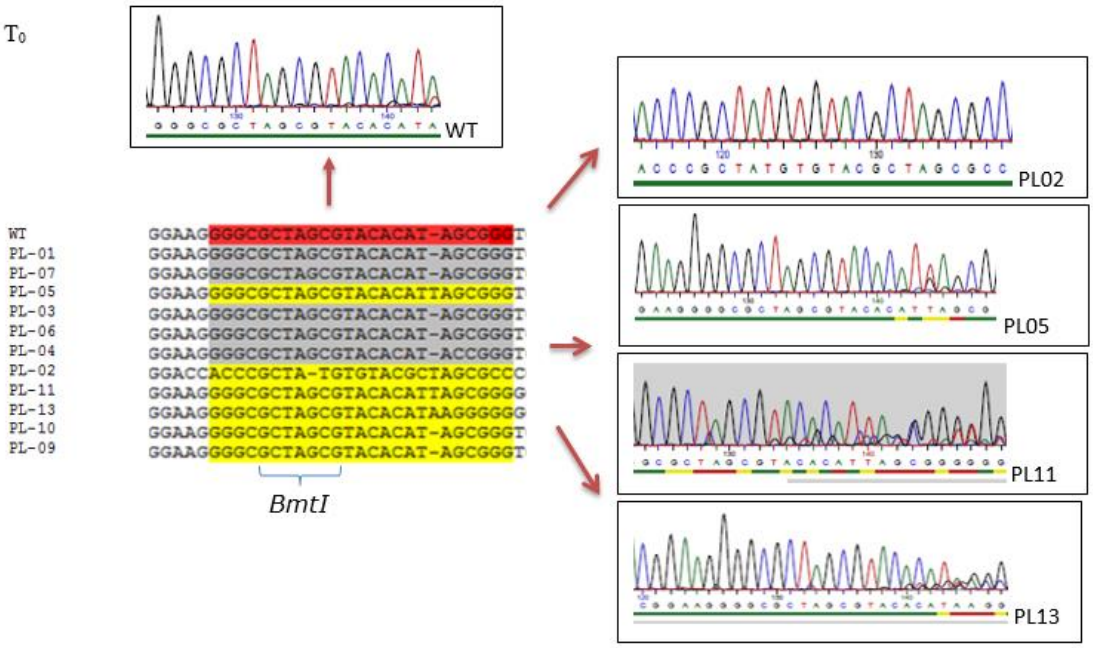


Figure 6-9 Confirmation of inheritance of a modified or non-modified *PL* gene. Sequence alignment and DNA sequencing traces from PCR purification of CRISPR::PL as target sites which there has a Cas9/sgRNA-mediated gene modification including of insertion of an A nucleotide (line PL13) and a T nucleotide in (line PL05 and PL11) and deletion a G nucleotide in line PL02.

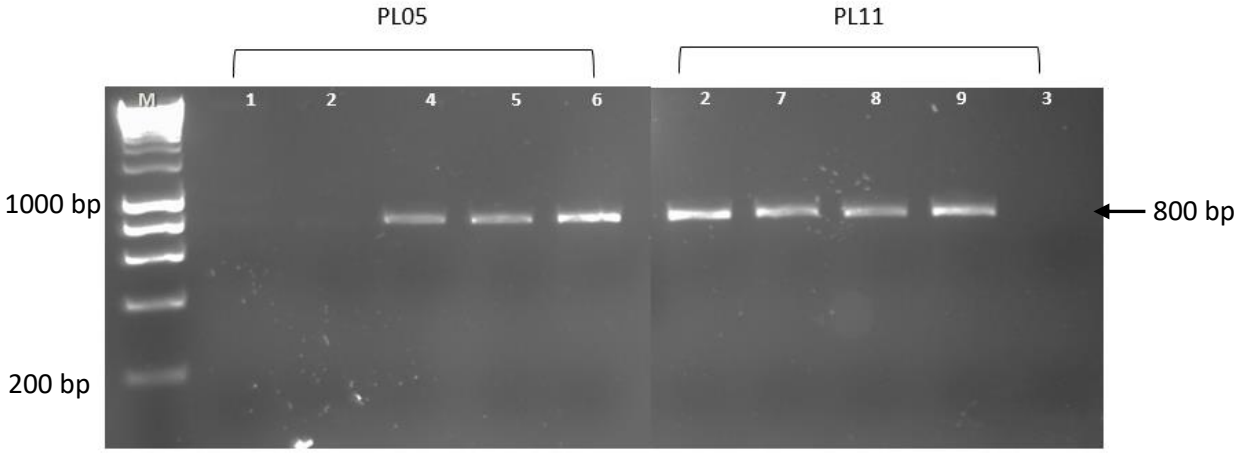


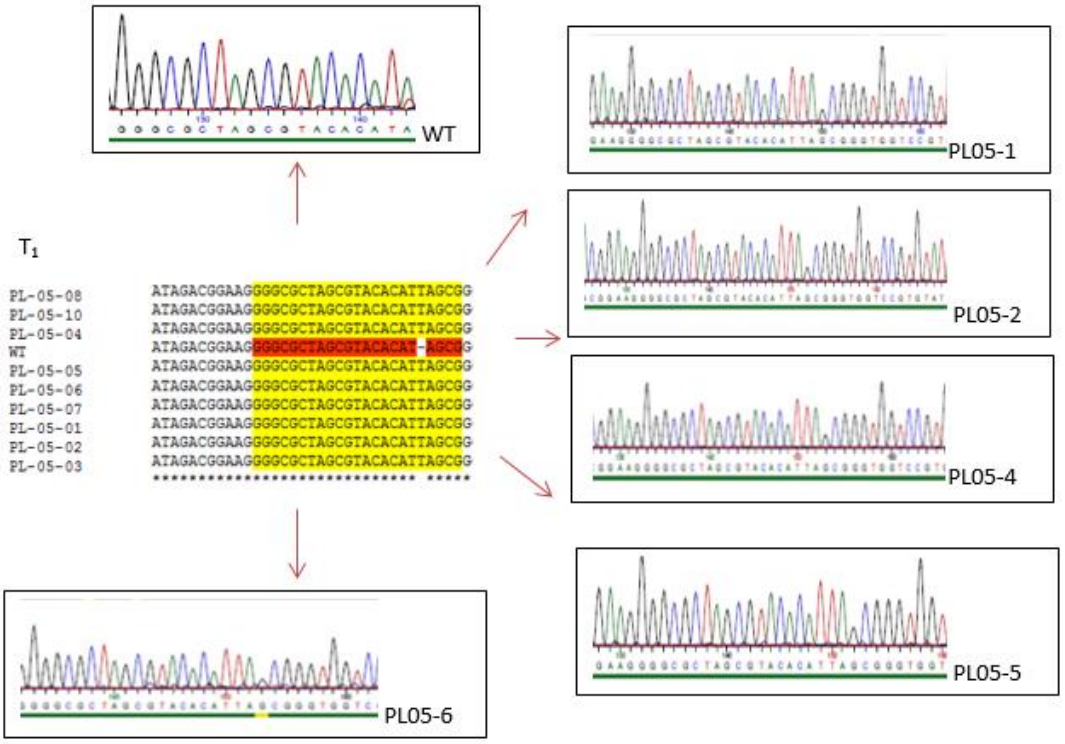
Figure 6-10 Confirmation of inheritance of modified PL gene in each of T1 generation plants. Out of the five progeny from plant line PL05, two plants lacked the CRISPR/Cas9 transgene. Lane M: Hyperladder Marker, Lanes 1, 2, 4 5 and 6 was PCR products of CRISPR::PL05, Lanes 2, 7, 8, 9 and 3 was PCR products of CRISPR::PL11.

6.3.3 Identification of Homozygous lines

Twelve T_1 seeds were sown from the PL05 and PL11. As described in Chapter 2, Section 2.2, DNA was extracted from the leaves of transgenic plant using DNeasy Plant Mini Kit (Qiagen). Two sets of primers NPTII/Cas9 and PL01 gene specific primers, (Appendix 9-8) were used to identify the presence of Cas9 and also to amplify the guide RNA targeted region to validate the mutation at the target area. For identification of the presence or absence of the Cas9 gene a PCR product was expected of around 800 bp (Figure 6-10). Five T_1 progeny from line PL05 were positive for Cas9 and two (PL-05-01 and PL-05-02) were negative. For line PL11 all T_1 lines except line PL-11-03 inherited the Cas9 gene (Figure 6-10).

Inspection of the sequence data and traces from the T_1 lines indicated that 100% of the CRISPR::PL05 progeny contained an additional T nucleotide at the target site (Figure 6-11). For progeny of the CRISPR::PL11 T_0 line both wild type, heterozygous (PL-11-01, PL-11-04 and PL-11-10) and homozygous lines were identified (Figure 6-12). The base changes in the CRISPR::PL lines result in stop codons in the coding sequence of *PL* (Figure 6-11B and 6-12B) and therefore we predict that the mutant lines will have no *PL* enzyme activity resulting from the expression of Solyc03g111690.

(A)



(B)

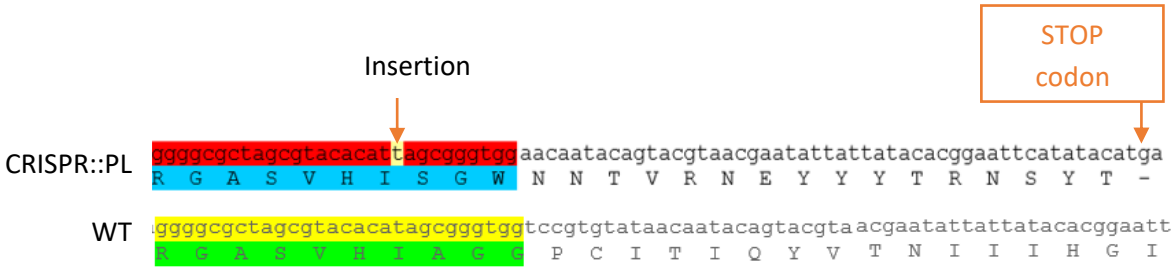
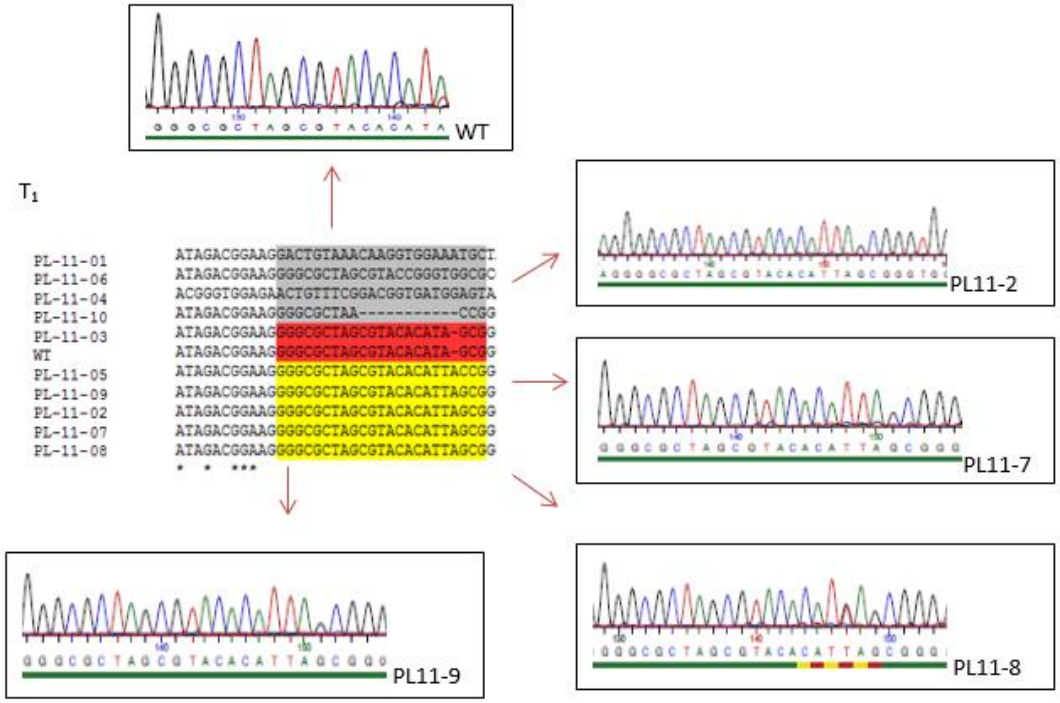


Figure 6-11 Inheritance of modified *PL* gene in each of T_1 progeny**.** (A) DNA sequence alignment and sequencing traces of CRISPR::PL-05 with insertion A base near the PAM site although some of them still inherited CRISPR/Cas9 from T_0. (B) Effect on protein sequence of transgenic lines i.e PL-05-01

(A)



(B)

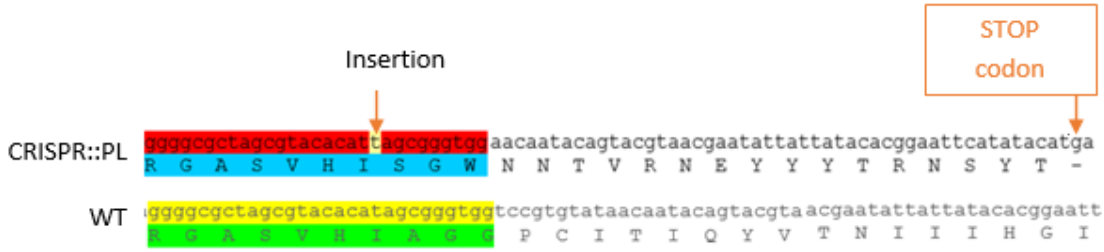


Figure 6-12 Inheritance of modified *PL* gene in each of T_1 progeny. (A) DNA sequence alignment and sequencing traces of CRISPR::PL11 with varied of mutation rate at the target regions although some of them still inherited CRISPR/Cas9 from T_0. Effect on protein sequence of transgenic lines i.e PL-11-02

6.3.4 Gene expression of *PL*

The target *PL* expression in ripening fruits from a selected T_1 homozygous CRISPR::PL05 line was investigated by QPCR. Levels of *PL* expression in the PL05 were substantially reduced in comparison to wild type Ailsa Craig (Figure 6-13). Elevated levels of *PL* expression were similar in wild type to those reported by the Tomato Genome Consortium (2012). The levels of PL gene expression were significantly different at B+4 ($P=0.005$) and the B+7 stage ($P<0.001$) compared to control fruits (Figure 6-13). This is consistent with RT-PCR from other *PL* transgenic lines which gave a similar result (Appendix 9.10). The CRISPR mutations generate a stop codon and delivered a non-functional PL protein.

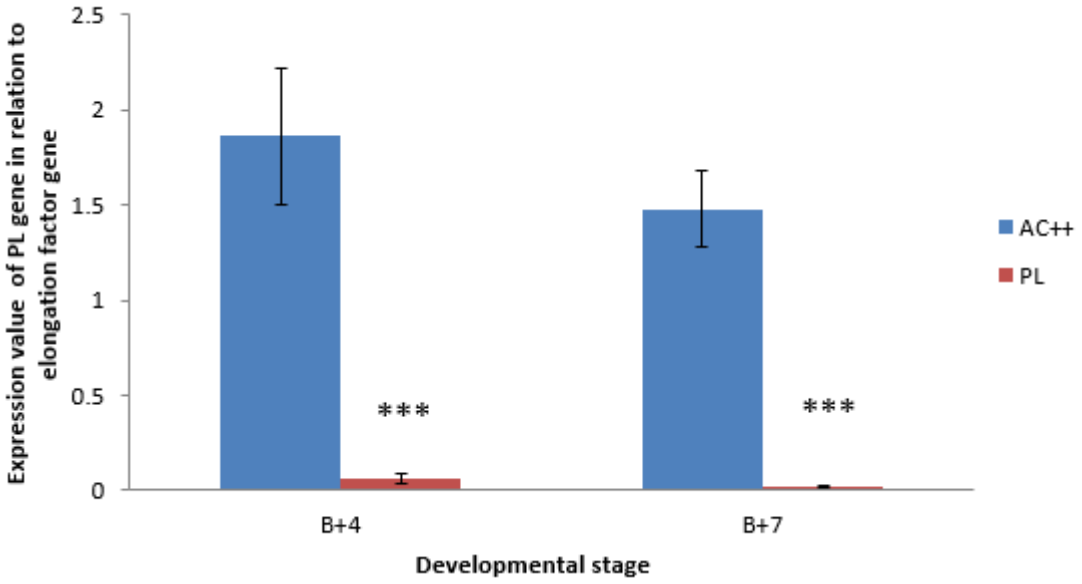


Figure 6-13 Comparison of *PL* gene expression during ripening in wild-type Alisa Craig and CRISPR::PL05 at B+4 (breaker +4) and B+7 (breaker+7). The standard error of the mean (SEM) represented by vertical error bars with n=3. Significant differences to wild-type are indicated by *** (P<0.001) from T test analysis.

6.3.5 *PL* enzyme activity

To verify that that the CRISPR mutation resulted in reduced *PL* enzyme activity we made a crude cell wall protein extract by extracting cell wall proteins from fresh tissue using 1M sodium chloride (Brooks et al., 1990; Collmer et al., 1988; Tucker et al., 1980). However, no *PL* activity could be detected in these extracts in either WT or *PL* fruits.

In related work in the Seymour lab, Selman Uluisik was able to recover *PL* activity in tomato acetone powders and this protocol was used to investigate *PL* activity in CRISPR::PL lines. In this experiment, *PL* enzyme assays were performed using acetone insoluble solids (AIS) from tomato pericarp tissue and determining the level of reaction products by incubating these in an alkaline buffer (Tris-HCL). Unsaturated galactosyluronic acid products were measured spectrophotometrically at 235 nm. This is the same approach that was reported by Uluisik et al., (2016). The *PL* enzyme activity for CRISPR::*PL* line PL05 was measured and was significantly lower ($P<0.01$) compare to control fruits (Figure 6-14). The limited residual *PL* enzyme activity may arise from *PL* enzymes generated through transcription from other *PL* ripening related genes which are expressed at a very low level, e.g. Solyc02g093580 (Uluisik et al., 2016)

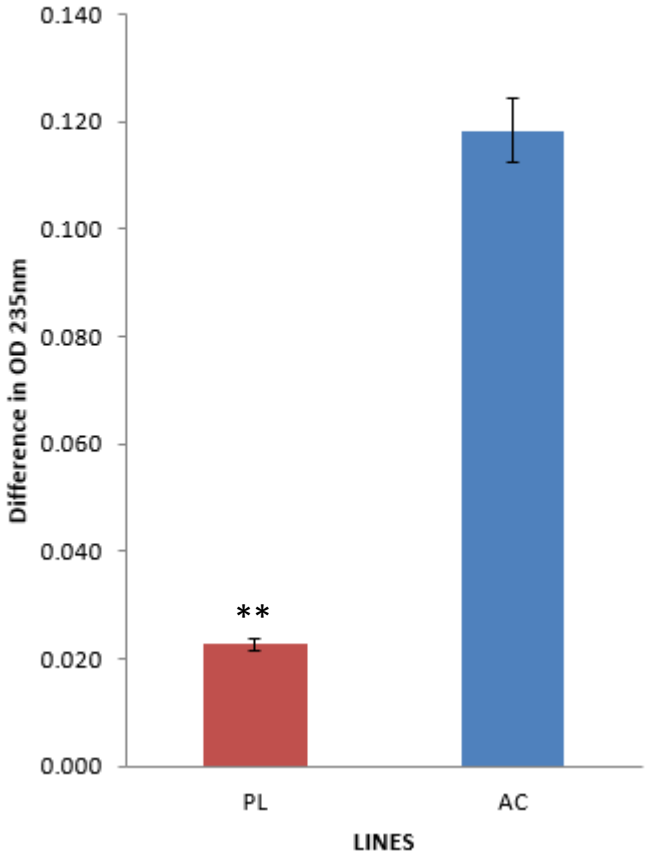


Figure 6-14 *PL* enzyme activity using acetone powder that prepared from red ripe fruits pericarp of Ailsa Craig and CRISPR::PL05 line. 0.08 g of AIS in 50 mM Tris-HCL, pH8 buffer measure after 120 minutes of stirring at room temperature. The standard error of mean (SEM) is represented by vertical error bars, n=4. Value significantly different to the control line is indicated by ** ($P<0.01$) using T test analysis.

6.3.6 Phenotypic Analysis

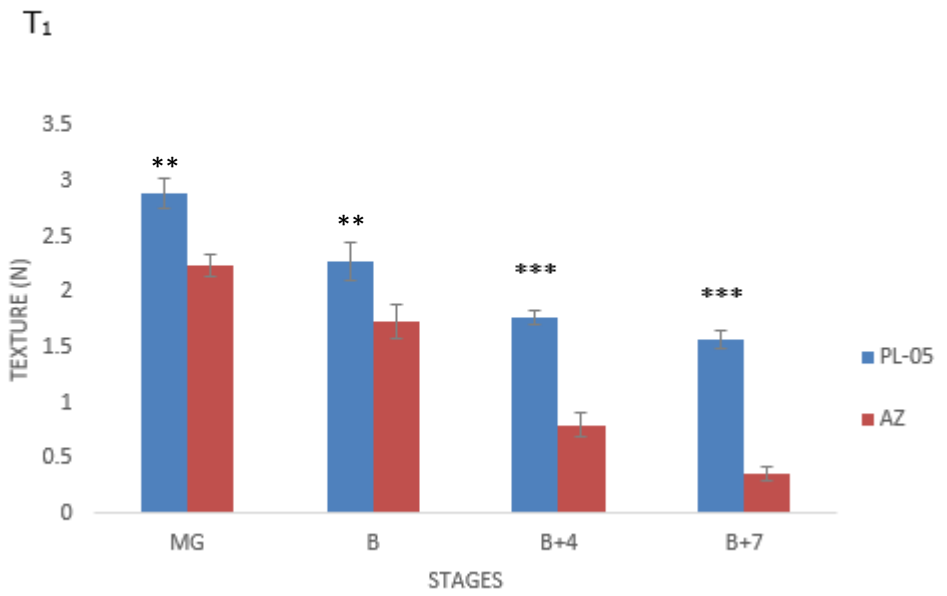
Fruit phenotypes from the CRISPR::PL lines were examined over two seasons. In the first season (winter 2015) fruits were collected at MG, B, B+4 and B+7 stages. For the second season (spring 2016) fruits were collected at B+4 and B+7 stages only. These fruits were used in fruit texture, colour and total soluble solids analysis.

The texture of transgenic fruits was measured on two independent homozygous lines (PL-05 and PL-11) from T_1 (Figure 6-15) and T_2 (Figure 6-16) generations. All transgenic plants showed the normal growth habit and produced ripening fruits that were visually indistinguishable from the controls. The results for a representative line (CRISPR::PL05) are described and data for CRISPR::PL11 are shown in the Appendix 9.13 and in Supplementary Fig. 11 (Uluisik et al., 2016). In all experiments, the firmness of the CRISPR::PL lines in both of outer and inner fruit pericarp was significantly enhanced ($P<0.005$) in comparison to the wild type especially at the orange ripe (breaker+4) and red ripe (breaker+7) stages (Figures 6-15 and 6-16).

The *PL* data for the CRISPR lines (Supplementary Fig. 11) are entirely consistent with the data from silencing the *PL* gene using RNAi (Uluisik et al., 2016). This is the first time a major effect on fruit softening has been reported by silencing a cell wall remodelling enzyme. The ability of *PL* to modulate softening has been reported in other fruits such as

strawberry (Medina-Escobar et al., 1997; Benítez-Burraco et al., 2003; Shulaev et al., 2011) and apple (Atkinson et al., 2012). However in these fruits, *PG* was reported to be much more important in cell wall disassembly (Quesada et al., 2009). There were no significant ($P>0.05$) differences between azygous and CRISPR lines with respect to pericarp colour or brix (Figure 6-17).

(A)



(B)

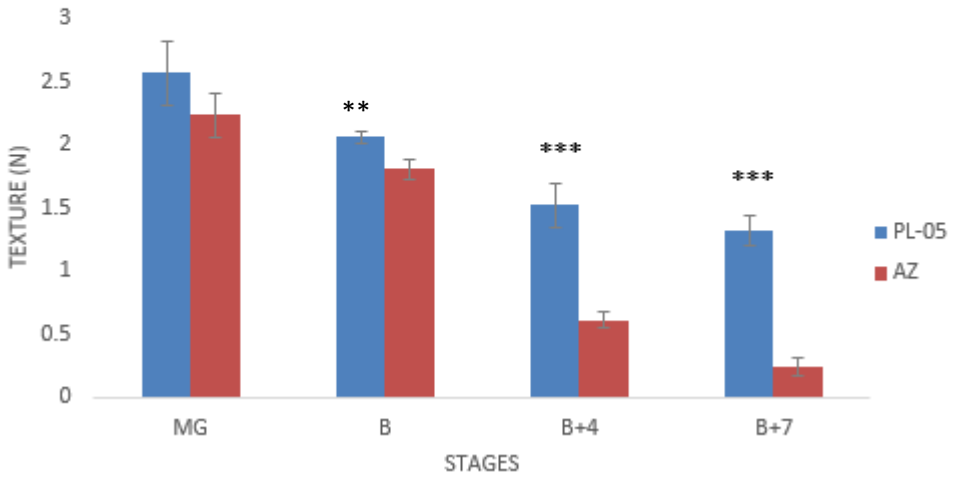
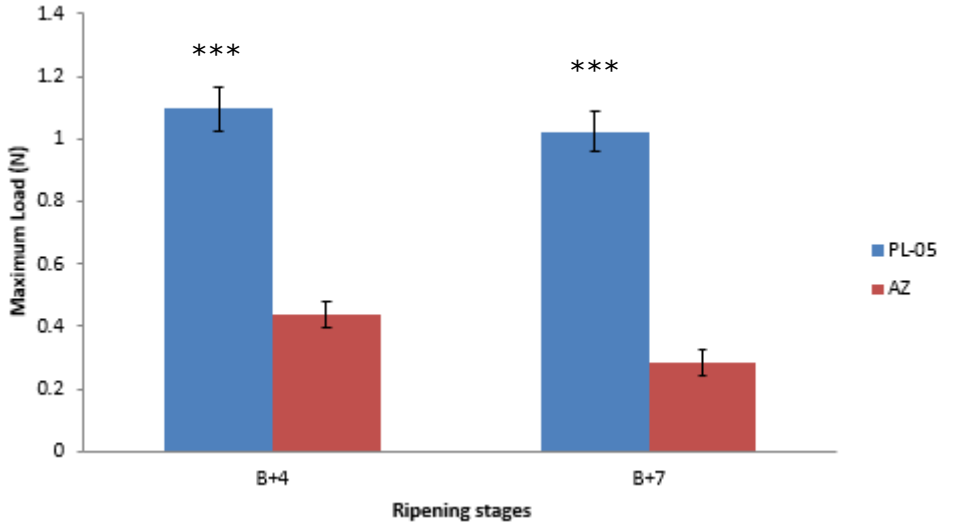


Figure 6-15 Pericarp firmness as determine by measurement of maximum load in CRISPR::PL05 T_1 fruits. (A) Outer pericarp and (B) Inner pericarp maximum load at stages MG (Mature green), B (Breaker), B+4 (Orange ripe) and B+7 (Red ripe) in T_1 generation. The standard error of mean (SEM) is represented by vertical error bars, n=13. Significant differences to control line are indicated by ** ($P<0.01$) and *** ($P<0.005$) from T test analysis.

(A)

T_2



(B)

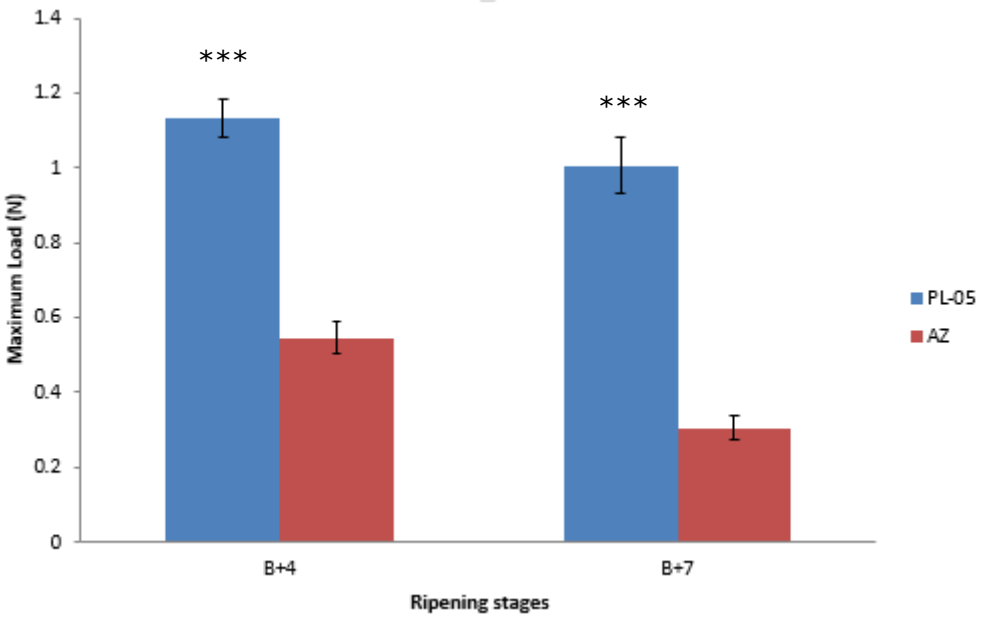
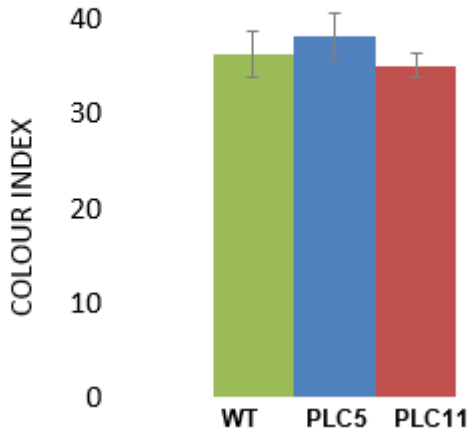


Figure 6-16 Pericarp firmness as determine by measurement of maximum load in CRISPR::PL05 T_2. (A) Outer pericarp and (B) Inner pericarp maximum load at stage B+4 (Orange ripe) and B+7 (Red ripe) in both of CRISPR::PL lines in T_2 generation. The standard error of mean (SEM) is represented by vertical error bars, n=4. Value significantly different to the control line are indicated by *** ($P<0.005$) using T test analysis

(A)



(B)

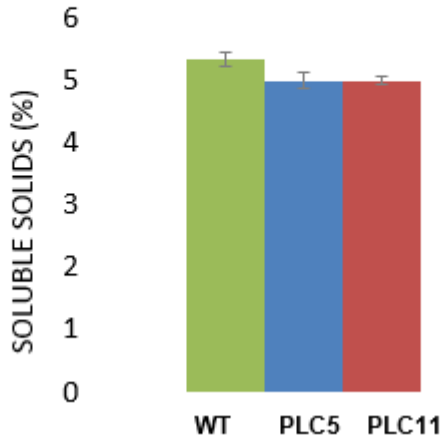


Figure 6-17 Fruit characteristics with respect to (A) Colour index values from azygous and transgenic fruits containing CRISPR::PL transgene and (B) brix measurement in red ripe fruits. The standard error of mean (SEM) is represented by vertical error bars, n=3. No significant difference to the control line with ($P>0.05$) using T test analysis.

6.3.7 Viscosity behaviour in *PL* tomato purees

Analysis of tomato paste of the selected lines was performed using rotational rheology, as described in Chapter 5, Section 5.2.10. The results of Control Share Rate (CSR) test are shown in Figure 6-18 and Table 6-2. This analysis indicated that CRISPR::PL puree shows thinning behaviour, with a higher η than control samples at all share rate. The η of the puree samples at low (1 s^{-1}) and high (100 s^{-1}) shear rate is shown in Table 6-2. The η is significantly influenced by plant line at both 1 s^{-1} (P=0.03) and 100 s^{-1} (P=0.01), with CRISPR::PL having a significantly higher η than azygous at 1 s^{-1} and 100 s^{-1} (P<0.05). This indicated that PL puree has stronger gel like properties than azygous puree. Sesmero et al. (2009) also indicated that strawberry juices of anti-pectate lyase fruits are more viscous and the content of large particles enhanced then the wild-type.

Table 6-2 Shear viscosity values at low (1 s^{-1}) and high (100 s^{-1}) shear rates for the CSR test of the tomato puree of CRISPR::PL and azygous line in Figure 6-19. Means are displayed with the error value as SD. Value significantly different to the control line are indicated by * (P<0.05). Significant values are displayed in the final row of the relevant column of each table.

Line	η at 1 s^{-1}	η at 100 s^{-1}
Azygous	3.26 (±0.43)	0.14 (±0.006)
CRISPR::PL	5.00 (±0.26)*	0.18 (±0.004)*
	P=0.03	P=0.01

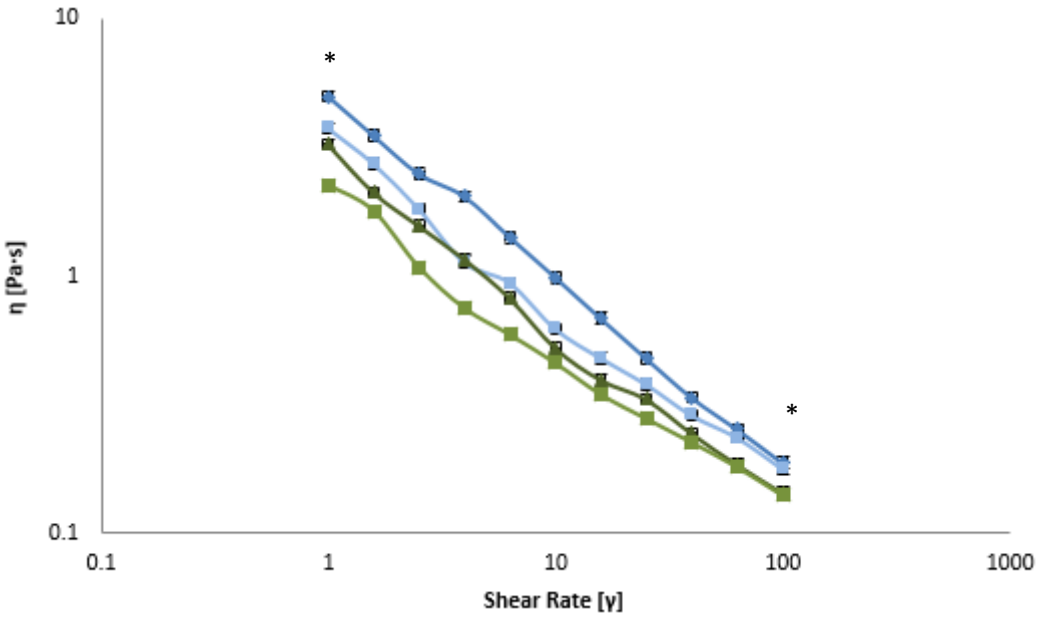


Figure 6-18 The CSR test data from CRISPR::PL and azygous line puree. Symbol represent ◆ CRISPR::PL and ▲ azygous line. ——◆—— and ——▲—— are the viscosity value at 1-100 s^{-1} share rate. ···■··· and -■- are the viscosity value at 100-1 s^{-1} share rate. The standard error of mean (SEM) is represented by vertical error bars and n=4 (biological replicates). Values significantly different to the control line are indicated by * ($P<0.05$).

6.4 Conclusion

The data from analysis of CRISPR::PL transgenic plants demonstrated that CRISPR/Cas9 can induce mutations in *PL* gene. The analysis also indicated that CRISPR::PL fruit had reduced *PL* expression, enzyme activity and substantially increased fruit firmness compared to the azygous control line and subsequently published these results along with the RNAi data in *Nature Biotechnology* (Uluisik et al, 2016). The transgenic CRISPR::*PL* plants show a large effect on the fruit softening during ripening but in terms of normal ripening phenotypes, the fruits still look similar to the controls.

Chapter 7

General Discussion and Future Work

CHAPTER 7: GENERAL DISCUSSION AND FUTURE WORK

7.1 General Discussion

Fruit ripening is a complex developmental process unique to plants which has evolved to aid seed dispersal and involves multiple biochemical changes that can include the conversion of starch to sugar, accumulation of acids, pigments and volatile compounds as well as cell wall disassembly leading to texture changes (Klee and Giovannoni, 2011).

The biochemical and molecular events leading to the softening of fleshy fruits has been the subject of intensive research effort over the last 50 years. It is well established that the cell walls are remodelled during the ripening process and especially the pectic polysaccharides that can make up more than more than 40% of cell wall of fleshy fruits (Brummell 2006; Cantu et al., 2008). The aim of the project was to identify the common and unique subsets of genes encoding cell wall enzymes across fruit bearing species using new information from genome sequencing studies. Then target selected genes for mutation by CRISPR-Cas9 in transgenic tomatoes. The information from our study could then provide a new resource in breeding programmes to slow the rate of fruit softening and improve the postharvest life consistent with our modern supply chain.

Gene hunting for CRISPR-Cas9 targets by comparative genomics

Fleshy fruit genomes that were readily accessible for analysis by comparative genomics approaches at the start of this project included melon (Garcia-Mas et al., 2012), grape (Wang et al., 2014), banana (D'Hont et al., 2012) and tomato (Tomato Genome Consortium, 2012). Several other fruit genome sequences had been published, but their annotation was of low quality or fruit expression data was limited.

Comparative genomics approaches have provided connections between function and evolution in complex biological structures and systems (Caicedo et al., 2005). Previous work on plant genome sequencing and annotation of *Oryza sativa* and *Arabidopsis thaliana* also adopted this approach to identify genomic regions of functional importance (Garcia-Hernandez et al., 2002, Ware et al. 2002, Haas et al., 2003). In the current study, the comparison between our fleshy fruit species indicated that only small numbers of genes encoding fruit-related cell wall degrading enzymes were identified as orthologous in all genomes. These were a *CELLULOSE SYNTHASE* (*CesA-like*) and *GLUCAN ENDO-BETA-GLUCOSIDASE*. However, all the fruit species studied showed an expression of a range of genes family members related to degradation of the main fruit cell wall polysaccharides and especially the pectin components. The role of GLUCAN ENDO-BETA-GLUCOSIDASES is unknown in relation to cell wall disassembly so *CesA-like* gene was selected as our first candidate.

Polysaccharide-degrading enzymes that contribute to cell wall disassembly in tomato fruit have been thoroughly studied both at the genetic and biochemical level (Brummell and Harpster 2001; Giovannoni 2001, Brummell 2006). The principal focus has been on genes that are involved in degrading pectic polysaccharides such as *POLYGALACTURONASE* (*PG*), *PECTIN METHYLESTERASE* (*PME*), *PECTATE LYASE* (*PL*) and *BETA-GALACTOSIDASE* (*β-GALase*). For our gene editing study we selected *PG* and *PL* as targets for silencing in tomato. *PG* (Solyc10g080210) was selected because there are orthologs in melon and grape that are expressed during ripening (Table 3-3) and early experiments only ever reduced the expression of the gene in tomato rather than delivering a full knockout (Sheehy et al., 1988; Smith et al., 1988).

These early transgenic lines used an RNAi approach that suppressed the expression of *PG* to 0.5-1% in comparison to wild-type, but the silencing of *PG* had no detectable effect on fruit firmness (Smith et al., 1990). Interestingly, the viscosity of juice and paste prepared from antisense *PG* lines significantly increased due to low levels of soluble pectin during processing (Schuch et al., 1991; Kramer et al., 1992; Brummell and Labavitch 1997; Kalamaki et al., 2003).

In the case of *PL*, three genes were expressed during fruit ripening in tomato (Uluisik et al., 2016). Two of the *PLs* (Solyc05g014000 and Solyc06g083580) showed orthologous relationships with other fruits like banana, melon, grape, apple and strawberry. However, (Solyc03g111690)

was identified as uniquely fruit expressed in tomato. There were also strong indications of the importance of *PL* in softening in fruits other than tomato at the onset of the study, e.g. in banana (Pua et al., 2001; Marín-Rodríguez et al., 2003), strawberry (Medina-Escobar et al., 1997; Benítez-Burraco et al., 2003; Shulaev et al., 2011) and apple (Velasco et al., 2010. For this reasons, *PL* (Solyc03g111690) was selected as our third candidate gene.

Our last target for CRISPR/Cas9 knocked-out was PHYTOENE SYNTHASE (*PSY1*) which produces yellow fruits when silenced. We selected this gene because the absent of this gene can immediately recognize as *yellow flesh* phenotype in tomato fruits.

Efficiency of generating CRISPR/Cas9 -based mutations in tomato fruit

The development of methods for genome editing has progressed rapidly in the last two decades. Several methods have been developed including the use of zinc finger nucleases (Kim et al., 1996) and transcription activator-like effector nucleases (TALENs) (Christian et al., 2010). More recently the clustered regularly interspaced short palindromic repeat (CRISPR) CRISPR-associated 9 endonuclease has been used (Jinek et al., 2012). These modifications will induce the double-strand breaks (DSB) at the specific target regions and then the mutation takes place during genome repair via non-homologous end joining (NHEJ). According to Chen and Gao (2014), the most efficient and high specificity mutation can be

achieved with the CRISPR/Cas9 system involving an RNA guided DNA nucleases; compared to other systems that used protein-DNA interaction.

The CRISPR/Cas9 system has now been used in a variety of plant species to direct mutations in target genes. These include *Arabidopsis thaliana* (Li et al., 2013), *Nicotiana benthaminan* (Nekrasov et al., 2013), *Oryza sativa* (Zhang et al., 2014), *Sorghum bicolor* (Jiang et al., 2013), *Triticum aestivum* (Wang et al., 2013). Very recently the system has also been tested in tomato. The first report for the use of CRISPR/Cas9 in tomato was by Ron et al., (2014) and Brooks et al. (2014). In 2015, Ito et al. also reported the efficiency of CRISPR/cas9 method to silence the *RIN* gene where the mutations that contain insertion or deletion which result in inhibition of fruit ripening.

The CRISPR::*PSY1* lines demonstrated effective silencing by CRISPR/Cas9. There is a natural mutant of *PSY1* in tomato known as *yellow-flesh,* which was initially described around 100 years ago. The mutant can be recognized by its pale yellow fruit (Fleming et al., 1937; Jenkins, 1948). This mutation that contributed to carotenoid biosynthesis in plants was linked to a single locus *r* that was mapped to chromosome 3 (http://solgenomics.net/locus/1599/view) (Liu et al., 2003). The CRISPR::*PSY1* lines produced fruit with a pale yellow appearance. This fruit had a phenotype was identical that of fruits where *PSY1* was silenced by antisense RNA as reported by Fray et al. (1993) which contained reduced levels of *PSY1* compare to WT tomato fruit. Moreover, some

CRISPR::*PSY1* fruits showed the phenotype in the T_0 generation and the mutant alleles were transmitted stably in T_1 progeny.

In this project four sets of transgenic plants were generated, but only three showed successful induction of mutations by CRISPR/Cas9 system in the target genes. There were no mutations recovered in the *CesA-like* gene and this indicates that this gene product is likely essential for regeneration of plantlets from tissue culture. RNAi technology is therefore an effective option where the removal of a gene product has lethal consequences.

What do the CRISPR / Cas9 Modification in Tomato Pectin Enzymes tell us about softening in tomato?

In higher plants, pectins play a major role in cell wall architecture and are also important in cell-to-cell adhesion in plant tissues (Willats et al., 2001). The first pectin degrading enzyme whose role in ripening was investigated by gene silencing was *PG*. Smith et al. (1988) demonstrated how antisense RNAi could be used to inhibit gene *PG* expression in tomato fruits. Although pectin depolymerisation was significantly reduced, there was no significant effect on pectin solubilisation or fruit softening between the wild type and antisense *PG* fruits (Smith et al., 1990). These data indicated that *PG* plays a main role in polyuronide depolymerisation but, other factors must be involved in pectin solubilisation from the cell wall and softhening (Smith et al., 1990). However the elevated molecular

weights of the pectins from the *PG* silenced fruit (Smith et al, 1990) resulted in more viscous paste after processing. The antisense experiments described by Smith et al (1988) reduced the levels of *PG* activity to 1%. In this study the CRISPR mutants resulted in a mutation in the *PG2a* gene (Solyc10g080210) and therefore a non-functional enzyme. However, the *PG*::CRISPR mutants still showed a significant degree of softening.

During this project, work in the Seymour lab (Uluisik et al, 2016), demonstrated that silencing the expression of a ripening-related *PL* gene (Solyc03g111690) by RNAi had a major effect on tomato fruit softening. The fruit had a substantially firmer texture at the red ripe stage. However other aspects of ripening such as changes in colour and taste metabolites proceeded as normal. By generating *PL* knockout lines, I was able to show with CRISPR / Cas 9 identical effects on fruit texture and ripening to those observed using RNAi lines and these observations are reported in Uluisik et al (2016).

The mechanism of action of *PL* in tomato fruits involves the degradation of pectic polysaccharides in the middle lamella and tricellular junctions (Figure 7-2). Uluisik et al. (2016) found that silencing *PL* inhibited both pectin solubilisation and degradation and *PL* is likely responsible for releasing high molecular weight pectins from the cell wall that are then degraded by *PG*. Pectin solubilisation is commonly associated with ripening in tomato (Seymour et al. 1987). Uluisik et al. (2016) investigated the differences in WT and PL::RNAi cell walls using

light and transmission electron microscopy. The results were consistent with *PL* being involved in 'dissolving' pectin between adjacent cells at the cell junctions. This is likely to leads to cell separation and softening (Willats et al., 2001). The enhancement of viscosity in CRISPR::*PL* pastes is entirely consistent with the presence of pectin polymers of high molecular size and wider size distribution in serum. Similar effects have been reported in juices from strawberry in *PL* antisense plants (Sesmero et al., 2009).

In our CRISPR::*PL* mutant fruits the maximum load of the pericarp tissue decreases from around 2.5 N in unripe fruits to around 1.4 N in red ripe fruits. In normal wild type fruits the maximum load at red ripe can be below 0.2 N. Thus although the effects of silence *PL* are impressive they account for perhaps 50% of the texture changes that are observed at breaker +7. A major question remains as to the identity of other cell wall enzymes that account for the remaining 50% of the softening effects.

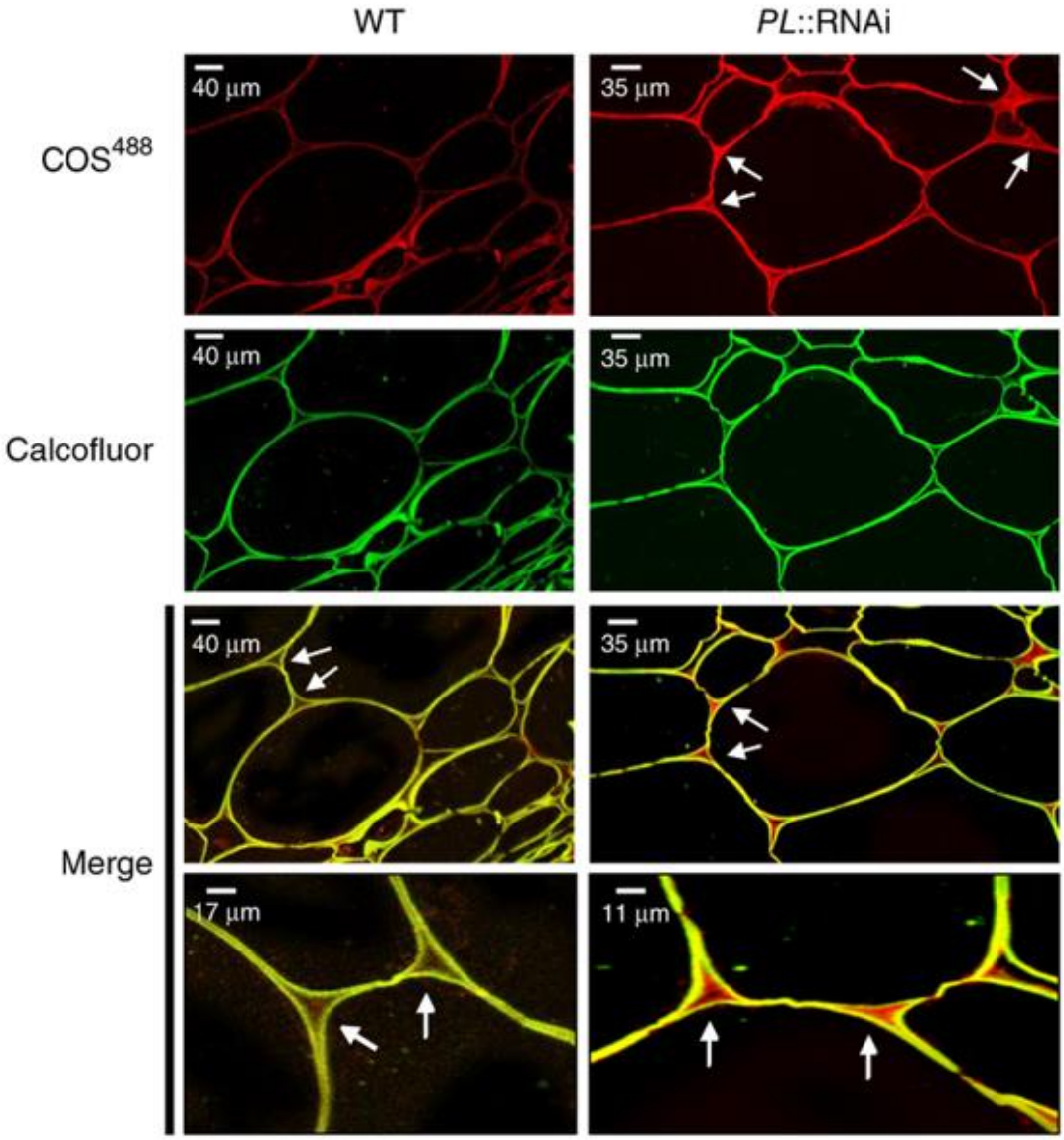


Figure 7.2 Pericarp cell walls from Azygous (5-6) and PL:RNAi (5-5) lines at Br+7 stage, labelled stained with calcoflour or with COS^{488}. This probe recognises the deesterified homogalacturonan domain of pectic polysaccharides (taken from Uluisik et al., 2016).

Apart from *PG*, other transgenic studies on cell wall enzymes have focused on pectin esterase (PME). Hall et al (1993) examined the effects of suppressing (Solyc02g077840.1), but silencing these genes had no measurable effects on fruit softening. In a later study Phan et al (2007) silenced the *Pmeu1* gene which induced more rapid softening in the transgenic lines. Chapman et al (2012) also reported enhanced firmness associated with low PME activity. These data indicate that the relationship between PME and fruit firmness in tomato is complex. Reduced PME activity may result in less sites for the development of calcium bridges between adjacent pectic polysaccharides in the cell wall (Chapman et al, 2012). Silencing of β-galactosidase (*TBG4*) (Solyc02g084720) gene and expansin (*EXP1*) (Solyc06g051800) show an effect on fruit texture in tomato although the enhancement was small (Smith et al., 2002; Brummell et al., 1999). In both cases the effects accounted for less than a 10% change in fruit firmness.

The tomato genome project (The Tomato Genome Consortium, 2012) investigated the cell wall structure-related genes expressed during ripening and a simplified Supplementary Table from the paper is reproduced in Table 7.1 showing genes that are up-regulated during the ripening process. These data clearly show that there are 12 cell wall enzyme encoding genes showing substantial changes during ripening and many of them have yet to be investigated for their effects on fruit softening. This is especially true for the xyloglucan endotransglycosylases and xyloglucan hydrolases.

Table 7-1 List of genes that shows large changes in expression during the ripening process in tomato (The Tomato Genome Consortium, 2012).

Gene	Expression			Description	Short name
	MG	Breaker	Breaker+10		
Solyc03g093130	243.72	34.90	156.91	Xyloglucan endotransglucosylase/hydrolase 9	SlXTH3a
Solyc03g093080	272.88	37.94	146.02	Xyloglucan endotransglucosylase/hydrolase 9	SlXTH3d
Solyc03g093120	218.86	29.60	141.28	Xyloglucan endotransglucosylase/hydrolase 9	SlXTH3c
Solyc01g081060	66.13	384.24	743.34	Xyloglucan endotransglucosylase/hydrolase 14	SlXTH5
Solyc12g017240	158.84	98.15	258.69	Xyloglucan endotransglucosylase/hydrolase 7	SlXTH11
Solyc10g080210	3.92	1130.56	4907.76	Polygalacturonase A	PG-2a
Solyc07g064170	4702.95	4391.96	1339.68	Pectinesterase	PME1.9
Solyc07g064180	1229.70	1091.79	434.82	Pectinesterase	PME2.1
Solyc03g111690	11.67	871.60	1930.12	Pectate lyase	PL
Solyc01g008710	1.55	804.53	119.45	Mannan endo-1 4-beta-mannosidase	LeMAN4a
Solyc02g092790	548.34	171.30	397.20	Arabinogalactan	AGP-1c
Solyc08g061100	428.25	128.37	103.96	Cellulose synthase	CESA

7.2 Conclusions and Future Work

The major output from this study was to demonstrate the usefulness of implementing CRISPR/ Cas9 technology in genome editing to generate cell wall mutant in tomato. Part of the study on the CRISPR / Cas9 *PL* mutant has now been published (Uluisik et al, 2016).

Genome editing has recently been the subject of an investigation by the US government to ensure that the food that derived from GM crop does not pose greater risk. Furthermore, recently studies have indicated that CRISPR techniques only used *Agrobacterium* as a medium to introduce the gene for CRISPR-Cas9 where the cell itself were then used its own machinery to form a complex with guide RNA and Cas9 protein. Based on GMO definition, this activity was considered legal which fall into a regulatory grey area and can be used in commercial practice (Parisi et al., 2016).

Our novel knockout mutations in *PG* and *PL* not only add to our understanding of cell wall disassembly in ripening fruits, but could also be utilised as a new resource in breeding programmes to improve fruit quality. This is because they provide a rapid and targeted approach to alter fruit texture and or paste properties without affecting other ripening processes. This is a substantial advance on current technologies where slow ripening mutants are used to enhance shelf-life but adversely impact flavour and colour (Uluisik et al, 2016).

This work has provided the foundation for a more thorough study of the CRISPR lines; which was not possible due to time constraints in the present project. These future investigations could include analysis of the effect of silencing PL and PG on fruit metabolism and ripening related gene expression i.e carotenoids and flavonoid analysis and determine if the results match those reported for the RNAi lines. Furthermore, it will be important to better understand the mechanism of PL and PG action on the tomato cell walls by determination of molecular weight of soluble pectin and the changes in cell wall structure by chemical analysis and light microscopy.

The effect of the CRISPR mutations on the processing properties of the fruits would also be an interesting avenue to explore further. This includes the influence of processing on the texture or rheology behaviour of products such as in tomato puree and paste. Furthermore, changes in the nutritional content in CRISPR tomato processed product should be analysed. Processing exposes foods to high levels of heat, light or oxygen and can cause nutrient loss. It will be interesting to see if the cell wall changes can help reduce any of these negative effects.

The work on PG and PL CRISPR lines is an important breakthrough that should help tomato breeders generate great tasting tomatoes that also have an extended shelf life to improve fruit quality and quantity.

8 REFERENCES

Abersheim, P. 1976. The primary cell wall. In: BONNER, J. (ed.) *Plant Biochemistry*. New York: Academic Press.

Agrotis, A. and Ketteler, R., 2015. A new age in functional genomics using CRISPR/Cas9 in arrayed library screening. *Frontiers in genetics*, 6.

Al-Dous, E.K., George, B., Al-Mahmoud, M.E., Al-Jaber, M.Y., Wang, H., Salameh, Y.M., Al-Azwani, E.K., Chaluvadi, S., Pontaroli, A.C., DeBarry, J. and Arondel, V., 2011. De novo genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). *Nature Biotechnology, 29*, pp.521-527.

Alexander, L. And Grierson, D. 2002. Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening. *Journal of Experimental Botany, 53*, 2039-2055.

Alexeyenko, A., Tamas, I., Liu, G. and Sonnhammer, E.L., 2006. Automatic clustering of orthologs and inparalogs shared by multiple proteomes. *Bioinformatics, 22*, pp.e9-e15.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J., 1990. Basic local alignment search tool. *Journal of Molecular Biology, 215*, pp.403-410.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research, 25*, 3389-3402.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S. and Eppig, J.T. (2000) Gene Ontology: Tool For The Unification Of Biology. The Gene Ontology Consortium. *Nature Genetics, 25*, 25–29

Atkinson, R.G., Gunaseelan, K., Wang, M.Y., Luo, L., Wang, T., Norling, C.L., Johnston, S.L., Maddumage, R., Schröder, R. and Schaffer, R.J., 2011. Dissecting the role of climacteric ethylene in kiwifruit (*Actinidia chinensis*) ripening using a 1-aminocyclopropane-1-carboxylic acid oxidase knockdown line. *Journal of Experimental Botany, 62*, pp.3821-3835.

Atkinson, R.G., Sutherland, P.W., Johnston, S.L., Gunaseelan, K., Hallett, I.C., Mitra, D., Brummell, D.A., Schröder, R., Johnston, J.W. and Schaffer, R.J., 2012. Down-regulation of POLYGALACTURONASE1 alters firmness, tensile strength and water loss in apple (*Malus x domestica*) fruit. *BMC plant biology, 12*, p.1.

Ayub, R., Gris, M., Amor, M.B., Gillot, L., Roustan, J.P., Latché, A., Bouzayen, M. and Pech, J.C., 1996. Expression of ACC oxidase antisense gene inhibits ripening of cantaloupe melon fruits. *Nature Biotechnology, 14*, pp.862-866.

Baker, K.E. and Parker, R., 2004. Nonsense-mediated mRNA decay: terminating erroneous gene expression. *Current Opinion in Cell Biology, 16*, pp.293-299.

Ballester, A.R., Molthoff, J., de Vos, R., te Lintel Hekkert, B., Orzaez, D., Fernández-Moreno, J.P., Tripodi, P., Grandillo, S., Martin, C., Heldens, J. and Ykema, M., 2010. Biochemical and molecular analysis of pink tomatoes: deregulated expression of the gene encoding transcription factor SlMYB12 leads to pink tomato fruit color. *Plant Physiology, 152*, pp.71-84.

Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., Albert, V.A., Aono, N., Aoyama, T., Ambrose, B.A., Ashton, N.W. and Axtell, M.J., 2011. The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science, 332*, pp.960-963.

Barnes, H.A., Hutton, J.F. and Walters, K., 1989. An introduction to rheology. Elsevier.

Barnes, M.F. and PATCHETT, B.J., 1976. Cell wall degrading enzymes and the softening of senescent strawberry fruit. *Journal of Food Science, 41*, pp.1392-1395.

Barry, C.S., Llop-Tous, M.I. and Grierson, D., 2000. The regulation of 1-aminocyclopropane-1-carboxylic acid synthase gene expression during the transition from system-1 to system-2 ethylene synthesis in tomato. *Plant Physiology, 123*, pp.979-986.

Barry, C. S. And Giovannoni, J. J. 2007. Ethylene and Fruit Ripening. *Journal of Plant Growth Regulation, 26*, 143-159.

Beerli, R.R. and Barbas, C.F., 2002. Engineering polydactyl zinc-finger transcription factors. *Nature Biotechnology, 20*, pp.135-141.

Bemer, M., Karlova, R., Ballester, A.R., Tikunov, Y.M., Bovy, A.G., Wolters-Arts, M., de Barros Rossetto, P., Angenent, G.C. and de Maagd, R.A., 2012. The tomato FRUITFULL homologs TDR4/FUL1 and MBP7/FUL2 regulate ethylene-independent aspects of fruit ripening. *The Plant Cell, 24*, pp.4437-4451.

Benítez-Burraco, A., Blanco-Portales, R., Redondo-Nevado, J., Bellido, M.L., Moyano, E., Caballero, J.L. and Muñoz-Blanco, J., 2003. Cloning and characterization of two ripening-related strawberry (Fragaria× ananassa cv. Chandler) pectate lyase genes. *Journal of Experimental Botany, 54*, pp.633-645.

Berglund, A.C., Sjölund, E., Östlund, G. & Sonnhammer, E. L. L. 2008. InParanoid 6: eukaryotic ortholog clusters with inparalogs. *Nucleic Acids Research, 36,* 263-266.

Bergougnoux, V., 2014. The history of tomato: from domestication to biopharming. *Biotechnology Advances, 32*, pp.170-189.

Bortesi, L. and Fischer, R., 2015. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances, 33*, pp.41-52.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.*Analytical Biochemistry, 72*, pp.248-254.

Bramley, P. M. 2002. Regulation of carotenoid formation during tomato fruit ripening and development. *Journal of Experimental Botany, 53*, 2107-2113.

Brooks, A.D., He, S.Y., Gold, S., Keen, N.T., Collmer, A. and Hutcheson, S.W., 1990. Molecular cloning of the structural gene for exopolygalacturonate lyase from Erwinia chrysanthemi EC16 and characterization of the enzyme product. *Journal of Bacteriology, 172*, pp.6950-6958.

Brooks, C., Nekrasov, V., Lippman, Z.B. and Van Eck, J., 2014. Efficient gene editing in tomato in the first generation using the clustered regularly interspaced short palindromic repeats/CRISPR-associated9 system. *Plant Physiology, 166*, pp.1292-1297.

Brownleader, M. D., Jackson, P., Mobasheri, A., Pantelides, A. T., Sumar, S., Trevan, M. & Dey, P. M. 1999. Molecular Aspects of Cell Wall Modifications during Fruit Ripening*. Critical Reviews in Food Science and Nutrition, 39*, 149-164.

Brummell, D. & Harpster, M. 2001. Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. *Plant Molecular Biology, 47*, 311-339.

Brummell, D. A. & Labavitch, J. M. 1997. Effect of Antisense Suppression of Endopolygalacturonase Activity on Polyuronide Molecular Weight in Ripening Tomato Fruit and in Fruit Homogenates. *Plant Physiology, 115*, 717-725.

Brummell, D.A., 2006. Cell wall disassembly in ripening fruit. *Functional Plant Biology, 33*, pp.103-119.

Brummell, D. A., Catala, C., Lashbrook, C. C. & Bennett, A. B. 1997. A membrane-anchored E-type endo-1,4-β-glucanase is localized on

Golgi and plasma membranes of higher plants. *Proceedings of the National Academy of Sciences of the United States of America, 94*, 4794-4799.

Brummell, D.A., Hall, B.D. and Bennett, A.B., 1999. Antisense suppression of tomato endo-1, 4-β-glucanase Cel2 mRNA accumulation increases the force required to break fruit abscission zones but does not affect fruit softening. *Plant Molecular Biology, 40*, pp.615-622.

Brummell, D. A., Lashbrook, C. C. & Bennett, A. B. 1994. Plant endo-1,4-β-glucanases: Structure, properties and physiological function. In: HIMMEL, M. E., BAKER, J. O. & OVEREND, R. P. (eds.) *Enzymatic Conversion of Biomass for Fuels Production*. Washington, DC: American Chemical Society.

Brummell, D.A. and Labavitch, J.M., 1997. Effect of antisense suppression of endopolygalacturonase activity on polyuronide molecular weight in ripening tomato fruit and in fruit homogenates. *Plant Physiology, 115*, pp.717-725.

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

Caicedo, A.L. and Purugganan, M.D., 2005. Comparative plant genomics. Frontiers and prospects. *Plant Physiology, 138*, pp.545-547.

Campos-Vargas, R. and Orellana, A., 2012. Assessment of Prunus persica fruit softening using a proteomics approach. *Journal of Proteomics, 75*, 1618-1638.

Cantu, D., Vicente, A.R., Greve, L.C., Dewey, F.M., Bennett, A.B., Labavitch, J.M. and Powell, A.L.T., 2008. The intersection between cell wall disassembly, ripening, and fruit susceptibility to Botrytis

cinerea.*Proceedings of the National Academy of Sciences, 105*, pp.859-864.

Carey, A. T., Holt, K., Picard, S., Wilde, R., Tucker, G. A., Bird, C. R., Schuch, W. & Seymour, G. B. 1995. Tomato exo-(1-->4)-beta-D-galactanase. Isolation, changes during ripening in normal and mutant tomato fruit, and characterization of a related cDNA clone. *Plant Physiology, 108*, 1099-1107.

Carpita, N. C. And Gibeaut, D. M. 1993. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *The Plant Journal, 3*, 1-30.

Carpita N., Mccann, M. 2000. The Cell wall In: BUCHANAN, B. B., GRUISSEM, W., JONES, R.L., (ed.) *Biochemistry and Molecular Biology of Plants*. Rockville MD: American Society of Plant Physiologists.

Carrari, F. and Fernie, A.R., 2006. Metabolic regulation underlying tomato fruit development. *Journal of Experimental Botany, 57*, pp.1883-1897.

Carrington, C.S. and Pressey, R., 1996. ß-Galactosidase II activity in relation to changes in cell wall galactosyl composition during tomato ripening*. Journal of the American Society for Horticultural Science, 121*, pp.132-136

Chagné, D., Crowhurst, R.N., Pindo, M., Thrimawithana, A., Deng, C., Ireland, H., Fiers, M., Dzierzon, H., Cestaro, A., Fontana, P. and Bianco, L., 2014. The draft genome sequence of European pear (*Pyrus communis L.'Bartlett'*). *PloS one, 9*, p.e92644.

Chapman, N.H., Bonnet, J., Grivet, L., Lynn, J., Graham, N., Smith, R., Sun, G., Walley, P.G., Poole, M., Causse, M. and King, G.J., 2012. High-resolution mapping of a fruit firmness-related quantitative trait

locus in tomato reveals epistatic interactions associated with a complex combinatorial locus. *Plant Physiology, 159*, pp.1644-1657.

Chahar, M.K., Sharma, N., Dobhal, M.P. and Joshi, Y.C., 2011. Flavonoids: a versatile source of anticancer drugs. *Pharmacognosy reviews*, *5*, p.1.

Chen, K. and Gao, C., 2014. Targeted genome modification technologies and their applications in crop improvements. *Plant Cell Reports, 33*, pp.575-583.

Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B. and Jaenisch, R., 2013. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Research, 23*,1163-1171.

Che Radziah, C.M.Z., Nurul Shahnadz, A.H., Naziratul Ain, A.N. and Zainal, Z., 2011. Genetic transformation of antisense ACC oxidase in *Carica papaya L. cv. Sekaki* via particle bombardment. *Malaysian Applied Biology, 40*, pp.39-45.

Cho, H. T. And Kende, H. 1997. Expansins in deepwater rice internodes. *Plant Physiology, 113*, 1137-1143.

Chothia, C. and Lesk, A.M., 1986. The relation between the divergence of sequence and structure in proteins. *The EMBO Journal, 5*, p.823.

Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J. and Voytas, D.F., 2010. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics, 186*, pp.757-761.

Collmer, A. and Ried, J.L., 1988. Construction and characterization of an Erwinia chrysanthemi mutant with directed deletions in all of the pectate lyase structural genes. *Molecular Plant-Microbe Interactions, 1*, pp.32-38.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M. & Robles, M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics, 21*, 3674-3676.

Cosgrove, D.J., 2005. Growth of the plant cell wall. *Nature Reviews Molecular Cell Biology, 6*, pp.8

Deepak, S., Shailasree, S., Kini, R.K., Muck, A., Mithöfer, A. and Shetty, S.H., 2010. Hydroxyproline-rich Glycoproteins and Plant Defence. *Journal of Phytopathology, 158*, pp.585-593.

Delmer, D. P. 1999. CELLULOSE BIOSYNTHESIS: Exciting Times for A Difficult Field of Study. *Annual Review of Plant Physiology and Plant Molecular Biology, 50*, 245-276.

Delmer, D.P. and Haigler, C.H., 2002. The regulation of metabolic flux to cellulose, a major sink for carbon in plants. *Metabolic Engineering, 4*, pp.22-28.

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*, 602-607.

Devos, D. and Valencia, A., 2000. Practical limits of function prediction. *Proteins: Structure, Function, and Bioinformatics, 41*, pp.98-107.

Domingo, C., Roberts, K., Stacey, N.J., Connerton, I.A.N., Ruíz-teran, F. and McCann, M.C., 1998. A pectate lyase from Zinnia elegans is auxin inducible. *The Plant Journal, 13*, pp.17-28.

Doolittle, R.F., 1981. Similar amino acid sequences: chance or common ancestry. *Science, 214*, pp.149-159.

D'Hont, A., Denoeud, F., Aury, J.M., Baurens, F.C., Carreel, F., Garsmeur, O., Noel, B., Bocs, S., Droc, G., Rouard, M. and Da Silva, C., 2012.

The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature, 488*, pp.213-217.

Egea, I., Barsan, C., Bian, W., Purgatto, E., Latché, A., Chervin, C., Bouzayen, M. And Pech, J.-C. 2010. Chromoplast Differentiation: Current Status and Perspectives. *Plant and Cell Physiology, 51*, 1601-1611.

Eklöf, J.M. and Brumer, H., 2010. The XTH gene family: an update on enzyme structure, function, and phylogeny in xyloglucan remodeling. *Plant Physiology, 153*, pp.456-466.

Elitzur, T., Yakir, E., Quansah, L., Zhangjun, F., Vrebalov, J.T., Khayat, E., Giovannoni, J.J. and Friedman, H., 2016. Banana MaMADS transcription factors are necessary for fruit ripening and molecular tools to promote shelf-life and food security. *Plant Physiology,* pp.pp-01866.

Enright, A.J., Van Dongen, S. And Ouzounis, C.A. 2002. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Research, 30*, 1575-1584.

Fantini, E., Falcone, G., Frusciante, S., Giliberto, L. and Giuliano, G., 2013. Dissection of tomato lycopene biosynthesis through virus-induced gene silencing. *Plant Physiology, 163*, pp.986-998.

Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D.L., Wang, Z., Zhang, Z., Zheng, R., Yang, L. and Zeng, L., 2014. Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. *Proceedings of the National Academy of Sciences, 111*, pp.4632-4637.

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.L., Wei, P., Cao, F., Zhu, S., Zhang, F., Mao, Y. and Zhu, J.K., 2013. Efficient genome editing in plants using a CRISPR/Cas system.

Fils-Lycaon, B. And Buret, M. 1991. Changes in glycosidase activities during development and ripening of melon. *Postharvest Biology and Technology, 1*, 143-151.

Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L. And Postlethwait, J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics, 151*, 1531-1545.

Fraser, P.D., Bramley, P. and Seymour, G.B., 2001. Effect of the Cnr mutation on carotenoid formation during tomato fruit ripening.*Phytochemistry, 58*, pp.75-79.

Fraser, P.D., Enfissi, E.M., Halket, J.M., Truesdale, M.R., Yu, D., Gerrish, C. and Bramley, P.M., 2007. Manipulation of phytoene levels in tomato fruit: effects on isoprenoids, plastids, and intermediary metabolism. *The Plant Cell, 19*, pp.3194-3211.

Fraser, P. D., Enfissi, E. M. A. and Bramley, P. M. 2009. Genetic engineering of carotenoid formation in tomato fruit and the potential application of systems and synthetic biology approaches. *Archives of Biochemistry and Biophysics, 483*, 196-204.

Fraser, P.D., Romer, S., Shipton, C.A., Mills, P.B., Kiano, J.W., Misawa, N., Drake, R.G., Schuch, W. and Bramley, P.M., 2002. Evaluation of transgenic tomato plants expressing an additional phytoene synthase in a fruit-specific manner. *Proceedings of the National Academy of Sciences, 99*, pp.1092-1097.

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

Fray, R.G. and Grierson, D., 1993. Molecular genetics of tomato fruit ripening. *Trends in Genetics, 9*, pp.438-443.

Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M. and Dubchak, I., 2004. VISTA: computational tools for comparative genomics. *Nucleic Acids Research, 32*, pp.W273-W279.

Fujisawa, M., Shima, Y., Nakagawa, H., Kitagawa, M., Kimbara, J., Nakano, T., Kasumi, T. and Ito, Y., 2014. Transcriptional regulation of fruit ripening by tomato FRUITFULL homologs and associated MADS box proteins. *The Plant Cell, 26*, pp.89-101.

Gapper, N., Mcquinn, R. and Giovannoni, J. 2013. Molecular and genetic regulation of fruit ripening. *Plant Molecular Biology, 82*, 575-591.

Garcia-Hernandez, M., Berardini, T., Chen, G., Crist, D., Doyle, A., Huala, E., Knee, E., Lambrecht, M., Miller, N., Mueller, L.A. and Mundodi, S., 2002. TAIR: a resource for integrated Arabidopsis data. *Functional & Integrative Genomics, 2*, pp.239-253.

Garcia-Mas, J., Benjak, A., Sanseverino, W., Bourgeois, M., Mir, G., González, V.M., Hénaff, E., Câmara, F., Cozzuto, L., Lowy, E. and Alioto, T., 2012. The genome of melon (*Cucumis melo L.*). *Proceedings of the National Academy of Sciences, 109*, pp.11872-11877.

García-Gago, J.A., Posé, S., Muñoz-Blanco, J., Quesada, M.A. and Mercado, J.A., 2009. The polygalacturonase FaPG1 gene plays a key role in strawberry fruit softening. *Plant Signaling & Behavior, 4*, p.766.

German, M.A., Kandel-Kfir, M., Swarzberg, D., Matsevitz, T. and Granot, D., 2003. A rapid method for the analysis of zygosity in transgenic plants. *Plant Science, 164*, pp.183-187.

Gillaspy, G., Ben-David, H. and Gruissem, W., 1993. Fruits: a developmental perspective. *The Plant Cell, 5*, p.1439.

Giorio, G., Stigliani, A.L. and D'Ambrosio, C., 2008. Phytoene synthase genes in tomato (Solanumlycopersicum L.)–new data on the structures, the deduced amino acid sequences and the expression patterns. *Febs Journal,275*, pp.527-535.

Giovannoni, J. J., Dellapenna, D., Bennett, A. B. and Fischer, R. L. 1989. Expression of a chimeric polygalacturonase gene in transgenic rin

(ripening inhibitor) tomato fruit results in polyuronide degradation but not fruit softening. *The Plant Cell Online, 1*, 53-63.

Giovannoni, J. J. 2004. Genetic Regulation of Fruit Development and Ripening. *The Plant Cell Online, 16*, S170-S180.

Giovannoni, J. J. 2007. Fruit ripening mutants yield insights into ripening control. *Current Opinion in Plant Biology, 10*, 283-289.

Glover, H. and Brady, C. 1994. Purification of three pectin esterases from ripe peach fruit. *Phytochemistry, 37*, 949-955.

Goff, S.A. and Klee, H.J., 2006. Plant volatile compounds: sensory cues for health and nutritional value?. *Science, 311*, pp.815-819.

Goulao, L.F. and Oliveira, C.M., 2008. Cell wall modifications during fruit ripening: when a fruit is not the fruit. *Trends in Food Science & Technology,19*, pp.4-25.

Greenberg, A.J., Moran, J.R., Fang, S. and Wu, C.I. 2006. Adaptive loss of an old duplicated gene during incipient speciation. *Molecular Biology and Evolution, 23*, 401-410.

Grierson, D., Slater, A., Speirs, J. and Tucker, G.A., 1985. The appearance of polygalacturonase mRNA in tomatoes: one of a series of changes in gene expression during development and ripening. *Planta, 163*, pp.263-271.

Gross, K.C., 1982. A rapid and sensitive spectrophotometric method for assaying polygalacturonase using 2-cyanoacetamide. *HortScience, 17*, pp.933-934.

Guo, S., Zhang, J., Sun, H., Salse, J., Lucas, W.J., Zhang, H., Zheng, Y., Mao, L., Ren, Y., Wang, Z. and Min, J., 2013. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions.*Nature genetics, 45*, pp.51-58.

Haas, B.J., Volfovsky, N., Town, C.D., Troukhan, M., Alexandrov, N., Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L., 2002.

Full-length messenger RNA sequences greatly improve genome annotation.*Genome Biology, 3*, p.1.

Hadfield, K. A., Rose, J. K. C., Yaver, D. S., Berka, R. M. and Bennett, A. B. 1998. Polygalacturonase Gene Expression in Ripe Melon Fruit Supports a Role for Polygalacturonase in Ripening-Associated Pectin Disassembly. *Plant Physiology, 117*, 363-373.

Hadfield, K. A. and Bennett, A. B. 1998. Polygalacturonases: Many Genes in Search of a Function. *Plant Physiology, 117*, 337-343.

Hall, L.N., Tucker, G.A., Smith, C.J., Watson, C.F., Seymour, G.B., Bundick, Y., Boniwell, J.M., Fletcher, J.D., Ray, J.A., Schuch, W. and Bird, C.R., 1993. Antisense inhibition of pectin esterase gene expression in transgenic tomatoes. *The Plant Journal, 3*, pp.121-129.

Hamilton, A. J., Lycett, G. W. and Grierson, D. 1990. Antisense gene that inhibits synthesis of the hormone ethylene in transgenic plants. *Nature, 346*, 284-287.

Harb, J., Gapper, N. E., Giovannoni, J. J. and Watkins, C. B. 2012. Molecular analysis of softening and ethylene synthesis and signaling pathways in a non-softening apple cultivar, 'Honeycrisp' and a rapidly softening cultivar, 'McIntosh'. *Postharvest Biology and Technology, 64*, 94-103.

Hardison, R.C., 2003. Comparative genomics. *PLoS Biol, 1*, p.e58.

Harpster, M.H., Brummell, D.A. and Dunsmuir, P., 2002. Suppression of a ripening-related endo-1, 4-β-glucanase in transgenic pepper fruit does not prevent depolymerization of cell wall polysaccharides during ripening*. Plant Molecular Biology, 50*, pp.345-355.

Harpster, M. H., Brummell, D. A. And Dunsmuir, P. 1998. Expression Analysis of a Ripening-Specific, Auxin-Repressed Endo-1,4-β-Glucanase Gene in Strawberry. *Plant Physiology, 118*, 1307-1316.

Harriman, R. W., Tieman, D. M. And Handa, A. K. 1991. Molecular Cloning of Tomato Pectin Methylesterase Gene and its Expression in Rutgers, Ripening Inhibitor, Nonripening, and Never Ripe Tomato Fruits. *Plant Physiology, 97*, 80-87.

He, X. And Zhang, J. 2005. Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. *Genetics, 169*, 1157-1164.

Hileman, L.C., Sundstrom, J.F., Litt, A., Chen, M., Shumba, T. and Irish, V.F., 2006. Molecular and phylogenetic analyses of the MADS-box gene family in tomato. *Molecular Biology and Evolution, 23*, pp.2245-2258.

Hobson, G. and Grierson, D., 1993. Tomato. In *Biochemistry of fruit ripening* (pp. 405-442). Springer Netherlands.

Hobson, G.E., 1964. Polygalacturonase in normal and abnormal tomato fruit.*Biochemical Journal, 92*, p.324.

Hobson, G. E. 1962. Determination of Polygalacturonase in Fruits. *Nature, 195*, 804-805.

Hobson, G.E. and Davies, J.N., 1971. The tomato. *The biochemistry of fruits and their products.* Vol. 2: 437-482.(Hulme, AC ed.).

Horiguchi, G. 2004. RNA silencing in plants: a shortcut to functional analysis. *Differentiation, 72*, 65-73.

Huang, S., Ding, J., Deng, D., Tang, W., Sun, H., Liu, D., Zhang, L., Niu, X., Zhang, X., Meng, M. and Yu, J., 2013. Draft genome of the kiwifruit Actinidia chinensis. *Nature communications, 4*.

Huang, S., Li, R., Zhang, Z., Li, L., Gu, X., Fan, W., Lucas, W.J., Wang, X., Xie, B., Ni, P. and Ren, Y., 2009. The genome of the cucumber, Cucumis sativus L. *Nature genetics, 41*, pp.1275-1281.

Huber, D. J. 1983. The Role of Cell Wall Hydrolases in Fruit Softening. *Horticultural Reviews*. John Wiley & Sons, Inc.

Huber, D. J. and O'donoghue, E. M. 1993. Polyuronides in Avocado (Persea americana) and Tomato (Lycopersicon esculentum) Fruits Exhibit Markedly Different Patterns of Molecular Weight Downshifts during Ripening. *Plant Physiology, 102*, 473-480.

Ito, Y., Nishizawa-Yokoi, A., Endo, M., Mikami, M. and Toki, S., 2015. CRISPR/Cas9-mediated mutagenesis of the RIN locus that regulates tomato fruit ripening. Biochemical and Biophysical Research Communications,467, pp.76-82.

Jaakola, L., Poole, M., Jones, M. O., Kämäräinen-Karppinen, T., Koskimäki, J. J., Hohtola, A., Häggman, H., Fraser, P. D., Manning, K., King, G. J., Thomson, H. and Seymour, G. B. 2010. A SQUAMOSA MADS Box Gene Involved in the Regulation of Anthocyanin Accumulation in Bilberry Fruits. *Plant Physiology, 153*, 1619-1629.

Jaillon, O., Aury, J.M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C. and Vezzi, A., 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature, 449*, pp.463-467.

Jensen, R.A., 2001. Orthologs and paralogs-we need to get it right. *Genome Biology, 2*, p.1.

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B. and Weeks, D.P., 2013. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Research*, p.gkt780.

Jiménez-Bermúdez, S., Redondo-Nevado, J., Muñoz-Blanco, J., Caballero, J. L., López-Aranda, J. M., Valpuesta, V., Pliego-Alfaro, F., Quesada, M. A. & Mercado, J. A. 2002. Manipulation of Strawberry Fruit Softening by Antisense Expression of a Pectate Lyase Gene. *Plant Physiology, 128*, 751-759.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E. 2012. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science, 337*, 816-821.

Jones, C.D. & Begun, D.J. 2005. Parallel evolution of chimeric fusion genes. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 11373–11378.

Jones D.T., Taylor W.R., and Thornton J.M. 1992. The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences 8*, 275-282.

Kachanovsky, D.E., Filler, S., Isaacson, T. and Hirschberg, J., 2012. Epistasis in tomato color mutations involves regulation of phytoene synthase 1 expression by cis-carotenoids. *Proceedings of the National Academy of Sciences, 109*, pp.19021-19026.

Kalamaki, M.S., Harpster, M.H., Palys, J.M., Labavitch, J.M., Reid, D.S. and Brummell, D.A., 2003. Simultaneous transgenic suppression of LePG and LeExp1 influences rheological properties of juice and concentrates from a processing tomato variety. *Journal of Agricultural and Food Chemistry,51*, pp.7456-7464.

Kalamaki, M.S., Powell, A.L., Struijs, K., Labavitch, J.M., Reid, D.S. and Bennett, A.B., 2003. Transgenic overexpression of expansin influences particle size distribution and improves viscosity of tomato juice and paste.*Journal of agricultural and food chemistry, 51*, pp.7465-7471.

Kalamaki, M.S., Stoforos, N.G. and Taoukis, P.S., 2006. Pectic enzymes in tomatoes. *Food Biochemistry and Food Processing*, pp.271-90.

Kanellis, A. K., Solomos, T. & Mattoo, A. K. 1989. Changes in Sugars, Enzymic Activities and Acid Phosphatase Isoenzyme Profiles of Bananas Ripened in Air or Stored in 2.5% O(2) with and without Ethylene. *Plant Physiology, 90*, 251-258.

Keegstra, K., Talmadge, K. W., Bauer, W. D. & Albersheim, P. 1973. The Structure of Plant Cell Walls: III. A Model of the Walls of Suspension-cultured Sycamore Cells Based on the Interconnections of the Macromolecular Components. *Plant Physiology, 51*, 188-197.

Keller, E. and Cosgrove, D. J. 1995. Expansins in growing tomato leaves. *The Plant Journal, 8*, 795-802.

Kevany, B. M., Taylor, M. G. and Klee, H. J. 2008. Fruit-specific suppression of the ethylene receptor LeETR4 results in early-ripening tomato fruit. *Plant Biotechnology Journal, 6*, 295-300.

Kevany, B. M., Tieman, D. M., Taylor, M. G., Cin, V. D. and Klee, H. J. 2007. Ethylene receptor degradation controls the timing of ripening in tomato fruit. *The Plant Journal, 51*, 458-467.

Kim, Y.G., Cha, J. and Chandrasegaran, S., 1996. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences, 93*, pp.1156-1160.

Klee, H.J., 2004. Ethylene signal transduction. Moving beyond Arabidopsis.*Plant Physiology, 135*, pp.660-667.

Klee, H.J., 2010. Improving the flavor of fresh fruits: genomics, biochemistry, and biotechnology. *New Phytologist, 187*, pp.44-56.

Klee, H. J. and Giovannoni, J. J. 2011. Genetics and Control of Tomato Fruit Ripening and Quality Attributes. *Annual Review of Genetics, 45*, 41-59.

Knee, M., Sargent, J. A. and Osborne, D. J. 1977. Cell Wall Metabolism in Developing Strawberry Fruits. *Journal of Experimental Botany, 28*, 377-396.

Kramer, M., Sanders, R., Bolkan, H., Waters, C., Sheeny, R.E. and Hiatt, W.R., 1992. Postharvest evaluation of transgenic tomatoes with reduced levels of polygalacturonase: processing, firmness and

disease resistance.*Postharvest Biology and Technology, 1*, pp.241-255.

Kumar, S., Stecher, G. and Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, p.msw054.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. and Funke, R., 2001. Initial sequencing and analysis of the human genome. *Nature, 409*, pp.860-921.

Lashbrook, C. C., Gonzalez-Bosch, C. and Bennett, A. B. 1994. Two divergent endo-beta-1,4-glucanase genes exhibit overlapping expression in ripening fruit and abscising flowers. *The Plant Cell, 6*, 1485-1493.

Lee, E.R., Kang, G.H. and Cho, S.G., 2007. Effect of flavonoids on human health: old subjects but new challenges. *Recent patents on biotechnology,1*, pp.139-150.

Leseberg, C.H., Eissler, C.L., Wang, X., Johns, M.A., Duvall, M.R. and Mao, L., 2008. Interaction study of MADS-domain proteins in tomato. *Journal of Experimental Botany, 59*, pp.2253-2265.

Lester, D. R., Sherman, W. B. and AtwelL, B. J. 1996. Endopolygalacturonase and the Melting Flesh (M) Locus in Peach. *Journal of the American Society for Horticultural Science, 121*, 231-235.

Levine, M.T., Jones, C.D., Kern, A.D., Lindfors, H.A. and Begun, D.J. 2006. Novel genes derived from noncoding DNA in Drosophila melanogaster are frequently x-linked and exhibit testis-biased expression. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 9935-9939.

Li, J.F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M. and Sheen, J., 2013. Multiplex and homologous

recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nature Biotechnology, 31*, pp.688-691.

Li, L., Stoeckert, C.J. and Roos, D.S., 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Research, 13*, pp.2178-2189.

Li, Z.-C., Durachko, D. and Cosgrove, D. 1993. An oat coleoptile wall protein that induces wall extension in vitro and that is antigenically related to a similar protein from cucumber hypocotyls. *Planta, 191*, 349-356.

Lietzke, S.E., Yoder, M.D., Keen, N.T. and Jurnak, F., 1994. The three-dimensional structure of pectate lyase E, a plant virulence factor from Erwinia chrysanthemi. *Plant Physiology, 106*, pp.849-862.

Litt, A. and Irish, V. F. 2003. Duplication and diversification in the APETALA1/FRUITFULL floral homeotic gene lineage: implications for the evolution of floral development. *Genetics, 165,* 821-833.

Liu, J. and Rost, B. 2003. Domains, motifs and clusters in the protein universe. *Current Opinion in Chemical Biology, 7*, 5-11.

Liu, Y., He, Z., Appels, R. and Xia, X., 2012. Functional markers in wheat: current status and future prospects. *Theoretical and Applied Genetics,125*, pp.1-10.

Long, M. and Langley, C.H. 1993. Natural selection and the origin of jingwei, a chimeric processed functional gene in *Drosophila*. *Science, 260*, 91-95.

Lynch, M. and Conery, J.S. 2003. The evolutionary fate and consequences of duplicate genes. *Science, 290*, 1151-1155.

Lynch, M. and Force, A. 2000. The probability of duplicate gene preservation by subfunctionalization. *Genetics, 154*, 459-473.

Maclachlan, G. and Brady, C., 1992. Multiple forms of 1, 4-β-glucanase in ripening tomato fruits include a xyloglucanase activatable by xyloglucan oligosaccharides. *Functional Plant Biology, 19*, pp.137-146.

Maclachlan, G. and Brady, C., 1994. Endo-1, 4-[beta]-glucanase, xyloglucanase, and xyloglucan endo-transglycosylase activities versus potential substrates in ripening tomatoes. *Plant Physiology, 105*, pp.965-974.

Maclachlan, G. and Brady, C. 1994. Endo-1,4-[beta]-Glucanase, Xyloglucanase, and Xyloglucan Endo-Transglycosylase Activities Versus Potential Substrates in Ripening Tomatoes. *Plant Physiology, 105*, 965-974.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M., 2013. RNA-guided human genome engineering via Cas9. *Science, 339*, pp.823-826.

Manning, K., Tör, M., Poole, M., Hong, Y., Thompson, A.J., King, G.J., Giovannoni, J.J. and Seymour, G.B., 2006. A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nature Genetics, 38*, pp.948-952.

Martel, C., Vrebalov, J., Tafelmeyer, P. and Giovannoni, J. J. 2011. The Tomato MADS-Box Transcription Factor RIPENING INHIBITOR Interacts with Promoters Involved in Numerous Ripening Processes in a COLORLESS NONRIPENING-Dependent Manner. *Plant Physiology, 157*, 1568-1579.

Martel, C. and Giovannoni, J. J. 2007. Fruit Ripening. In: ROBERTS, J. and GONZALEZ-CARRANZAS, Z. (eds.) *Plant Cell Separation and Adhesion*. Oxford, UK:: Blackwell Publishing Ltd.

Marín-Rodríguez, M. C., Smith, D. L., Manning, K., Orchard, J. And Seymour, G. B. 2003. Pectate lyase gene expression and enzyme

activity in ripening banana fruit. *Plant Molecular Biology, 51*, 851-857.

Marín-Rodríguez, M.C., Orchard, J. and Seymour, G.B., 2002. Pectate lyases, cell wall degradation and fruit softening. *Journal of Experimental Botany, 53*, pp.2115-2119.

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

Matsunaga, T., Ishii, T., Matsumoto, S., Higuchi, M., Darvill, A., Albersheim, P. and O'Neill, M.A., 2004. Occurrence of the primary cell wall polysaccharide rhamnogalacturonan II in pteridophytes, lycophytes, and bryophytes. Implications for the evolution of vascular plants. *Plant Physiology, 134*, pp.339-351.

Mbéguié-A-Mbéguié, D., Hubert, O., Baurens, F.C., Matsumoto, T., Chillet, M., Fils-Lycaon, B. and Sidibé-Bocs, S., 2009. Expression patterns of cell wall-modifying genes from banana during fruit ripening and in relationship with finger drop. *Journal of Experimental Botany*, p.erp079.

Mcqueen-Mason, S., Durachko, D. M. and Cosgrove, D. J. 1992. Two endogenous proteins that induce cell wall extension in plants. *The Plant Cell, 4*, 1425-1433.

Mcqueen-Mason, S. J. and Cosgrove, D. J. 1995. Expansin mode of action on cell walls. Analysis of wall hydrolysis, stress relaxation, and binding. *Plant Physiology, 107*, 87-100.

Medina-Escobar, N., Cárdenas, J., Moyano, E., Caballero, J.L. and Muñoz-Blanco, J., 1997. Cloning, molecular characterization and expression pattern of a strawberry ripening-specific cDNA with sequence

homology to pectate lyase from higher plants. *Plant Molecular Biology, 34*, pp.867-877.

Ming, R., Hou, S., Feng, Y., Yu, Q., Dionne-Laporte, A., Saw, J.H., Senin, P., Wang, W., Ly, B.V., Lewis, K.L. and Salzberg, S.L., 2008. The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya Linnaeus*).*Nature, 452*, pp.991-996.

Mitcham, E.J., Gross, K.C. and Ng, T.J., 1989. Tomato Fruit Cell Wall Synthesis during Development and Senescence In Vivo Radiolabeling of Wall Fractions Using [14C] Sucrose. *Plant physiology, 89*, pp.477-481.

Moctezuma, E., Smith, D. L. and Gross, K. C. 2003. Antisense suppression of a β-galactosidase gene (TB G6) in tomato increases fruit cracking. Journal of Experimental Botany, 54, 2025-2033.

Nakamura, A., Maeda, H., Mizuno, M., Koshi, Y. and Nagamatsu, Y. 2003. β-Galactosidase and Its Significance in Ripening of "Saijyo" Japanese Persimmon Fruit. *Bioscience, Biotechnology, and Biochemistry, 67*, 68-76.

Nei, M. and Roychoudhury, A.K., 1973. Probability of fixation of nonfunctional genes at duplicate loci. *American Naturalist*, pp.362-372.

Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D. and Kamoun, S., 2013. Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. *Nature biotechnology, 31*, pp.691-693.

Nguyen-Quoc, B. and Foyer, C.H., 2001. A role for 'futile cycles' involving invertase and sucrose synthase in sucrose metabolism of tomato fruit.*Journal of Experimental Botany, 52*, pp.881-889.

Nilo, P. R., Campos-Vargas, R., Orellana, A. 2012. Assessment of Prunus persica fruit softening using a proteomics approach. *Journal of Proteomics, 75*, pp. 1618–1638.

O'brien, K. P., Remm, M. and Sonnhammer, E. L. L. 2005. Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Research. 33*, 476-480.

O'Malley, R.C., Rodriguez, F.I., Esch, J.J., Binder, B.M., O'Donnell, P., Klee, H.J. and Bleecker, A.B., 2005. Ethylene-binding activity, gene expression levels, and receptor system output for ethylene receptor family members from Arabidopsis and tomato. *The Plant Journal, 41*, pp.651-659.

O'Neill, M.A., Ishii, T., Albersheim, P. and Darvill, A.G., 2004. Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annual Review of Plant Biology, 55*, pp.109-139.

Ochoa-Villarreal, M., Aispuro-Hernández, E., Vargas-Arispuro, I. and Martínez-Téllez, M.Á., 2012. Plant cell wall polymers: function, structure and biological activity of their derivatives. *Polymerization, 4*, pp.63-86.

Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H. and Kanehisa, M. 1999. KEGG: Kyoto Encyclopedia of Genes and Genomes, *Nucleic Acids Research, 27*, 29–34.

Ohno, S. 1970. Evolution by gene duplication. New York: Springer-Verlag.

Osorio, S., Muñoz, C. and Valpuesta, V. 2010. Physiology and Biochemistry of Fruit Flavors. *Handbook of Fruit and Vegetable Flavors*. John Wiley & Sons, Inc.

Ossowski, S., Schwab, R. and Weigel, D. 2008. Gene silencing in plants using artificial microRNAs and other small RNAs. *The Plant Journal, 53*, 674-690.

Parisi, C., Tillie, P. and Rodríguez-Cerezo, E., 2016. The global pipeline of GM crops out to 2020. *Nature biotechnology, 34*, pp.31-36.

Percy, A.E., O'Brien, I.E., Jameson, P.E., Melton, L.D., MacRae, E.A. and Redgwell, R.J., 1996. Xyloglucan endotransglycosylase activity during fruit development and ripening of apple and kiwifruit. *Physiologia Plantarum,96*, pp.43-50.

Petrov, D.A. and Hartl, D.L., 2000. Pseudogene evolution and natural selection for a compact genome. *Journal of Heredity, 91*, pp.221-227.

Phan, T. D., Bo, W., West, G., Lycett, G. W. and Tucker, G. A. 2007. Silencing of the Major Salt-Dependent Isoform of Pectinesterase in Tomato Alters Fruit Softening. *Plant Physiology, 144*, 1960-1967.

Pickersgill, R., Jenkins, J., Harris, G., Nasser, W. and Robert-Baudouy, J., 1994. The structure of Bacillus subtilis pectate lyase in complex with calcium. *Nature Structural & Molecular Biology, 1*, pp.717-723.

Pirhonen, M., Saarilahti, H., Karlsson, M.B. and Palva, E.T., 1991. Identification of pathogenicity determinants of Erwinia carotovora subsp. carotovora by transposon mutagenesis. *Molecular Plant-Microbe Interactions, 4*, pp.276-83.

Posé, S., Paniagua, C., Cifuentes, M., Blanco-Portales, R., Quesada, M.A. and Mercado, J.A., 2013. Insights into the effects of polygalacturonase FaPG1 gene silencing on pectin matrix disassembly, enhanced tissue integrity, and firmness in ripe strawberry fruits. *Journal of Experimental Botany, 64*, pp.3803-3815.

Powell, A. L. T., Kalamaki, M. S., Kurien, P. A., Gurrieri, S. and bennett, A. B. 2003. Simultaneous Transgenic Suppression of LePG and LeExp1 Influences Fruit Texture and Juice Viscosity in a Fresh Market Tomato Variety. *Journal of Agricultural and Food Chemistry, 51*, 7450-7455.

Pressey, R. 1983. β-Galactosidases in Ripening Tomatoes. *Plant Physiology, 71*, 132-135.

Pressey, R. and Avants, J. K. 1972. Multiple forms of pectinesterase in tomatoes. *Phytochemistry, 11*, 3139-3142.

Pressey, R. and Avants, J. K. 1982. Solubilization Of Cell Walls by Tomato Polygalacturonases: Effects Of Pectinesterases. *Journal of Food Biochemistry, 6*, 57-74.

Pua, E.C., Ong, C.K., Liu, P. and Liu, J.Z., 2001. Isolation and expression of two pectate lyase genes during fruit ripening of banana (Musa acuminata). *Physiologia Plantarum, 113*, pp.92-99.

Qin, C., Yu, C., Shen, Y., Fang, X., Chen, L., Min, J., Cheng, J., Zhao, S., Xu, M., Luo, Y. and Yang, Y., 2014. Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum domestication and specialization. *Proceedings of the National Academy of Sciences, 111*, pp.5135-5140.

Quesada, M.A., Blanco-Portales, R., Posé, S., García-Gago, J.A., Jiménez-Bermúdez, S., Muñoz-Serrano, A., Caballero, J.L., Pliego-Alfaro, F., Mercado, J.A. and Muñoz-Blanco, J., 2009. Antisense down-regulation of the FaPG1 gene reveals an unexpected central role for polygalacturonase in strawberry fruit softening. *Plant Physiology, 150*, pp.1022-1032.

Ramírez-Tapias, Y.A., Rivero, C.W., Britos, C.N. and Trelles, J.A., 2015. Alkaline and thermostable polygalacturonase from Streptomyces halstedii ATCC 10897 with applications in waste waters*. Biocatalysis and Agricultural Biotechnology, 4*, pp.221-228.

Redgwell, R.J. and Fry, S.C., 1993. Xyloglucan endotransglycosylase activity increases during kiwifruit (Actinidia deliciosa) ripening (implications for fruit softening). *Plant Physiology, 103*, pp.1399-1406.

Remm, M., Storm, C.E. and Sonnhammer, E.L., 2001. Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *Journal of Molecular Biology, 314*, pp.1041-1052.

Rhodes, M.J.C. and Wooltorton, L.S.C., 1971. The effect of ethylene on the respiration and on the activity of phenylalanine ammonia lyase in swede and parsnip root tissue. *Phytochemistry, 10*, pp.1989-1997.

Ringli, C., Hauf, G. and Keller, B., 2001. Hydrophobic interactions of the structural protein GRP1. 8 in the cell wall of protoxylem elements. *Plant Physiology, 125*, pp.673-682.

Ron, M., Kajala, K., Pauluzzi, G., Wang, D., Reynoso, M.A., Zumstein, K., Garcha, J., Winte, S., Masson, H., Inagaki, S. and Federici, F., 2014. Hairy root transformation using Agrobacterium rhizogenes as a tool for exploring cell type-specific gene expression and function using tomato as a model. *Plant Physiology, 166*, pp.455-469.

Rose, J. K. C., Braam, J., Fry, S. C. and Nishitani, K. 2002. The XTH Family of Enzymes Involved in Xyloglucan Endotransglucosylation and Endohydrolysis: Current Perspectives and a New Unifying Nomenclature. *Plant and Cell Physiology, 43*, 1421-1435.

Rose, J. K. C., Cosgrove, D. J., Albersheim, P., Darvill, A. G. and Bennett, A. B. 2000. Detection of Expansin Proteins and Activity during Tomato Fruit Ontogeny. *Plant Physiology, 123*, 1583-1592.

Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Guldener, U., Mannhaupt, G., Munsterkotter, M. and Mewes, H.W. 2004. The Funcat, A Functional Annotation Scheme for Systematic Classification Of Proteins From Whole Genomes. *Nucleic Acids Research, 32*, 5539–5545.

Ruiz-May, E. and Rose, J.K., 2013. Cell wall architecture and metabolism in ripening fruit and the complex relationship with softening. *The molecular biology and biochemistry of fruit ripening*, pp.163-187.

Saladié, M., Rose, J. K. C., Cosgrove, D. J. and Catalá, C. 2006. Characterization of a new xyloglucan endotransglucosylase/hydrolase (XTH) from ripening tomato fruit

and implications for the diverse modes of enzymic action*. The Plant Journal, 47*, 282-295.

Sander, J.D. and Joung, J.K., 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology, 32*, pp.347-355.

Santiago-Doménech, N., Jiménez-Bemúdez, S., Matas, A.J., Rose, J.K., Muñoz-Blanco, J., Mercado, J.A. and Quesada, M.A., 2008. Antisense inhibition of a pectate lyase gene supports a role for pectin depolymerization in strawberry fruit softening. *Journal of Experimental Botany, 59*, pp.2769-2779.

Saurabh, S., Vidyarthi, A.S. and Prasad, D., 2014. RNA interference: concept to reality in crop improvement. *Planta, 239*, pp.543-564.

Scheller, H.V. and Ulvskov, P., 2010. Hemicelluloses. *Plant Biology, 61*, p.263.

Schomburg, I., Chang, A., and Ebeling, C. 2004. BRENDA, The Enzyme Database: Updates and Major New Developments, *Nucleic Acids Research, 32*, D431–D433.

Schroeder, J. W. 2004. Forage nutrition for ruminants. [Online]. Available: www.ext.nodak.edu/extpubs/ansci/dairy/as1250w.htm

Schröder, N.W., Heine, H., Alexander, C., Manukyan, M., Eckert, J., Hamann, L., Göbel, U.B. and Schumann, R.R., 2004. Lipopolysaccharide binding protein binds to triacylated and diacylated lipopeptides and mediates innate immune responses. *The Journal of Immunology, 173*, pp.2683-2691.

Schuch, W., Kanczler, J., Robertson, D., Hobson, G., Tucker, G., Grierson, D., Bright, S. and Bird, C., 1991. Fruit quality characteristics of transgenic tomato fruit with altered polygalacturonase activity. *HortScience, 26*, pp.1517-1520.

Schuch, W., Kanczler, J., Robertson, D., Hobson, G., Tucker, G., Grierson, D., Bright, S. and Bird, C., 1991. Fruit quality characteristics of transgenic tomato fruit with altered polygalacturonase activity. *HortScience, 26*, pp.1517-1520.

Seifert, G.J. and Roberts, K., 2007. The biology of arabinogalactan proteins. *Annual Review of Plant Biology, 58*, pp.137-161.

Seki, K. and Mostafa, J. 2008. Gene Ontology Annotation As Text Categorization: An Empirical Study. *Information Processing And Management. 44*, 1754–1770

Sesmero, R., Mitchell, J.R., Mercado, J.A. and Quesada, M.A., 2009. Rheological characterisation of juices obtained from transgenic pectate lyase-silenced strawberry fruits. *Food Chemistry, 116*, pp.426-432.

Seymour, G.B., 1993. Banana. *In Biochemistry of fruit ripening* (pp. 83-106). Springer Netherlands.

Seymour, G.B., Colquhoun, I.J., Dupont, M.S., Parsley, K.R. and Selvendran, R.R., 1990. Composition and structural features of cell wall polysaccharides from tomato fruits. *Phytochemistry, 29*, pp.725-731

Seymour, G.B., Ryder, C.D., Cevik, V., Hammond, J.P., Popovich, A., King, G.J., Vrebalov, J., Giovannoni, J.J. and Manning, K., 2011. A SEPALLATA gene is involved in the development and ripening of strawberry (Fragaria× ananassa Duch.) fruit, a non-climacteric tissue. *Journal of Experimental Botany, 62*, pp.1179-1188.

Seymour, G.B., Østergaard, L., Chapman, N.H., Knapp, S. and Martin, C., 2013. Fruit development and ripening. *Annual Review of Plant Biology, 64*, pp.219-241.

Seymour, G.B. and Gross, K.C., 1996. Cell wall disassembly and fruit softening. *Postharvest News and Information* (United Kingdom).

Shalem, O., Sanjana, N.E. and Zhang, F., 2015. High-throughput functional genomics using CRISPR-Cas9. *Nature Reviews Genetics, 16*, pp.299-311.

Shalem, O., Sanjana, N.E. and Zhang, F., 2015. High-throughput functional genomics using CRISPR-Cas9. *Nature Reviews Genetics, 16*, pp.299-311.

Sheehy, R.E., Kramer, M. and Hiatt, W.R., 1988. Reduction of polygalacturonase activity in tomato fruit by antisense RNA. *Proceedings of the National Academy of Sciences, 85*, pp.8805-8809.

Shulaev, V., Sargent, D.J., Crowhurst, R.N., Mockler, T.C., Folkerts, O., Delcher, A.L., Jaiswal, P., Mockaitis, K., Liston, A., Mane, S.P. and Burns, P., 2011. The genome of woodland strawberry (Fragaria vesca). *Nature Genetics, 43*, pp.109-116.

Shulaev, V., Sargent, D.J., Crowhurst, R.N., Mockler, T.C., Folkerts, O., Delcher, A.L., Jaiswal, P., Mockaitis, K., Liston, A., Mane, S.P. and Burns, P., 2011. The genome of woodland strawberry (Fragaria vesca). *Nature Genetics, 43*, pp.109-116.

Sitrit, Y. and Bennett, A. B. 1998. Regulation of Tomato Fruit Polygalacturonase mRNA Accumulation by Ethylene: A Re-Examination. *Plant Physiology, 116*, 1145-1150.

Smith, C.J., Watson, C.F., Morris, P.C., Bird, C.R., Seymour, G.B., Gray, J.E., Arnold, C., Tucker, G.A., Schuch, W., Harding, S. and Grierson, D., 1990. Inheritance and effect on ripening of antisense polygalacturonase genes in transgenic tomatoes. *Plant Molecular Biology, 14*, pp.369-379.

Smith, C. S., Watson, C., Morris, P., Bird, C., Seymour, G., Gray, J., Arnold, C., Tucker, G., Schuch, W., Harding, S. and Grierson, D. 1990. Inheritance and effect on ripening of antisense

polygalacturonase genes in transgenic tomatoes. *Plant Molecular Biology, 14*, 369-379.

Smith, D. L., Abbott, J. A. and Gross, K. C. 2002. Down-Regulation of Tomato β-Galactosidase 4 Results in Decreased Fruit Softening. *Plant Physiology, 129*, 1755-1762.

Smith, D. L. and Gross, K. C. 2000. A Family of at Least Seven β-Galactosidase Genes Is Expressed during Tomato Fruit Development. *Plant Physiology, 123*, 1173-1184.

Smith, L.G., 2001. Plant cell division: building walls in the right places. *Nature Reviews Molecular Cell Biology, 2*, pp.33-39.

Smith, S.E. and Gianinazzi-Pearson, V., 1988. Physiological interactions between symbionts in vesicular-arbuscular mycorrhizal plants. *Annual Review of Plant Physiology and Plant Molecular Biology, 39*, pp.221-244.

Snel, B., Bork, P. and Huynen, M.A., 2002. Genomes in flux: the evolution of archaeal and proteobacterial gene content. *Genome Research, 12*, pp.17-25.

Sorek, R., Lawrence, C. M. and Wiedenheft, B. 2013. CRISPR-Mediated Adaptive Immune Systems in Bacteria and Archaea*. Annual Review of Biochemistry, 82*, 237-266.

Sozzi, G. O., Greve, L. C., Prody, G. A. and Labavitch, J. M. 2002. Gibberellic Acid, Synthetic Auxins, and Ethylene Differentially Modulate α-l-Arabinofuranosidase Activities in Antisense 1-Aminocyclopropane-1-Carboxylic Acid Synthase Tomato Pericarp Discs. *Plant Physiology, 129*, 1330-1340.

Spjut, R.W., 1994. A systematic treatment of fruit types (Vol. 70). New York Botanical Garden.

Stacewicz-Sapuntzakis, M. and Bowen, P.E., 2005. Role of lycopene and tomato products in prostate health. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, *1740*, pp.202-205

Stein, L., 2001. Genome annotation: from sequence to biology. *Nature Reviews Genetics, 2*, pp.493-503.

Su, L., Diretto, G., Purgatto, E., Danoun, S., Zouine, M., Li, Z., Roustan, J.P., Bouzayen, M., Giuliano, G. and Chervin, C., 2015. Carotenoid accumulation during tomato fruit ripening is modulated by the auxin-ethylene balance. *BMC plant biology, 15*, p.1.

Tacken, E.J., Ireland, H.S., Wang, Y.Y., Putterill, J. and Schaffer, R.J., 2012. Apple EIN3 BINDING F-box 1 inhibits the activity of three apple EIN3-like transcription factors. AoB plants, 2012, p.pls034.

Tadakittisarn, S., Songpim, M. and Vaithanomsat, P., 2009. Polygalacturonase and Pectate Lyase Activity During Ripening of Kluay Hom Thong Fruit. *In Proceedings of the 47th Kasetsart University Annual Conference, Kasetsart*, 17-20 March, 2009. Subject: Plants. (pp. 153-161). Kasetsart University.

Tateishi, A., Inoue, H., Shiba, H. and Yamaki, S. 2001. Molecular Cloning of β-Galactosidase from Japanese Pear (*Pyrus pyrifolia*) and its Gene Expression with Fruit Ripening. *Plant and Cell Physiology, 42*, 492-498.

Tateishi A, Inoue H, Yamaki S., 2001b. Fluctuations in Activities of Three. BETA.-Galactosidase Isoforms from Ripening Avocado (*Persea americana*) Fruit and their different Activities against its Cell Wall Polysaccharides as Substrates*. 園芸学会雑誌, 70*, pp.586-592.

Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov,D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Smirnov, S., Sverdlov, A.V., Vasudevan, S., Wolf, Y.I., Yin, J.J. and Natale, D.A. 2003. The COG

Database: An Updated Version Includes Eukaryotes. *BMC Bioinformatics, 4*,1-41.

Tatusov, R.L., Koonin, E.V. and Lipman, D.J., 1997. A genomic perspective on protein families. *Science, 278*, pp.631-637.

Thompson, A. J., Tor, M., Barry, C. S., Vrebalov, J., Orfila, C., Jarvis, M. C., Giovannoni, J. J., Grierson, D. and Seymour, G. B. 1999. Molecular and Genetic Characterization of a Novel Pleiotropic Tomato-Ripening Mutant. *Plant Physiology, 120*, 383-390.

Tieman, D. M., Harriman, R. W., Ramamohan, G. and Handa, A. K. 1992. An Antisense Pectin Methylesterase Gene Alters Pectin Chemistry and Soluble Solids in Tomato Fruit. *The Plant Cell Online, 4*, 667-679.

Tieman, D.M., Zeigler, M., Schmelz, E.A., Taylor, M.G., Bliss, P., Kirst, M. and Klee, H.J., 2006. Identification of loci affecting flavour volatile emissions in tomato fruits. *Journal of Experimental Botany, 57*, pp.887-896.

Tieman, D. M. and Klee, H. J. 1999. Differential Expression of Two Novel Members of the Tomato Ethylene-Receptor Family. *Plant Physiology, 120*, 165-172.

Tomato Genome Consortium, 2012. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature, 485*, pp.635-641.

Trainotti, L., Zanin, D. and Casadoro, G. 2003. A cell wall-oriented genomic approach reveals a new and unexpected complexity of the softening in peaches. *Journal of Experimental Botany, 54*, 1821-1832.

Tucker, G.A., Robertson, N.G. and Grierson, D., 1980. Changes in polygalacturonase isoenzymes during the 'ripening'of normal and mutant tomato fruit. *European Journal of Biochemistry, 112*, pp.119-124.

Tucker, G.A., Robertson, N.G. and Grierson, D., 1981. The Conversion of Tomato-Fruit Polygalacturonase Isoenzyme 2 into Isoenzyme 1 in vitro. *European Journal of Biochemistry, 115*, pp.87-90.

Tucker, G. A., Robertson, N. G. and Grierson, D. 1982. Purification and changes in activities of tomato pectinesterase isoenzymes. *Journal of the Science of Food and Agriculture, 33*, 396-400.

Tucker, M. L., Sexton, R., Del Campillo, E. and Lewis, L. N. 1988. Bean Abscission Cellulase : Characterization of a cDNA Clone and Regulation of Gene Expression by Ethylene and Auxin. *Plant Physiology, 88*, 1257-1262.

Tucker, M. L. and Milligan, S. B. 1991. Sequence Analysis and Comparison of Avocado Fruit and Bean Abscission Cellulases. *Plant Physiology, 95*, 928-933.

Uluisik, S., Chapman, N.H., Smith, R., Poole, M., Adams, G., Gillis, R.B., Besong, T.M., Sheldon, J., Stiegelmeyer, S., Perez, L. and Samsulrizal, N., 2016. Genetic improvement of tomato by targeted control of fruit softening.*Nature Biotechnology*.

Vandenbussche, F., Vancompernolle, B., Rieu, I., Ahmad, M., Phillips, A., Moritz, T., Hedden, P. and Van Der Straeten, D. 2007. Ethylene-induced Arabidopsis hypocotyl elongation is dependent on but not mediated by gibberellins. *Journal of Experimental Botany, 58*, 4269-4281.

Van Der Pijl, L. 1982. Principle of Dispersal in Higher Plant. 3rd ed. New York: Springer-Verlag.

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., Fontana, P., Bhatnagar, S.K., Troggio, M., Pruss, D. and Salvi, S., 2010. The genome of the domesticated apple (*Malus domestica Borkh*.). *Nature genetics, 42,* pp.833-839.

Velasco, R., Zharkikh, A., Troggio, M., Cartwright, D.A., Cestaro, A., Pruss, D., Pindo, M., FitzGerald, L.M., Vezzulli, S., Reid, J. and

Malacarne, G., 2007. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. *PloS one, 2*, p.e1326.

Verde, I., Abbott, A.G., Scalabrin, S., Jung, S., Shu, S., Marroni, F., Zhebentyayeva, T., Dettori, M.T., Grimwood, J., Cattonaro, F. and Zuccolo, A., 2013. The high-quality draft genome of peach (Prunus persica) identifies unique patterns of genetic diversity, domestication and genome evolution.*Nature Genetics, 45*, pp.487-494.

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

Vicente, A. R., Saladié, M., Rose, J. K. C. and Labavitch, J. M. 2007. The linkage between cell wall metabolism and fruit softening: looking to the future. *Journal of the Science of Food and Agriculture, 87*, 1435-1448.

Voytas, D.F. and Gao, C., 2014. Precision genome engineering and agriculture: opportunities and regulatory challenges. *PLoS Biol, 12*, p.e1001877.

Vrebalov, J., Ruezinsky, D., Padmanabhan, V., White, R., Medrano, D., Drake, R., Schuch, W. and Giovannoni, J., 2002. A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus.*Science, 296*, pp.343-346.

Wang, M., Vannozzi, A., Wang, G., Liang, Y.H., Tornielli, G.B., Zenoni, S., Cavallini, E., Pezzotti, M. and Cheng, Z.M.M., 2014. Genome and transcriptome analysis of the grapevine (Vitis vinifera L.) WRKY gene family.*Horticulture Research, 1*, 14016

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C. and Qiu, J.L., 2014. Simultaneous editing of three homoeoalleles in hexaploid

bread wheat confers heritable resistance to powdery mildew. *Nature Biotechnology, 32*, pp.947-951.

Ware, D.H., Jaiswal, P., Ni, J., Yap, I.V., Pan, X., Clark, K.Y., Teytelman, L., Schmidt, S.C., Zhao, W., Chang, K. and Cartinhour, S., 2002. Gramene, a tool for grass genomics. *Plant Physiology, 130*, pp.1606-1613.

Watson, J.M., Fusaro, A.F., Wang, M. and Waterhouse, P.M., 2005. RNA silencing platforms in plants. *Febs Letters, 579*, pp.5982-5987.

Weber, E., Gruetzner, R., Werner, S., Engler, C. and Marillonnet, S., 2011. Assembly of designer TAL effectors by Golden Gate cloning. *Plos One, 6*, p.e19722.

Willats, W.G., Orfila, C., Limberg, G., Buchholt, H.C., van Alebeek, G.J.W., Voragen, A.G., Marcus, S.E., Christensen, T.M., Mikkelsen, J.D., Murray, B.S. and Knox, J.P., 2001. Modulation of the Degree and Pattern of Methyl-esterification of Pectic Homogalacturonan in Plant Cell Walls IMPLICATIONS FOR PECTIN METHYL ESTERASE ACTION, MATRIX PROPERTIES, AND CELL ADHESION. *Journal of Biological Chemistry,276*, pp.19404-19413.

Wing, R.A., Yamaguchi, J., Larabell, S.K., Ursin, V.M. and McCormick, S., 1990. Molecular and genetic characterization of two pollen-expressed genes that have sequence similarity to pectate lyases of the plant pathogen Erwinia. *Plant Molecular Biology, 14*, pp.17-28.

Wu, J., Wang, Z., Shi, Z., Zhang, S., Ming, R., Zhu, S., Khan, M.A., Tao, S., Korban, S.S., Wang, H. and Chen, N.J., 2013. The genome of the pear (Pyrus bretschneideri Rehd.). *Genome Research, 23*, pp.396-408.

Wóycicki, R., Witkowicz, J., Gawroński, P., Dąbrowska, J., Lomsadze, A., Pawełkowicz, M., Siedlecka, E., Yagi, K., Pląder, W., Seroczyńska, A. and Śmiech, M., 2011. The genome sequence of the North-

European cucumber (*Cucumis sativus L.*) unravels evolutionary adaptation mechanisms in plants. *PloS one, 6*, p.e22728.

Yadav, S., Yadav, P.K., Yadav, D. and Yadav, K.D.S., 2009. Pectin lyase: a review. *Process Biochemistry, 44*, pp.1-10.

Yoder, M. D., Keen, N. T. and Jurnak, F. 1993. New domain motif: the structure of pectate lyase C, a secreted plant virulence factor. *Science, 260*, 1503-1507.

Yoder, M.D. and Jurnak, F., 1995. Protein motifs. 3. The parallel beta helix and other coiled folds. *The FASEB journal, 9*, pp.335-342.

Youssef, S.M., Jiménez-Bermúdez, S., Bellido, M.L., Martín-Pizarro, C., Barceló, M., Abdal-Aziz, S.A., Caballero, J.L., López-Aranda, J.M., Pliego-Alfaro, F., Muñoz, J. and Quesada, M.A., 2009. Fruit yield and quality of strawberry plants transformed with a fruit specific strawberry pectate lyase gene. *Scientia Horticulturae, 119*, pp.120-125.

Zhang, H., Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z., Mao, Y., Yang, L., Zhang, H., Xu, N. and Zhu, J.K., 2014. The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnology Journal, 12*, pp.797-807.

Zhang, H., Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z., Mao, Y., Yang, L., Zhang, H., Xu, N. and Zhu, J.K., 2014. The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnology Journal, 12*, pp.797-807.

Zhang, J., 2003. Evolution by gene duplication: an update. *Trends In Ecology & Evolution, 18*, pp.292-298.

Zhang, Q., Chen, W., Sun, L., Zhao, F., Huang, B., Yang, W., Tao, Y., Wang, J., Yuan, Z., Fan, G. and Xing, Z., 2012. The genome of Prunus mume. *Nature communications, 3*, p.1318.

Zhong, S., Fei, Z., Chen, Y.R., Zheng, Y., Huang, M., Vrebalov, J., McQuinn, R., Gapper, N., Liu, B., Xiang, J. and Shao, Y., 2013. Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nature Biotechnology, 31*, pp.154-159.

Östlund, G., Schmitt, T., Forslund, K., Köstler, T., Messina, D.N., Roopra, S., Frings, O. and Sonnhammer, E.L., 2010. InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Research, 38*, pp.D196-D203.

9 APPENDIX

9.1 Speed DNA Extraction Buffer

1 M Tris HCL (pH 7.5)	200	mL
4 M NaCl	62.5	mL
0.5 M EDTA	50	mL
10% SDS	50	mL

(The recipe for 1 L)

9.2 Lysis Buffer

Triton lysis buffer (complete)

25 mM HEPES

100 mM NaCl

1mM EDTA → Preparation EDTA at 0.5 M (pH 8.0): Add 186.1 g of disodium EDTA•$2H_2O$ to 800 mL of H_2O. Stir vigorously on a magnetic stirrer. Adjust the pH to 8.0 with NaOH (~20 g of NaOH pellets). Dispense into aliquots and sterilize by autoclaving. The disodium salt of EDTA will not go into solution until the pH of the solution is adjusted to ~8.0 by the addition of NaOH

10% (v/v) glycerol

1% (v/v) Triton X-100

9.3 Buffer

Plasmid extraction kit

Buffer P1 (cell suspension buffer)

50mM Tris-HCl pH 8.0 (25℃)

10mM EDTA50ug/ml of RNase A

Buffer P2 (lysis buffer)

0.2M NaCl and 1% SDS

Buffer N3 (neutralization and binding buffer)

4M guanidine hydrochloride (GuHCl)

0.5M potassium acetate pH 4.2

Buffer PB (wash buffer)

5M guanidine hydrochloride (GuHCl)

20mM Tris-HCl pH 6.6 (25℃)

38% ethanol

Buffer PE (wash buffer)

20mM NaCl

2mM Tris-HCl pH 7.5 (25℃)

80% ethanol

Buffer EB (elution buffer)

10mM Tris-HCl pH 8.5 (25℃)

Gel extraction kit (DNA fragment recovery from agarose gel)

Buffer QG (neutralization and binding buffer)

5.5M guanidine thiocyanate (GuSCN)

20mM Tris-HCl pH 6.6 (25℃)

Buffer PE (wash buffer)

20mM NaCl

2mM Tris-HCl pH 7.5 (25℃)

80% ethanol

Buffer EB (elution buffer)

10mM Tris-HCl pH 8.5 (25℃)

PCR purification kit (Direct purification of PCR products withoutelectrophoresis)

Buffer PB (DNA binding buffer)

5.5 M guanidine hydrochloride (GuHCl)

20mM Tris-HCl pH 6.6 (25℃)

Buffer PE (wash buffer)

20mM NaCl

2mM Tris-HCl pH 7.5 (25℃)

80% ethanolBuffer

EB (elution buffer)

10mM Tris-HCl pH 8.5 (25℃)

9.4 Media

1: Luria Broth media (LB) following Sambrook and Russell,

2001 protocol.

Tryptone 10 g

Yeast Extract 5 g

NaCl 10 g

Dissolve components in 1 liter of distilled or deionized water.

For LB agar* add agar to a final concentration of 1.5%.

Heat the mixture to boiling to dissolve agar and sterilize by autoclaving at 15 psi, from 121-124°C for 15 minutes.

2: Media for plant tissue culture

Chemicals / Volume	KCMS medium (liquid) 1L	KCMS medium (solid) 1L	2Z medium 1L	Rooting medium 1L	1/2 MS medium 1L
MS (Including vitamins, Duchefa MO222)	4.4 g	4.4 g	4.4 g	4.4 g	2.2 g
Sucrose	20 g	20 g	30 g	30 g	15 g
KH2PO4	200 mg	200 mg			
PH	5.8	5.7	5.8	5.8	5.9
Agar		8 g	8 g	8 g	8 g
	Autoclaving	Autoclaving	Autoclaving	Autoclaving	Autoclaving
Thiamine	0.9 mg L^{-1}	0.9 mg L^{-1}			
Acetosyringone	100 µM	100 µM			
2,4 D	200 µg L^{-1}				
Kinetin	100 µg L^{-1}				
R3 vitamins					500 µL
Nitsch vitamin × 1000			1 mL	1 mL	
Zeatin			2 mg L^{-1}	1 mg L^{-1}	
IAA				0.87 mg L^{-1}	
Kanamycin			75 mg L^{-1}	50 mg L^{-1}	
Timentin			150 mg L^{-1}	150 mg L^{-1}	

R3 vitamins

Thiamine	1 g L^{-1}
Nicotinic acid	0.5 g L^{-1}
Pyridoxine	0.5 g L^{-1}

9.5 Bioinformatics Script

Local BLAST

```
blastall –p <type of blast> -d <directory of database> -i
<input file in fasta format> -o <output file>
```

Format DB

```
formatdb –p F -i all_seqs.fasta -n customBLASTdb
```

Inparanoid

```
tar –zxvf inparanoid.tar

formatdb –i <genome protein sequences>

perl inparanoid.pl <genome sequences> <genome sequences
that we want to compare with>
```

Multiparanoid

```
perl multiparanoid.pl -species
tomato_slycopercium.fasta+musa_banana.fasta+Vitis_vinifer
a_grape.fasta
```

9.6 List of genes expression in all fleshy fruits

Gene Name	Expression Tomato	Expression Melon	Expression Grape	Expresion Banana
acid beta-fructofuranosidase-like			GSVIVT01018625001	GSMUA_Achr5P19880_001
alpha-l-arabinofuranosidase 1-like		MELO3C009722	GSVIVT01025847001	
Arabinogalactan	Solyc02g092790			
beta-galactosidase 3-like			GSVIVT01022339001	
Beta-galactosidase	Solyc12g008840			
beta-galactosidase 3-like	Solyc02g084720		GSVIVT01018853001	
beta-galactosidase 8-like			GSVIVT01025170001	
beta-galactosidase 1-like			GSVIVT01016849001	
beta-galactosidase 9-like			GSVIVT01020995001	
beta-xylosidase alpha-l-arabinofuranosidase 2-like		MELO3C018667	GSVIVT01013328001	
			GSVIVT01022541001	
beta-xylosidase	Solyc05g005080			
endo- -beta-glucanase			GSVIVT01023102001	
endoglucanase 6-like	Solyc08g082250		GSVIVT01019523001	
endo- -beta-glucanase			GSVIVT01019664001	
endo- -beta-glucanase			GSVIVT01028042001	
endoglucanase 9-like		MELO3C022160	GSVIVT01037210001	
expansin	Solyc06g051800		GSVIVT01024946001	
expansin 2			GSVIVT01032681001	GSMUA_Achr5P07480_001
expansin protein			GSVIVT01008762001	
Expansin	Solyc02g088100			
Expansin	Solyc09g010860			
alpha-expansin 20 precursor family protein	Solyc03g031840		GSVIVT01023857001	

alpha-expansin 1			GSVIVT01029411001	
beta-expansin precursor			GSVIVT01018163001	
			GSVIVT01018168001	
expansin			GSVIVT01024946001	
expansin	Solyc06g051800		GSVIVT01024946001	
expansin-like a1-like			GSVIVT01023919001	
probable beta- -xylosyltransferase irx14h-like			GSVIVT01022064001	
probable beta- -xylosyltransferase irx9h-like			GSVIVT01009245001	
mannan endo- -beta-mannosidase 5-like			GSVIVT01037212001	
			GSVIVT01037214001	
Mannan endo-1 4-beta-mannosidase	Solyc01g008710			
Pectinesterase	Solyc07g064170			
Pectinesterase	Solyc07g064180			
Pectinesterase	Solyc03g123630			
polygalacturonase pectinase			GSVIVT01038241001	
polygalacturonase-like		MELO3C013129	GSVIVT01038243001	
probable polygalacturonase at1g80170-like			GSVIVT01017279001	GSMUA_Achr3P13670_001
polygalacturonase pectinase			GSVIVT01038241001	
polygalacturonase-like		MELO3C013129	GSVIVT01038243001	
polygalacturonase pectinase			GSVIVT01038241001	
polygalacturonase-like		MELO3C013129	GSVIVT01038243001	
polygalacturonase	Solyc10g080210	MELO3C023556	GSVIVT01033303001	
			GSVIVT01033364001	
polygalacturonase-inhibiting protein		MELO3C016384	GSVIVT01033370001	
bifunctional inhibitor lipid-transfer protein seed storage 2s albumin		MELO3C010312	GSVIVT01001298001	

superfamily protein				
kda proline-rich			GSVIVT01015895001	
probable xyloglucan endotransglucosylase hydrolase protein 23-like	Solyc03g093080		GSVIVT01029162001	
xyloglucan endo-1 family protein		MELO3C017481		
probable xyloglucan endotransglucosylase hydrolase protein 30-like	Solyc01g081060		GSVIVT01013055001	
probable xyloglucan endotransglucosylase hydrolase protein 32-like			GSVIVT01031486001	
probable xyloglucan endotransglucosylase hydrolase protein 32-like			GSVIVT01033658001	
exgt1 family protein			GSVIVT01001124001	
Xyloglucan endotransglucosylase/hydrolase	Solyc02g091920			
Xyloglucan endotransglucosylase/hydrolase	Solyc12g011030			
Xyloglucan endotransglucosylase/hydrolase	Solyc07g056000			
Xyloglucan endotransglucosylase/hydrolase	Solyc12g017240			
xyloglucan endotransglucosylase hydrolase protein 9	Solyc07g052980		GSVIVT01000416001	
Xyloglucan endotransglucosylase/hydrolase	Solyc01g099630			
alpha-glucan-branching enzyme-like			GSVIVT01008673001	
4-alpha-glucanotransferase dpe2-like		MELO3C019135	GSVIVT01025810001	GSMUA_Achr3P08780_001
alpha fucosyltransferase		MELO3C018453	GSVIVT01031736001	
alpha-glucosidase 2-like		MELO3C013786	GSVIVT01022815001	GSMUA_Achr11P06330_001
			GSVIVT01036946001	
probable alpha-mannosidase i mns4-like			GSVIVT01018097001	
gdp-man:man c -pp-dol alpha- -mannosyltransferase-like		MELO3C018365	GSVIVT01032791001	

alpha- -mannosyltransferase alg2-like			GSVIVT01014186001	
lactosylceramide 4-alpha-			GSVIVT01002627001	
alpha- -glucan-protein synthase		MELO3C002047	GSVIVT01008552001	
alpha- -glucan-protein synthase		MELO3C011416		
alpha- -glucan-protein synthase		MELO3C002047	GSVIVT01008552001	
alpha- -glucan-protein synthase		MELO3C011416		
probable beta-d-xylosidase 6-like			GSVIVT01006776001	GSMUA_Achr5P12120_001
probable beta-d-xylosidase 7-like		MELO3C011366	GSVIVT01009003001	
alpha-l-fucosidase 1			GSVIVT01031495001	
alpha-mannosidase 2x-like			GSVIVT01032586001	
arabinogalactan peptide 20-like		MELO3C006062	GSVIVT01027049001	GSMUA_Achr8P27950_001
			GSVIVT01028741001	
uncharacterized loc101208147			GSVIVT01001144001	GSMUA_Achr9P28420_001
polygalacturonase qrt3-like			GSVIVT01029528001	
polygalacturonase qrt3-like		MELO3C009962		
glycoside hydrolase family 2 family protein		MELO3C012472	GSVIVT01013489001	
probable beta-d-xylosidase 2-like			GSVIVT01028104001	
beta-amylase 7-like			GSVIVT01026920001	
lysosomal beta glucosidase-like			GSVIVT01024793001	
lysosomal beta glucosidase-like			GSVIVT01022483001	
lysosomal beta glucosidase-like		MELO3C026609		
beta-galactosidase 16-like			GSVIVT01008835001	
beta-galactosidase 13-like		MELO3C013360	GSVIVT01024031001	
beta-galactosidase 7-like			GSVIVT01015038001	
beta-galactosidase 17-like		MELO3C025840	GSVIVT01016865001	
uncharacterized loc101203100		MELO3C020910	GSVIVT01000105001	
beta-glucosidase 18-like			GSVIVT01032009001	

hydroxyisourate hydrolase family protein		MELO3C011430	GSVIVT01012191001 GSVIVT01012192001	
beta-glucosidase 11-like		MELO3C017964	GSVIVT01028001001	GSMUA_Achr9P16420_001
beta-glucosidase 47-like			GSVIVT01012650001	
beta-glucosidase 12-like			GSVIVT01032005001	
vicianin hydrolase-like			GSVIVT01032007001	
beta-mannosidase		MELO3C015324	GSVIVT01000073001	
eg45-like domain containing			GSVIVT01035758001 GSVIVT01035760001	
eg45-like domain containing protein			GSVIVT01035834001	
cellulose synthase family protein		MELO3C008022	GSVIVT01013471001	
cellulose synthase a catalytic subunit 6		MELO3C025029	GSVIVT01034552001	
cellulose synthase 4			GSVIVT01028402001	
Cellulose synthase	Solyc04g071650			
cellulose synthase a catalytic subunit 1	Solyc08g061100	MELO3C003689	GSVIVT01035830001	GSMUA_Achr5P06050_001
cellulose synthase-like protein g3-like			GSVIVT01019568001	
cellulose synthase-like protein h1-like			GSVIVT01030456001	
cellulose synthase-like protein h1-like			GSVIVT01030462001	
cellulose synthase-like protein e1-like			GSVIVT01014703001	
cellulose synthase		MELO3C002672	GSVIVT01021248001	
cellulose synthase catalytic subunit	Solyc01g087210	MELO3C023114	GSVIVT01033297001	
cellulose synthase a catalytic subunit 7			GSVIVT01023643001	
cellulose synthase-like protein d4-			GSVIVT01021798001	

like			GSVIVT01023850001	
Cellulose synthase	Solyc07g043390			
probable xyloglucan glycosyltransferase 5-like			GSVIVT01032523001	
probable xyloglucan glycosyltransferase 12-like			GSVIVT01009290001	
glucomannan 4-beta-mannosyltransferase 9-like			GSVIVT01033767001	
cellulose synthase-like protein d5-like	Solyc11g066820		GSVIVT01028071001	
cellulose synthase-like protein d3-like			GSVIVT01024398001	
endo- -beta-xylanase c		MELO3C018522	GSVIVT01037825001	GSMUA_Achr1P12500_001
			GSVIVT01037832001	
glycosyl hydrolase family 10 protein			GSVIVT01024049001	
endo- -beta-xylanase a-like			GSVIVT01009272001	
endoglucanase 5-like		MELO3C023598	GSVIVT01012043001	
glycosyl hydrolase family 9 family protein	Solyc04g081300		GSVIVT01009881001	
endoglucanase 14-like			GSVIVT01032799001	
endoglucanase 11-like			GSVIVT01037709001	
endoglucanase 24-like			GSVIVT01024179001	
exostosin family protein			GSVIVT01036850001	
probable glucuronoxylan glucuronosyltransferase irx7-like		MELO3C004998	GSVIVT01031485001	
probable glucuronoxylan glucuronosyltransferase irx7-like		MELO3C026947		
alpha-expansin 20 precursor family protein			GSVIVT01023857001	
expansin-a12-like			GSVIVT01014362001	
expansin-a7-like			GSVIVT01019889001	
fasciclin-like arabinogalactan protein 11-like			GSVIVT01025167001	
fasciclin-like arabinogalactan protein 2-like	Solyc07g045440		GSVIVT01014684001	

fasciclin-like arabinogalactan protein 1-like	Solyc10g005960		GSVIVT01030085001	
putative fasciclin-like protein			GSVIVT01020616001	
fasciclin-like arabinogalactan protein 4-like			GSVIVT01024973001	
galactoside 2-alpha-l-fucosyltransferase-like			GSVIVT01003228001	
galactoside 2-alpha-l-fucosyltransferase-like			GSVIVT01007978001	
probable beta- -xylosyltransferase irx9-like			GSVIVT01025636001	
hydrolyzing o-glycosyl		MELO3C015243	GSVIVT01031317001	
glucan endo- -beta-glucosidase-like protein 2-like	Solyc03g115200	MELO3C017220	GSVIVT01007873001	GSMUA_Achr4P32790_001
callose synthase 11-like			GSVIVT01005204001	
callose synthase 12-like		MELO3C013621		
callose synthase 3-like		MELO3C015659	GSVIVT01025362001	
callose synthase 3-like		MELO3C015661	GSVIVT01019235001	
callose synthase 9-like isoform x2			GSVIVT01007560001	
probable polygalacturonase-like	Solyc05g049980		GSVIVT01019405001	
probable polygalacturonase-like			GSVIVT01028057001	
			GSVIVT01035089001	
probable polygalacturonase-like		MELO3C023234	GSVIVT01026477001	
probable polygalacturonase-like		MELO3C005746	GSVIVT01019516001	
hydrolyzing o-glycosyl			GSVIVT01019433001	GSMUA_Achr2P06020_001
			GSVIVT01000573001	
beta- -mannosyl-glycoprotein 4-beta-n-acetylglucosaminyltransferase-like		MELO3C011944	GSVIVT01000535001	GSMUA_Achr10P16790_001
		MELO3C021589	GSVIVT01032386001	
glycosyltransferase family protein 2		MELO3C004898	GSVIVT01025204001	
probable galacturonosyltransferase 7-like		MELO3C025283	GSVIVT01029827001	
probable		MELO3C025283		

galacturonosyltransferase 7-like				
nucleotide-diphospho-sugar transferases superfamily protein		MELO3C003305	GSVIVT01024323001	
ndh-dependent cyclic electron flow 1 isoform 2			GSVIVT01015957001	
glycosyl transferase family 1 family protein		MELO3C023411	GSVIVT01009775001	
glycosyl transferase family 1 protein		MELO3C012336	GSVIVT01034635001	
glycosyltransferase family protein			GSVIVT01027020001	
glycosyl transferase family 1 family protein		MELO3C012953	GSVIVT01017132001	
gdp-mannose-dependent alpha-mannosyltransferase-like		MELO3C009186	GSVIVT01033998001	
probable galacturonosyltransferase 4-like		MELO3C007991	GSVIVT01006978001	
polygalacturonate 4-alpha-galacturonosyltransferase-like			GSVIVT01027574001	GSMUA_Achr9P24290_001
galacturonosyltransferase 8-like			GSVIVT01003922001	
probable galacturonosyltransferase 6-like isoform x1			GSVIVT01000744001	
probable galacturonosyltransferase 3-like			GSVIVT01023883001	
galacturonosyltransferase			GSVIVT01011500001	
probable galacturonosyltransferase 12-like		MELO3C025427	GSVIVT01014592001	
probable galacturonosyltransferase 11-like		MELO3C017247	GSVIVT01008143001	
probable galacturonosyltransferase 10-like			GSVIVT01030583001	
glycosyltransferase-like domain-containing protein 2-like			GSVIVT01037699001	GSMUA_Achr10P20100_001
glycosyltransferase family 61 protein		MELO3C019188	GSVIVT01033248001	
o-glucosyltransferase rumi homolog			GSVIVT01028972001	

o-glucosyltransferase rumi-like			GSVIVT01024736001	
udp-d-xylose:l-fucose alpha- -d-xylosyltransferase 2-like			GSVIVT01009666001	
probable galacturonosyltransferase-like 1-like			GSVIVT01019033001	
probable galacturonosyltransferase-like 7-like		MELO3C015949	GSVIVT01028072001	
beta- -n-acetylglucosaminyltransferase lunatic isoform 1		MELO3C008234	GSVIVT01016494001	
fringe-related family protein			GSVIVT01034119001	
PREDICTED: uncharacterized protein LOC101222721			GSVIVT01030631001	
nucleotide-diphospho-sugar transferase family protein		MELO3C011951	GSVIVT01005754001	
fringe-related family protein		MELO3C005701	GSVIVT01019400001	
transferring glycosyl		MELO3C026852	GSVIVT01025084001	
xyloglucanase 113 isoform 1			GSVIVT01036881001	
mannan endo- -beta-mannosidase 7-like	Solyc02g084990	MELO3C007842	GSVIVT01009746001	
			GSVIVT01018923001	
mannan endo- -beta-mannosidase 2-like		MELO3C013622	GSVIVT01020553001	
mannan endo- -beta-mannosidase 6-like		MELO3C019796	GSVIVT01030163001	
pectate lyase		MELO3C009891	GSVIVT01011348001	
pectate lyase			GSVIVT01020068001	
pectate lyase	Solyc05g014000		GSVIVT01000592001	GSMUA_Achr6P28260_001
			GSVIVT01007582001	
pectate lyase	Solyc03g111690			
probable pectate lyase 4-like		MELO3C003093	GSVIVT01005974001	
			GSVIVT01025593001	
probable pectate lyase 8-like			GSVIVT01028548001	
probable pectate lyase 15-like		MELO3C023087		
major pollen allergen-like protein		MELO3C001958		

probable pectate lyase 13-like		MELO3C019840	GSVIVT01016208001	
protein notum homolog			GSVIVT01031342001	GSMUA_Achr5P18410_001
notum-like protein		MELO3C006917	GSVIVT01028921001	GSMUA_Achr11P03970_001
protein notum homolog			GSVIVT01036914001	
protein notum homolog			GSVIVT01034178001	
Pectinacetylesterase	Solyc08g005800			
pectinesterase pectinesterase inhibitor ppe8b-like	Solyc07g017600		GSVIVT01023135001	
pectinesterase pectinesterase inhibitor 24-like			GSVIVT01016667001	
pectinesterase 3-like		MELO3C022701	GSVIVT01027356001	
probable pectinesterase pectinesterase inhibitor 7-like		MELO3C015962	GSVIVT01028040001	
pectinesterase 2-like			GSVIVT01037349001	
probable pectinesterase pectinesterase inhibitor 34-like			GSVIVT01038546001	
pectinesterase pectinesterase inhibitor 22-like			GSVIVT01031164001	
probable pectinesterase 53-like		MELO3C015419	GSVIVT01015290001	
probable pectinesterase pectinesterase inhibitor 25-like			GSVIVT01027359001	
pectinesterase pectinesterase inhibitor-like		MELO3C007375	GSVIVT01012639001	
pectinesterase pectinesterase inhibitor 28-like			GSVIVT01017687001	
probable pectinesterase pectinesterase inhibitor 40-like	Solyc06g009190		GSVIVT01028041001	
21 kda			GSVIVT01018598001	
21 kda		MELO3C006821	GSVIVT01018606001	
pectinesterase pectinesterase inhibitor 28-like			GSVIVT01031165001	
Pectinesterase	Solyc07g042390			
probable pectinesterase 67-like		MELO3C006718	GSVIVT01034985001	
probable pectinesterase 29-like			GSVIVT01000525001	

probable pectinesterase 8-like			**GSVIVT01015852001**	
pectinesterase 31-like			**GSVIVT01011459001**	
Pectinesterase	**Solyc12g009270**			
probable pectinesterase 68-like		**MELO3C005532**	**GSVIVT01001327001**	**GSMUA_Achr1P01520_001**
probable pectinesterase 53-like			**GSVIVT01028170001**	
pectinesterase ppme1-like			**GSVIVT01013827001**	
pectinesterase ppme1-like		**MELO3C015550**	**GSVIVT01015022001**	
probable pectinesterase pectinesterase inhibitor 51-like			**GSVIVT01026518001**	
pectinesterase 11-like		**MELO3C010953**	**GSVIVT01009729001**	
			GSVIVT01023786001	
pollen-specific protein c13-like			**GSVIVT01008996001**	**GSMUA_Achr11P13000_001**
olee1-like		**MELO3C005385**	**GSVIVT01021211001**	
		MELO3C007213		
pollen ole e 1 allergen and extensin family protein			**GSVIVT01036543001**	
uncharacterized loc101212998			**GSVIVT01037612001**	
probable polygalacturonase-like			**GSVIVT01015044001**	
polygalacturonase at1g48100-like			**GSVIVT01028189001**	
			GSVIVT01031298001	
polygalacturonase at1g48100-like			**GSVIVT01017644001**	
exopolygalacturonase clone gbge184		**MELO3C006686**	**GSVIVT01028044001**	
			GSVIVT01035103001	
pectin lyase-like superfamily			**GSVIVT01012232001**	
polygalacturonase at1g48100-like			**GSVIVT01013800001**	
probable polygalacturonase at1g80170-like			**GSVIVT01017279001**	
polygalacturonase			**GSVIVT01013771001**	
pectin lyase-like superfamily protein			**GSVIVT01033349001**	**GSMUA_Achr3P11090_001**

Polygalacturonase	**Solyc09g075460**			
probable pectinesterase pectinesterase inhibitor 34-like			**GSVIVT01038546001**	
protein vernalization insensitive 3-like		**MELO3C004546**	**GSVIVT01015423001**	
protein vernalization insensitive 3-like		**MELO3C004546**	**GSVIVT01035732001**	
Xyloglucan endotransglucosylase/hydrolase	**Solyc02g080160**			
probable xyloglucan endotransglucosylase hydrolase protein 8-like	**Solyc03g031800**		**GSVIVT01012635001**	
probable xyloglucan endotransglucosylase hydrolase protein 28-like		**MELO3C005245**	**GSVIVT01011603001**	
xyloglucan endotransglucosylase hydrolase protein 33-like	**Solyc05g005680**	**MELO3C012004**	**GSVIVT01020228001**	
probable xyloglucan endotransglucosylase hydrolase protein 23-like			**GSVIVT01029162001**	
xyloglucan endo-1 family protein		**MELO3C017481**		
xyloglucan endotransglucosylase hydrolase protein 9-like			**GSVIVT01029372001**	
probable xyloglucan endotransglucosylase hydrolase protein 10-like			**GSVIVT01037018001**	

9.7 List of expression gene in banana

Locus ID	Function	Expression level 60 days after flowering	Expression level 90 days after flowering
GSMUA_Achr6P18730_001	GSMUA_Achr6T18730_001~ Beta-galactosidase 3~ Os01g0875500~ complete	3	3
GSMUA_Achr9P16910_001	GSMUA_Achr9T16910_001~ Endoglucanase 9~ GLU1~ complete	9	4
GSMUA_Achr6P24820_001	GSMUA_Achr6T24820_001~ expressed protein~ At1g28695~ complete	10	10
GSMUA_Achr5P07110_001	GSMUA_Achr5T07110_001~ pectinacetylesterase domain containing protein, expressed~ Notum~ complete	4	3
GSMUA_Achr5P19880_001	GSMUA_Achr5P19880_001~ Beta-fructofuranosidase 1~ MAVIN2~ missing_completeness	26	21
GSMUA_Achr10P04520_001	GSMUA_Achr10T04520_001~ Probable pectinesterase/pectinesterase inhibitor 51~ PME51~ complete	22	20
GSMUA_Achr5P12120_001	GSMUA_Achr5T12120_001~ Probable beta-D-xylosidase 6~ BXL6~ complete	53	25
GSMUA_Achr2P02550_001	GSMUA_Achr2T02550_001~ Polygalacturonase QRT3~ QRT3~ complete	9	23
GSMUA_Achr9P16420_001	GSMUA_Achr9T16420_001~ Beta-glucosidase 22~ BGLU22~ complete	140	262
GSMUA_Achr9P12430_001	GSMUA_Achr9T12430_001~ Putative Polygalacturonase inhibitor~ PGIP~ complete	4	5
GSMUA_Achr4P19460_001	GSMUA_Achr4T19460_001~ Probable galacturonosyltransferase 4~ GAUT4~ modules	5	9
GSMUA_Achr3P08780_001	GSMUA_Achr3T08780_001~ Putative 4-alpha-glucanotransferase~ malQ~ modules	24	32
GSMUA_Achr6P26410_001	GSMUA_Achr6T26410_001~ Probable pectinesterase 53~ PME53~ complete	5	1
GSMUA_Achr6P31810_001	GSMUA_Achr6T31810_001~ Cellulose synthase A catalytic subunit 4 [UDP-forming]~ CESA4~ complete	1	1
GSMUA_Achr4P32790_001	GSMUA_Achr4T32790_001~ X8 domain containing protein, expressed~ At5g61130~ complete	57	40
GSMUA_Achr10P20100_001	GSMUA_Achr10T20100_001~ glycosyltransferase, putative, expressed~ Ago61~	119	43

	complete		
GSMUA_Achr3P13670_001	GSMUA_Achr3T13670_001~ Probable polygalacturonase At1g80170~ At1g80170~ complete	18	38
GSMUA_Achr9P18190_001	GSMUA_Achr9T18190_001~ Putative Probable polygalacturonase~ GSVIVT00026920001~ complete	6	12
GSMUA_Achr3P03120_001	GSMUA_Achr3T03120_001~ Putative Major pollen allergen~ uvrA~ modules	10	14
GSMUA_Achr9P15830_001	GSMUA_Achr9T15830_001~ Probable xyloglucan endotransglucosylase/hydrolase protein 28~ XTH28~ complete	5	1
GSMUA_Achr8P24350_001	GSMUA_Achr8T24350_001~ Beta-glucosidase 24~ BGLU24~ modules	2	2
GSMUA_Achr1P17440_001	GSMUA_Achr1T17440_001~ Beta-amylase 7~ BAM7~ complete	1	1
GSMUA_Achr3P11090_001	GSMUA_Achr3T11090_001~ polygalacturonase, putative, expressed~ pehX~ complete	23	21
GSMUA_Achr6P35850_001	GSMUA_Achr6T35850_001~ Beta-galactosidase 2~ Os01g0580200~ complete	2	1
GSMUA_Achr6P32940_001	GSMUA_Achr6T32940_001~ Arabinogalactan protein 2~ FLA11~ complete	1	1
GSMUA_Achr6P08940_001	GSMUA_Achr6T08940_001~ glycosyl transferase, group 1 domain containing protein, expressed~ lpsB~ complete	5	3
GSMUA_Achr10P30870_001	GSMUA_Achr10T30870_001~ Xyloglucan endotransglucosylase/hydrolase protein 22~ XTH22~ complete	12	17
GSMUA_Achr1P27410_001	GSMUA_Achr1T27410_001~ Callose synthase 12~ CALS12~ complete	2	1
GSMUA_Achr1P01520_001	GSMUA_Achr1T01520_001~ Probable pectinesterase 68~ PME68~ complete	78	118
GSMUA_Achr7P04580_001	GSMUA_Achr7T04580_001~ Probable pectate lyase 15~ At4g13710~ complete	17	6
GSMUA_Achr11P13000_001	GSMUA_Achr11T13000_001~ Putative Pollen-specific protein C13~ MGS1~ complete	47	61
GSMUA_Achr3P15900_001	GSMUA_Achr3T15900_001~ Hypothetical protein~ unknown_gene~ missing_functional_completeness	1	1
GSMUA_Achr5P06050_001	GSMUA_Achr5T06050_001~ Probable cellulose synthase A catalytic subunit 1 [UDP-forming]~ CESA1~ complete	32	23

GSMUA_Achr1P10090_001	GSMUA_Achr1T10090_001~ expressed protein~ waaU~ complete	8	5
GSMUA_Achr9P24290_001	GSMUA_Achr9T24290_001~ Alpha-1,4-galacturonosyltransferase 1~ GAUT1~ complete	31	29
GSMUA_Achr2P19670_001	GSMUA_Achr2T19670_001~ Putative glycosyl transferase, group 1 domain containing protein, expressed~ ytcC~ modules	3	2
GSMUA_Achr6P33210_001	GSMUA_Achr6T33210_001~ Putative Probable pectate lyase 4~ At1g30350~ fragment	5	4
GSMUA_Achr1P15040_001	GSMUA_Achr1T15040_001~ Callose synthase 3~ CALS3~ fragment	2	1
GSMUA_Achr6P11120_001	GSMUA_Achr6T11120_001~ Putative Asparagine-linked glycosylation protein 11 homolog~ ALG11~ complete	9	9
GSMUA_Achr6P15100_001	GSMUA_Achr6T15100_001~ Putative Periplasmic beta-glucosidase~ bglX~ fragment	1	1
GSMUA_Achr3P08140_001	GSMUA_Achr3T08140_001~ Mannosylglycoprotein endo-beta-mannosidase~ EBM~ complete	5	4
GSMUA_Achr2P05580_001	GSMUA_Achr2T05580_001~ Cellulose synthase-like protein D4~ CSLD4~ modules	5	4
GSMUA_Achr4P20830_001	GSMUA_Achr4T20830_001~ Beta-galactosidase 6~ Os03g0255100~ complete	6	3
GSMUA_Achr5P09680_001	GSMUA_Achr5T09680_001~ Beta-galactosidase 8~ Os05g0539400~ complete	3	1
GSMUA_Achr5P03530_001	GSMUA_Achr5T03530_001~ Putative Alpha-1,3-mannosyltransferase ALG2~ ALG2~ complete	7	8
GSMUA_Achr8P27950_001	GSMUA_Achr8T27950_001~ Arabinogalactan peptide 16~ AGP16~ complete	54	50
GSMUA_Achr11P20220_001	GSMUA_Achr11T20220_001~ Putative Beta-galactosidase~ lacZ~ complete	5	4
GSMUA_Achr7P15620_001	GSMUA_Achr7T15620_001~ Pectinesterase/pectinesterase inhibitor PPE8B~ PME32~ complete	7	20
GSMUA_Achr10P16790_001	GSMUA_Achr10T16790_001~ glycosyl transferase family 17 protein, putative, expressed~ Mgat3~ complete	52	7
GSMUA_Achr6P27800_001	GSMUA_Achr6T27800_001~ Probable xyloglucan glycosyltransferase 6~ CSLC6~ complete	3	2
GSMUA_Achr6P28260_001	GSMUA_Achr6T28260_001~ Probable pectate lyase 22~ At5g63180~ complete	1061	1092
GSMUA_Achr11P07300_001	GSMUA_Achr11T07300_001~ Probable galacturonosyltransferase 3~ GAUT3~ complete	2	2

GSMUA_Achr11P05430_001	GSMUA_Achr11T05430_001~ Pectinesterase 3~ PECS-1.1~ complete	10	12
GSMUA_Achr5P18410_001	GSMUA_Achr5T18410_001~ pectinacetylesterase domain containing protein, expressed~ Notum~ complete	37	61
GSMUA_Achr5P07480_001	GSMUA_Achr5T07480_001~ Expansin-A2~ EXPA2~ complete	80	60
GSMUA_Achr3P06980_001	GSMUA_Achr3T06980_001~ Glycoprotein 3-alpha-L-fucosyltransferase A~ FUT11~ complete	2	1
GSMUA_Achr2P05240_001	GSMUA_Achr2T05240_001~ exostosin family domain containing protein, expressed~ At5g25310~ complete	6	3
GSMUA_Achr11P06330_001	GSMUA_Achr11T06330_001~ Putative Alpha-glucosidase 2~ GAA~ modules	30	28
GSMUA_Achr4P01640_001	GSMUA_Achr4T01640_001~ Fibronectin type III domain containing protein, expressed~ VIN3~ complete	3	3
GSMUA_Achr11P03970_001	GSMUA_Achr11T03970_001~ Putative pectinacetylesterase domain containing protein, expressed~ Notum~ complete	16	34
GSMUA_Achr3P00830_001	GSMUA_Achr3T00830_001~ Putative Predicted protein~ TPRP-F1~ complete	2	1
GSMUA_Achr1P11850_001	GSMUA_Achr1T11850_001~ Probable galacturonosyltransferase 6~ GAUT6~ complete	4	4
GSMUA_Achr5P11660_001	GSMUA_Achr5T11660_001~ Callose synthase 9~ CALS9~ complete	2	2
GSMUA_Achr4P04690_001	GSMUA_Achr4T04690_001~ 1,4-alpha-glucan-branching enzyme, chloroplastic/amyloplastic~ SBE1~ complete	15	15
GSMUA_Achr11P02870_001	GSMUA_Achr11T02870_001~ Probable polygalacturonase~ GSVIVT00026920001~ complete	14	11
GSMUA_Achr4P02940_001	GSMUA_Achr4T02940_001~ Mannan endo-1,4-beta-mannosidase 1~ MAN1~ complete	6	3
GSMUA_Achr7P02230_001	GSMUA_Achr7T02230_001~ Putative Alpha-mannosidase 2~ MAN2A1~ complete	4	5
GSMUA_Achr10P09360_001	GSMUA_Achr10T09360_001~ glycosyl transferase, group 1 domain containing protein, expressed~ lpsB~ complete	2	1
GSMUA_Achr7P26770_001	GSMUA_Achr7T26770_001~ Pectinesterase QRT1~ QRT1~ complete	2	1
GSMUA_Achr1P04970_001	GSMUA_Achr1T04970_001~ Mannan endo-1,4-beta-mannosidase 2~ MAN2~ complete	1	1

GSMUA_Achr5P08790_001	GSMUA_Achr5T08790_001~ Putative Lysosomal beta glucosidase~ gluA~ fragment	5	1
GSMUA_Achr2P21230_001	GSMUA_Achr2T21230_001~ Beta-galactosidase 15~ Os12g0429200~ complete	4	3
GSMUA_Achr6P22080_001	GSMUA_Achr6T22080_001~ Putative Polygalacturonase At1g48100~ At1g48100~ complete	1	1
GSMUA_Achr1P15030_001	GSMUA_Achr1T15030_001~ Callose synthase 3~ CALS3~ fragment	24	17
GSMUA_Achr9P28420_001	GSMUA_Achr9T28420_001~ arabinogalactan protein, putative, expressed~ SPAC29B12.11c~ complete	64	112
GSMUA_Achr1P22480_001	GSMUA_Achr1T22480_001~ expressed protein~ xynB~ fragment	5	2
GSMUA_Achr6P12000_001	GSMUA_Achr6T12000_001~ Galacturonosyltransferase 8~ GAUT8~ complete	9	6
GSMUA_Achr2P00950_001	GSMUA_Achr2T00950_001~ expressed protein~ PCDHB9~ complete	8	17
GSMUA_Achr2P06020_001	GSMUA_Achr2T06020_001~ glycosyl hydrolase family 5 protein, putative, expressed~ Acel_0614~ complete	35	51
GSMUA_Achr11P18030_001	GSMUA_Achr11T18030_001~ Putative Probable polygalacturonase~ GSVIVT00026920001~ modules	3	1
GSMUA_Achr11P22060_001	GSMUA_Achr11T22060_001~ Putative alpha-L-fucosidase 1~ Os04g0560400~ complete	3	1
GSMUA_Achr3P28000_001	GSMUA_Achr3T28000_001~ Probable xyloglucan glycosyltransferase 5~ CSLC5~ complete	5	2
GSMUA_Achr6P23600_001	GSMUA_Achr6T23600_001~ fringe-related protein, putative, expressed~ B3GALTL~ modules	2	1
GSMUA_Achr1P12500_001	GSMUA_Achr1T12500_001~ Putative Endo-1,4-beta-xylanase C~ xlnC~ modules	74	13

No Expression (0)
Extremely low expression (1)
Very low expression (2 - 5)
Low expression (6 - 10)
Moderate expression (11 - 25)
Moderately high expression (26 - 100)
High expression (101 - 500)
Very high expression (501 - 5000)
Extremely high expression (> 5000)

9.8 List of primer

NPTII/Cas9F	ACGTGGATCATATCGTGCCC
NPTII/Cas9R	GGTGGCCTTGCCTATTTCCT
PL01F	AAACGATAGACGGAAGGGGC
PL01R	CCGTCCGAAACAGTTCTCCA
CESA01F	TTTGGCATTGCAGAGGGAGT
CESA01R	ATCAGTGGCCTACCAAAGCA
PL02_fwd	ACGCGTTTAATTGGACGTATG
PL02_rev	AACAGAGGAAGTGCCCATTG
PL02_probe	TTGAGTTTGAGTGCAAGGCCGTC
PG2A001_fwd	AAGACTTGGCAGGGAGGATC
PG2A001_rev	TATGGCCACCTTTGTTGCAC
PG2A001_prob	TGGACAAGCTAGCAACATCA
PSY_fwd	GGCCATTTGACATGCTCGAT
PSY_rev	CGCTCTCTGTTGTTGCCTTT
PSY_prob	TGGTACGGTTGGGTTGATGA
ELF-fwd	ACCTTTGCTGAATACCCTCCATTG
ELF-rev	CACAGTTCACTTCCCCTTCTTCTG
ELF-probe	TCGTTTTGCTGTGAGGGACATGAGGCA
Cas902F	GACAATGGAAGCATCCCCCA
Cas902R	CAGTTTTCTTGACAGCCGCC

9.9 Sequence of CRISPR::*PG/PL/PSY1*

PSY1 GENOMIC MARKED INTRONS

GATTTCACTATATTGTAATATTAACTTGAGGTCACTATAGGAGCTCAAAAACTTCTAATTTTGAATCAATGTCTGGTTATACTTTT
TTTGTCATAACTGTATCTCAAATGTGGTGTTTGGTTTATCTCATTTTGCAGAAGTCAAGAAACAGGTTACTCCTGTTTGAGTGAGG
AAAAGTTGGTTTGCCTGTCTGTGGTCTTTTTATAATCTTTTTCTACAGAAGAGAAAGTGGGTAATTTTGTTTGAGAGTGGAAATAT
TCTCTAGTGGGAATCTACTAGGAGTAATTTATTTTCTATAAACTAAGTAAAGTTTGGAAGGTGACAAAAAGAAAGACAAAAATCTT
GGAATTGTTTTAGACAACCAAGGTTTTCTTGCTCAGAATGTCTGTTGCCTTGTTATGGGTTGTTTCTCCTTGTGACGTCTCAAATG
GGACAAGTTTCATGGAATCAGTCCGGGAGGGAAACCGTTTTTTTGATTCATCGAGGCATAGGAATTTGGTGTCCAATGAGAGAATC
AATAGAGGTGGTGGAAAGCAAACTAATAATGGACGGAAATTTTCTGTACGGTCTGCTATTTTGGCTACTCCATCTGGAGAACGGAC
GATGACATCGGAACAGATGGTCTATGATGTGGTTTTGAGGCAGGCAGCCTTGGTGAAGAGGCAACTGAGATCTACCAATGAGTTAG
AAGTGAAGCCGGATATACCTATTCCGGGGAATTTGGGCTTGTTGAGTGAAGCATATGATAGGTGTGGTGAAGTATGTGCAGAGTAT
GCAAAGACGTTTAACTTAGGTTAGCTTCTTCAATCTATTCATTCGTTTACCAAATATTATTTGGTAAGCACTAATTATGAATATAT
ATATGTTCATGTTATTGATGAAGACAAAATTTGATCTTTGTTTGTTTATTCAGGAACTATGCTAATGACTCCCGAGAGAAGAAGGG
CTATCTGGGCAATATATGGTGAGGTTTCTAGCCATTTAATAACAGTTACGCGCACAAACACATATGATTAATCGGGGACGAGAAAA
AAAGAAATGAAGTTTGAGTTTTGAGGGTCATATGTAATAGGTAAATCCGAGCTTGACTAGCTTGAGATGTTTATTGTCATATCATG
CTCAATACTAACCAAAACACTGAAAAAGAACTTGATTATATTTACATACTAATATTTTCATTTGCGTTGCTGTTCACATTTTTACC
TATGGAACTGGTTTTTGTGATTTGTTATACTTCATATTCGATGTTAATAAAATATATCATTCCTCCCTTTTTCTCCACTTCAAGCT
TTACTGTAGTGTTGAAAGGGGAAACTCCTTTTAATGATTGCATATATAAACGAACTTCTTGAGTTGAATAGTTTCTCATTATGATC
TGTTTAAACAGTATGGTGCAGAAGAACAGATGAACTTGTTGATGGCCCAAACGCATCATATATTACCCCGGCAGCCTTAGATAGGT
GGGAAAATAGGCTAGAAGATGTTTTCAATGGGCGGCCATTTGACATGCTCGATGGTGCTTTGTCCGATACAGTTTCTAACTTTCCA
GTTGATATTCAGGTTAGTCTACCAATTCTATGGTCTTTATATTTGTTCAATTTGCGTTTGATGTCACTTTTGCTGAGGGCTTTTCT
AATAGCTTACTTCAGCCTAGCGGAAATGTTTGTAGTTGAATCTCTAGTTCTGTCTCCTATATCTGTTTCTCTCGTCCTAGATACTA
CACATACTTCATTTCTGTTTTAACATTTTATTCGTCTTTTGGTGTTGTTTTGTATGTGAATCATATATTTGGAACAGAATCATTAT
TAGTTCACATGATTTCATTTGCTTTCTTCAATAGCGTAATTGTCTAACCTTCCAATATATGTTGCAGCCATTCAGAGATATGATTG
AAGGAATGCGTATGGACTTGAGAAAATCGAGATACAAAAACTTCGACGAACTATACCTTTATTGTTATTATGTTGCTGGTACGGTT
GGGTTGATGAGTGTTCCAATTATGGGTATCGCCCCTGAATCAAAGGCAACAACAGAGAGCGTATATAATGCTGCTTTGGCTCTGGG
GATCGCAAATCAATTAACTAACATACTCAGAGATGTTGGAGAAGAGTAAGTACAAAGCTGTGTTTTACGCACATAATTTTTTTTGC
TAATATTTACATATCAAAATATAGGAAAATGAGCTCTTCGGTTATCCGGTTTATATTTTTTTTATGTCAACATAATAGTATAAAGT
AATTAGTATCAGTCGTTCTGGGAATAAAATTGCAGAACTCAATTTAGCCGTGTTGTGAAATCCTGCTTGTTTTGAGAGCTTAAAGC
TCATTAGTTAGTCGTTAGAGACGAAGAAATTCTTCGTTGTCCATCTTTATTCCACCTTAAAGTTGTGATATTTTCATTATTGGTAC
ATTTGGCAAAAACACCTGAACAAATTTATGACGGATGCCTTTTGAAAGTCACTATACCTGTCTAGTCGGCGTTTATCACATTTCTT
TGACATATTGAACTTTGAAACATGATATCAGCTCTAGACAGTGACGAGCCATGATCAATTTCTTTCCTTTATTCTTTCTTTGGAAG
TGCCGTATTTAGGCTTCCGTTGTTCTTATATATTGCTTTCCCTGCAGTGCCAGAAGAGGAAGAGTCTACTTGCCTCAAGATGAATT
AGCACAGGCAGGTCTATCCGATGAAGATATATTTGCTGGAAGGGTGACCGATAAATGGAGAATCTTTATGAAGAAACAAATACATA
GGGCAAGAAAGTTCTTTGATGAGGCAGAGAAAGGCGTGACAGAATTGAGCTCAGCTAGTAGATTCCCTGTAAGCATTCGTAAACTC
TTTAGTTTTATGAAATGATTCTTTTTTCGCGTTATTAGATGAATATGGTTGCTTGTGTTGAGTATTTCTAGGTCGATGAAGTTGAG
ACAAGGGTTTTTAAGTTTTAACGACTTTTACGGGGTGCCATGTTATCTGCTACCTAATCTTAGGTAGTTGACCGGAAGGGCTAGAA
TTTTAACCTCATGTTCACCCTACCAACCAAGAAATGAACCTCGCATAGAGCTCGTAGTTATGAATATTTGCTTTGGCATGACATTG
TGCGGATCATGAAATGTCTTAGATTATATGGAAAAATCATTCTATTACATCGAATAGATACATTAGATCTAAGAAGCACGCCGTGT
TGTAAATGAGAAATTCTATAGCTCAGATCTTTAGTTTTCTCTGAACGACCTACAAACCAACGGATAACCTTGTATTGAGCTTGTCG
TTCTCAGTATTTGCACTAACATTACGTCGTGTGGATCCTGAAATGGCTTGGATTGCTATTATTCTGGATATGGCAAAACCATTTTA
TTAGTACTAGATATCGAATAACTACATTTGACCCTACAAGTACCCTGGGTTGGAGTTACAATATCCCATACCTCGTATCTTTAGTG
TTCTCTTATTTATCACCTTTGTCTACTATTCTGGCAAAATAACCTCACTCGTTACTCGGTGTTTTCCAGGTATGGGCATCTTTGGT
CTTGTACCGCAAAATACTAGATGAGATTGAAGCCAATGACTACAACAACTTCACAAAGAGAGCATATGTGAGCAAATCAAAGAAGT
TGATTGCATTACCTATTGCATATGCAAAATCTCTTGTGCCTCCTACAAAAACTGCCTCTCTTCAAAGATAAAGCATGAAATGAAGA
TATATATATATATATATATAGCAATATACATTAGAAGAAAAAAAGGAAGAAGAAATGTTGTTGTATTGATATAAATGTATATCATA
AATATTAGGTTGTAGTAACATTCAATATAATTATCTCTTGTAGTTGTTGTATCTTCACTTTATCTCAACTCCTTTGAGAGAACTTT
CCGTAGTTATCTGCTTTGCACTTGGTTACTCAGAATTTTACTGTGGGCATGATAATTGATATACCAAATTCAGTTTTGATTCTATC
GAAAAATTTGTTATTACATTTTTTTGGGGGGAAAGGAA

PG GENOMIC MARKED INTRONS

AATCTTTTTCAATAGACAAGTTTAAAAACCATACCATATAACAATATATCATGGTTATCCAAAGGAATAGTATTCT
CCTTCTCATTATTATTTTTGCTTCATCAATTTCAACTTGTAGAAGCAATGTTATTGATGACAATTTATTCAAACAA
GTTTATGATAATATTCTTGAACAAGAATTTGCTCATGATTTTCAAGCTTATCTTTCTTATTTGAGCAAAAATATTG
AAAGCAACAATAATATTGACAAGGTTGATAAAAATGGGATTAAAGTGATTAATGTACTTAGCTTTGGAGCTAA
GGGTGATGGAAAAACATATGATAATATTGTAAGTATTTAAATATTGGAATATATTTGTGGGGATGAAAATGATAGA
GAATATAAGAATTATTTGGAAGGATGAAAAGTTATATTTTATAAAGTAGAAAATTATTTTCTCGTTTTTAGTATTA
AGGTGAAAATGAGTTTCTCGTTAAGCGAGGAAAAGCTATTTTCCATGGTAACTGTATTTTTTTTTTACTTTTADAT
AACGTCATAGTATTTGCTATACTCAAGAATAAGACACTTATTATTGATGATTTAGTGCTCGAAAAGAAATTGATAG
TAATTTTGCTTAATATAACTATCAATTTCTTATATGTATATTTTTCAACCAAAATAACAAAGCGTAATCCAATAAG
TGGGCCTCTAGAATAAAGAGTAAGTTCTATTCAATTCTTAACCTTATTTAATTTTAGTGGAAACCTCGACAAAAAC
GAACAAACGTATTCAAACTTTTATATTCGGAATTCGAGACCAACCATATGAACAACCTCACACATGCATATAGTCC
TAATATATATAATTTTTCTAAAAAATATCTTCAATCTACCATATTGAAATATTGAAAAATGACTTTTATCCTATCG
AACACATAATCAAGAGTTTCTTTTAAGAATTTACCACTACATTTGGTATGTTTCTTATCGTGTTAAAATTATCTTT
CAGGCATTTGAGCAAGCATGGAATGAAGCATGTTCATCTAGAACACCTGTTCAATTTGTGGTTCCTAAAAACAAGA
ATTATCTTCTCAAGCAAATCACCTTTTCAGGTCCATGCAGATCTTCTATTTCAGTAAAGGTTAGCATATTGATTAT
TTATATCCTCTTTGTTAGCAATATATTATCTGGTTTATGACAAAATTTAAGAAAGTAATCAAAGATAGATAAACAA
TGAATTTTCGTCACTAATTTAGCGGATTAGTGAGGAATTATCAAAATGTTATGTTAGCTATGAGCAACTTAGCTAT
GAATTAGCTAGTGAAGAAGTTTGATGCTAATTCTATTTTTTTTTTGTAGAGTAAAGATATTTGAAACACATGTATT
AATTATTAATTATGTCTTAATTAATATGTCAATGGATAGTTCAAACTAAAGAACTGTCAAAAGAAAATAAGAAAGA
AATATTTATTTTTAAAATAAATTAAAAAGAAAAATATGAGAAATAAATTCAAAGCGAGAAGGTATTACATAATCTA
TGGGGATAAAAGGATATTATATATGTAAGAAAACAGCACTACACATATCTAATAAAGTCTCATAAATGGATATAAA
AAATAGTGTGTAAGCAACAGTTATCCCTACAAAAACTTTTGTGGGGTAGATCGATCCAGAGGTTGTTTCCAGACTC
TTGCTTAAAAAAAATGTTTTTTCTAAATAAGTTTGAAAGAAATGTTATATGATGAAAATATGAAGAAAAACATATC
AATATTAAAAATAATAAAGTAATCAAAGTAAACGAAATAACAATAGGAATAATACTCATAAATGAAAATTTAGTGG
CTTTTCGTTAACATAATCTTAGTTTATTCATTGTTTCTTTAATTTCCCTTCTTATTTTTTTTGAAATTACTAATGC
AGATTTTTGGATCCTTAGAAGCATCTAGTAAAATTTCAGACTACAAAGATAGAAGGCTTTGGATTGCTTTTGATAG
TGTTCAAAATTTAGTTGTTGGAGGAGGAGGAACTATCAATGGCAATGGACAAGTATGGTGGCCAAGTTCTTGCAAA
ATAAATAAATCACTGGTAATTTTATAACCTTGCTTATAAGTTTTACGCTATGTTGCTCGAATTCTTTAAACTTGTT
CTAAAGATATTATATATTTGAAGGAGGTGTCACAAATGCATCACATTTTTAGAGATTCCGACCAATATTAGTTTTA
TGTAATCTAATTTTCAGAGCATCTTTGCCTTGTACTGATCATTGTTACCCTTTTTTTCTTCATGCAGCCATGCAGG
GATGCACCAACGGTACGTTAATTGCATTTGATTTGATAAAAAAAAAAAAGCCTAAAATATATTTGAATTTTAATTGA
AAGGTTATAATAATTCTTAACTTTGGGCAGGACCTATTACCCCTTGCACTATTTAATAGTGTATTTTAAAGATATA
AAAGTGTTTAGTTGAAACAAAAATTTAGATATTCAAAAACTATTTGAAAATTACTATAAATTGCAATTTTTTTGCA
TATCAATATGATTAAAAAATATTAGTTAAAGTTCTTATGATTTGATTCTAAAAATAAAAATCATGACAAACAATAG
TAGACGGAGAAAGTATATAACAATACCTCTTCAAGTAGAATCGATTTGTACACACACCTCAAAACCTACGTTTTCT
TCGATTTATATTTCCTATTTCTTTTAATAGTAATCAAAGGCTATTAGTTCTGTCAAAATCTATACATTGGAAACTC
TATCTTTGACGCCTCGTACATTCGAGATCGTTGAACAATGGATGAATGATTATTTAACTTTGTATTTAAATATTAA
AACTAATATTGTTTAATTTTCAGGCCTTAACCTTCTGGAATTGCAAAAATTTGAAAGTGAATAATCTAAAGAGTAA
AAATGCACAACAAATTCATATCAAATTTGAGTCATGCACTAATGTTGTAGCTTCAAATTTGATGATCAATGCTTCA
GCAAAGAGCCCAAATACTGATGGAGTCCATGTATCAAATACTCAATATATTCAAATATCTGATACTATTATTGGAA
CAGGTTTATTTATTTAATTTTTATTTATCCAATTTAATTAGAAAAAAAAAGGAGTATTTTTATTTGATAACTAAAT
TATTAATTTTTAATTTTTTTTTTATAGGTGATGATTGTATTTCAATTGTTTCTGGATCTCAAAATGTGCAGGCCACA
AATATTACTTGTGGTCCAGGTCATGGTATAAGGTACTCTATTTTACAAATATACTTGTTTCCATTTTCTCTATTTC
ATAAAAGGTAGTATGATATAATAATTACTTTAAATCCTTTAATTAATTTATTGGCAAATTTTTTCTCTTGTCTTTA
TGGTTAATGACTTAGCACAATAATTAGGGCCGTTTGGATGGGCGAATAAAAGCAGCTTTAAAAAAGTACTTTTAAA
AGTGTTGAAACTTATTTTTAAAATAAGCAGTTATCGGTTTGGATAAAAGTGCTGAAGTTGTTATGTCAAACGTGAA
AAGGGAAAAATGGAAGAAAGAAATGTTAGGGTTATATGGGTTATTTGTATAAAAATATTAAGCACAAAAAGATAAA
AATGTGGTCAACTTAAAACAACTTATAAGCTACCCTACCCTACCCCAGCTTTTAACTTTTGGCTTAAAATAAGTTT
TTTTTTTTAAAACTTAAAATAAGTTGTTTTGAGTATTGCCAAAGAGCTAAATAATGCAAAAACCAGCTTTTAAGTC
AGTTTGACCAGCTTTTAAGCTGAGCCAAACAGGCTCTTAAAATGTCTGCTTAGATGTGCTATATATATTTGAGCTT
TTTTTGAAGTAGTATATTATCCTTAAGTTCAACATAAAATACATGCTTTAACATAGCACATATAGTTAATCAAAAG
ACGAAATGATGAATAATTTTGCGAATTTGATTATTCACAAGAAAAGGGATAGTTCAAAGTGTACATTTCAATGAAT
TGAAGATATCATAAAGACTAAAATTAGAAGAATCAATAATTGAGGGATCAAAAATGTTATTACCTTATTAAAATAC
TATTCCATTTTCATATTAAATTAACTAATTAAGAGTGTTTTATAATCTAATAAAACATGCAATAATTATTGACGAA

ATGTGGTTTTGGTACCTATAATCTTTCTGAATATTTGCTCTATTTTTTCTCTTTTTATTTTTCCATGGATTACTAT
TGGAAGCTTAGGATCTGGAAATTCAGAAGCTTATGTGTCTAATGTTACTGTAAATGAAGCCAAAATTATCGGTGCC
GAAAATGGAGTTAGGATCAAGACTTGGCAGGTACCCTCCCCCCCCCCCCCCCCCCCACAGGCCCATTTTTTTAATT
TTTTTTAAATTTTTATTCGAATATCAATATTAAAGATTAATTTGATTTCATGTTTGAAATTTATATTTGGATAAAG
TATGTATTTTACTAGCTTTCTATGTTATATAGAAAAAAAAATGTTCAGAACTTCAGATTATTGTACTCGTACTAAG
TGTAAATGTGTTGCTTTGTTTAGAAGTTTGGTTTATCCAGTTTTGGGTCATGATTAAACCAAACTTATAATGAAAA
GGGGCTGCAACGGCCGGCCCACTAGTGCTAGTATCAATAGGAAGATCTCACGTCTGTTTATTCAGATGGACGTTCT
TGGTTGAATGTTAATAATTATAAATTTAATTAACATGTAATTAAGCATTATATAAATTAATGTGGTTTAATAATGT
AGGGAGGATCTGGACAAGCTAGCAACATCAAATTTCTGAATGTGGAAATGCAAGACGTTAAGTATCCCATAATTAT
AGACCAAAACTATTGTGATCGAGTTGAACCATGTATACAACAGGTAATTTTTTATTAACGAACAATTTATTATATT
TTATTACTTCTTAAATCACCTTACATCATTAAAACTTTGAGATTCTTTTCACTAGTTAGTAACTTTTTGAATAGAT
TTTTAGTAAATGATATTCATTATTCCTTTTATTTTTCTTCTAATTTATGGATCTTTTGGACTATGGTCTAAAAATC
TTGTTAAAGTAAACTGAATATCATAAGAAAAAATGTTAGATTATAATCTAAATTTTTTATAAATTATTAGACGTTA
TCTAATATTTTGTATGTAAGATTGAGAAACATATACATAAAACATTAGATTCAAATTTAATAATATCTAAAATATT
GATTCAAATCAATCATGACTACACAAACGAATACATGCAGATTCTCAAACATATAGATGAAGTCATTTCAAAACGA
ATCAAATATAGTAGAGTATATCCTTAAAAGAGAGCATTTGGGTAAATAAGTAAAAATCATTAAGTTATAAAAAAAA
TTCTAACTCGATCTCTCACGATTATTTAATCACTTTGTTCCAGTTTTCAGCAGTTCAAGTGAAAAATGTGGTGTAT
GAGAATATCAAGGGCACAAGTGCAACAAAGGTGGCCATAAAATTTGATTGCAGCACAAACTTTCCATGTGAAGGAA
TTATAATGGAGAATATAAATTTAGTAGGGGAAAGTGGAAAACCATCAGAGGCTACGTGCAAAAATGTCCATTTTAA
CAATGCTGAACATGTTACACCACACTGCACTTCACTAGAAATTTCAGAGGATGAAGCTCTTTTGTATAATTATTAA
TTTATACTATAGATCTTCAATATATAGCAGATATGATATATCACAATAAACAAATCTATATCTATGTATTGAATAA
TTATTATTAATATGTACGGATTGAAGTTTTAATAAGACTACTATGTATTTCTATTTTCTAGTCAAAAGTTTGACGA
TTGTACTTTTTAATGTACAAAAATAATAAAATGGTTATTTATATGATGTATATATCCCTTTGGTATTTCTTGTTGA
ACTATAATGTCATTATTTAATAACTATTATCTGTGCAATGATTGTATTTGTTAATGATACATAATATATCTTTCAT
CATTGATAATAAGAATAAAATATTTTACGTCTATTACTTTGTGAATTATATGTAGATTTTAGTTTTTGTTTTATTT
TTAATTAAACCGAGTGAAATATAAAGAG

Pectate Lyase – 20bp

GGGCGCTAGCGTACACATAGCG

>Solyc03g111690.2 SL2.50ch03:62343287..62345438

TATATAACTTAAACGCGTTTAATTGGACGTATGCTAGAGTAGTAATCATATGCTAACAAGAGATAGATACTGTTAC
TTGAACAGACAAAAACTACCGTAGTCTTCTTCTCTCTATCATCTAGGTTAGAGGAATCTCCGTAAAACATGTAATC
ACGAAGGAACAAGGGACCCATCCCTTTTGTTTAATTACGCTAATTTCTCTACTTTTTCCTTCTTTATAAACTCCTC
CATTTCCCTCTCTTTTCTTCATAACCCAAAATCCTCTGTTTTTAAAGCTAATTTATAAAACAAACAATGGGCACTT
CCTCTGTTTTTCTACTATTCCTTCTTTCTTTTCTTCTCCTTCTCCCGTCCCTCCTTGCCTCCTCTAATCCTCAACA
AGTTGTCGATGAAGTACACAGGTACGTGCATTTTTAATTTTATTTTAATTCAACATTGCATATTTGAAAAGGGACA
GAGTCAGATAGGAACAAGGAGTTCATTCGTATTAAAATAATTGCATATGGTAATACTCATGTGTACATTATTATTA
TAATAATGCAGGAGTATAAATGGTTCAAGGAGGAATTTAGGGTATTTATCGTGTGGTACCGGAAATCCAATCGACG
ATTGTTGGCGTTGTGATCCGAATTGGGAGAAAAACCGTCAGCGGTTAGCTGACTGTGCAATTGGATTTGGGAAAAA
CGCTATTGGTGGTAGAGACGGTAAAATCTACGTGGTGACCGATTCGGGTGACGACAATGCCGTCACTCCGAAGCCG
GGGACTCTCCGGCACGCAGTGATTCAGACTGAGCCACTGTGGATTATTTTCGCGAGGGATATGGTTATACAGTTGA
AAGAGGAATTAATTATGAATTCGTTTAAAACGATAGACGGAAG**GGGCGCTAGCGTACACATAGCG**GGTGGTCCGTG
TATAACAATACAGTACGTAACGAATATTATTATACACGGAATTCATATACATGACTGTAAACAAGGTGGAAATGCT
ATGGTGCGGAGCAGTCCATCGCATTACGGGTGGAGAACTGTTTCGGACGGTGATGGAGTGTCCATATTCGGTGGGA
GCCATGTTTGGGTGGATCATTGTTCTTTGTCCAACTGTAAGGACGGTCTGATTGATGCGATTATGGGGTCTACAGC
AATAACCATTTCAAATAATTACATGACACATCATGATAAAGTTATGCTCTTGGGACATAGTGATACATATACTCAA
GATAAAAACATGCAAGTAACTATAGCTTTTAATCACTTTGGTGAAGGTCTTGTCCAAAGAATGCCAAGGTATGAGC
CTGAGTTATTTAACTTTTATATACTGACAATTTTGACAATGTGAATTTTATGAATTATTATTTGGAATAAATGCAG
ATGTAGACATGGTTACTTCCATGTGGTGAACAATGACTACACACATTGGGAAATGTATGCTATTGGTGGAAGTGCT
GATCCCACTATCAATAGTCAAGGGAACCGATTTCTTGCCCCTGATATCAGGTTTAGCAAAGAGGTTAGTTAGTTAT
TTCACTTACATTTTTCTATCTAGTCTGTTAATGAGTTAACGGGTAATTTTATTATATTTATGTAGGTTACAAAGCA
CGAGGATGCACCGGAGAGTGAATGGAAGAATTGGAATTGGAGGACCGATGGGGACCTAATGTTGAATGGTGCATTT
TTTACGCGATCAGGAGTTAGAACTGGTTCATCAAGTTATGCTAAAGCTTCGAGTTTGAGTGCAAGGCCGTCTTCAT
TAGTAGCCAATCTTGTGTCTAGTTCTGGTGCACTCAACTGTAAAAAGGGTTCTCGTTGCTAATTATTATTATTATT
TCCTAAAAAAGACACAAGAACCAAAGCAAAAGGGGATTAATTATTTAACTTGGTTTTGAAAAAAGAAAAAGGATGT
AATTTGAAGAAAAAAACCATCTAGTGAGGGTGAAAAAGTAAGGGACAAGAAGTTTAGGGTTGTGCTTTTAATATCA
TTTTTTCACTCCCTTTATTTTTATTTTTACTTTTTAAGTTTTTTTTGACTTGATTGTCAAATACAACCTCGGCCTG
TTATTCTCAGCCATTGGGACTTTTAGTTCTTGTTGGTTTTAGAGAGAGGCAAGCTAAGAAGTGTTGATATATACAT
CAATATATATGTATTTATTTTTGA

Cellulose synthase -20bp GTTTCTCGTGAGAAGCGACCAGG

>Solyc08g061100.2 SL2.50ch08:46667016..46658543 (sequence from reverse strand)

TACCGGTTAAGCCGGTTAAAATTCCTAATAATGTAATTTAATGTAATAAATTAGTGAAACTAAATCTGTCAAAGAG
ATGATAAAAGAGAGTACCACTGGCATTAGAATATCAGAGAAAGAAGGTGAAGTAGAGGAGTGATGTTCAGTTGAGG
AAAATGCACTGTTTCTTCACATTGGCTTAACCCCTATCAAGATTGATTAAGCTGTGTTTTAAGGTACCCTCTGTTT
CTTTGAGGAGAAGAAATGGAGGCAGGTGGTGGAATGGTAGCTGGATCTCACAAGCGTAATGAGTTGGTTAGGATTC
GACATGATTCTGATAGTGGGGTATCATTCTCAAATTCTCTGCTTTATCAATTTTTTTTTTTGTTCTTGATTCAATA
CACTTGTTTTAGATCTGATTTTGATGAAATTCTACATCTTATCTGGATTTAGGTGTTCATGTTGACTCAGATTTCT
GATAAAGTTTTGATCTTTTTGTCTATTTGTACCTTAATTTAGTTTTCTTTTGGAGCATTTTGGGTTTTGAGTTAGT
CTTCTCGATTCAGTGTTGGGGTTTAGATCTGACCTGGTGATAGTCTGTTGTTGAGATTTGTGATAAAGTTTTGATC
TATTCATCAATTTGGACTTTAAAGTTTTCTATTACTACTTCTTTAAGCATATTAGGGCTTGTAATTTGGTGGTTGA
AGTAGAGTGAGTTCTTACACCTAGACTGGGGTTCTCCCATGAATTTAGAAGTTTATTATGTTAATCTTGTTTTAGA
TTTGATTCCTATATCAGGATTTGAGTACTTTTTGGAATGAAATGGATGTATCAAGATTTGAGCTCTCTATAATTAT
TAACATTTTGCTTTTTCCCCTTTCATTCACTTTTCAACACCCCCACCTCCCATCTGATGTGGTTGGACATCTGATG
CATGATCAAATTAAATTTGGGGCCCAACATTGGGTGTAACGGGGTCTTACTATAATACCATGTGAAAATAGGCATT
GGCTTTACTCACTCCTCCAAAGCTAAGCCCAAAAGGGAGGAGGATTGTCCAAGCCCGATACGAGACCATCCATCTC
ATTGATCACCAATGTGGAATAACTTTTTAACTCTTCAACATATATATATATATATATATATATATATACCTGAT
GTTAATTTGTGACGGAAAGTTCTTCAAAAATATAATTTTTTTGCTGCTACAAATTTGAGCAGTTTTGCCTCTTGTG
TTATCCTAATGTAGTAACTGATGGCTGATTAGGAGAAAAATAAGCATGTTCTATTTTGGGGTTTCATTTTAGATCC
ACATATTAATTCCAGAGTATGGATCATGAATATTTCATTGATCTTTCCTTTTATGTTTGGCATAAGAATGTCAATC
AAAATTGAAGTCTATATATTTGGTTTAAGGCCTGCATTTCACATTTCGAGCATTCAATTCCTTCTTGGATGAACTG
AGATGCAATTCTTCCTAAGCACTTTTATGGACAAATATAAGTTGGTGAAATTTTTCAATGTCAAAAGCAAAAAGAT
AATCGCTTCTAAACAGACTTGTCTAGTTCATTCACTTTAAGAGCCTTCTTGGAATGGTTGAATCTACTGATTATTA
GTTTCCAGTACCCTCCATCCAAATGAGTAAATGACTCCTCTGCCAGTAGTTTATTAGACTGAAGCTAACGACTCCT
CTGCCAGTAGTTTATTAGACTGAAGCGTGAGTAGGTTTGCATGTATCTATTGGTTCATCTTAATTCTGTTAATTTA
AGTGGACATTTTGGTGCAGCCGAAACCCTTGAAGCCTCTCAATAGCCAAATATGCCAAATTTGTGGTGATACTGTT
GGTTCGACTGCTAGTGGTGATGTATTCATTGCTTGTAATGAGTGTGCCTTTCCTGTTTGCCGGGCTTGTTATGAGT
ACGAGCGTAAAGATGGAAATCAATCATGTCCTCAATGCAAGACCAGATACAAGAGGTTCAAGGGTGAGTGAGTTCA
ATATTTTTTTCTGTTACTTTTTTGGTTCCTTGCGCTTTCGTTTTTGCTTCTAAACTGAACTTTATAAACTAAGTCA
TTGCTTGGTCAATGTATTAGGGAGTCCACGAGTGGATGGAGATGATGACGAGGAAGATGTTGATGATATCGACAAT
GAGTTCAATTATGCTCAAGGAAACAGCAAGGCCAGACAGCAATGGCAGGGTGATGATGCTGGTCTCTCTTCATCCT
CTAGACACGAATCTCAGCAACCGATTCCTCTTCTTACAGATGGCCAGCCAGTGAGTGATACTGTGCTTATGTCGTC
ATTTTGTGCTTTGGCTGTATTTTTATTTCATTAAGTTGTAGTTGTAGAACCTCCTGGGAGGTCCTTGATTTGCATC
TTTAATTGGGAGTGAAGGGATTTCTGCAGGAAACTCAGTTTTGAATGCTGATAAAAAGTGTGGCTTAAAATGAATT
GATTTTTAATGTATTATTAAAATAATCACTTGTTTAGTCAATCAGACATGGTCTTGGGTCCTGCAAGTGTCCCATG
AAGCAAAAAGTAGCCATCGTTTTAAGTCCTTTTCCTTGCACTAAACTTTACAAAAATCGCTACAAATTTTGATAGT
TGACTATCCTGCTTGTCTTTCTGATTCTTTTTGAATAAAAACTTATGGACATCTGACTGTGGGTTTACGTCTGATG
GATGAAAATTGAAACAGACCCTGTTTCAATTAGATTGGCTGCGTTCCGTATACTAGTATTATCCACTGGGAGATTT
TCCTGCTTTTGTGCTTCGACTTTGAAAAGCAGATTTGGAGCATGGTGTTATATATGGTGCATGAACAAACTAGTGT
GTTCATTTGATAATTTATTAACCAAATATGATGGATTTTTGGTGAAGTCCTTACCTATTGGCATGATATTACTTTT
TGTATTTCTCTGGCAAATGTACTGAGCTCTTCTAAGCATTGTTATTTTCTAAAAAAAAGTCTTTCTGCTTGTCTTA
GGTGTCTGGTGATTTTCCTAGTGCAACAACTGATACACAATCTGTAAGAAGTATGTCAGGACCCTTAGGCCCTGGA
GACAAGCATTCATCCCTGTCTTATGTTGATCCAAGACAGCCTGGTACTGGCCTACTTGAATATATTCTTAGTTGGC
TGCAATAATGTATTAAAACACTTTGTTCGTCCTCTAAACTTCTGATAATTTCATGCTTTTGCAGTTCCTGTAAGAA
TTGTGGACCCCTCAAAGGACTTGAATTCTTATGGGCTTGGAAGTGTTGACTGGAAGGAGAGAGTAGAGGGTTGGAA
ACTGAAACAGGAGAAAAATCTGGTGCATTCGACAAACAGATATTCTGAAGGGAAAGGCGGGGACATTGAAGGGACA
GGGTCGAATGGAGAAGAGCTACAAATGTAAGCATCTTATTGAATAAATGAAAAATATTTAGGGAACATACGATATA
TTATAGATGACTGGATCAAACTTTTGCTTCTTCCTTTAATAACCCCAGCACTGATATCCATCTCTCTCTCACTGCA
AAAATCTTGTAATACTTGGTTTGAGAAGTGAAAAGTTTTTTAAAAAATCCTTTTTAATGTCTTGACACTGTTATGA
TGTCCACTGCACCATCCCTGCTTAAATTTGTAGTACTTAGTTTGGGAAATGAATTTCCTTTCTATTGTCCTGGAAC
TATTTCTCTGTTTATTAGTCAATTTTCTTGGTTTTATAATACTGCAAGAAAATTTAATTTAAAACTATGATACTTA
TTGGTTTTCTCTTGACAGGGCTGACGATGCTCGCCAGCCTATGAGCCGTGTGGTGCCTATTCCTTCATCCCACCTT
ACCCCATATCGGGTTGTTATCATTTTGCGGTTGATCATCTTGGGATTTTTCATGCAATACCGTCTAACTCATCCAG
TGAATGATGCATATCCTCTGTGGCTAGTATCAGTTATTTGTGAGGTTTGGTTTGCCTTGTCCTGGCTCCTCGATCA
GTTTCCAAAATGGTCACCAGTCAATCGTGAGACATTCCTTGACAGGCTAGCTTTGAGGTTAGTTGCTATGCTACTC
TAGGGAGGAATTATTTGAGTAGTGGAAGGTTTCAAATATTTCACTTCCATTTAGGATAAATACTATGTTTCAAATT

CCGGGGTCAAGTTTTCAGTCTTGAAAGTCTTTTCCTCACTTTCTCAGACATGATAGAGAAGGGGAGCCTTCGCAGC
TTGCACCTGTTGACATCTTTGTCAGTACAGTGGATCCTTTGAAAGAACCTCCTCTCATTACTGCAAACACAGTGTT
GTCCATCCTTGCTGTTGATTATCCTGTGGACAAGGTCTCATGCTATGTGTCAGATGATGGTTCAGCAATGCTAACA
TTTGAAGCCCTTTCTGAGACTGCAGAGTTTGCAAGGAGATGGGTTCCCTTCTGCAAGAAATTCAGCATTGAGCCTC
GGGCCCCTGAATTTTATTTTGCTCAAAAGATAGATTATTTGAAGGACAAGGTTCAGCCTTCTTTTGTGAAAGAGCG
CAGGGCAATGAAGGTACAAATGAAACTTTAATAAGCTGAATTTAATGGGAATTTATTGGTATGTAGGTGACTAGGC
ATCAGGTTTTAGATATAAAACTTCTGTTTGATGGTAGTTGAGTTTTATTCACCAAATGCTCCATTAGTTAAACTCT
ATTATTTTTGATAGGATTGTGGAAGGCTGAATGTCTAAGATCCCTTGATAAATCAATGTATACAGGATAAGTAGGT
CTGGAGATAACTTTAGCCTGAAAGTTTGTTGGTTACCTGCGGAGTGATTTACATAAATCATTTGATAAGTTCTACT
AGTTGGGAAAGTTCAGGCCTCTCTGTTGCTAGGATTTATGCTGATAATACTGCCAGAATCCTTTTTCTCTCTAATA
AGCAAACAATTACAATAATCTGATAATTTTGGCATTGCAGAGGGAGTATGAAGAATTTAAAATACGAATCAATGCT
CTTGTTGCAAAAGCGCAAAAAATGCCTGAAGAAGGTTGGACAATGCAAGATGGAACAGCTTGGCCTGGAAATAACC
CTAGGGATCATCCAGGAATGATCCAGGTATTCTTTATTTATCTTATAATTATGTGAAAGTTGATTCACTGTACATC
TTCTCCTCTTTTAGTTAATTACTTCGGATTGGATGTAGGTCTTTTTGGGGCACAGCGGGGGCCTTGATACTGATGG
AAATGAACTTCCTCGTCTGGTGTAT**GTTTCTCGTGAGAAGCGACCAGG**TTTTCAACACCACAAGAAAGCTGGAGCT
ATGAATGCTTTGGTAGGCCACTGATAATTACAATATGAGATCTTCCAATTTTATTGAGCATGCCATCAAGTTGAAT
AGCTCTAATTTATATACACACATTATGGTTGCGATGATGGAGAATGTCTCATAGCATACTTAAATCATACAGCTCT
CCTTATGCCATCTGATATTATTATATTGTGTGATGTTCCATTTATTTAAGCTGTACCTTCAATTTTCTGCAGATTC
GTGTTTCTGCAGTTCTTACCAATGGAGCATATCTTTTGAATGTCGATTGTGATCACTACTTTAACAACAGCAAAGC
ACTTAAAGAAGCAATGTGTTTCCTCATGGATCCTGTTCTTGGAAAGAAGACATGCTATGTTCAGTTCCCTCAGCGA
TTTGATGGCATTGACTTGCACGATCGATATGCTAACCGAAACATTGTGTTCTTTGATGTAAGTCTCACATACACAT
TCCCCAGATTGTTTATCAACCATATTTGGTATGTCAGAGTCATGTTTAAGTTAGAGTGCTCATTTCATTATGTTCA
TGTTTCTCAGATCAACTTGAAAGGGTTGGATGGCCTTCAGGGTCCTATGTATGTGGGAACTGGGTGTTGTTTTAAC
AGGCAGGCCCTTTACGGGTATGATCCTGTATTAACTGAGGCTGATTTGGAGCCAAACATCATTGTTAAGAGCTGTT
GTGGTGGGTCAAGGAAGAAGGGAAGGAGTGGAAACAAGAAATACATTGATAAGAAGAGGGCAGTAAAAAGAACAGA
ATCCACTGTCCCAATTTTCAATATGGAAGACATTGAAGAAGGCGTAGAAGGTAATGTCCTTAGAAAATCTTGGTTC
ATCCTCTATGGTGATATTAAGTGTTTTACTTAGATAAGTCAATATCGTCTTCAAGCATCTAACTTCTGATTACAAC
TAAGATATACTTAGGAAGTATCATACTTTTGACATACTCTTAATGTTCTTTTCTATCAGATATATGATGGGTTCCG
AAGTGAACTGAGAAAACCTTTTGATGATGTTTGAAAAGTTAGTTTAGCTGAATGTGATTATTGAACTCTCAACTCT
TCTTCTGTTCTGTTTTTATTTGACAAACCGATTTGTTCAGTCCTCCCAACTTCTCAACCTGTCTCATACCTGAGGC
TGCTGAATATATGAATACCCTCATCACAGTGTTGTTTGTTGACTGTTTTCTTTGGATTTGCAGGATATGATGAGGA
GAAGTCACTGCTTATGTCACAGAGGAGCTTGGAGAAGCGATTTGGTCAATCACCAGTGTTTATTGCTGCCACATTT
ATGGAACAAGGAGGCATTCCTGCATCCACTAATCCTGCATCTCTTTTGAAAGAAGCAATTCATGTTATTAGCTGTG
GCTACGAGGACAAGACTGAATGGGGAAAAGAGGTACAACAATGTTGAGTTATATACTGTTATGACTTAGATATCCA
TTTTGCTTGGCTGTATTGATCTTACTGTGGCACAAAAACCCTTATTAAGCATTTTTGGTTTCCTATCAGATTGGGT
GGATATATGGGTCTGTCACTGAAGATATTTTGACTGGGTTTAAGATGCATGCACGAGGATGGTATTCCCTCTATTG
CATGCCTCCTCGACCAGCGTTTAAAGGGTCTGCTCCTATCAATCTTTCTGATCGTTTAAACCAGGTGCTCCGGTGG
GCTTTGGGTTCTGTTGAAATCCTGTTGAGTAGGCATTGCCCAATATGGTATGGCTACAGTGGCAGATTGATGCTTC
TGGAGAGATTAGCATATATTAATACAATCGTCTATCCTCTCACATCTTTACCCTTGCTTGCTTACTGTACGCTTCC
TGCCATCTGTCTTCTGACCGGCAAATTTATCATTCCTGAGGTATGTTAATAGTCTCATTCAGATTGCTTTATGCTG
TTAAGTTTAAATAGTTGATTGAAACATTGAATACCAACAGAATGATTTCCATGCCATACGTGTCCAATTAGAAAAG
GCTGGGTATACTAACTTCTCACACTGAAATGATCCCCTGGTGTTGTGAAAAGGGAGCCTGTAAAGAAAAAAAAGAG
TTGAAAATAGGAAAAAACAAAGAAGTAAATATTCAAATTGTTTCCTCTACCTTATATTATCCTCCTCTTAGTGCCT
GTCAAAATTCTGTTCCGTGTCTTCATGCTGAGAACTATTTTGCTGTTCTGATTTTAGTCTTTTCTTTGGATGAGAC
TGTTTGTTGTTCACCCCTTGAGATATATGTATATACTCAAGCTTGCCCATGATTATGGTCTTGATCCCCTTAGCTT
GTATATACATCTATGAATAATTCTGAATTTACCTGCAGATAAGCAACTATGCTGGCATGTGGTTCATTCTACTCTT
TCTTTCCATTTTCGCAACTGGTATATTGGAGCTTAGGTGGAGTGGTGTGAGTATTGAAGACTGGTGGAGAAATGAA
CAGTTTTGGGTCATTGGTGGTACATCGGCTCACTTGTTTGCAGTTTTCCAAGGTCTCCTGAAAGTGCTTGCTGGGA
TTGATACCAACTTTACAGTCACATCAAAGGCAAACGATGAGGATGGAGATTTTGCTGAGCTTTATGTGTTCAAATG
GACAACTCTCCTTATTCCTCCTACAGCAATCCTTATCGTGAATCTGGTAGGAATTGTGGCAGGTGTTTCGTATGCC
ATAAACAGTGGATATCAATCATGGGGTCCTCTATTCGGGAAGTTGTTCTTTGCCATTTGGGTCATTGTCCACTTGT
ACCCTTTCCTCAAAGGTTTGTTGGGACGTCAGAATCGTACTCCTACCATTGTCATTGTCTGGGCTGTTCTCCTTGC
ATCCATATTCTCATTGCTTTGGGTGCGTATCGACCCCTTCACATCAGACGCGTCAAAGACTGCTGCAAGAGGTCAA
TGTGGTATCAACTGTTAGTGAGAAGTTTATTTATTTTAAGGTAAGATGTTTGCTTATCCATTCTCATGTTGTATAT
AAGAAAACTATGCCGGTTCGGTTAAGTCCTGTAGTCATTTTGTACCAATGTTGTAAGTCGGAAGGAGTGTGATGCA
TGAAAGACATCAAAACTCATGGATTTACGCCTTTTTGTTGTATACCTTTTGCCCTATCTCACATGTTCATGCAAAA
ACTGTGACTTATGATGTGAAGAATTTTAAATTTAATTTTGTTATCTTATCGTTGGTGCAGAAGTATCACTGATATG
TTTGTGAGTTTGAGAATGGAGCTGTTAATGTGTTTTTA

Sequencing results from PCR Purification

>psy1

TTGGTTTGCCTGTCTGTGGTCTTTTTATAATCTTTTTCTACAGAAGAGAAAGTGGGTAATTT
TGTTTGAGAGTGGAAATATTCTCTAGTGGGAATCTACTAGGAGTAATTTATTTTCTATAAAC
TAAGTAAAGTTTGGAAGGTGACAAAAAGAAAGACAAAAATCTTGGAATTGTTTTAGACAACC
AAGGTTTTCTTGCTCAGAATGTCTGTTGCCTTGTTATGGGTTGTTTCTCCTTGTGACGTCTC
AAATGGGACAAGTTTCATGGAATCAGTCCGGGAGGGAAACCGTTTTTTTGATTCATCGAGGC
ATAGGAATTTGGTGTCCAATGAGAGAATCAATAGAGGTGGTGGAAAGCAAACTAATAATGGA
CGGAAATTTTCTGTACGGTCTGCTATTTTGGCTACTCCATCTGGAGAACGGACGATGACATC
GGAACAGATGGTCTATGATGTGGTTTTGAGGCAGGCAGCCTTGGTGAAGAGGCAACTGAGAT
CTACCAATGAGTTAGAAGTGAAGCCGGATATACCTATTCCGGGGAATTTGGGCTTGTTGAGT
GAAGCATATGATAGGTGTGGTGAAGTATGTGCAGAGTATGCAAAGACGTTTAACTTAGGTTA
GCTTCTTCAATCTATTCATTCGTTTACCAAATATTATTTGGTAAGCACTAATTATGAATATA
TATATGTTCATGTTATTGATGAAGACAAAATTTGATCTTTGTTTGTTTATTCAGGAACTATG
CTAATGACTCCCGAGAGAAGAAGGGCTATCTGGGCAATATATGGTGAGGTTTCTAGCCATTT
AATAACAGTTACGCGCACAAACACA

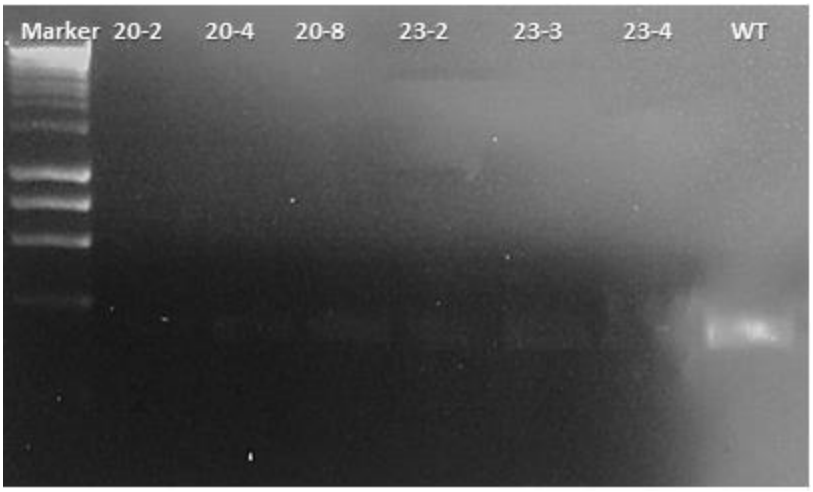
>PG2A

ATGGTTATCCAAAGGAATAGTATTCTCCTTCTCATTATTATTTTTGCTTCATCAATTTCAAC
TTGTAGAAGCAATGTTATTGATGACAATTTATTCAAACAAGTTTATGATAATATTCTTGAAC
AAGAATTTGCTCATGATTTTCAAGCTTATCTTTCTTATTTGAGCAAAAATATTGAAAGCAAC
AATAATATTGACAAGGTTGATAAAAATGGGATTAAAGTGATTAATGTACTTAGCTTTGGAGC
TAAGGGTGATGGAAAAACATATGATAATATTGTAAGTATTTAAATATTGGAATATATTTGTG
GGGATGAAAATGATAGAGAATATAAGAATTATTTGGAAGGATGAAAAGTTATATTTTATAAA
GTAGAAAATTATTTTCTCGTTTTTAGTATTAAGGTGAAAATGAGTTTCTCGTTAAGCGAGGA
AAAGCTATTTTCCATGGTAACTGTATTTTTTTTTTACTTTTADATAACGTCATAGTATTTGC
TATACTCAAGAATAAGACACTTATTATTGATGATTTAGTGCTCGAAAAGAAATTGATAGTAA
TTTTGCTTAATATAACTATCAATTTCTTATATGTATATTTTTCAACCAAAATAACAAAGCGT
AATCCAATAAGTGGGCCTCTAGAATAAAGAGTAAGTTCTATTCAATTCTTAACCTTATTTAA
TTTTAGTGGAAACCTCGACAAAAACGAACAAACGTATTCAAACTTTTATATTCGGAATTCGA
GACCAACCATATGAACAACCTCACACATGCATATAGTC

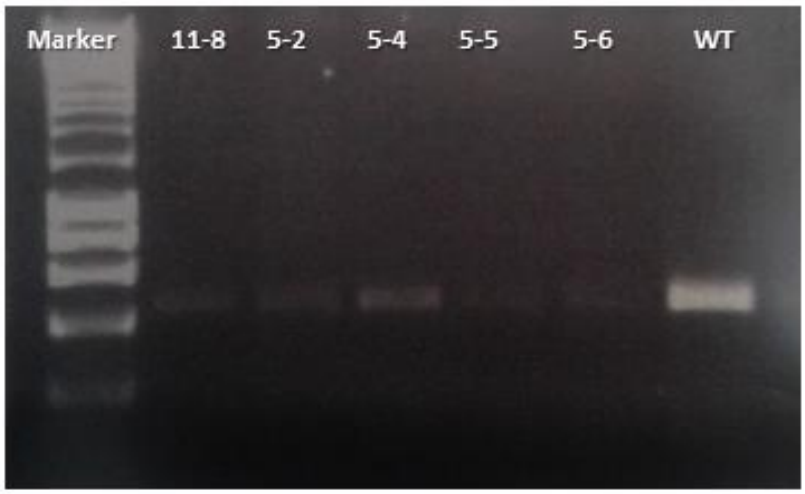
>Pectate lyase

ATTGGGAGAAAAACCGTCAGCGGTTAGCTGACTGTGCAATTGGATTTGGGAAAAACGCTATT
GGTGGTAGAGACGGTAAAATCTACGTGGTGACCGATTCGGGTGACGACAATGCCGTCACTCC
GAAGCCGGGGACTCTCCGGCACGCAGTGATTCAGACTGAGCCACTGTGGATTATTTTCGCGA
GGGATATGGTTATACAGTTGAAAGAGGAATTAATTATGAATTCGTTTAAAACGATAGACGGA
AGGGGCGCTAGCGTACACATAGCGGGTGGTCCGTGTATAACAATACAGTACGTAACGAATAT
TATTATACACGGAATTCATATACATGACTGTAAACAAGGTGGAAATGCTATGGTGCGGAGCA
GTCCATCGCATTACGGGTGGAGAACTGTTTCGGACGGTGATGGA

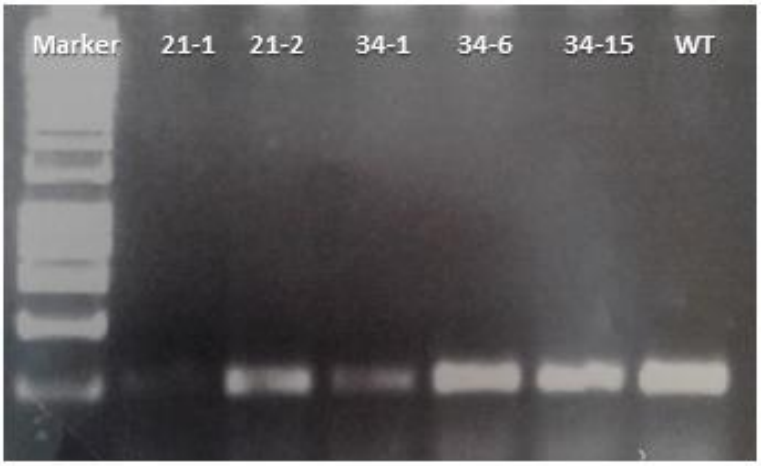
9.10 RT-PCR – gene expression



(A) PSY1 gene expression in wild type and PSY1 tomato fruit pericarp at B+4 (breaker +4) and B+7 (breaker+7).

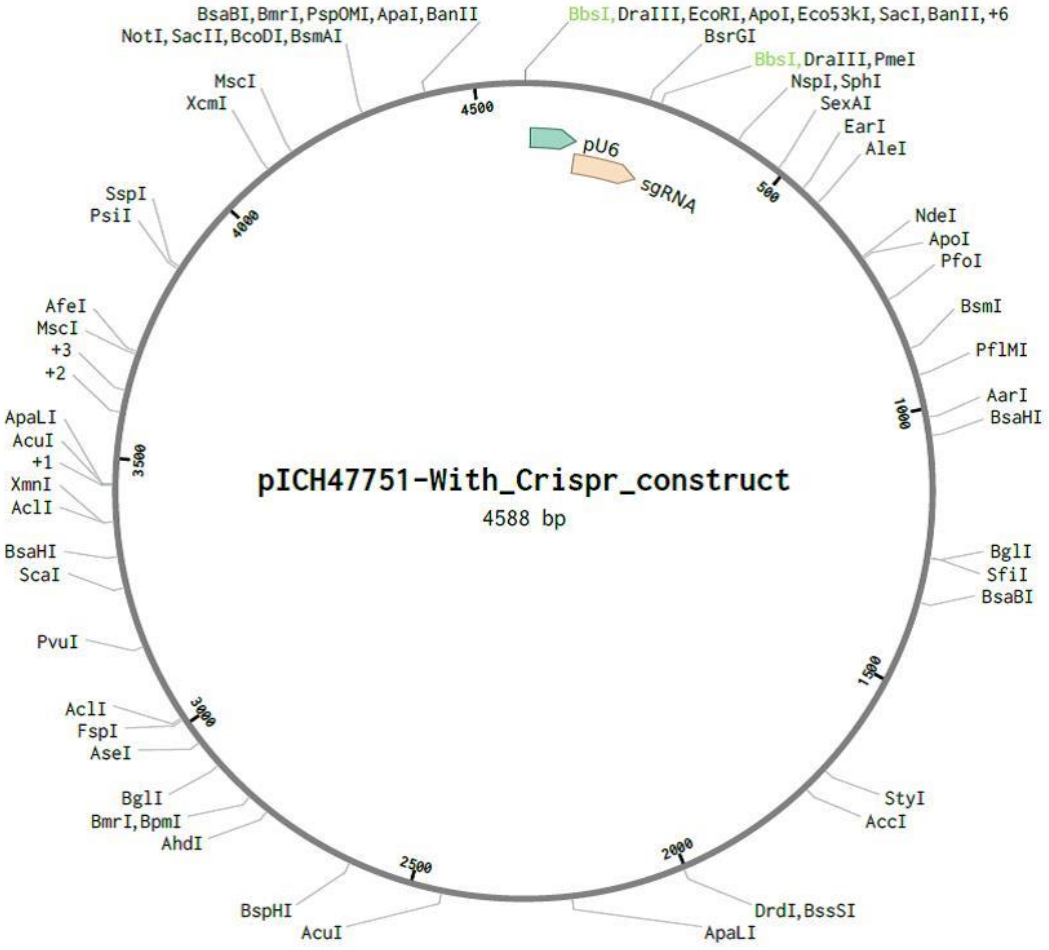


(B) PL gene expression in wild type and PSY1 tomato fruit pericarp at B+4 (breaker +4) and B+7 (breaker+7).

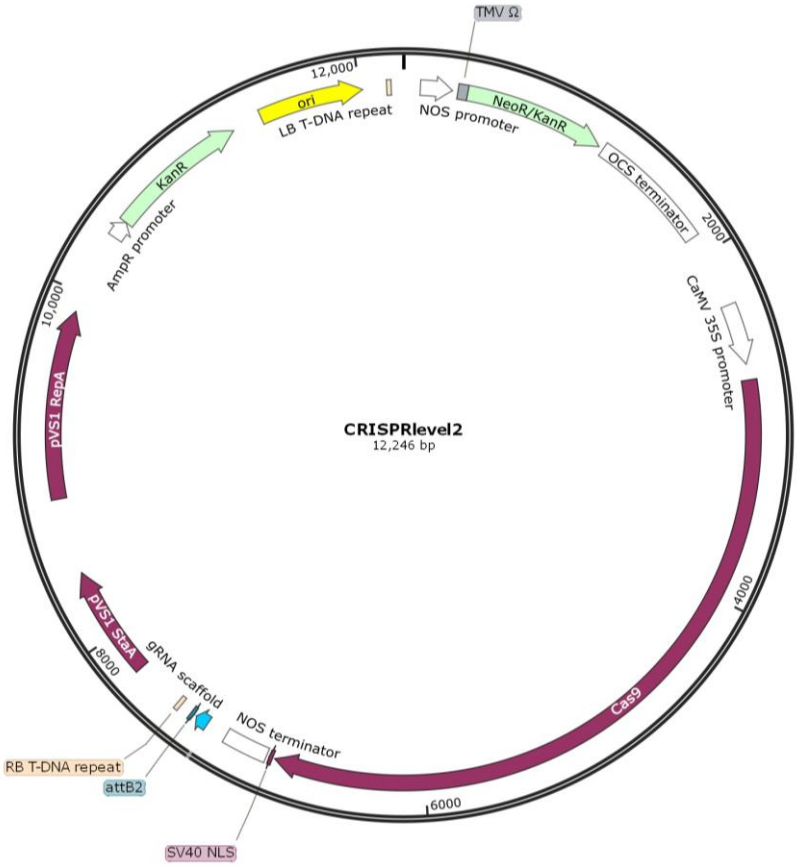


(C) PG gene expression in wild type and PSY1 tomato fruit pericarp at B+4 (breaker +4) and B+7 (breaker+7).

9.11 Level 1 CRISPR construct

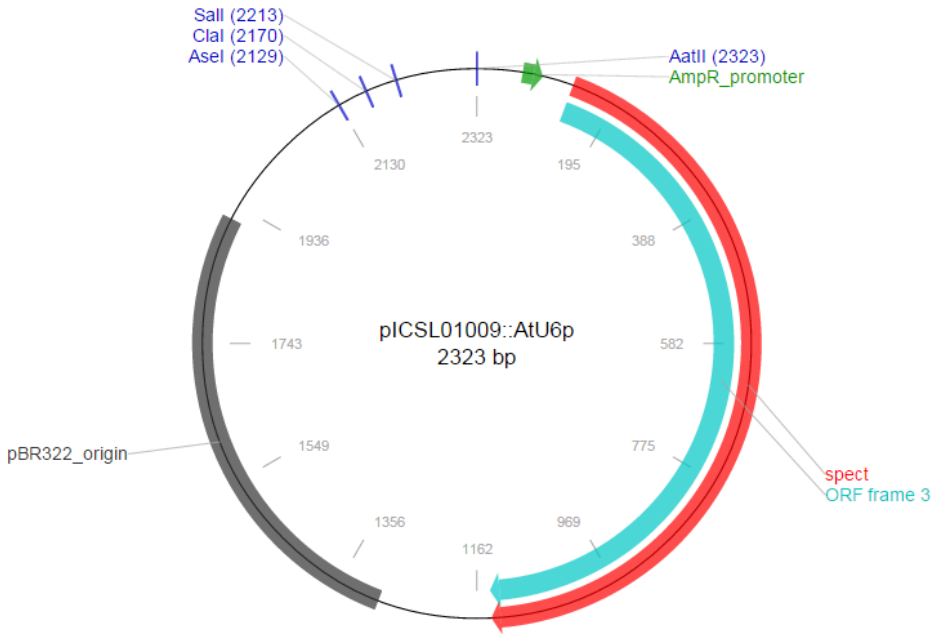


9.12 Level 2 CRISPR construct

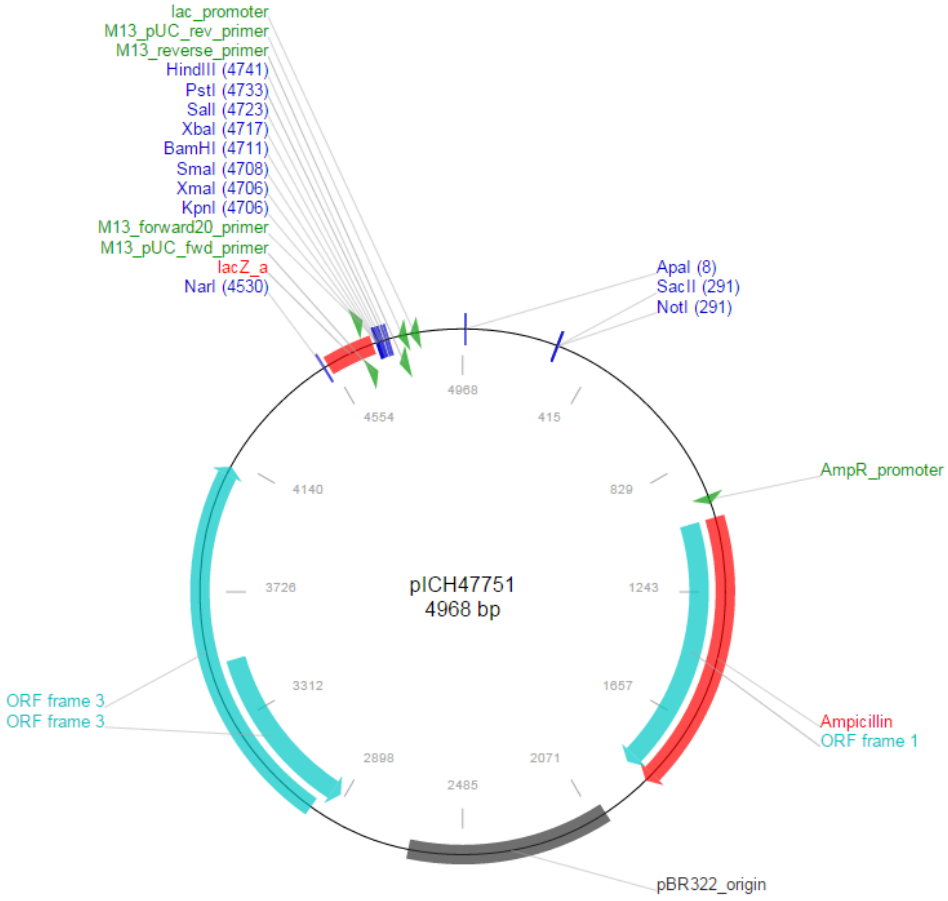


9.13 Vector in CRISPR contructs

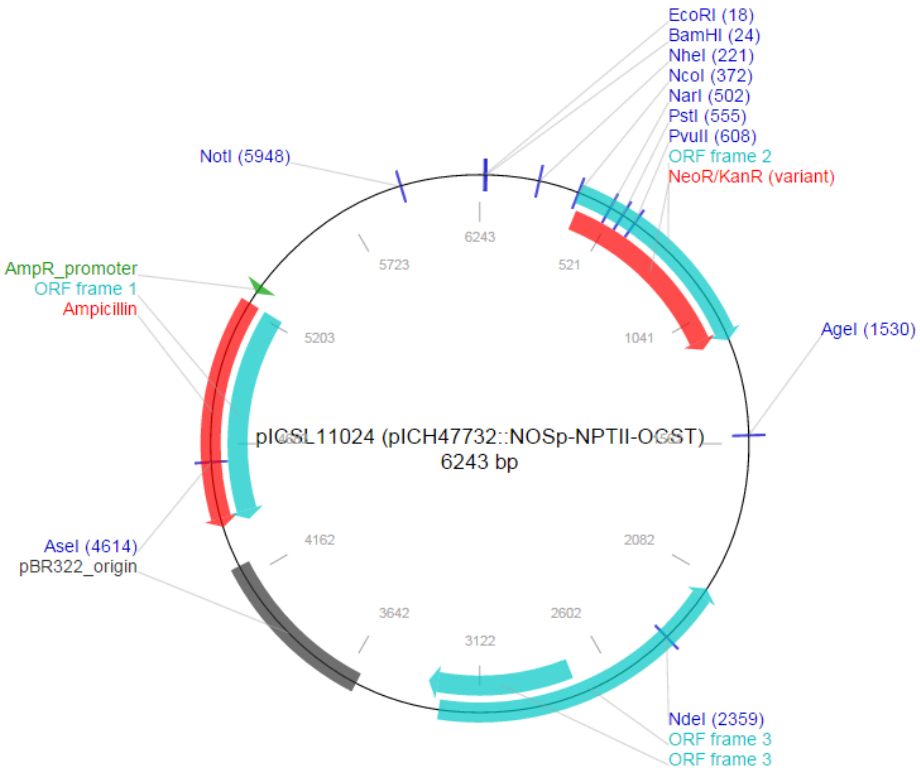
Vector Maps for pICSL01009::AtU6p



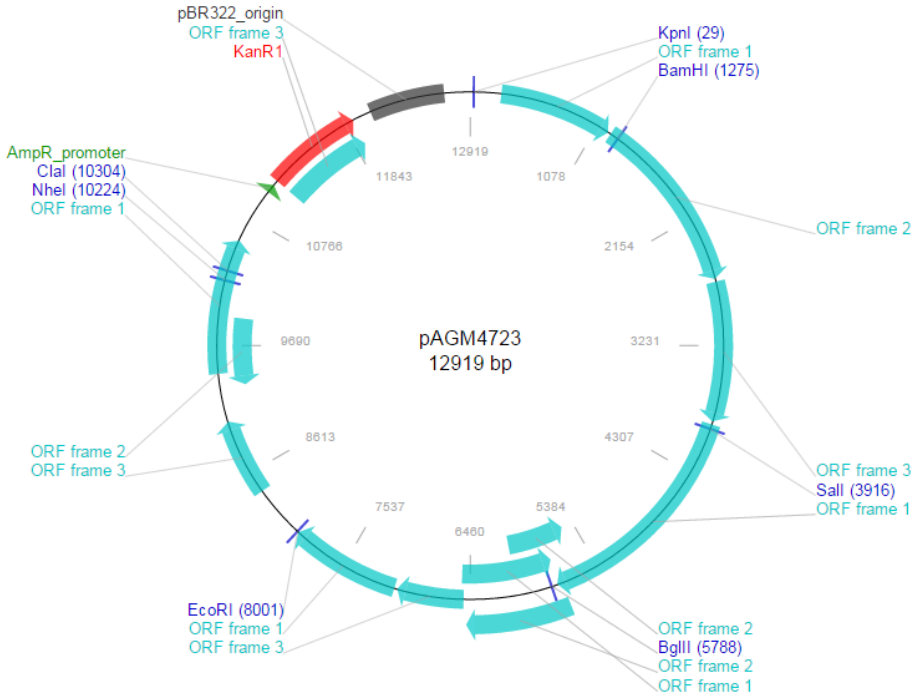
Vector Maps for pICH47751



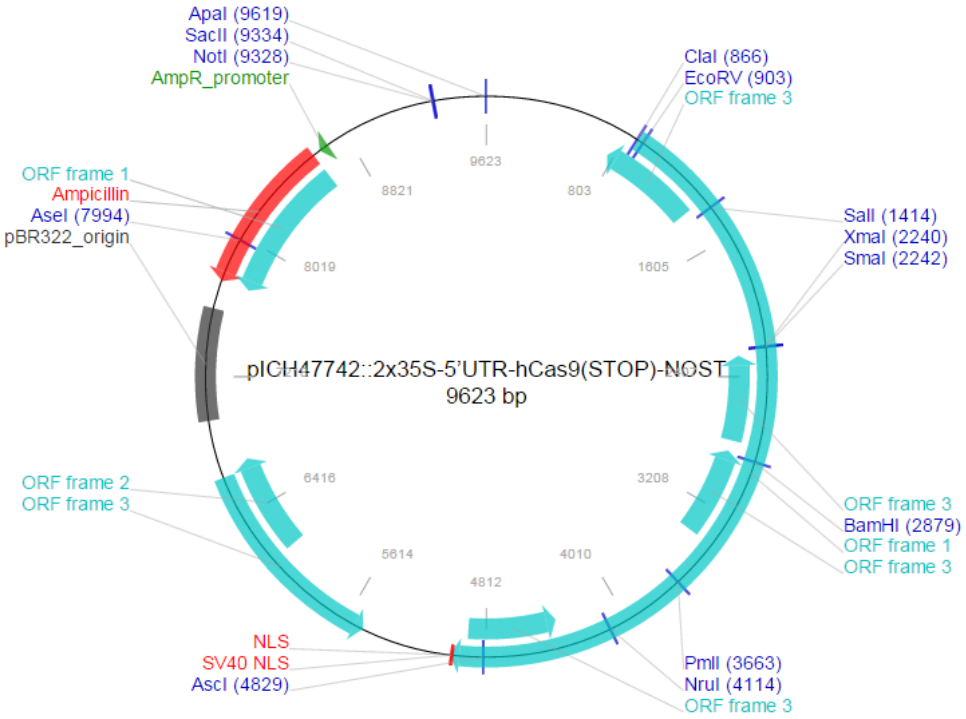
Vector Maps for pICH47732::NOSp-NPTII-OCST



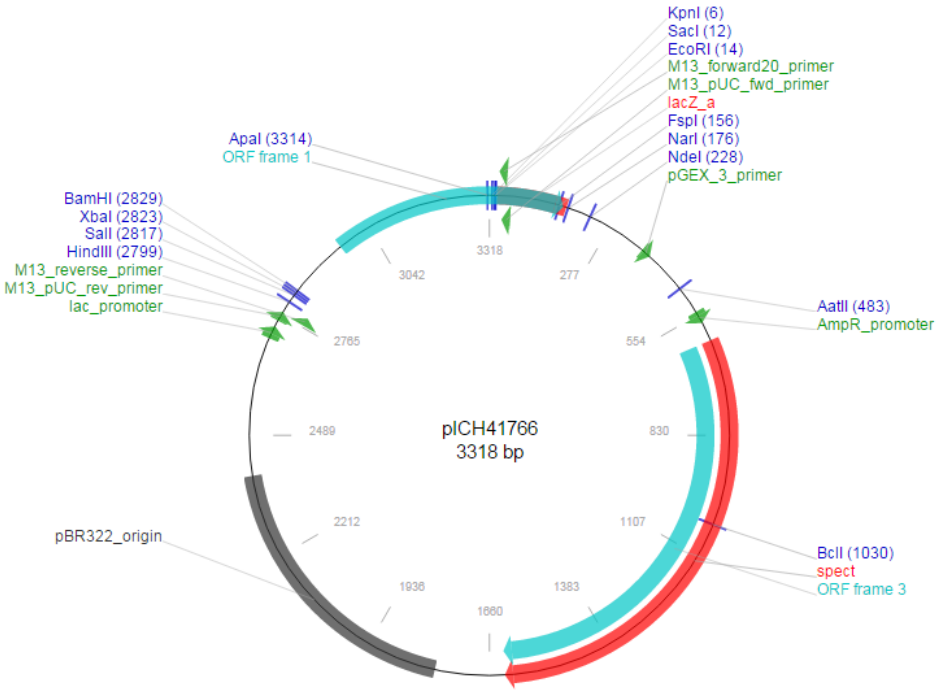
Vector Maps for pIAGM4723



Vector Maps for pICH47742::2x35S-5’UTR-hCas9(STOP)-NOST

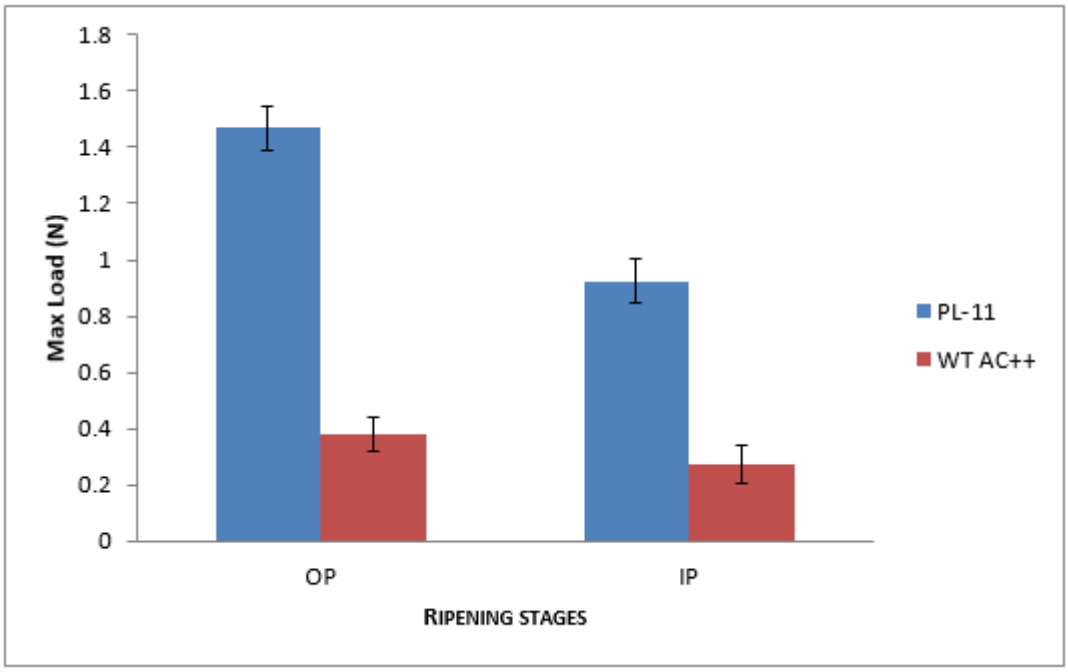


Vector Maps for pICH41766

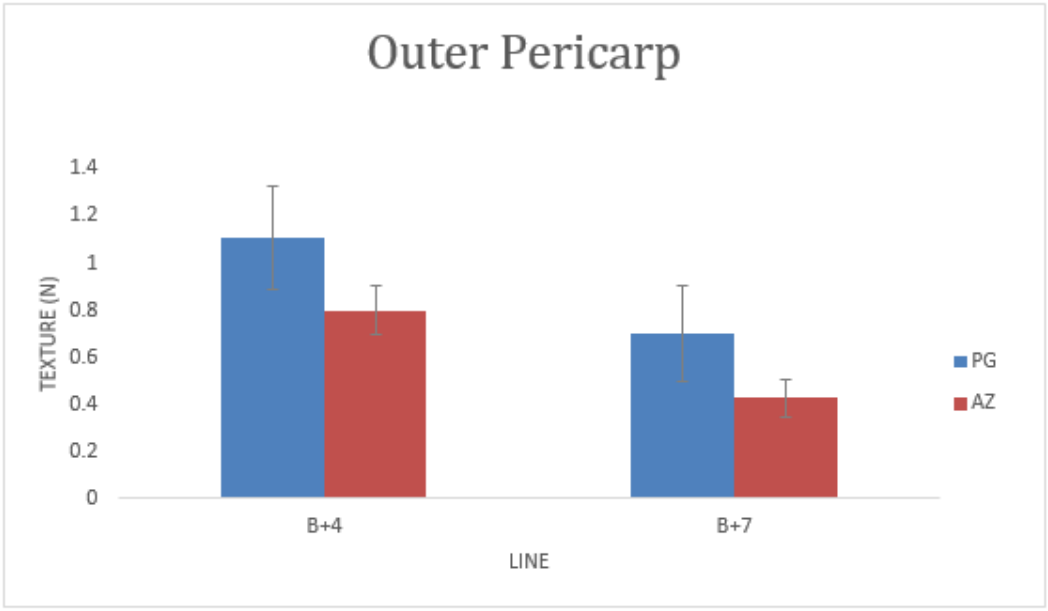


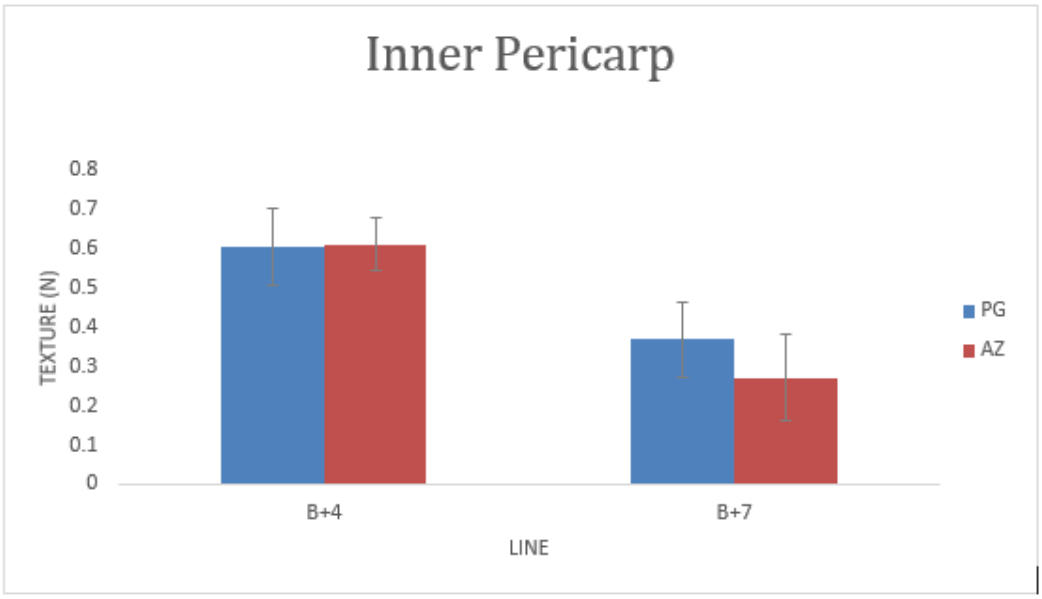
9.14 Texture measurement of Transgenic Lines

CRISPR::PL11



CRISPR::PG21


Outer Pericarp
1.4
1.2
1
0.8
0.6
0.4
0.2
0
TEXTURE (N)
B+4
B+7
LINE
PG
AZ


Inner Pericarp
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
TEXTURE (N)
B+4
B+7
LINE
PG
AZ