12-2015

Analysis of Ac/Ds Activation Tagged Mutants in Tomato (Solanum lycopersicum)

Ipeleng Randome
University of Arkansas, Fayetteville

Analysis of Ac/Ds Activation Tagged Mutants in Tomato (*Solanum lycopersicum*)

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Cell and Molecular Biology

by

Ipeleng Randome
University of Botswana
Bachelor of Science in Biological Sciences, 2009

December 2015
University of Arkansas

This thesis is approved for recommendation to the Graduate Council.

_____

Dr. Andy Pereira
Thesis Director

_____ _____

Dr. Fiona Goggin
Committee member

Dr. Vibha Srivastava
Committee member

_____

Dr. Ainong Shi
Committee member

Abstract

Tomato (*Solanum lycopersicum*) is a crop of immense economic and nutritional importance worldwide and also a good model organism for genomic studies of other dicot species. The recent completion of the tomato genome sequence is a great milestone towards learning about the tomato genome. Elucidation of the function of the different genes using different functional genomic tools is therefore important in adding to this resource. To this end, we have developed an Ac-Ds transposon 'activation tagging' (ATag) system to be able to transpose transposon inserts, bearing a strong 35S-enhancer element, all around the genome. An Ac-Ds ATag construct was used to generate transformants in the tomato cultivar M82 that has an erect determinate habit, suitable for greenhouse and field screening. The progeny of putative tomato transformants were germinated and grown to maturity in the greenhouse. Plants with obvious mutant phenotypes were identified, which included dwarfism, altered leaf morphology and necrotic spots on leaves. Presence of the ATag transformed construct was confirmed in the plants by genomic PCR using primers specific to different parts of the Ac/Ds cassette. Activity of the transposon system was also tested by excision PCR using primers flanking the Ds insert in the construct. Insertion sites of the Ds ATag were determined using TAIL-PCR for the progeny plants, and the tagged genes in two mutants were identified by alignment of the flanking sites to the tomato genome. With the availability of the tomato genome sequence, the mutants described will be a good resource for the identification of genes for plant development and tomato breeding.

Acknowledgement

My first and special appreciation goes to the Almighty God for how far He has led me. Second, I extend my profound gratitude to Dr. Andy Pereira, my advisor, and all my committee members for their wonderful contribution and support to the success of my study. Third, a special thank-you goes to Dr. Supratim Basu who selflessly worked with me throughout my research; and all my lab mates for their support. Finally, to all my friends and loved ones, I appreciate you being a part of my experience in the course of my education.

Dedication

Dedicated to my son, Mmusi Jaden Randome.

Table of Contents

1.0 Introduction

1.1 Tomato as a plant genetic system

Tomato (*Solanum lycopersicum*) belongs to the family *Solanaceae*. It is a perennial plant that typically grows up to 3 meters and has a weak stem that sometimes sprawls on the ground or climbs over other plants. Tomato is grown worldwide because of its fruit that is of great nutritional importance and also forms a major part of human diet. Tomato is a rich source of antioxidants such as lycopene, beta-carotene, flavonoids and vitamins A, C and K. Consumption of this fruit has been linked to reduced cases of cancers as well as heart disease in humans (El-Gaied *et al*, 2013).

In addition to its nutritional and economic importance, tomato serves as a good model plant for genomic studies of other dicot species (Carter *et al*, 2013). It has a moderately sized genome, approximately 950Mb (Tomato Genome Consortium, 2012) which is relatively easy to analyse and manipulate. The life span of the tomato plant also makes it easy for researchers to follow its life cycle. Additionally, the tomato genome is more readily amenable to transformation than other dicot species (Matthews et al, 2009), hence its usefulness in genomic and biotechnology studies. The recent completion of the genome sequencing of tomato (Tomato Genome Consortium, 2012), ensures that its role as a model organism will continue to grow. To add to this useful resource however, elucidation of gene function for the different genes is vital.

Various tools have been applied towards elucidation of gene function in different plants species, including use of chemical mutagens (Greene *et al,* 2003), transposon tagging (Aarts *et al*, 1993), gene trapping (Sundaresan, 1995) and Agrobacterium meditated T-DNA mutagenesis (Alonso *et al,* 2003). The most common method has been insertional mutagenesis leading to gene knock out

and then observing the phenotype (Krysan *et al*, 1999; Bouche and Bouchez, 2001; Thorneycroft *et al*, 2001). In eukaryotic systems however knock out mutagenesis does not always work due to functional redundancy of genes. Adoption of gain-of–function strategy such as activation tagging (Kakimoto, 1996, Weigel *et al*, 2000) overcomes the problems associated with establishing functions of duplicate genes, which are common in many higher eukaryotes.

In activation tagging a strong enhancer element, such as a tetrameric repeat of the cauliflower mosaic virus (CaMV) 35S gene enhancer, is used to direct the transcription of adjacent genes, and in the process create gain-of-function dominant mutations (Weigel *et al*, 2000). The enhancer may be introduced into a host plant using T-DNA mediated transformation or using a mobile transposon system, allowing ease in identification of the site of integration via PCR based protocols. Transposon insertions have been found to be more ideal than simple T-DNA insertions because transposable elements have an endogenous role in gene regulation within the plant. Also new insertions can be achieved through crossing and propagation as the transposon keeps changing location and therefore allowing different genes to be tagged (Sundaresan and Ramachandran, 2001). The activation tagging strategy has previously been employed in genomic studies of several plants species such as rice (Wan *et al*, 2009), Arabidopsis (Martinez *et al*, 2002) and tomato (Matthews *et al*, 2003, Carter *et al*, 2013).

The current study was initiated because of the economic and genomic research importance of tomato. The objective was to contribute towards the functional annotation of tomato genes following the recent publication of the complete tomato genome (Tomato genome Consortium, 2012). The elucidation of the functions of the different tomato genes will go a long way in establishing the tomato as a good model plant for genomic studies and will also aid in tomato breeding and understanding tomato plant development.

1.2 Mechanisms of gene regulation

Regulation of gene expression is a very important phenomenon in eukaryotic cells. In living cells, genes are transcribed differently and RNA transcripts are variably utilized (Brown, 1981). In plant cells regulation of gene expression may be influenced by tissue specificity, for example genes needed for photosynthesis have to be expressed in the leaves and are of little necessity in the roots. Gene expression also differs at different developmental stages of the plant and in response to different environmental stimuli.

Various mechanisms are employed by plant cells in controlling gene expression. The simplest form of regulation is controlling the amount of transcript produced. To achieve this, the cell makes use of regulatory sequences in the form of promoters, silencers, and enhancers. A promoter is a sequence located upstream of the transcription initiation site and is a site where the RNA polymerase binds to initiate transcription of a gene. Enhancers and silencers are cis-acting elements that bind to proteins which in turn affect the binding of RNA polymerase to the promoter and may either negatively or positively affect gene expression (Griffiths, 2000). An enhancer binds to activator proteins and increases expression of a gene, while a silencer binds to repressor proteins to reduce transcription of a gene, and in some cases may shut down the expression of the gene. By utilizing these regulatory sequences therefore a cell may either increase or decrease the amount of transcript produced for a particular gene or even shut down expression of a particular gene.

Gene regulation may also take place post-transcriptionally, that is, after the gene has been expressed. Post-transcriptional regulation can be in the form of either degrading the mRNA produced or inhibiting translation of the mRNA. In the process of post transcriptional modification the cell makes use of small RNAs in a process termed RNA interference (RNAi)

(Agrawal *et al*, 2003). In this process double-stranded RNA molecules are cleaved by proteins called dicers transforming them into single stranded RNA. The single stranded RNA then binds to another class of proteins called ARGONAUTE proteins. The ARGONAUTE RNA and small interference RNA (siRNA) complex then combines with the mRNA, this process then attracts nucleases which cleave the ARGONAUTE-bound RNA and in the process degrading the mRNA. In some cases the binding may not cause destruction of the mRNA but rather inhibit the translation of the mRNA. Gene regulation may also occur after translation (Phillips, 2008). After the production of the polypeptide a protein called ubiquitin binds the newly formed polypeptide. The binding of ubiquitin to the protein serves as a tag for the destruction of the protein. The tagged protein is then taken to a proteasome where it is destroyed (Phillips, 2008). In these ways plant cells are able to control which genes are expressed, the amount of transcript produced and which proteins are needed or not needed for the cell.

In addition to regulatory sequences such as enhancers and silencers, post transcriptional and post translational strategies for gene regulation, plant cells have other mechanisms of gene regulation in the form of additional regulatory sequences termed transposable elements or transposons. Since the discovery of transposable elements in the 1940s research on these elements has blossomed alongside their use in molecular biology studies.

1.3 Transposable Elements

Transposable elements (TEs) or transposons are mobile genetic elements found in all eukaryotic organisms. They are also called ‘jumping genes’ owing to their ability to transpose or change location in the genome. They were first discovered in 1944 by Barbara McClintock who received the Nobel Prize for this in 1983, and was the first woman to receive the Medicine and Physiology Nobel Prize alone. McClintock considered transposons as agents of gene regulation, as part of a

plants response to change in the environment (McClintock, 1984). She first discovered transposons while working on maize kernels and realizing that they were the genetic agents responsible for sectors of altered pigmentation on the maize kernels (McClintock, 1951). TEs were later documented in the fruit fly (*Drosophila melanogaster*), yeast and humans. To date they are known as the single largest component of the genetic material in most eukaryotes (Feschotte *et al*, 2002). The proportion ranges from 4% in the yeast *Saccharomyces cerevisiae* to 44% in humans (Mills *et al*, 2007) and more than 60% in *Zea mays* (maize) (Biemont and Vieria, 2006).

1.3.1 Classes of Transposable elements

TEs are classified into two categories; Class I and Class II based on their mechanism of transposition. Class I TEs are also called retrotransposons while Class II elements are also known as DNA elements. In both classes of transposons there are autonomous and non-autonomous elements. Autonomous elements are able to transpose on their own while the non-autonomous elements are not. The autonomous elements have open reading frames that code for the proteins required for transposition, mainly the enzyme transposase. By using this enzyme the autonomous elements are able to facilitate their own transposition as well as that of the non-autonomous elements. The non-autonomous elements do not have the coding capacity but they maintain in-cis sequences that are necessary for transposition (such as the binding sites for the transposase). Mechanisms of transposition differ between the two classes by way of the transposition intermediate; which may either be RNA or DNA.

Class I elements have what is termed as a ‘copy and paste’ mechanism of transposition and have RNA as the transposition intermediate. In this category of TEs, the TE DNA is transcribed to RNA and then reverse transcribed to DNA. The copy then inserts itself at another place in the

genome. Reverse transcription is catalyzed by an enzyme named reverse transcriptase which is often encoded within the TE itself. The Class I retrotransposons are further divided into two categories; Long Terminal Repeats (LTR) retrotransposons and Non-Long Terminal Repeat retrotransposons. The LTRs are characterized by long terminal repeats in direct orientation (Feschotte *et al*, 2002). The autonomous elements of LTRs have at least two genes in their open reading frames that aid in transposition namely the *gal* and *pol* genes (Figure 1). The *gal* genes encode a capsid like protein and the *pol* genes encode proteins responsible for protease, reverse transcription and integration at a new location. The non-autonomous LTRs do not possess any of the genes responsible for transposition. Non-LTRs are further characterized into Short Interspersed Nuclear Elements (SINEs) and Long Interspersed Nuclear Elements (LINEs). SINEs are non-autonomous and are characterized by an internal RNA *pol III* promoter and a poly-A tail the 3' end. The LINEs are autonomous and have coding genes for the reverse transcriptase (RT) and an endonuclease (EN) (Figure 1). The ORF1 gene of the LINEs is similar to the *gag* gene of the LTR retrotransposons.

Class II TEs have what is termed as a 'cut and paste' mechanism. In this category, the transposon completely excises from one location and inserts into another, leaving behind a footprint or small sequence rearrangement at site of excision. Due to the fact that in this category of transposons an RNA intermediate is not required for transposition, the transposons in this class are also referred to as DNA transposons. In this category of TEs the autonomous elements code for the transposase, while the non- autonomous elements are usually derivatives of the autonomous elements with some deletions on the transposase coding region (Figure 1).

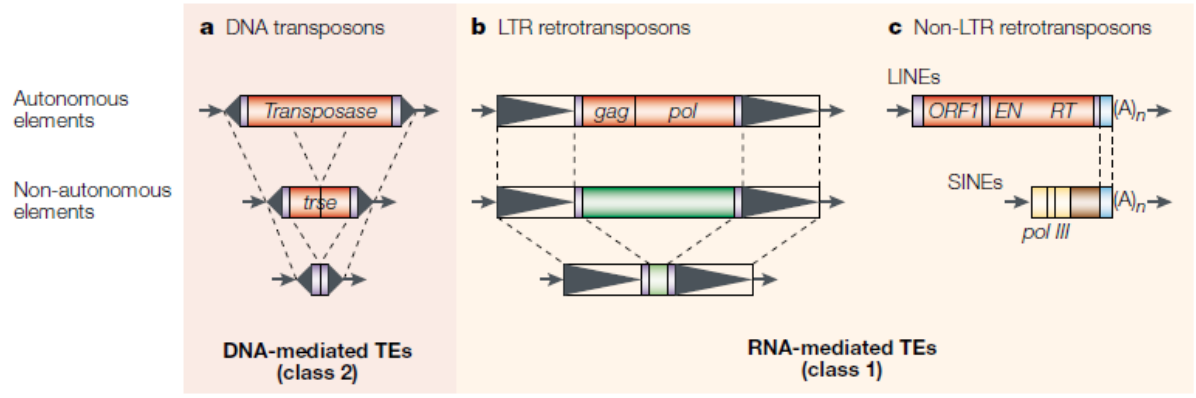


Figure 1: Classes of transposable elements and their structure (adapted from Feschotte *et al*, 2002).

1.3.2 Functions of Transposable Elements

Movement of transposable elements can result in varying consequences on the genome, including expansions to the genome size of the host (Dooner and Weil, 2007). The 'copy and paste' mechanism of transposition of LTR elements renders them the capability of reaching extremely high copy number in plant genomes (Kumar and Bennetzen, 1999). As such LTR retrotransposons are the most abundant type of transposon in plants genomes (Kumar and Bennetzen, 1999). The composition of LTR in genomes of different plants has been documented at 15% in rice (Jiang and Wessler, 2001), as high as 80% in maize (Meyers et al, 2001) and 70% in barley (Vicient et al, 1999).

In addition to contributing to genome sizes TEs are also regulatory elements, that is they have the capacity to alter gene expression. A transposon can insert itself within an open reading frame of a gene and in the process shut down the expression of the particular gene. TEs can also cause deletions and insertions of DNA sequences (Bennetzen, 2005). Integration of almost all TEs results in duplication of short genomic sequences at the site of insertion leading to genetic

variation among members of the same species as well as members of different species (SanMiguel and Bennetzen, 1998). Transposons also provide variation by causing chromosomal rearrangements (Biemont and Vieria, 2006).

1.3.3 Applications of TEs in genomic research

1.3.3.1 *Insertional mutagenesis*

Owing to their ability to induce mutations, transposable elements have been employed in studies aimed at elucidation of gene function as well as discovery of novel genes. One of the early applications of transposons is through insertional mutagenesis. Insertional mutagenesis with a transposon can cause disruption of a gene leading to a mutant phenotype (Aarts *et al*, 1999). The gene can then be isolated using the DNA insert as a molecular probe (Figure 2).

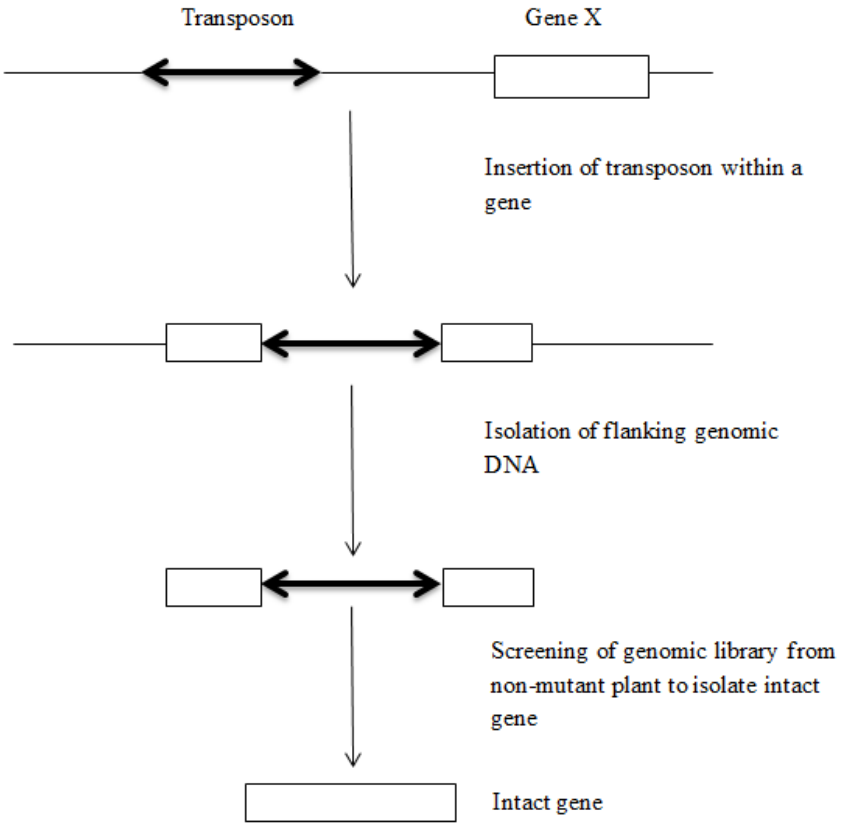


Figure 2: Insertional mutagenesis using a transposon system and isolation of tagged gene

Several types of transposable elements have been used as insertional mutagens. The most common are the *Activator (Ac)* and *Ds (Dissociation)* elements, the *Suppressor-mutator/Enhancer (Spm/En)* and *Mutator (Mu)* which were originally discovered and characterized in maize (Robertson et al, 1978; Fedoroff et al, 1986) but their use has since been extended to other plants such as rice and potato (Table 1). The Ac and En/Spm elements are both autonomous elements belonging to different families (Aarts et al 2000). The Ac elements are capable of activating their own transposition as well as the transposition of members of the Ds elements while the *En/Spm* activate the transposition of the non-autonomous *defective Spm (dSpm)* elements (Aarts et al 2000).

Table 1: Transposon activity in heterologous plants (Adapted from Sundaresan *et al*, 2001)

Plant of Origin	**Transposon Type**	**Activity in Heterologous Plants**
Maize	*Ac/Ds*	Arabidopsis Rice Tomato *Petunia* Flax Tobacco Carrot Lettuce Potato
Maize	*Spm/En*	Arabidopsis Potato Tobacco

Although insertional mutagenesis by transposon tagging has been widely utilized, the method only works when gene knock-out leads to an obvious phenotype. In plants however, as is the case with other eukaryotic systems, disrupting a gene often does not result in obvious phenotypes due to gene redundancy (Bouche and Bouchez, 2001). Many genes in plant cells have duplicates and as such if one gene is mutated, other redundant genes can still carry out the same function and therefore there will be no net phenotypic change. Moreover genes that are required early in development cannot be studied as well because their knock-out might lead to embryonic lethality. Other set of genes that cannot be studied by insertional mutagenesis are those that take part in metabolic pathways that have alternatives. To overcome these problems other strategies for studying gene functions using transposons have been devised.

1.3.3.2 *Enhancer Trapping*

In this method a promoter and a reporter gene (usually the glucuronidase gene) are inserted into a transposon system and inserted into a host genome. The reporter gene expression will then take place only if the transposon inserts next to or within a gene (Figure 3). The expression of the reporter gene makes use of the endogenous enhancer sequence next to the gene of interest and the expression of the reporter gene reflects the activity of the disrupted gene.

1.3.3.3 *Gene Trapping*

Gene trapping also makes use of the expression of a reporter gene. In gene trapping however the promoter region is excluded from the transposon construct. This ensures that the reporter gene gets expressed only if the transposon inserts downstream of an active endogenous plant promoter (Kumar and Nayaranan, 1998). A splice acceptor is included to the immediate upstream of the reporter gene (Figure 3). Addition of the splice acceptor ensures that the reporter gene is transcribed and not spliced out if the transposon were to insert into an intron region of the gene.

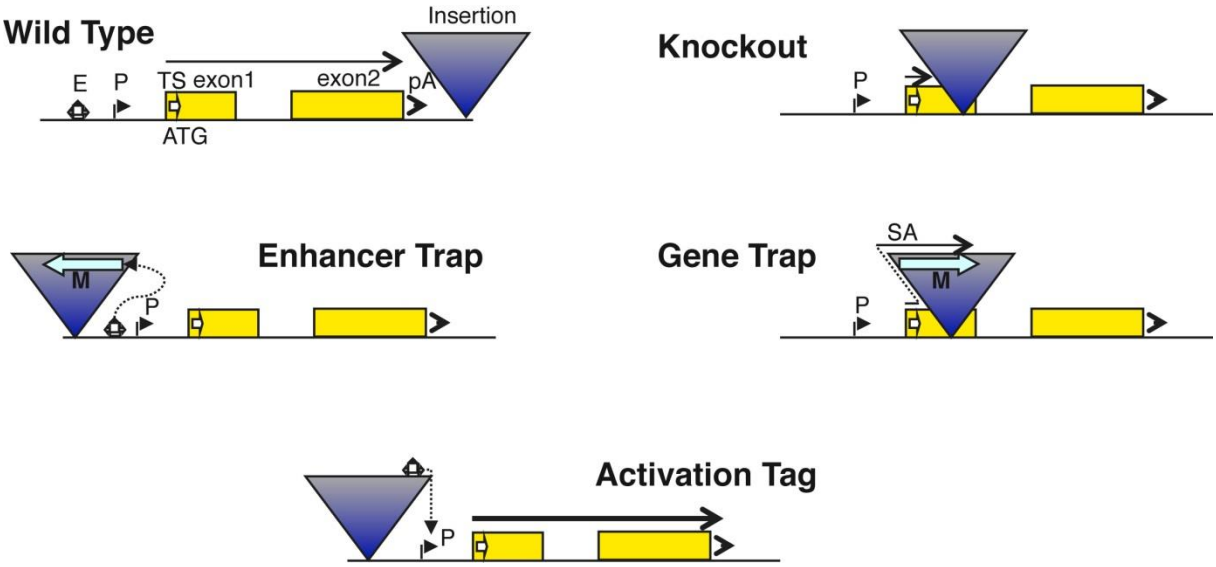


Figure 3: Transposons for Gene Function analysis in different gene constructs. Wild-type gene showing regulatory sequence: E -enhancer elements, P –promoter sequence with TATA box, TS-transcription start site, ATG –translation start site, pA –polyadenylation signal, SA -splice acceptor site. Knockout insertion, within a gene terminating transcription/translation of functional gene product. Enhancer Trap, with transposon bearing a minimal promoter and marker gene (e.g. GUS, GFP). Gene trap insertion within a gene, bearing a splice acceptor site and making a fusion product between host gene and marker gene (GUS/GFP), displaying expression pattern of tagged gene. Activation tag insertion, within close proximity to gene in genome for the activity of the enhancer sequence to drive higher expression of tagged gene (Adapted from Pereira, 2002; Ramchandran and Sundaresan, 2001)

1.4 Activation tagging

Activation tagging is a gain-of-function method in which mutations are created by insertion of activation tag (ATag) elements into the plant genome. The inserts are designed to carry strong activating enhancer sequences that are able to act on genes adjacent to the insertion site and

change their expression (Martinez *et al*, 2002). In activation tagging an enhancer is usually inserted at random into the genome of a host plant. An enhancer is a short stretch of DNA (50-1500 bp) that aids in the activation of the transcription of genes adjacent to its location. An enhancer affects genes both downstream and upstream of its location usually by interacting with activator proteins and promoters upstream of genes and in the process enhancing the transcription of the genes.

Activation tagging normally uses a tetrameric repeat of the enhancer element of the cauliflower mosaic virus (CaMV) 35S gene to direct the transcription of adjacent genes and in the process creating dominant mutations (Robinson *et al*, 2001). The inserts can be incorporated into the genome either through Transfer DNA (T-DNA) or using a transposon system. In activation tagging the presence of the enhancer can elevate the endogenous expression of the genes adjacent to it, and therefore the phenotypic changes that result from the increased gene expression most likely reflect the normal role of the gene (Wang *et al*, 2009). With the use of activation tagging the function of genes that act redundantly can be elucidated, a task that is difficult with normal loss-of-function mutagenesis. Activation tagging is also useful for identification of genes that take part in metabolic pathways that have redundant genes. In loss-of-function mutagenesis, if a gene in one metabolic pathway is affected, and the pathway has an alternative gene copy, there will be no resultant phenotypic change. Genes that function at different stages of the lifecycle of a plant can also be elucidated through activation tagging, for example genes that function at the embryo stages of a plant, whose knock-out would result in embryonic lethality (Tani *et al*, 2004). In activation tagging, instead of mutating a gene and causing death of the plant, expression of the gene is elevated, resulting in an overexpression mutant phenotype and hence the function of the gene can be studied.

1.5 The Ac/Ds Transposon System

The *Activator* (*Ac*)/*Dissociation* (*Ds*) system is a maize transposon system consisting of the autonomous Ac transposable element that encodes the transposase, and a non-autonomous Ds element. The Ac element is 4565 bp long and has a 11 bp terminal inverted repeat (TIR) (Muller-Neumann et al, 1984). The region encoding the transposase in the Ac element is 3.5kb in size and produces a protein (the transposase) that is 807 amino acids long (Kunze et al 1987). For transposition to occur the transposase binds to a 200 bp region on both ends of the Ac element which also encompasses the 11 bp repeat motif.

The Ds element is dependent upon the transposase encoded by the Ac element for transposition. This element also harbors the same 11 bp TIR found in the Ds element (Kumar and Nayaranan, 1998). In studies aimed at elucidating gene function such as activation tagging, the Ds element is usually modified to carry an enhancer that is to be inserted into the genome. Since the Ds element can only move in the presence of the Ac element, the Ds element therefore can be stabilized by segregation of the Ac transposon thereby creating an activation tag. In studies utilizing the Ac/Ds system to create mutations, the Ac portion in the system is usually truncated so as to immobilize it thus allowing only the Ds portion to move. This allows the Ds element and the gene of interest to be fixed once established itself at a new location on the genome.

Selectable marker genes such as kanamycin resistance gene and hygromycin resistance genes are usually also included in the system to allow for selection of the transformants either through tissue culture or by application of the herbicide to the plants once they are grown (Carter *et al*, 2013). The selectable marker genes may also be the target for PCR reactions aimed at identifying positive transformants. Separate selectable marker genes are usually included for both the Ac and Ds portions of the system. This is so as to identify those plants that contain only the Ds element,

which would be the stable mutants since in the absence of the Ac element, the Ds element cannot transpose.

In activation tagging, the Ds element is normally modified to carry the enhancer sequence usually the strong and well characterized CaMV 35s enhancer. Use of the Ac/Ds transposon system allows many independent mutants to be generated from a single transformed line because of the ability of the Ds element to transpose to different parts of the genome. Rapid identification of the mutagenized or affected sequences is also possible through the isolation of the Ds insertion flanking regions using the polymerase chain reaction (Kuromori *et al*, 2004). The Ds flanking regions can then be purified and sequenced. The Ac/Ds system has been used to study gene function in plant species such as Arabidopsis and rice (Altmann *et al*, 1995; Chi *et al* 1999).

1.6 Activation tagging in tomato

In this study activation tagging has been used to create mutations in order to elucidate gene function in the tomato cultivar M82. Tomato was selected for study because of its economic and nutritional importance worldwide. Consumption of tomato has been related to reduced cases of coronary heart disease and prostate, breast and colon cancers (Weisburger, 2002). This is because tomato is rich source of antioxidants such as lycopene and beta-carotene. Tomato also serves as a good model organism for other dicot species as well as for other plants in the Solanaceae family (Carter *et al*, 2013). The cultivar M82 in particular was chosen because it is conducive for greenhouse conditions and has prolific seed production (Carter *et al*, 2013).

Several studies have been carried out to identify gene function in tomato using activation tagging. A gene responsible for anthocyanin biosynthesis was identified using activation tagging in the tomato cultivar Micro-Tom (Mathews *et al*, 2003). The mutant had an intense purple

coloration resulting from overexpression of a gene that encodes a MYB transcription factor. The function of the gene was verified by overexpressing it in tobacco and the same intense purple coloration was observed (Mathews *et al,* 2003). A gain-of-function principle was also used to identify a gene involved in flower and fruit development in tomato. Ectopic overexpression of a MADS-box domain related gene led to the absence of sepals in tomato and a fleshy ripening organ was observed instead in place of the sepals (Gimenez *et al,* 2010). The mutation came about as a result of rearrangement of the T- DNA during the integration process, resulting in a truncated 35S promoter placed in reverse orientation to the tagged gene. As a result of the phenotype observed, the gene was identified as one of the genes responsible for regulation of flower and fruit development in tomato.

Many other genes remain uncharacterized in the tomato genome. The economic and nutritional importance of tomato as well as its importance in molecular research makes it an attractive target for continued study. With the recent completion of the tomato whole genome sequencing project and subsequent publication of the tomato genome (Tomato Genome Consortium, 2012), the present study will add to knowledge of the tomato genome by identifying putative functions of genes not previously characterized.

2.0 Objectives

Objective 1.

Characterize putative tomato transformant lines with an Ac-Ds Activation Tag construct by PCR, for evidence of transformation, and transpositional activity of the Ds A-Tag. Identify actively transposing lines that would be useful to generate a population of Ds A-Tag inserts.

Objective 2.

Screen the active A-Tag lines for gain-in-function mutant phenotypes and identify the tagged genes, overexpressed in the mutant responsible for the phenotype.

3.0 Materials and Methods

3.1 Optimization of germination process for tomato seeds

Seeds of the tomato cultivar M82 were subjected to 5 different treatments with 10 seeds per treatment. The treatments used comprised of a) 1/2 strength MS medium with no hormone supplement, no sugar and in the light b) Seeds were imbibed in water by keeping on moist filter paper and kept in the dark c) water imbibed seeds were kept in the light d) water imbibed seeds were kept in the dark and supplemented with 1µM gibberellic acid (GA_3) e) water imbibed seeds were kept in the light and supplemented with 1µM GA_3. The germination percentages in each case were recorded and the treatment with the highest germination efficiency was selected for germination of tomato seeds for further experiments.

3.2 Plant Growth Conditions

Seeds of putative tomato cultivar M82 transformants obtained from previous research at Virginia Tech were germinated by imbibition with water in the dark for approximately 48 h. Pre-germinated seeds were transplanted into small pots and grown for 10 days in the greenhouse with day/night temperature of $22/26^0$C, light intensity of 600 $\mu molm^{-2}s^{-1}$ with light/dark cycles of 10/14 h. After 10 days the plants were transferred to medium size pots and grown to maturity. The pots were placed in water filled trays, with periodic fertilizations using 24-8-16 Miracle-Gro (Scotts Miracle-Gro Product).

3.3 Genomic DNA isolation

Leaf samples were collected from young seedlings (15-20 days) and DNA was isolated from the leaves using a modified CTAB protocol (Doyle and Doyle, 1987). Fresh green leaf tissue was homogenized in 1ml of 2X CTAB buffer (pre-warmed at 65^0C) with addition of 2% BME. The homogenate was incubated at 65^0C for 45 min with occasional swirling. An equal amount (1 ml)

of chloroform:isoamyl alcohol (24:1) was added, and the tube was inverted several times to mix the contents for 10-15 min. Samples were centrifuged at 8000 rpm for 10 min at room temperature. The upper aqueous layer was collected and the chloroform isoamyl alcohol extraction was repeated. Afterwards, $1/10^{th}$ volume of sodium acetate and 2.5 vol ice-cold absolute ethanol were added. The mixture was incubated at 4^0C for 1h. The nucleic acid was pelleted down by centrifugation at 8000 rpm for 10 min. The pellet was washed with 70 % ethanol and air-dried overnight. The pellet was then suspended in 70 µl sterile distilled water. The amplified PCR products were run in 1% agarose gel using 1X TBE buffer stained with 0.5X gel red stain, visualized and photographed using UVP Gel Doc-it ^{TS2} Imager.

3.4 Polymerase Chain Reaction (PCR)

Genomic PCR was carried out using a PCR mixture that consisted of 3µl template DNA (20 ng), 4 µl 10x PCR buffer, 3 µl (25mM) $MgCl_2$, 0.5 µl of each of the reverse and forward primers (10 µM), 1 µl dNTPs (10mM), 0.25 µl Taq polymerase and 6.75 µl water. The reactions were carried out using a standard PCR program of initial denaturation at 94^0C for 60 seconds, followed by 30 cycles of 94^0C for 30 seconds, T_m (melting temperature) at 55^0C for 60 seconds, 72^0C for 90 seconds and final extension of 72^0C for five min followed by storage at 4^0C.

3.5 Excision PCR

To demonstrate the ability of the Ds element to transpose in the tomato genome, primers from sequences flanking the Ds element were designed at the Ac/Ds promoter junction and the Right Border (RB) of the T-DNA construct, and used to amplify the empty donor site (EDS) to reveal excision of the Ds element. Genomic PCR was carried out using these primers to test if the Ds element excised from its original location in the construct. The reaction mixture contained 3 µl template DNA, 0.5 µl of each of the forward and reverse primers, 10 µl of PCR buffer and 6 µl

sterile distilled water. The PCR cycle was set as follows: initial denaturation at 94^0C for 1 min, followed by 34 cycles of 94^0C for 2 min, 58^0C for 1 min, 72^0C for five min and final extension at 72^0C for 10 min followed by storage at 4^0C. The products of the PCR reaction products were then run on a 1.2% agarose gel and the bands cut out and purified.

3.6 Gel Purification

The purification process was carried out using the EZNA Gel Extraction Kit from Omega Bio-Tek Inc. USA following the manufacturer's protocol with slight modification. Briefly, the extracted gel pieces were weighed and an equal volume of binding buffer was added. The mixture was incubated at 65^0C for 10 min with vortexing at 2-3 min intervals. About 700 µl of the suspension was then added to a column and spun at 13000 rpm for 1 min. The process was repeated until all the sample was transferred to the column. The supernatant was then discarded and 300 µl binding buffer was added followed by centrifuging at 13000 rpm for 1 min. The supernatant was discarded and 500 µl Wash Buffer was added. Spinning was then carried out at 13000 rpm for 1 min, the supernatant discarded and the empty tube was spun again at 13000 rpm for 2 min. The collection tube was discarded and 20 µl of nuclease free water was added to the column. Centrifugation was done at 13000 rpm for one min and the column was discarded. The DNA was then stored at -4^0C prior to cloning.

3.7 Cloning of Excision PCR products

3.7.1 Ligation of DNA

5X Ligase Reaction Buffer 2 µl

Insert: Vector Molar Ratio 3:1

Vector: 4 µl

Insert: 2 µl

T4 DNA Ligase: 0.5 µl

Autoclaved distilled water: 1.5µl

Total Volume: 10 µl.

The reaction mix was incubated at room temperature for 2h and then stored at -20^{0}C till further use.

3.7.2 Preparation of competent cells

0.01% of DH5α cells from glycerol stock were inoculated in 5ml LB medium and incubated overnight at 37^{0}C with shaking. Next day 20ml LB medium was inoculated with saturated overnight culture (0.1%) and incubated at 37C for 2^1/2 h or until the OD reached 0.6. Cells were centrifuged at 8,000 rpm 5 min, 4^{0}C. Supernatant was discarded and resuspended in 1ml, 100mM $CaCl_2$. Centrifuged at 8000 rpm, 5 min 4C. Supernatant was discarded and pellet resuspended in 200µl $CaCl_2$ and incubated in ice for 30 min. Next 2µl ligation mix was added followed by incubation in ice for ten minutes. Heat shock was then carried out at 42^{0}C for 90s and immediately put in ice. Then 800 µl of LB medium was added followed by incubation at 37^{0}C for 90 min. The mixture was then spun down at 8000 rpm for 5 min at 4^{0}C. The supernatant was discarded and 100 µl of LB medium was added and then spread on agar plates with antibiotic and incubated overnight at 37^{0}C.

3.7.3 Isolation of Plasmid DNA

Isolated single colonies were picked up and incubated in 5ml LB medium with antibiotic (Ampicillin) and incubated overnight with shaking at 37^{0}C. Cells from overnight culture were harvested at 8000 rpm for 5 min. The supernatant was discarded and the process repeated. Next 100 µl of Solution 1 was added followed by vortexing. Then 200 µl Solution II was added

followed by incubation for 2 min. Afterwards, 350 µl Solution III was added followed by centrifugation at 13000 rpm for 10 min. The supernatant was then added to a column followed by spinning at 13000 rpm for 1 min. The flow-through was discarded and 500 µl of HBC buffer was added to the column and then spun at 13000 rpm for 1 min. Next 300 µl of DNA Wash Buffer was added followed by spinning at 13000 rpm for 1 min. The supernatant was discarded and the empty column was spun at 13000 rpm followed by final elution of the DNA in 35µl sterile distilled water. The DNA was then stored at -4^0C prior to sequencing.

3.8 DNA Sequencing

The reaction mix was prepared as per the combination listed below and the volume was made up to 13µl with autoclaved double distilled water. Sequencing was carried out using the ABI Sequencer 11300 from ABI, USA at the University core lab.

Template	**Amount**	**Primer (pmol)**
Double Stranded DNA (plasmid)	300-500 ng	3.4 pmol
PCR Product:	PCR products (<100 bp do not usually work)	
100-200 bp	1-3 ng	3.4 pmol
200-500 bp	3-10 ng	3.4 pmol
500-1000 bp	5-20 ng	3.4 pmol
1000-2000 bp	10-40 ng	3.4 pmol
>2000 bp	40-100 ng	3.4 pmol

3.9 Thermal Asymmetric Interlaced PCR (TAIL-PCR)

Three rounds of TAIL-PCR (Carter et al, 2013) were performed to identify the tagged genes. The Primary TAIL PCR reaction consisted of 2 µl template DNA 0.5 µl forward primer, 3 µl degenerate primer, 10 µl PCR buffer and 4.5 µl sterile distilled water. The Primary Tail PCR product was then diluted 1: 60 and the diluted product was used as the template for the secondary TAIL PCR reaction. The Secondary Tail Reaction consisted of 2 µl template DNA, 0.5 µl forward primer, 3 µl degenerate primer, 10 µl PCR buffer and 4.5 µl sterile distilled water. The Secondary TAIL-PCR reaction was then diluted 1:60 and used as the template for the tertiary reaction which comprised 2 µl template DNA, 1 µl forward primer, 5 µl degenerate primer, 20 µl PCR buffer and 12 µl sterile distilled water. The tertiary TAIL-PCR products were then run on 1% agarose gel and the bands excised out and purified as described above. Sequencing was then carried out as described above and the sequences were aligned to tomato genome using Phytozome 9.1 (http://phytozome.jgi.doe.gov/pz/portal.html) for identification.

3.10 Gene Expression Analysis

3.10.1 RNA Isolation

Total RNA was isolated from the leaf or fruit samples using Trizol (Invitrogen). To isolate RNA, the tissue was ground in liquid nitrogen and 1 ml of Trizol was added and incubated at room temperature for 5 min. Next, 200 µl of chloroform was added and the solution was mixed well and then incubated at room temperature for 2 min, and the tubes centrifuged at 12000 rpm for 15 min at 4^0C. The supernatant was collected and 500 µl of chilled isopropanol was added. The reaction mix was incubated at room temperature for 10 min and the tubes centrifuged at 12000 rpm for 10 min at 40^0C. The supernatant was discarded, 500 µl of 75% ethanol added, and tubes

centrifuged at 7500 rpm for 5 min at 4^0C. The pellet was allowed to semi-dry, 35 µl of nuclease free water added, and the RNA was stored at -80^0C.

3.10.2 cDNA synthesis

For cDNA synthesis, 2 µg of RNA was taken for each reaction. RNA was made DNA-free by incubating with 1 µl (25 units/µl) DNAse 1 at room temperature for 15 min, 1 µl of 25 mM EDTA was added to inactivate DNAse 1 and incubated at 65^0C for 15 min. The reaction mix was then kept in ice for 15-20 mins. The RNA was then mixed with 4 µl (25mM) $MgCl_2$, 2 µl (10 mM) dNTPs, 1 µl (500ng/µl) oligo-dT, 2 µl 10xPCR buffer, 0.5 µl Reverse Transcriptase (25 units/µl), and 0.5 µl (40 units/µl) RNasin. The Reverse Transcriptase reaction was carried out at 42^0C for 1 h, 95^0C for 5 min and then stored at 4^0C.

3.10.3 Quantitative PCR (qPCR)

To set up the qPCR, 10 µl of reaction mixture was used which comprised 2 µl cDNA, 5 µl qPCR buffer (GoTaq® qPCR Master Mix, Promega), 0.5 µl of each of the forward and reverse primers and 2 µl H_2O. The qRT-PCR experiments were conducted using GoTaq® qPCR Master Mix (Promega), gene-specific primers, and Ubiquitin as standard with three biological replicates in a CFX-96 Bio-Rad thermocycler (Bio-Rad). Increasing temperature (0.5°C 10 s^{-1}) from 55°C to 95°C was used for melt curve analysis. Un-transcribed RNA was also run as negative control2. The relative difference in expression for each sample in individual experiments was determined by normalizing the Ct value for each gene against the Ct value of Ubiquitin and was calculated relative to the calibrator using the equation $2^{-\Delta\Delta Ct}$

3.11 Stress Phenotyping

Stress response of the tomato plants was tested by first germinating the seeds for 48 h in the dark. The seedlings were then transferred to MS medium containing a) 100 mM mannitol b) 100 mM NaCl c) MS medium with no supplements. The seedlings were allowed to grow in the different treatments for a period of 10 d. After 10 days, the decrease in shoot length was measured as well as the increase in root length and the results were presented as percentage reduction.

3.12 Biochemical Assays

3.12.1 Preparation of Methanolic Extract

100 mg of fruit tissue was ground in liquid nitrogen and 1ml of absolute methanol was added. The mixture was centrifuged at 12000 rpm for 15 min at 4^0C. The supernatant was collected and used for subsequent biochemical analyses.

3.12.2 Estimation of Total Phenolic Content

To obtain an estimate of the total phenolic content in the fruit tissues, 0.5 ml of methanolic extract was mixed with 2.5 ml of 10-fold diluted Folin- Ciocalteu reagent and 2 ml of 7.5% (w/v) sodium carbonate. The mixture was allowed to stand for 30 min at room temperature (25^0C) and the absorbance was measured at 760 nm.

3.12.3 Determination of carotenoid content

To determine carotenoid content the tissue samples were thawed in the dark in a refrigerator at 4^0C to avoid carotenoid oxidation. 16 ml of acetone: hexane (4:6) solvent were added to 1.0 g of tomato homogenate and mixed in a test tube. Automatically, two phases separated and an aliquot was taken from the upper solution for measurement of absorbance at 663, 645, 505 and 453 nm

in a spectrophotometer. Lycopene and beta-carotene contents were calculated according to the equations:

Lycopene (mg/100 ml of extract) = $-0.0458 \times A_{663} + 0.204 \times A_{645} + 0.372 \times A_{505} - 0.0806 \times A_{453}$

β-carotene (mg/100 ml of extract) = $0.216 \times A_{663} - 1.22 \times A_{645} - 0.304 \times A_{505} + 0.452 \times A_{453}$

3.12.4 Statistical Analysis

The Tukey's Test and One Way ANOVA in Graph Pad Prism 5 (www.graphpad.com/scientific-software/prism/) were used to estimate significant differences in the level of lycopene, beta-carotene and total phenolic content between the control and the different mutants.

3.13 Harvesting of fruits and storage of seeds

To collect the seeds for storage, the seeds were first separated from the pulp and fruit. An equal volume of 3N HCL was then added to the seeds in the beaker and left to stand for 15 min. The beaker was then filled with water to dilute the acid. Afterwards the seeds were rinsed in a strainer under a sink tap. The seeds were returned to the beaker and an equal volume of 10 % Na_3PO_4 was added and the mixture was allowed to stand for 15 min. Water was then added to fill the beaker and the seeds were rinsed under a sink tap. The seeds were placed in an oven at 37^0C overnight to dry. The seeds were then collected in envelopes and appropriately labelled and stored.

4.0 Results

4.1 Screening of germination process for tomato seeds

For selecting the optimum method for tomato seed germination tomato seeds of the cultivar M82 were subjected to five different treatments; i) the seeds were grown in ½ strength MS medium with no hormone supplements, no sugar and in the light, ii. Water imbibed under light and dark conditions and iii. Water imbibed in light and dark conditions supplemented with 1µM GA_3. In the presence of light germination was low while the maximum germination was observed under water imbibition in the dark. The results are summarized in Table 2. Imbibition in water by keeping the seeds on moist filter paper under dark conditions at room temperature was therefore selected as the best method for germination of tomato seeds.

4.2 Phenotyping progeny of transformants

Ten different putative transformant lines (named M1-10), obtained from independent transformation events (Dragana Avirovik, 2013 MS thesis) were used in this study. The plants had been transformed with *Agrobacterium tumefaciens* T-DNA modified to harbor an Ac/Ds construct (Figure 4), and were from independent transformation events selected for Hygromycin in media and showed resistance to Basta herbicide in painting assays on adult plants, but no molecular experiments had been done. The construct had two components of the in-cis Ac/Ds system and harbored four copies of the Cauliflower Mosaic Virus CaMV 35S enhancer in the Ds element (Figure 4). The hygromycin resistance and herbicide (Basta) resistance genes served as the transformation and plant-selectable marker genes, respectively. Eight seeds from respective lines were pre-germinated as described above, and allowed to grow in soil. The different putative tomato transformants were allowed to grow to maturity in the greenhouse. From all the different transformant lines screened, more than half of the plants grown had mutant phenotypes (Table

3). Several different mutant phenotypes were observed. One plant from the M1 line (M1/5) showed necrotic spots on the leaf (Figure 5). The mutant was named *necrotic* leaf because of the spots observed. From the M4 line four out of six plants that reached maturity showed dwarf phenotypes (Figure 6). One of the plants in the M5 line (M5/3) showed stunted growth, leaf discoloration and elongated leaf shape (Figure 7). This mutant was named *brittle* mutant because the phenotype showed leaves that were very fragile and easily broke when handled. The plants in the M6 line also displayed a dwarf phenotype similar to the one observed in line M4 (Figure 8). In the M7 line one plant was observed that had severely stunted growth and hardly produced leaves (Figure 9). The plant however proliferated vegetatively throughout the life cycle. An irregularly shaped leaf was found in one of the lines (M8). The leaf had a palmatipartite shape as opposed to the normal lobate structure of the tomato leaf (Figure 10). The lines in which unique phenotypes were observed, *Necrotic* from M1 and *Brittle* from M5, were selected as the focus of the rest of the study. The first step was to verify that the plants in these lines were transformants.

4.3 Molecular Evidence of Transformants

Genomic DNA was isolated from 15-20 day old seedlings. PCR was then performed on the isolated genomic DNA using primers for the tomato squalene epoxidase (SQE) gene as a positive control to check the DNA. The PCR products were run on a 1% agarose gel (Figure 11). The gel shows eleven plants on which the genomic PCR was carried out. Lanes 1-6 are plants from the M1 line that reached maturity, and lanes 7-11 are the M5 plants that reached maturity. For some plants (M1-5 in lane 5 and M5-5 in lane 11) the PCR reaction did not work and had to be repeated. After the quality control check screening for transformation of the different plants by PCR was initiated.

4.4 Identification of putative transformants using PCR for components of construct

4.4.1 *Hygromycin phosphotransferase (HPT)*: The transformation marker *HPT* on the construct was used to select positively for transformants for resistance to hygromycin in selective media during shoot regeneration (Avirovik D., 2013). As no molecular evidence of transformation was available, DNA was isolated from progeny of putative transformants and PCR was carried out using primers specific to the hygromycin resistance gene (Table 4). The PCR products were run on 1% agarose gel and the results are shown in Figure 12. The gel represent 6 plants from the M1 line and 5 plants from the M5 line. 3 out of 6 plants from the M1 line tested positive for HPT and 4 out of 5 plants from the M5 line tested positive for HPT (Figure 12).

4.4.2 *BAR* gene: The *BAR* gene conferring resistance to the herbicide Basta/glufosinate was the plant selectable marker gene that was included in the construct (Figure 4). The gene was under the control of the maize Ubiquitin promoter (Figure 4) that confers resistance to the herbicide Basta/glufosinate. The *BAR* gene is also a very important part of the construct as it is located within the Ds element which also included four copies of the CaMV 35S enhancer. It is the CaMV 35S enhancer that was expected to affect expression of the genes adjacent to the Ds element in the different transformants. The putative transformants were tested with primers targeting the Bar gene and the results are shown in Figure 13. All the plants from the two lines (M1 and M5) tested positive for the Bar gene (Figure 13).

4.5 Validation of transformants

To further confirm the presence of different parts of the construct, and therefore for complete functionality of the transformants, an additional PCR reaction was set up targeting anther part of the construct:

4.5.1 ***Gos-2 Promoter (Gos)***: The final set of primers used targeted the maize *Gos-2* promoter. The *Gos -2* promoter is a constitutive promoter and was used to drive the expression of green fluorescent protein (GFP) gene (Figure 1). Under the control of this promoter the GFP is expected to be expressed giving a green pigmentation in plant tissue, which can be viewed under ultra violet radiation. Expression of GFP is also another way of identifying transformants. Primers designed specifically for the Gos promoter were used in the PCR reaction and the results run on a gel (Figure 14). Four plants from the M1 line tested positive for Gos promoter and four plants from the M5 line were positive for the Gos promoter (Figure 14).

4.6 Molecular Evidence of Excision by PCR

To demonstrate activity of the transposon construct in the transformant lines, PCR was performed to test excision of the Ds element from the construct in the 11 plants from the two selected lines. The excision PCR was used to amplify the empty donor site (EDS) between the right border and Ac promoter in the construct, to demonstrate the absence of the Ds element in its original site and showing evidence of excision. Amplification between the primers would occur only of the Ds had excised, as the Ds element in the original position would not reveal PCR products under the conditions used. Excision events of the Ac/Ds system were monitored by the amplification of the excision product footprint using a primer annealing to the Right Border (RB) of the T-DNA and another one annealing to the Ac promoter (AcPr). The PCR products were run on a gel shown in Figure 15. Three plants from the M1 line showed excision while all plants from the M5 line showed excision of the Ds element. The bands from the eight positive results were then cut out and purified. After purification the PCR products were re-amplified in preparation for cloning and then purified again.

The gel purified excision PCR products were ligated to pGEMT and transformed into *E. coli* (*DH5α*). After transformation, isolated single colonies were inoculated for plasmid isolation. The plasmid DNA isolated from transformants was used for PCR screening using the excision PCR primers and the product run on a gel for visualization (Figure 16). The plasmid DNA was then sent for sequencing but the quality and read length obtained were not good enough to make a substantial conclusion. The sequence flanking the Ds element comes from the *waxy-m7* mutant of maize bearing the Ac element, which has a GC content of over 70% that had made it difficult to obtain (Carter et al., 2013).

4.7 Isolation of Ds transposon flanking DNA by TAIL-PCR

TAIL-PCR was used to identify the insertion sites of the transposed element in the eight plants that were positive for the excision PCR. TAIL-PCR is a powerful method used for the recovery of DNA fragments adjacent to known sequences. The method utilizes three nested primers in consecutive reactions alongside an arbitrary degenerate (AD) primer. The degenerate primer has a lower melting temperature and binds loosely while the nested primers have increased specificity and they bind to the known sequences. TAIL-PCR proceeds in three steps; Primary TAIL-PCR, Secondary TAIL PCR and finally Tertiary PCR. The primary reaction uses the primer furthest from the tagged gene, the secondary reaction uses the primer in the middle of the known sequence and the tertiary reaction uses the primer that anneals towards the end of the known sequence and is also the closest to the tagged gene. In this way the specificity of the process is increased with each stage. In this study three primers were designed specifically for the Ds element (Ds3-1, Ds3-2 and Ds 3-3). These primers were used sequentially in three reactions alongside an arbitrary degenerate primer in order to amplify the DNA adjacent to the Ds element. The primary reaction utilized Ds 3-1, the secondary reaction utilized Ds 3-2 and

finally the tertiary reaction utilized the Ds-3-3 primer. The products of the secondary and tertiary TAIL-PCR reactions were run on 1% agarose gels (Figure 17). For the M1 line, all the three plants' insertion sites were successfully amplified at the tertiary TAIL-PCR stage and for the M5 line insertion sites were amplified for two of the plants.

The results of all the PCR reactions are summarized in Table 5. One plant from the M1 line (M1-5) tested positive for all the PCR reactions (Table 4) and also had a mutant phenotype. The mutant phenotype was the necrotic leaf observed at the phenotyping stage and hence the mutant had been named *necrotic*. In line M5 two of the plants tested positive for all the PCR reactions (M5-3 and M5-4) but only one plant (M5-3) had a mutant phenotype. The M5-3 plant had an abnormal shape that was discolored and easily broke. The mutant was named *brittle*. A decision therefore was made to pursue the two plants that had tested positive for all the PCR reactions and had mutant phenotypes. The bands for these plants were excised from the Tertiary TAIL-PCR gel and purified in preparation for sequencing of the tagged genes.

4.8 Sequencing of TAIL-PCR products

The purified TAIL-PCR product was sent for sequencing, that was carried out with the primer used for the tertiary TAIL-PCR reaction (Ds3-3) in an automated sequencer. A chromatogram was produced with the different sequences and the sequences were identified against the tomato genome using Phytozome 9.1. For the *brittle* mutant the Ds element was found to have inserted between two genes with locus IDs Solyc01g079070.2 and Solyc01g079080.2 respectively. The sequences are shown in a chromatogram (Figure 18). The genes were identified as those coding for a protein involved in vacuolar protein sorting and coiled coil protein respectively. For the *necrotic* leaf mutant the tagged gene was identified as a gene coding for a protein kinase, the output from the sequencer for the gene is shown in Figure 19.

4.9 Analysis of Expression of Tagged Genes by Quantitative PCR

To study the expression of the genes following tagging by the Ds element, total RNA was isolated from the mutants, cDNA was synthesized and qPCR was carried out. qPCR was done using the primers specific to the tagged genes which were designed after obtaining sequences of the genes. The results of the qPCR are shown in Figure 18. For the *Brittle* mutant the two tagged genes were highly induced by the presence of the Ds element (Figure 20A). The kinase gene that was tagged in the *necrotic* leaf mutant was also highly induced by the presence of the Ds element (Figure 20B). The induction of the tagged genes was expected due to the presence of the CaMV 35s enhancer embedded in the Ds element.

4.10 Stress Response Phenotype of Activation Tagged Mutant Lines

To test the performance of the two mutants under different stresses the plants were subjected to three different treatments; MS medium supplemented with a) 100 mM Mannitol b) 100 mM NaCl and c) MS medium with no supplement. These stresses were representative of the stresses commonly experienced by plants such as drought and salinity. The mannitol served as an osmoticum that conferred water deficit to the plant and sodium chloride provided salinity stress. The newly germinated seedlings were placed in the different media and allowed to grow for a period of ten days. After ten days the percentage increase in root length was noted as well as the percentage decrease in shoot length (Figure 21). Under salt stress and the *brittle* mutant exhibited tolerance. This was indicated by a decrease in shoot length that was below 40% (Figure 21A) and an increase in root length that was below 40% when the plant was subjected to 100 mM NaCl treatment (Figure 21B) The *necrotic* leaf mutant however proved susceptible to the stress. The mutant had a decrease in shoot length of more than 40% (Figure 21C) and an increase in root length of more than 40% when subjected to salt stress (Figure 21D). When subjected to

100mM mannitol both plants died almost immediately making the analysis of their response to the stress very difficult. This observation suggest that these mutants can be sensitive to osmotic stress more precisely dehydration.

4.11 Gene Expression Analysis of Tagged Genes in Response to Stress

Total RNA was isolated from the seedlings following stress application, cDNA was synthesized and qPCR was done to find out the effect of the stress on the expression of the tagged genes. Under both stresses all the tagged genes showed high induction after stress treatment. For the *brittle* mutant the coiled coil protein gene showed a 3X fold change increase while the same gene showed a 1X fold change increase in the wild-type following salt stress application. (Figure 22A). The gene involved in vacuolar protein sorting showed a 4X fold change following salt stress in the mutant while in the wild-type the gene showed a 1X fold change (Figure 22B). In the *necrotic* leaf mutant the calmodulin protein kinase gene showed a 5X fold increase following salt stress (Figure 22C).

4.12 Phenotypic Screen of Tagged Mutants for Tomato Fruit Quality

Upon reaching maturity, biochemical analyses were carried out on the fruits of the *necrotic* plant as well as other fruits from different lines that showed mutant phenotypes (with three different fruits from each plant serving as biological replicates). This was done so as to establish the impact if any, of the insertions on the tomato fruit quality. The *brittle* mutant did not bear any fruit and as such was not included in the biochemical assays.. The assays were carried out because tomato is a rich source of antioxidant and is very important nutritionally worldwide. Total Phenolic content, carotenoids and lycopene content of the mutants were estimated via spectrophotometry. The results of the biochemical assays are shown in Figure 23. For lycopene content the highest amount observed in the fruit from mutant line 9 (M9) with approximately

0.13 mg/g while the lowest was observed in the fruits from the *necrotic* leaf with approximately 0.06 mg/g (Figure 23A). The *necrotic* mutant had significantly lower lycopene content than the control (Figure 23A), as did the fruits from the M7 and M4 Lines. For beta-carotene content the highest amount was observed in the fruits from M7 line with an approximate value of 0.055 mg/g and the lowest was in the *necrotic* leaf with an approximate value of 0.025 mg/g (Figure 23B). The fruits from the *necrotic* mutant had significantly lower beta-carotene as compared to the control, while those from other lines did not have any significant difference from the control (Figure 23B). The highest total phenolic content was observed in the control with approximately 0.85 mg/g and the *necrotic* leaf mutant had about 0.7 mg/g while the lowest amount observed was in the fruits from the M6 line with 0.6 mg/g (Figure 23C). The fruits from M7, M6 and M4 lines had significantly lower total phenolic content than the control. The deviation of nutritional content between the mutants and the control suggests that the tagged genes have a role to play in nutrient accumulation pathways. The next step will be to isolate the DNA from the plants that were not characterized in this study (M4, M6, M7 and M9) and identify the tagged genes.

5.0 Discussion

Regulation of gene expression is very crucial to plants. Genes can be expressed in response to various stimuli, at certain stages of development and in different tissues depending on the function of the tissue. Plant cells, just like other eukaryotic cells have endogenous means of varying expression of genes. This can be achieved by use of regulatory sequences such as enhancers and silencers, RNA interference and post-translational modification by way of protein modification. One of the most interesting ways by which gene regulation is achieved endogenously is through transposable elements. Transposable elements can shut down gene expression as well as induce mutations. Owing to the natural ability of transposons to affect gene expression they have been used in the study of genomes as insertional mutagens. The most common use of transposon has been a reverse genetics approach where a gene is knocked out and the resulting phenotype observed. However due to high gene redundancy in eukaryotic genomes the knockout mutagenesis has not always been very effective. As such a different mechanism for studying gene function would prove to be of more use and one of such ways is through activation tagging.

Activation tagging is a very powerful method for elucidating gene function in plants. In this method a gene regulatory sequence usually an enhancer is inserted into a host genome. The enhancer then combines with activator proteins to elevate the expression of genes adjacent to it. The method overcomes the shortfalls observed in classical knockout mutagenesis where gene duplication and alternative metabolic pathways can mask a phenotype resulting from a knockout mutation of a specific gene. The method can create dominant mutations by causing overexpression of tagged genes, and the resultant phenotype can reflect the normal function of the gene.

In this study tomato plants of the cultivar M82 that had been transformed with *Agrobacterium* T-DNA harboring a strong enhancer element from the Cauliflower Mosaic Virus was studied. The T-DNA had been modified to include the maize Ac/Ds transposon system with the CaMV 35S enhancer on the non-autonomous Ds element. A truncated Ac element was also included in the construct that would facilitate transposition of the Ds element (Figure 4). Seeds of putative transformants from ten different lines (named M1-10) were germinated with germination efficiencies ranging from 50 to 100% for the different lines (Table 3). The germination method used had first been tested on wild-type tomato seeds and had proven to be the most successful in germination of tomato seeds (Table 2).

The seeds were then grown to maturity in the green house where different mutant phenotypes were observed. Numbers of plants showing mutant phenotypes varied in the different transformant lines (Table 3) however it was not less than 50% in any of the lines (Table 3). Some of the phenotypic mutants observed included dwarf stature and abnormal leaf morphology. The dwarfism and abnormal leaf morphology had been observed in an earlier study using the same transposon construct and in the same cultivar (Carter et al, 2013). Some unique phenotypes were also observed which included Necrotic spots on the leaf (Figure 5). This mutant from the line M1 had leaf necrosis and was termed *Necrotic*. The necrosis was observed on one leaf in the entire plant. One of the mutants from the line M5 exhibited stunted growth and leaf discoloration (Figure 7). The leaf was oblong-like in shape and easily broke when handled. Although small and weak-looking the plant proliferated throughout the life cycle of the tomato and reached maturity and given the texture of the leaf, this mutant was termed *Brittle*. Due to their unique phenotypes, the lines from which the two mutants (*Brittle* and *Necrotic*) were found (M1 and M5) respectively were selected as the focus of this study.

The two lines had a total of eleven plants that reached maturity (Table 3). When the seedlings were young (15-20 days) genomic PCR was carried out to determine which plants carried the Ac/Ds cassette. A genomic PCR using the tomato squalene epoxidase gene was first done as a quality control check. The SQE gene codes for the enzyme squalene epoxidase which catalyses the first step in the biosynthesis of sterols and triterpernoids and is suggested to be ones of the rate-limiting enzymes in this pathway (Stevenson *et al,* 2014). The gene was used as a control check both to confirm that the DNA being used was tomato DNA and also to optimize the PCR conditions.

To check for transformation primers specific to different parts of the construct (BAR gene, HPT, Gos promoter), were used in separate PCR reactions. Varying number of plants in each of the two lines tested positive for the different parts of the construct and the results are summarized in Table 4. The plants from the two lines were then subjected to excision PCR. An excision PCR demonstrates the activity of the transposon system in the plants.

An excision PCR amplifies the footprint left behind when a transposon excises from one region (or a donor molecule) to integrate at another region in the genome (Liu *et al*, 2010). Primers specific to the right border of the construct and the Ac promoter were used in the reaction in order to amplify the footprint. Amplification was to occur only if the Ds element had excised out. Eight out of the 11 plants tested showed successful amplification of the footprint left by the Ds element (Figure 15). This demonstrated that the transposon was active in the eight plants and that the Ds element had moved to a new location in the genome. Ac/Ds transposon activity has previously been demonstrated in the model plants Arabidopsis (Altmann et al, 1995, Kuromori et al, 2004) and rice (Chi et al, 1999). Transformation by this system was also shown in tomato by

Carter et al. 2013. They however did not demonstrate excision of the Ds element in the tomato plants.

After demonstration of Ac/Ds activity in the plants it was necessary to identify the site of insertion/integration in order to identify the tagged gene(s). This was achieved through TAIL-PCR. TAIL-PCR is a method used for the recovery of genes adjacent to inserts in functional genomic studies and has been identified as a high-thoroughput method for identifying insertion sites for T-DNA tags (Szabados *et al,* 2002) and Ds-Atags (Chi *et al,* 1999). In this study three primers specific to the Ds element were used alongside an arbitrary degenerate primer in three consecutive PCR reactions. The specificity of the reaction increased with each cycle and the final product of the tertiary TAIL-PCR was expected to give a more 'clean' and specific sequence of the tagged gene (Figure 17). The more prominent bands on the tertiary TAIL-PCR indicates that the specificity was increased leading to a more homogenous product.

Tertiary TAIL-PCR was carried out on seven plants that had tested positive for excision PCR. After the tertiary PCR the bands of the plants that had tested positive for all the previous reactions (M1-5 and M5-3) (Table 4) were cut out purified and sequenced. The M1-5 was the plant that had given the necrotic leaf spots and was therefore termed *necrotic* and M5-3 was the mutant with the irregularly shaped leaf that was termed *brittle*. The tertiary TAIL-PCR of these two plants was sequenced and the sequences aligned to tomato genome in Phytozhome 9.1 for identification.

In the *brittle* mutant the Ds element was found to have inserted between two genes; one coding for a protein involved in vacuolar protein sorting and another coding for a coiled coil protein (Figure 18). Vacuolar sorting proteins are involved in transport of proteins to the vacuole. The

vacuole serves a variety of functions in the plant including storage of nutrients and in physical and metabolic processes in the cells (Xiang et al, 2013). The vacuole also confers plant turgor by storing water in the cell. From the Phytozhome search the vacuolar protein gene was found to be homologous to the Arabidopsis gene with the locus ID AT3G62080.2. This gene has been putatively documented to be involved in protein transport but no known function of the gene has yet been established (https://www.arabidopsis.org/servlets/TairObject?id=39732&type=locus). In the *brittle* mutant the leaf was brown and easily broke. By tagging a protein that is responsible for transport of proteins to the vacuole, the function of the vacuole may have been impaired leading to loss of storage for nutrient (hence the brown color indicating nutrient deficiency) and loss of turgor (hence the leaf easily broke when handled).

The other gene that was tagged in the *brittle* mutant was for a coiled coil protein (Figure 18). The coiled coil protein gene was found to be homologous to the Arabidopsis gene that has a locus IDAT2G46980.2. The gene codes for a ASY3 protein that is required for the process of meiosis especially in cross over and chromosomal axis formation (Ferdous et al, 2012). When studying Arabidopsis, it was found that in mutants where the ASY3 protein was lacking the number of meiotic cross-overs was reduced and chromosomal axis formation was compromised (Ferdous et al, 2012). The tagged coiled coil protein gene in the *brittle* mutant could also have a function in meiosis since the plant never bore any fruit, this implies that reproductive development of the plant was affected. Coiled coil proteins are proteins that have 2-5 alpha helical structures wound around each other to form a supercoil (Buckhard *et al*, 2001). The proteins have a variety of functions in plant growth and development including regulation of gene expression as transcription factors, in molecular recognition and they are also movement proteins (Lupas, 1996). The mutant observed had brown and irregular shaped leaves, the coiled coil gene in this

case might be a transcription factor involved in leaf morphological development or accumulation of chlorophyll in the leaf.

The tagged gene in the *necrotic* mutant was identified as gene coding for a calmodulin protein kinase (Figure 19). Calmodulin proteins are calcium modulated proteins that are abundant in the cytoplasm of all eukaryotic cells. Calmodulin protein kinases have been suggested as crucial mediators of responses to a variety of endogenous and environmental cues in plants (Zhang et al, 2003). The proteins bind calcium, which is a signaling molecule and relays the signals to other calcium-sensitive enzymes, ion channels or other proteins. The signal could be in response to a variety of factors such as light, hormones, mechanical disturbances or abiotic stress. The phenotype observed was leaf necrosis, which is often observed under chemical disorder or nutrient deficiency. For a plant to sense nutrient deficiency or chemical disorder genes responsible for signaling have to be activated and signals have to be sent around the plant. It therefore makes sense that activation of a calmodulin protein kinase gene led to symptoms related to stress response in the plant.

The calmodulin protein kinase gene tagged in the *necrotic* leaf mutant was found to be homologous to the Arabidopsis gene with the locus ID AT1G05410.1. The gene codes for a protein whose function is not yet known (arabidopsis.org/servlets/TairObject?id=137182&type=locus). The gene however was found to be homologous to the McCAP1 gene in the ice plant (*Mesembryanthemum crystallinum*) (Pathakar and Cushman, 2006). The McCAP1 gene was found to be involved in cytoskeleton formation as an intermediate filament protein (Pathakar and Cushman, 2006). The cytoskeleton is a structure that helps maintain the shape and integrity of the cell. It also allows the cell to carry

out vital processes like cell division. Without the cytoskeleton cells would likely die leading to the necrosis such as was observed in the *necrotic* mutant.

Upon analysis of gene expression of the tagged genes it was found out that the genes were highly induced (Figure 20). The induction of the tagged genes was to be expected as a result of the strong CAMV 35s enhancer within the Ds element. An enhancer combines with activator proteins to facilitate binding of the RNA polymerase to the promoter of a gene and in the process elevating its expression. These results are consistent with other studies using the same method, such as activation tagging in poplar (Busov *et al*, 2010). The presence of the enhancer will drive the transcription of the genes adjacent to its location to levels higher than without the proximal enhancer. Transcript activation leads to creation of dominant mutations as opposed to knocking out a gene, and this process makes studying of duplicate genes or gene families possible. This method allows the functional annotation of many eukaryotic genes, which would otherwise be difficult to study using other methods of gene discovery and analysis.

Alongside being responsible for morphological development of plants, genes also control metabolic pathways such as those involved in stress signaling, response and tolerance. Such genes are highly induced in response to stress. Stress is a common phenomenon in many parts of the world and remains a threat to agricultural production worldwide. Salt stress in particular has been projected to cause 50% loss of arable land by the year 2050 (Wang et al, 2003). Due to the impact of stress globally stress tolerant crops are highly desirable, however in order to have stress tolerant plants, the genes that take part in stress response have to be first identified. To this end the two mutants in the study (*brittle* and *necrotic*) were subjected to different stresses (Mannitol, NaCl) in order to determine how they respond to the stresses and also how the tagged genes respond when exposed to the different stresses.

When exposed to 100 Mannitol which confers water deficiency to the plants both mutants died. This observation implies that both mutants were sensitive to water deficiency or drought. Measurements of increase in root length and decrease in shoot length of the *brittle* mutant after exposure to salt stress showed that the mutant was tolerant to both stresses (Figure 21). When plants are exposed to stress (water deficiency as conferred by salt stress or nutrient deficiency) there is reduction in shoot length and increase in root length. The roots increase in length in so as to reach areas where water/nutrients are available. As such a significant increase in root length or significant reduction in shoot length implies susceptibility to stress. The tagged genes in the *brittle* mutant (coiled coil protein and vacuolar protein sorting) showed high induction following exposure to the stresses (Figure 22). The high induction of the genes in response to the stresses implies that both these genes have a role to play in stress response. When a plant goes through stress, a cascade of signals is evoked eventually triggering increased expression of genes that code for products that help in the survival of the plant under the stressful conditions (Wang *et al*, 2003). High induction of these genes under stress therefore demonstrates a putative function of these genes in stress response and tolerance in tomato. The *Necrotic* leaf mutant however proved susceptible to salt stress (Figure 21) and the calmodulin protein kinase tagged in the plant was highly induced after salt stress. The high induction implies that the gene has a role in salt stress response even though the response was negative.

Biochemical analyses were carried out on fruits that gave mutant phenotypes. The basis for carrying out nutrient analysis on the fruits was the nutritional importance of tomato worldwide. Tomato is grown worldwide because of its fruit that forms a major part of human diet. Global production of tomato was estimated at over 153 million metric tonnes in 2009 (Krylod, 2012). Consumption of tomato has been related to reduced risks of cancer especially prostate cancer as

well as reduced occurrence of cardiovascular diseases (El-Gaied et al, 2013). This is because tomato contains high amounts of antioxidants such as vitamins A, C and E, carotenoids, flavonoids and phenolic acids (Sacco *et al*, 2013). The nutritional importance of tomato has seen it being named the second most important vegetable crop after potato (Liza *et al*, 2013). To increase the benefits derived from tomato, production of high quality tomato fruits can be implemented. To achieve this however identification of genes responsible for nutrient accumulation in tomato has to be carried out first.

Through activation tagging, a gene that takes part in a nutrient accumulation pathway may be tagged and that will have an impact on the tomato fruit quality and hence the function of the gene elucidated. Levels of lycopene, beta-carotene and phenolic content were estimated in fruits from one of the mutants (*necrotic* leaf) and some other fruits from other lines that showed mutant phenotypes. The fruit from *necrotic* leaf mutant had significantly lower lycopene and beta-carotene content than the control (Figure 23). Since the gene had shown high induction after qPCR analysis and the result was altered nutrient composition it can be inferred that the calmodulin protein kinase gene also has a role in pathways that are responsible for nutrient accumulation. Other fruits from other plants that were grown but not characterized in this study also showed altered nutrient content in comparison with the wildtype. The DNA from these plants can be isolated and the tagged genes also identified to obtain genes responsible for nutrient accumulation in tomato.

6.0 Conclusions and Future Direction

In this study putative transformants of the tomato cultivar M82 were germinated by water imbibition in the dark and then allowed to grow in the greenhouse. The seeds came from plants that had been transformed with an A/Ds transposon system harboring a CaMV 35S enhancer on the Ds element. Eight plants from each of ten different lines were germinated and with a germination percentage exceeding 50% for all lines. Two lines were selected for further consideration because of the unique phenotypes observed in them. In line M1 one of the plants (M1/5) had necrotic leaf spots and was termed *necrotic* and in the line M5 one plant (M5/5) had brown, deformed leaves and the plant was named *brittle*. Transformation was confirmed in these lines through PCR targeting different parts of the cassette and positive transformants were identified. Ac/Ds transposon activity was demonstrated in the form of an excision PCR where excision of the Ds element from its original location was shown. Insertion sites of the Ds element were also identified through TAIL-PCR.

In the *brittle* mutant two tagged genes were identified as genes coding for proteins involved in vacuolar protein sorting and coiled coil proteins respectively. In the *necrotic* mutant one tagged gene was identified as a gene coding for a calmodulin protein kinase. The tagged genes were found to be highly induced by the presence of the enhancer element. The *brittle* mutant also showed tolerance to salt stress while the *necrotic* mutant proved susceptible to the stress. In the *necrotic* mutant fruit quality was negatively affected shown by a reduction in the amount of lycopene and beta-carotene content when compared to non-transformed wildtype tomato.

The genes identified in this study represent a preliminary but very useful study on the possible functions of the genes. The identification of the different genes opens up possibility for further study. The seeds from the plants can be grown and then crossed to wild-type M82 plants to study

co-segregation with the mutant phenotypes to verify the involvement of the genes in causing the mutant phenotypes. Validation of the tagged gene can be done by selecting excisions of the Ds-ATag from the position near the gene, and finding a revertant phenotype (to wildtype). Alternatively, transformation of the tomato plants with an overexpression phenotype with the same gene will prove the causality of the phenotypes due to the gene. Further analysis may then be done to study the role of the different genes in causing the observed phenotypes and what pathways are involved.

References

Aarts MGM, Koncz C, Pereira A (2000) "Transposon and T-DNA mutagenesis." *ZA Wilson (Ed.), Arabidopsis: A practical approach.: Oxford University Press, Oxford, pp. 143-170*

Aarts MG, Dirkse WG, Steinem WJ, Pereira A (1993) Transposon tagging of a male sterility gene in Arabidopsis, *Nature* **363**: 715-717

Agrawal N, Dasaradhi PVN, Mohmmed A, Malhotra P, Bhatnagar RK, Mukherjee SK (2003) RNA interference: Biology Mechanism and Applications, *Microbiology and Molecular Reviews* **4**: 657-685

Alonso JM, Stepanova AN, *et al* (2003) Genome-Wide Insertional Mutagenesis of Arabidopsis thaliana, *Science* **301:**653-657

Altman T, Jesop A, Kauschman A, Uwer U, Pena-Cortez H, Willmitzer L (1995) Ac/Ds transposon mutagenesis in Arabidopsis thaliana: mutant spectrum and frequency of Ds insertion mutants, *Mol Gen Genet* **247**:646-652

Avirovik, D (2013) Transformation of a transposon construct into tomato for functional genomics studies. MS Thesis. Virginia Tech. (http://vtechworks.lib.vt.edu/handle/10919/24867)

Bennetzen JL (2005) Transposable elements, gene creation and genome rearrangement in flowering plants, *Current Opinion in Genetics and Development* **15**:621-627

Biemont C, Vieria C (2006) Junk DNA as an evolutionary force, Nature **443**:521-524

Bouche N, Bouchez D (2001) Arabidopsis gene knockout: phenotypes wanted, *Current Opinion in Plant Biology* **4**:111-117

Brown DD (1981) Gene Expression in Eukaryotes, *Science* **211**: 667-674

Buckhard P, Stetefeld J, Stelkov SV (2001) Coiled coils: a highly versatile protein folding motif, Trends in Cell Biology **11**: 83-88

Busov B, Yordanov Y, Gou J, Meilan R, Ma C, Regan S, Strauss S (2010) Activation tagging is an effective gene tagging system in populous, *Trees Genetics and Genomes* **7**: 91-101

Carter JD, Pereira A, Dickerman AW, Veilleux RE (2013) An active Ac/Ds transposon system for activation tagging in tomato cultivar M82 using clonal propagation. *Plant Physiology* **162:**145-156.

Chi GH, Choe MS, Lee S, Park, SH, Koo J, et al (1999) Molecular analysis of rice plants harbouring an Ac/Ds transposable element-mediated gene trapping system, *The Plant Journal* **19:**615-623

Dooner HK, Weil CF (2007) Give-and-take: interactions between DNA transposons and their host plant genomes, *Current Opinion in Genetics and Development* **17**:486-492

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue, *Phytochemical Bulletin* **19**:11-15

El-Gaied LF, El-Heba AA, El- Sherif AN (2013) Effect of growth hormones on some antioxidant parameters and gene expression in tomato, *GM Crops and Food: Biotechnology in Agriculture and the Food Chain* **4**:67-73

Fedoroff N (1989) Maize transposable elements, in Berg, DE and Howe MM (eds) Mobile DNA, American Society for Microbiology, Washington D.C. pp 375- 411

Ferdous, M, Higgins, JD, Osman, K, Lambing, C, Roitinger, E, Mechtler, K, et al (2012). Inter-homolog crossing-over and synapsis in Arabidopsis meiosis are dependent on the chromosome axis protein AtASY3. *PLoS Genet*, **8(2),** e1002507-e1002507.

Feschotte C, Jiang N, Wessler SR (2002) Plant Transposable Elements: where genetics meets genomics, Genetics **3**: 329-341

Gimenez E, Pineda B, Capel J, Anton MJ, Atarez A, et al (2010) Functional analysis of the Arlequin mutant corroborates the essential role of the Arlequin/TAGL1 gene during reproductive development of tomato, *Plos One* **5**:e14427

Greene EA, Christine AC, Nicholas ET, Jorja GH, Bradley JT, et al , (2003) Spectrum of Chemically Induced Mutations From a Large-Scale Reverse-Genetic Screen in Arabidopsis, *Genetics* **164**:731–740

Jo KR, Kim CJ, Kim TK, Bergevoet M, Jongsma MA, et al (2014) Development of late blight resistant potatoes by cisgene stacking, *BMC Biotechnology* **14**:50-60

Kakimoto T (1996) CKI1, a Histidine Kinase Homolog implicated in cytokinin signal transduction, *Science* **274**:792-801

Krylod A (2012) UNCTAD http://www.unctad.info/en/Infocomm/AACPproducts/. COMMODITY-PROFILE—Tomato/. Accessed on 17 April 2015

Krysan PJ, Young JC, Sussman MR (1999) T-DNA as an insertional mutagen in Arabidopsis, *The Plant Cell* **11**:2283-2290

Kumar A, Bennetzen JL (1999) Plant retrotransposons, *Annu. Rev. Genet.* **33**: 479–532

Kumar CS, Narayanan KK (1998) Plant Transposable elements and Functional Genomics, *Plant Biotechnology* **15**: 159-165

Kunze R, Stochaj U, Lauf J, Starlinger, P (1987). Transcription of transposable element Activator (Ac) of Zea mays L. EMBO Journal, **6**:1555-1563

Kuromori T, Hirayama T, Kiyosue Y, Takabe H, Mizukado S et al (2004) A collection of 11 800 single-copy Ds transposon insertion lines in Arabidopsis, *The Plant Journal* **37**:897-905

Liu G, Aronovich EL, Cui Z, Hackette PB (2004) Excision of Sleeping Beauty transposons: parameters and applications to gene therapy, *J. Gene. Med* **6**: 574-583

Livak KJ, Schmittgen TD (2000) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta-Delta C (T)) method, *Methods* **25**:402-408

Liza LN, Nasar ANM, Zinuah KMA, Chowdhury MNA, Ashrafuzzaman M (2013) In vitro growth media growth effect for regeneration of tomato (*Lycopersicon esculentum*) and evaluation of the salt tolerance activity of the callus, *Journal of Agriculture and Sustainability* **3:**132-143

Lupas I (1996) Coiled coils: New structures and new functions, *Reviews* **21**: 375-383

Martinez-Marsch N, Greco R, Arkel G, Estrella-Herrera L, Pereira A. (2002) Activation tagging using the En-I maize transposon system in Arabidopsis, *Plant Physiology* **129**:1544-1556

Matthews H, Clendennen SK, Caldwell CG, Liu XL, Connors K et al , (2003) Activation tagging in tomato identifies a transcriptional regulator of anthocyanin biosynthesis, modification and transport, *The Plant Cell* **15**: 1689-1703

McClintock B (1951) Chromosome organization and genic expression, Cold Spring Harbor Symp. Quant. **16**: 13-47

McClintock B (1984) The significance of responses of the genome to challenge, *Science* **226**:792-801

Meyers BC, Tingey SV, Morgante M (2001) Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. *Genome Res*. **11**:1660–1676

Mills RE, Bennett A, Iskow RC, Devine SE (2007) Which transposable elements are active in the human genome? *Trends in Genetics* **23** :183-191

Morgante M (2006) Plant genome organisation and diversity: the year of the junk!, *Current Opinion in Biotechnology* **17**: 168-173

Muller-Neumann M, Yoder IJ, Starlinger P (1984) The DNA sequence of the transposable element *Ac* of *Zea mays L.*, *Molecular and General Genetics* **198**: 19-24

Patharkar, OR, Cushman, JC (2006). A novel coiled-coil protein co-localizes and interacts with a calcium-dependent protein kinase in the common ice plant during low-humidity stress. *Planta*, **225:** 57-73.

Pereira A (2002) Transposon tagging methods in heterologous plants. In: Molecular Plant Biology; A Practical Approach Volume I (P. Gilmartin & C. Bowler eds.), Oxford University Press, Oxford, UK, pp 53-81.

Phillips T (2008) Regulation of transcription and Gene Expression in Eukaryotes, Nature Education **1**: 199-209

Ramchandran S, Sundaresan V (2001) Transposons as tools for functional genomics, Plant Physiol. Biochem **39**: 243-252

Robertson DS (1978) Characterization of a mutator system in maize. *Mutation Res.* **51**: 21–28

Robinson SJ, Tang LH, Mooney BA, Mc Kay SJ, Clarke W et al (2009) An archived activation tagged population of Arabidopsis thaliana to facilitate forward genetic approaches, *BMC Plant Biology* **9**:101-116

Sacco A, Matteo A, Lombardi N, Trotta N, Punzo B, Mari A, Barone A (2013) Quantitative trait loci pyramiding for fruit quality traits in tomato, *Molecular Breeding* **31**:217-222

SanMiguel P, Bennetzen JL (1998) Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. *Ann. Bot.* **81**:37–44

Simons GL, Napoli CA, Jansen BJ, Plummer KM, Snowden KC (2007) Analysis of the decreased apical dominance genes of Petunia in the control of axillary branching, *Plant Physiology* **143**:697-706

Stevenson J, Luu W, Kristiana I, Brown AJ (2014) Squalene mono-oxygenase, a key enzyme in cholesterol biosynthesis is stabilized by unsaturated fatty acids, *Biochem J.* **461**: 435-442

Sundaresan V, Springer P, Thomas V, Samuel H, Jonathan DGJ et al (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements, *Genes & Development* **9**:1797-1810

Szabados L, Koracs I, Oberschall A, Abraham E, Kerekes I et al (2002) Distribution of 1000 sequenced T-DNA tags in the Arabidopsis genome, *The Plant Journal* **32**:233-242

Tani H, Chen X, Numberg P, Grant JJ, SantaMaria M, et al, (2004) Activation tagging in plants: A tool for gene discovery, *Functional and Integrative Genomics* **4**:258-266

Thorneycroft D, Sherson SM, Smith SM (2001) Using gene knockouts to investigate plant metabolism, *Journal of Experimental Botany* **52**:1593-1601

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

Vicient CM, Suoniemi A, Anamthawat-Jonsson K, Tanskanen J, Beharav A et al. (1999) Retrotransposon BARE-1 and its role in genome evolution in the genus *Hordeum*. Plant Cell **11**:1769–1784

Wan S, Wu J, Zhang Z, Sun X, Lv Y et al (2009) Activation tagging, an efficient tool for functional analysis of the rice genome, *Plant Molecular Biology* **69**:69-80

Wang W, Vincour B, Altman A (2003) Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance, *Planta* **218**:1-4

Weigel D, Ahn JH, Miguel AB, Justin OB, Sioux KC, et al (2002) Activation tagging in Arabidopsis, *Plant Physiology* **122**:1003-1013

Weisburger JH (2002) Lycopene and tomato products in health promotion, *Experimental Biology and Medicine* **227**:924-927

Xiang L, Exteberria E, Van De Ende W (2013) Vacoular protein mechanisms in plants. FEBS Journal **280**: 979-993

Zhang L, Lu YT (2003) Calmodulin- binding protein kinases in plants. *Trends in Plant Science* **8**: 123-12

Tables

Table 2: Percentage germination of tomato seeds under different treatments

Treatment	Total Number of seeds	Number of Seeds Germinated	Percentage Germination
½ Strength MS Medium with no supplements, no sugar and kept in the light	15	5	33
Water imbibition and kept in the dark	15	13	87
Water imbibition and kept in light	15	6	40
Water imbibition, light + 1µM GA3	15	8	53
Water imbibition , dark + 1 µM GA3	15	11	73

Table 3: Putative transformant lines germinated, number reaching maturity and number showing mutant phenotypes

Mutant Line	Total Number of seeds	Number of seeds that germinated	Number of plants at maturity	Number showing mutant phenotype
M1	8	6	6	3
M2	8	8	8	4
M3	8	No germination		
M4	8	6	6	4
M5	8	5	5	4
M6	8	3	3	2
M7	8	5	5	3
M8	8	7	7	4
M9	8	5	5	3
M10	8	4	4	2
Control	8	7	7	0

Table 4: Different primers used and their sequences

Primer	Sequence
SQE F	5’-TGGGGTTCGTTGCAGTTTTC-3’
R	5’-CGCGTTGAGCATCAATTTTCTC-3’
BAR F	5’-CTGAAGTCCAGCTGCCAGAAACC-3’
R	5’-CTGCACCATCGTCAACCACTACAT-3’
HPT F	5’- AAA AGT TCG ACA GCG TCT CCG ACC-3’
R	5’- TCT ACA CAG CCA TCG GTC CAG ACG-3
Ac F	5’-ACCAAGACGATTGCGGTAAG-3’
R	5’-CCTTCAGCTCCAAAGACAAAGA-3’
Gos F	5’-GTGCGTAAGTACCTTGCATCTA-3’
R	5’-TAGTGGCAATCGGGCTAAATAA-3’
Ds3-1	ACCCGACCGGATCGTATCGGT
Ds3-2	ACGATGGACTCCAGTCCGGCCCGATTACCGT ATTTATCCCGTTC
Ds3-3	GTATTTATCCCGTTCGTTTTCGT
LAD1-1	ACGATGGACTCCAGAGCGGCCGCVNVNNNGGAA
LAD1-3	ACGATGGACTCCAGAGCGGCCGCVVNVNNNCCAA
RB	5’-GGAAACGACAATCTGATCTCTAGG-3’
Ac Promoter	5’-CTCAGTGGTTATGGATGGGAGTTG-3’

Table 5: Summary of the PCR results for plants in the putative transformant lines M1 and M5. M1-5 (*necrotic* mutant) and M5-3 (*brittle* mutant) tested positive for all the PCR reactions performed and the tagged genes were identified in these two plants.

Line	M1						M5				
Plant	1	2	3	4	5	6	1	2	3	4	5
Mutant Phenotype	x	x	✓	x	✓	✓	✓	✓	✓	x	✓
HPT	x	✓	✓	x	✓	x	✓	x	✓	✓	✓
Bar	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Gos Promoter	✓	x	✓	x	✓	✓	✓	✓	✓	✓	✓
Excision PCR	x	x	✓	x	✓	✓	✓	✓	✓	✓	✓
Secondary TAIL-PCR	x	x	x	x	✓	x	x	✓	✓	✓	x
Tertiary TAIL-PCR	x	x	x	x	✓	✓	x	✓	✓	✓	x

Figures

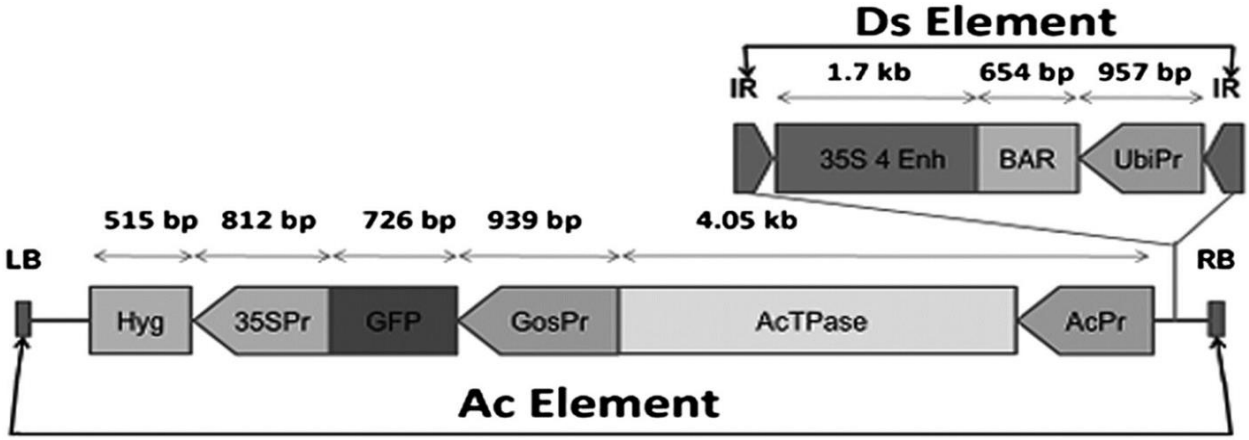


Figure 4: The Ac-Ds ATag construct. The elements of the construct are as follows IR= inverted repeats, BAR =glufosinate/Basta resistance gene, UbiPr =Ubiquitin Promoter, 35S 4 Enh =Cauliflower mosaic virus 35s enhancer, LB = left border. RB =right border, GosPr =Gos promoter, AcTPase =transposase, GFP =Green Fluorescent Protein, 35Pr =35S promoter, Hyg =Hygromycin resistance gene

Figure 5: Phenotype of the necrotic leaf mutant named *necrotic*. The mutant line contained all parts of the Ac/Ds construct by PCR, and the tagged gene was identified by TAIL-PCR.

Figure 6: Dwarf mutant (right) shown beside the M4 transformant progeny line without phenotype

Figure 7: *brittle* mutant of the M5 transformant line. The leaf showed discoloration and easily broke when handled.

Figure 8: Dwarf mutant from the M6 line (left) alongside a non-mutant from same line.

Figure 9: Dwarf mutant (right) from the M7 line. The plant was much smaller compared to the wild-type (left), and never flowered

Figure 10: Modified leaf shape and discolored leaf from the M8 line (right) compared to wild-type (left)

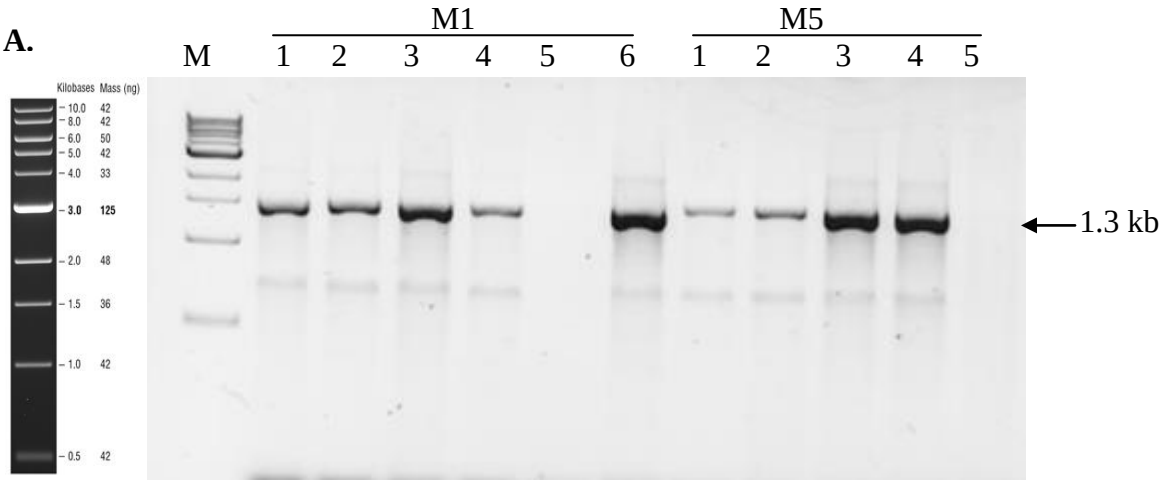


Figure 11: **A**. NEB 1 kb DNA ladder **B**. Genomic PCR analysis using tomato SQE primers for different tomato plants from two lines (M1 and M5).. M= 1 kb DNA ladder, Lanes 1-6 are the different plants from the M1 line that reached maturity, lanes 7-11 are the different plants from the M5 line that reached maturity. Lines M1 and M5 were selected based on the unique phenotypes observed in these lines.

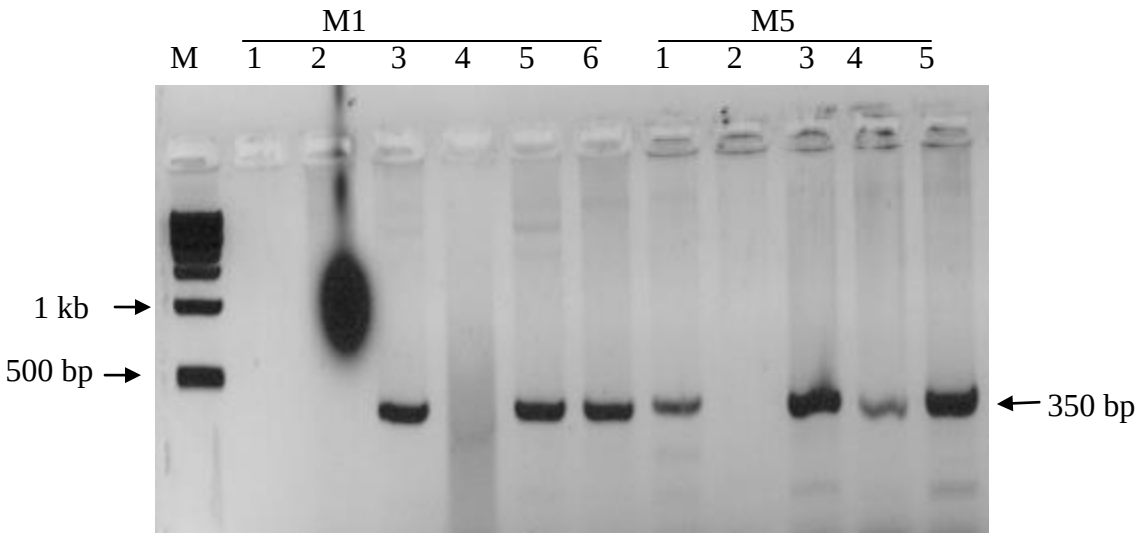


Figure 12: Genomic PCR analysis using HPT primers for different tomato plants from two lines (M1 and M5). The lanes show M=1kb DNA ladder, lanes M1 1-6 are the different plants from the M1 line that reached maturity, lanes 7-11 are M5 1-5 representing the different plants from the M5 line that reached maturity. Three plants (M1-1, M1-2 and M1-4) from the M1 line tested negative for the HPT gene and one plant from the M5 line (M5-2) was negative for the HPT gene

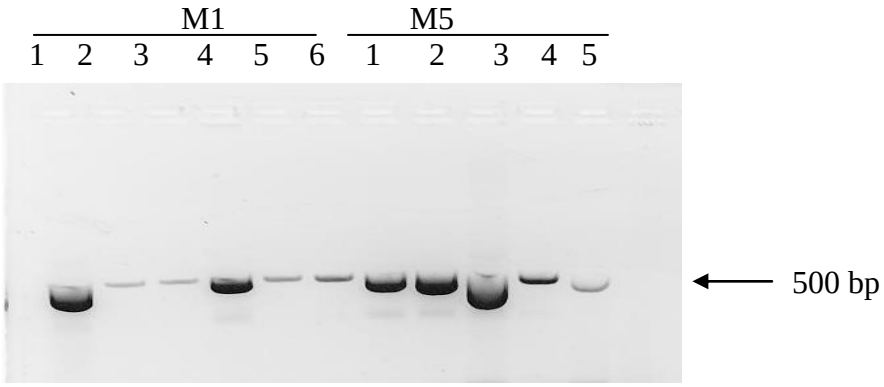


Figure 13: Genomic PCR analysis using BAR gene primers for the tomato plants from two lines (M1 and M5). Lanes 1-6 are the different plants from the M1 line and lanes 7-11 are the different plants from the M5 line that reached maturity.

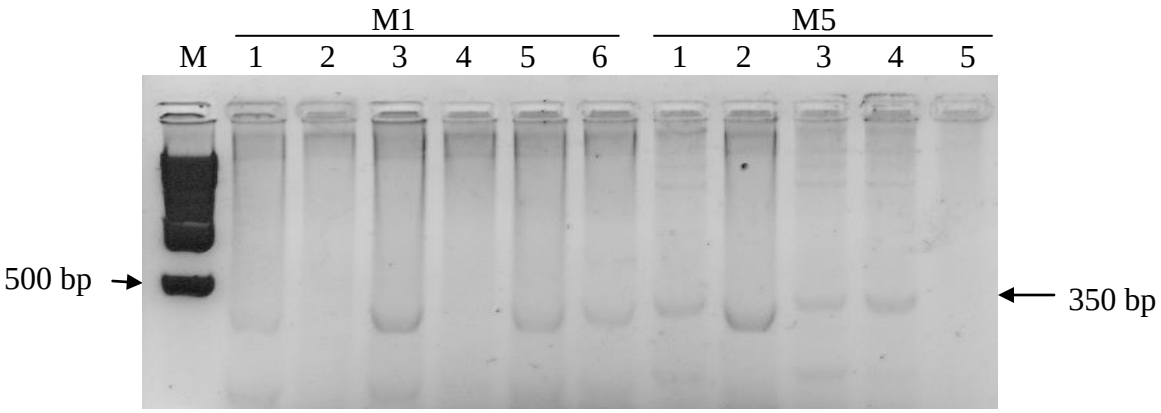


Figure 14: Genomic PCR analysis using *GOS* promoter primers for different tomato transgenic plants. M= 1kb DNA ladder, lanes 1-6 are the different plants from the M1 line that reached maturity, lanes 7-11 are the different plants from the M5 line that reached maturity. Lines M1 and M5 were selected based on their unique phenotypes.

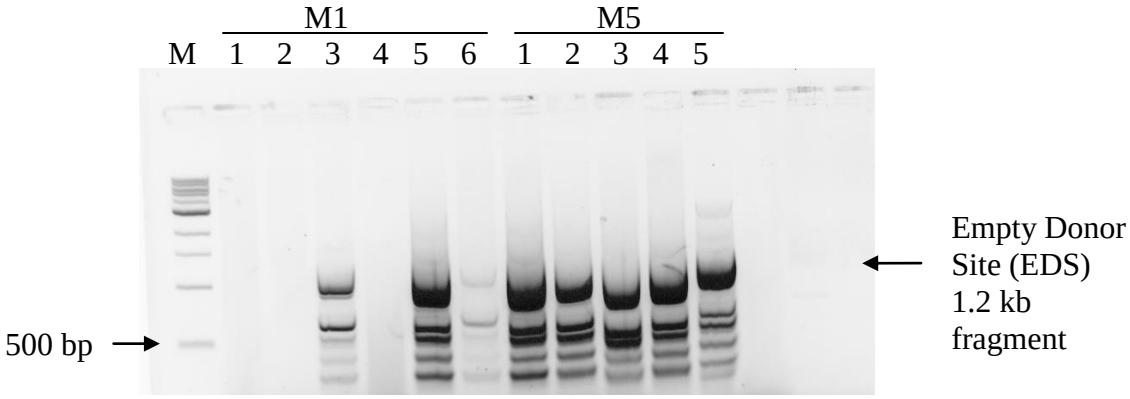


Figure 15: PCR analysis to check excision of Ds element in different tomato transgenic lines that were tested positive from preliminary screening. PCR reaction was done using the Right Border and Ac Promoter primers. Bands were obtained only when there is excision of the Ds element. M= 1kb ladder, Lanes 1-6 are the different plants from the M1 line that reached maturity, lanes 7-11 are the different plants from the M5 line that reached maturity.

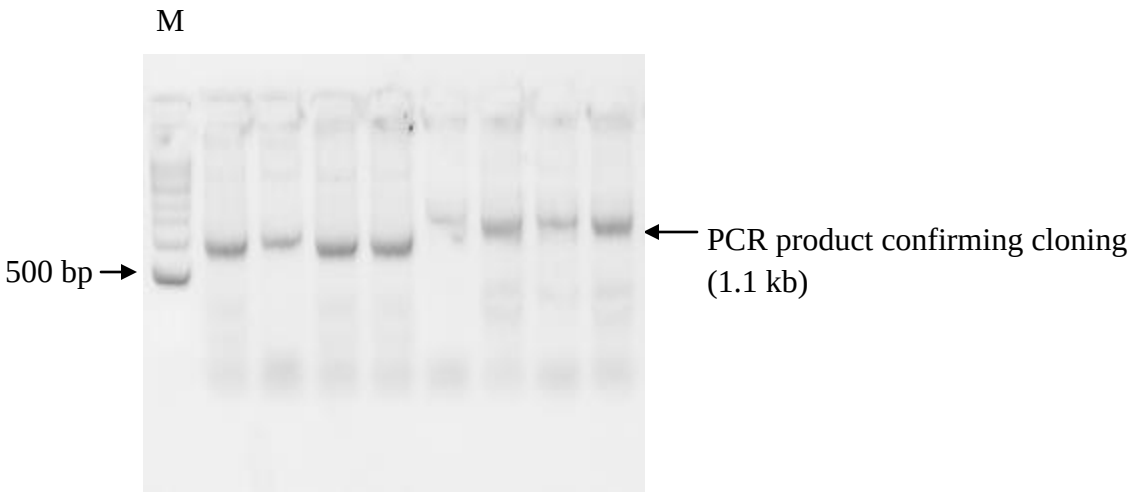
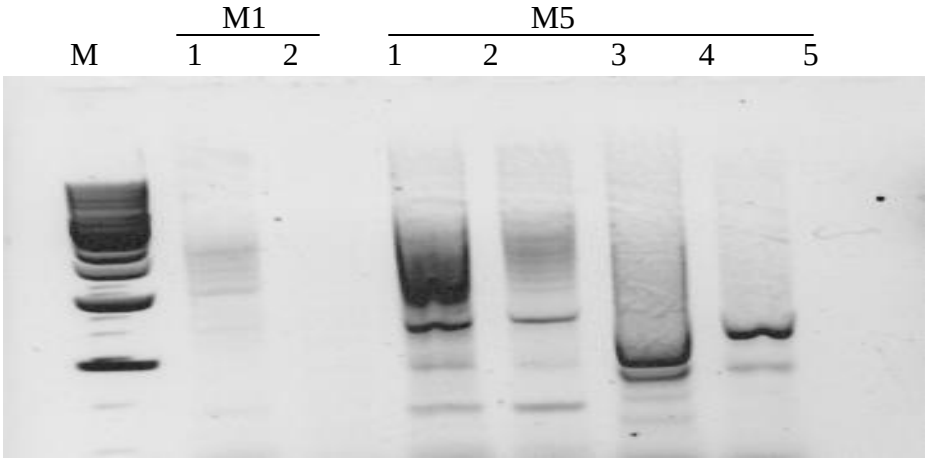


Figure 16: PCR analysis of Plasmid DNA obtained from cloning the excision PCR products. Single lanes representing putative clones were inoculated in medium for plasmid isolation. Isolated plasmids were used for PCR verification using RB and Ac Promoter primers. M= 1kb DNA ladder

Secondary TAIL-PCR



Tertiary TAIL-PCR

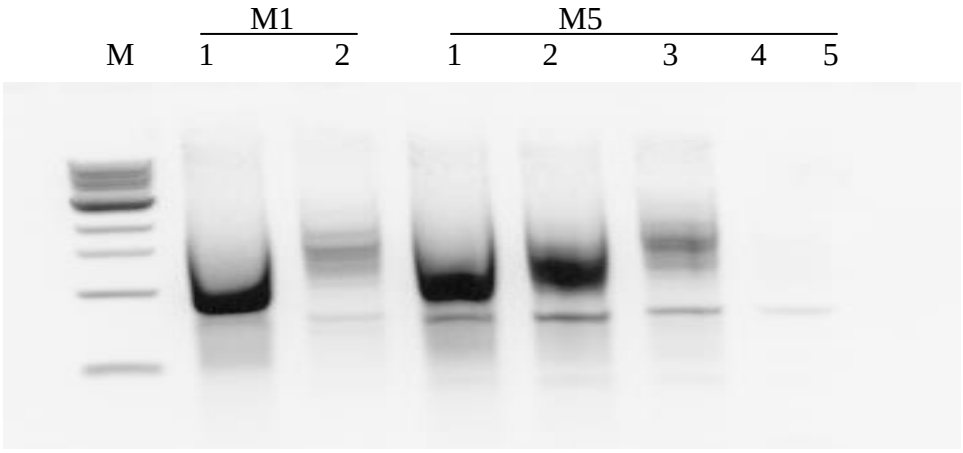
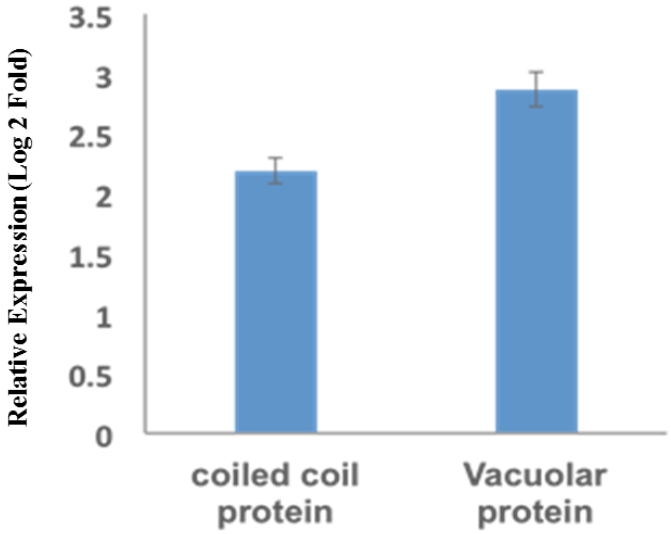


Figure 17: Secondary and Tertiary TAIL-PCR analysis of tomato transformants that were positive for excision PCR. Lanes 1-2 are different plants from M1 line that tested positive for excision PCR (M1/5 and M1/6) and lane 4 to 7 are the different plants from the M5 line that tested positive for excision PCR. The plants that had tested positive for all parts of the construct, and excision PCR were selected for identification of tagged genes by excising out the tertiary TAIL-PCR bands and sequencing them

Figure 18: *brittle* chromatogram

Figure 19: *necrotic* chromatogram

A.



B.

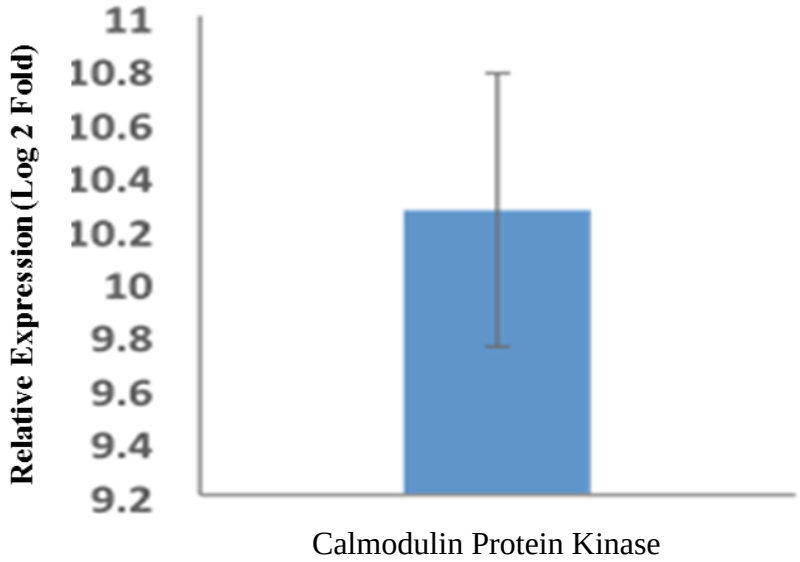
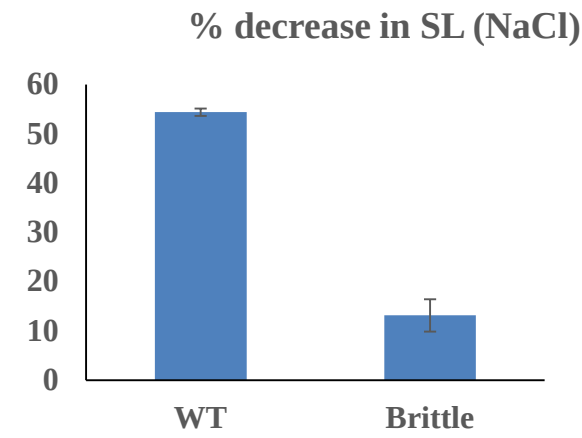
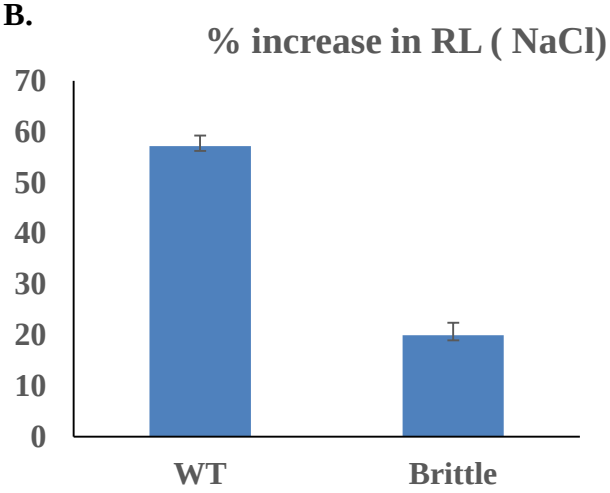


Figure 20: qPCR analysis of tagged genes in *brittle* and *necrotic* mutants **A**. Expression analysis of coiled coil protein and vacuolar protein sorting genes **B**. Expression analysis of Calmodulin protein kinase gene

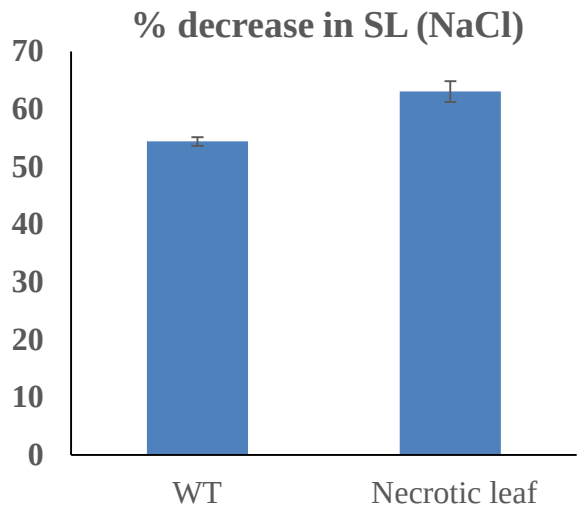
A.


% decrease in SL (NaCl)
60
50
40
30
20
10
0
WT
Brittle

B.


% increase in RL (NaCl)
70
60
50
40
30
20
10
0
WT
Brittle

C.



D.

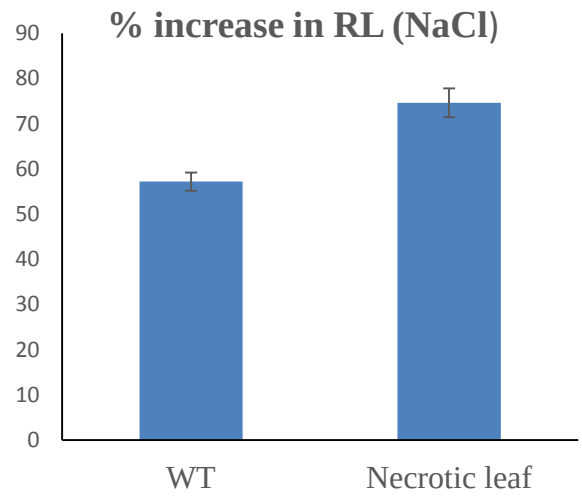
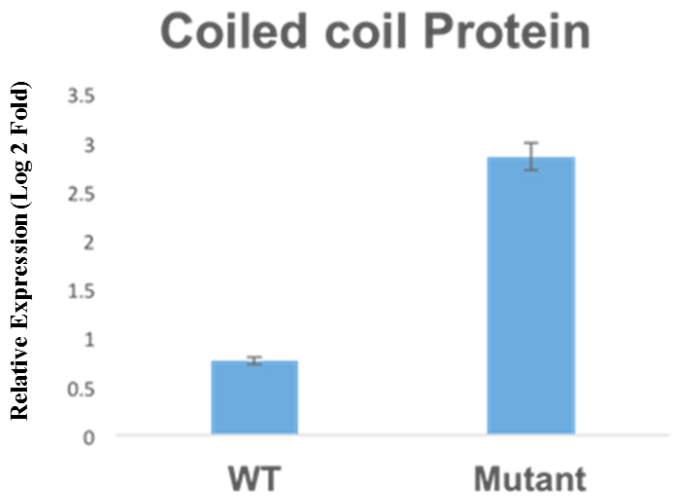
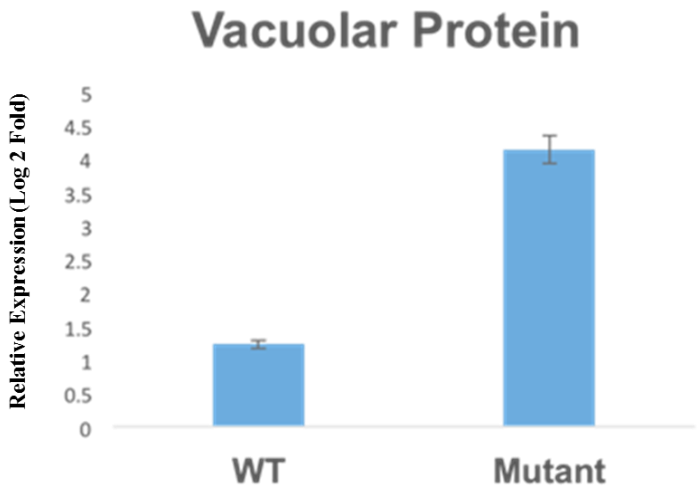


Figure 21: Stress response of *brittle* and *necrotic* mutants following exposure to 100 mM NaCl **A.** Percentage decrease in shoot length (SL) of *brittle* mutant **B**. Increase in root length (RL) of *brittle* mutant **C**. Percentage decrease in shoot length of *necrotic* mutant **D**. Percentage increase in root length of *necrotic* mutant

A.


Coiled coil Protein
Relative Expression (Log 2 Fold)
3.5
3
2.5
2
1.5
1
0.5
0
WT
Mutant

B.


Vacuolar Protein
Relative Expression (Log 2 Fold)
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
WT
Mutant

C.

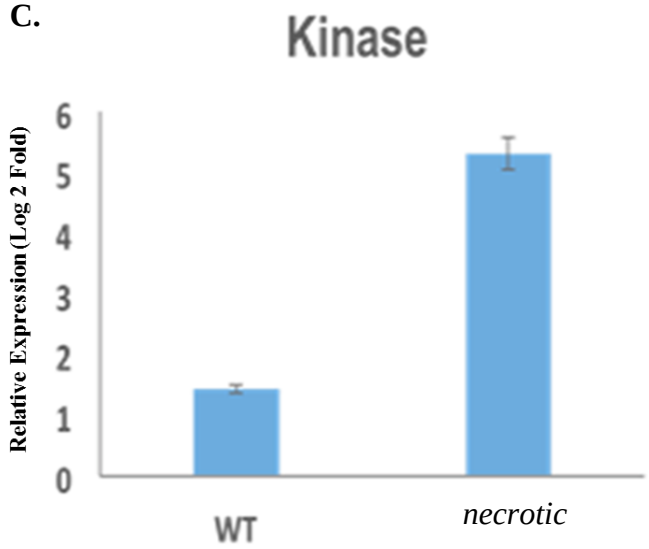
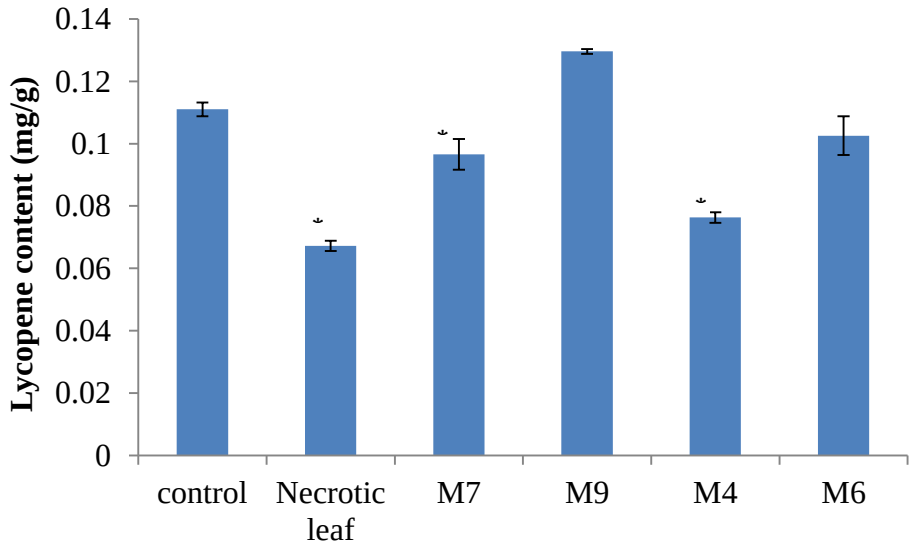
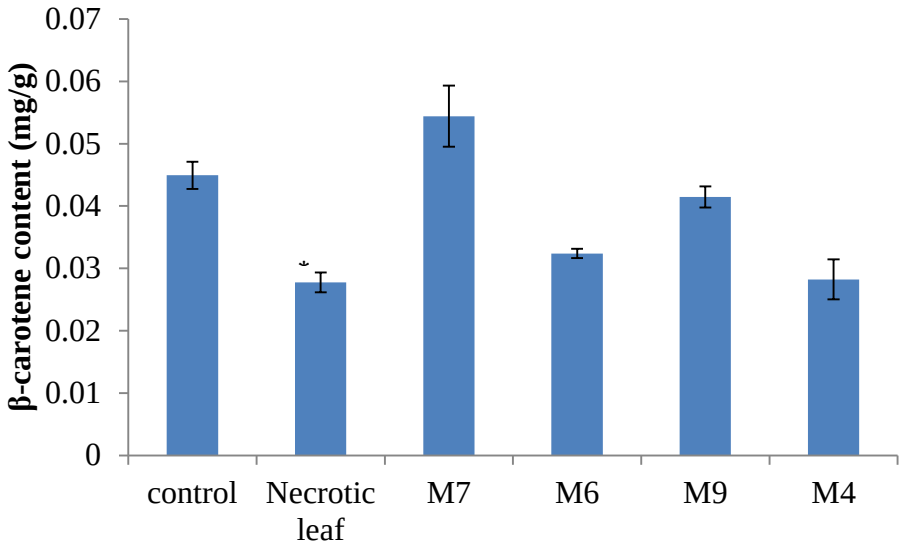


Figure 22: Expression analysis of tagged genes following salt stress **A.** Expression of coiled coil protein gene **B**. Expression of vacuolar protein sorting gene **C**. Expression of the calmodulin protein kinase gene

A.



B.



C.

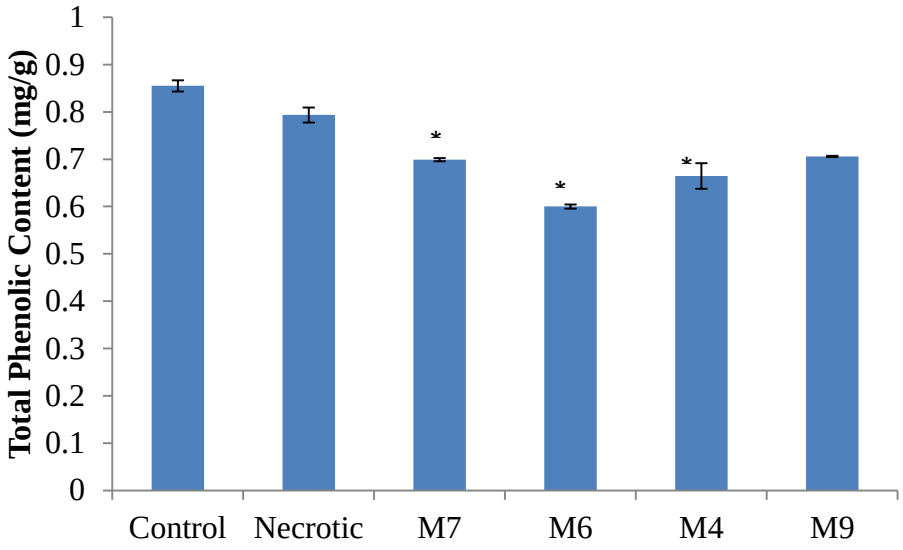


Figure 23: Lycopene Content **(A)**, Beta- carotene content **(B)** and Total Phenolic Content **(C)** of the mutant tomato fruits from different lines and control. The data represents the mean of three observations (n=3). The vertical bar at the top represents the standard error in each case. The asterisk (*) denotes significant difference to the control at P≥ 0.05.