Identification of Single Nucleotide Polymorphisms in COP1 and SPA3LIKE genes of tomato: Functional relevance for fruit ripening

Thesis submitted for the award of the degree of

Doctor of Philosophy

By

CHAITANYA CHARAKANA

Department of Plant Sciences

School of Life Sciences

University of Hyderabad

Hyderabad-500046, India

2015

Identification of Single Nucleotide Polymorphisms in COP1 and SPA3LIKE genes of tomato: Functional relevance for fruit ripening

Thesis submitted for the award of the degree of

Doctor of Philosophy

By

CHAITANYA CHARAKANA

Supervisor

Dr. Y. Sreelakshmi

Department of Plant Sciences

School of Life Sciences

University of Hyderabad

Hyderabad-500046, India

2015

Enrol. No. 08LPPH05

Department of Plant Sciences

School of Life Sciences

University of Hyderabad

DECLARATION

I hereby declare that the research work described in this thesis entitled "**Identification of Single Nucleotide Polymorphisms in *COP1* and *SPA3LIKE* genes of tomato: Functional relevance for fruit ripening**" has been carried out by me under the supervision of Dr. Y. Sreelakshmi, Department of Plant Sciences, School of Life Sciences, University of Hyderabad, Hyderabad- 500046, and this work has not been submitted for any degree or diploma of any other University.

Date: **Chaitanya Charakana**

Place: Hyderabad **Enrol. No. 08LPPH05**

Dr. Y. Sreelakshmi

Supervisor

Department of Plant Sciences

School of Life Sciences

University of Hyderabad

CERTIFICATE

This is to certify that the research work described in this thesis entitled "**Identification of Single Nucleotide Polymorphisms in *COP1* and *SPA3LIKE* genes of tomato: Functional relevance for fruit ripening**" is based on the results of the work done by Ms. Chaitanya Charakana for the degree of Doctor of Philosophy under my supervision. The work presented in this thesis is original and plagiarism free. No part of this thesis has been submitted for any degree or diploma of any other University.

Dr. Y. Sreelakshmi

Supervisor

Head

Dept. of Plant Sciences

Dean

School of Life Sciences

Dedicated to

My beloved Parents & Family

Acknowledgements

It gives me immense pleasure to thank all the people without whom this thesis might not have been written and I am greatly indebted to all.

Firstly, I express my deep sense of gratitude to my supervisor, ***Dr. Y. Sreelakshmi****, for her valuable guidance and patience during the entire period of my work. Her inexhaustible energy and encouragement to pursue new experiments, dedication, discussions and constructive criticism had helped me to have a better perspective of scientific research. I am deeply indebted of her unfailing care and unending support in every aspect of my professional and personal life that enabled me to successfully develop my skills and complete this Ph.D programme.*

With great pleasure, I want to extend my heartfelt thanks to ***Prof. R. P. Sharma*** *for his constant support and guidence throughout the course of work. Thank you Sir, for your encouragement, insightful comments, and constant inspiration.*

My sincere thanks to ***Prof. P. Reddanna*** *(Dean, School of Life Sciences),* ***Prof. A. S. Raghavendra, Prof. R. P. Sharma, Prof. Aparna Dutta Gupta, Prof. M. Ramanadham*** *(Former Deans, School of Life Sciences);* ***Prof. Ch. V. Ramana*** *(Head, Dept. of Plant Sciences), and* ***Prof A. R Reddy****,* ***Prof A. R. Podile****, (Former Heads, Dept. of Plant Sciences) for providing me the opportunity to utilize the facilities of the School and Department.*

I also would like to thank ***Prof. A. R. Reddy****, my doctoral committee member for his guidance and timely suggestions throught the research work.*

I am highly thankful to ***Prof. Jitendra P. Khurana*** *(University of Delhi South campus, New Delhi, India) and* ***Prof. Sudip Chattopadhyay*** *(NIT Durgapur, India) for their valuable suggestions in my research work.*

Thanks are due to all the lab members of RTGR (Repository of Tomato Genome Reseach) for their extended help and guidence. I sincerely thank the pioneer TILLING group of the lab- ***Dr. Soni Gupta, Dr. Vijee Mohan, Dr. Sherin, Dr. Mickey, Dr. Vineeta, Dr. Reddaiah*** *for their supprort, timely discussions and suggestions. I extend my gratitude to* ***Dr. Rahul, Dr. Vajir, Dr. Vineet and Dr. Bharati*** *for their help and suggestions. I specially thank* ***Dr. Arun Pandey and Dr. Sulabha*** *for their help during Southern blot experiments. I extend my gratitude to* ***Rakesh, Bindu, Kapil, Kamal, Supriya, Rachana, Sapna, Pallawi & Prateek*** *for their time to time suggestions and also for making my stay in the lab a memorable one. I am thankful to* ***Jayram, Kalyani, Hyma, Swati, Gayatri & Sumona*** *for their help during my research period. My sincere thanks to* ***Himabindu*** *and* ***Jayram,*** *without their help it was not*

possible to complete the work for Eco-TILLING lines (during tagging and tissue collection). I owe special thanks to ***Prateek*** *for his help in handling and analyzing U-HPLC data. I would be indebted to* ***Anusha*** *for her intimely support and company. My sincere thanks to* ***Anjana*** *for her help in modelling of* SPA3LIKE *protein structure. My special regards would go to my colleagues* ***Alka*** *and* ***Suresh*** *with whom I had daily discussions and debates (at times), which allowed me to gain better insight into my work.*

*I extend my thanks to field assistants (****Narsimha, Jamir, Yadagiri****) and lab accountant* ***Venkat Reddy*** *and* ***Anil*** *for maintenance of plants and clearance of bills.*

I have no words to put across my love and gratitude to all my friends (from school, college and so on..) who helped me stay happy and learn many things all through these years. Special thanks to ***Chinna, Jyothi, Debu, Manoj, Prasu, Bimo, Narmadha, Kishore, Venkat and Madhu****. Their support and care helped me overcome setbacks and stay focused on my graduate study. I greatly value their friendship.*

I would also thank all the members of former and present cricket teams –Daffodils and Bio Angles, also SLS badminton group for those unforgettable moments during my stay in the campus.

Most importantly none of this would have been possible without the love and support of my family members. I would like to extend my heart felt gratitude to ***my parents, sisters and in-laws****. A very special thanks to my father for his unending support. My special thanks to my better half* ***Ram*** *for supporting me and making my life more colorful.*

I thank the financial support provided by ***BBM*** *fellowship from University of Hyderabad and the* ***Council of Scientific and Industrial Research*** *(****CSIR****), New Delhi, in the form of* ***JRF*** *and* ***SRF****.*

I would also extend my heartfelt thanks to the Principal and staff of GDC, Adilabad for their timely support.

Last but not the least, I am highly grateful to the almighty for the blessings and spirit which helped me to successfully achieve this goal.

Chaitanya Charakana

2015

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AC	Ailsa Craig
At	*Arabidopsis thaliana*
AV	Arka Vikas
bHLH	basic helix-loop-helix
bp	base pair
cDNA	Complementary DNA
CO	Constans
COP1	Constitutive photomorphogenic1
CRY1	Cryptochrome 1
CRY2	Cryptochrome 2
CSN	Cop 9 Signalosome
CTR	Constitutive triple response
CUL4	Cullin 4
cv	Cultivar
CYCB	Chromoplast specific lycopene β-cyclase
D	Tajima's D
DDB1	DNA damage binding protein 1
DEPC	Diethyl pyrocarbonate
DET	De-etiolated
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
DPA	Days post anthesis
EMS	Ethyl Methyl Sulfonate
FKF1	Flavin-binding Kelch repeat F-box 1
F_L	Forward labeled primer

FR	Far-red light
F_{UL}	Forward unlabeled primer
FUS	*Fusca*
FW	Fresh weight
h	Hour
hp1	*high pigment 1*
hp2	*high pigment 2*
HPLC	High performance liquid chromatography
HT	haplotype
HY5	Elongated Hypocotyl 5
HYH	bZIP transcription factor
IIVR	Indian Institute of Vegetable Research
INDEL	Insertions and deletions
Kb	Kilobase
kD	Kilodalton
LAF1	Long After Far-red light1
LD	Long day
LKP2	LOV Kelch repeat Protein 2
Mb	Megabase
Md	*Malus domestica*
min	Minute
mRNA	Messenger RNA
NBPGR/NB	National Bureau of Plant Genetic resources
NGS	Next generation sequencing
PCR	Polymerase chain reaction
PHY	Phytochrome

PHYB	Phytochrome B
PHYA	Phytochrome A
PIF	Phytochrome-interacting factor
PSY1	Phytoene synthase 1
qRT-PCR	Quantitative real-time PCR
R_L	Reverse labeled primer
RNA	Ribonucleic acid
R_{UL}	Reverse unlabeled primer
RUP1	Repressor of UVR8 protein1
RUP2	Repressor of UVR8 protein 2
SD	Short day
SIFT	Sorting Intolerant from Tolerant
SNP	Single nucleotide polymorphism
SPA	Suppressor of Phytochrome A
TF	Transcription factor
TGRC	Tomato Genetic Resource Center
UPS	Ubiquitin- mediated 26S proteasome
UV	Ultra-violet light
UVR8	UV-B resistance protein 8, UV-B receptor
v/v	volume/volume
w/v	weight/volume
WT	Wild type
ZTL	Zeitlupe
π	Nucleotide diversity

Chapter 1

Introduction

Plant growth and development, right from seed germination to senescence involves a complex set of interactions between different gene products. Throughout the life cycle, many hormones and external factors (light, temperature etc.) also influence various physiological processes. In order to maintain the functional repertoire of proteins and their plasticity to adapt to ever changing surroundings, proteasome mediated degradation of proteins is majorly required, with few proteins degraded via non-proteasomal pathway. Recent advances in molecular genetics and introduction of high throughput technologies like transcriptomics, proteomics and metabolomics have further expanded the information regarding the different plant developmental processes linked with ubiquitin mediated proteasomal degradation (UPS) of cellular proteins (Moon et al., 2004). Moreover, it is estimated that about 5% of the Arabidopsis genome is encoded for UPS, modulating many internal and external targets (Vierstra, 2003). Ubiquitination is sequentially carried out by three enzymes, E1 - ubiquitin-activating enzyme, E2 - ubiquitin conjugating enzyme and E3 – ubiquitin protein ligase, which sequentially activates ubiquitin, conjugates and transfers it to target protein and finally coordinates to bind to target protein. This process goes on till a short chain of ubiquitin molecule is formed which acts as a signal for the proteasome mediated degradation of target protein. The substrate specificity is mainly defined by the E3 ligase and its multiple combinations with E2. E3 ligases have WD40 domains that bind to the substrates and are, therefore very specific. In nature there are very few E1's, many E2's and a large number of E3's (Hatfield et al., 1997; Kraft et al., 2005). With more than thousands of E3 ligases, plants have rich ubiquitination machinery (Mazzucotelli et al., 2006). Most of the proteins involved in plant hormonal signaling are regulated by targeted proteasome mediated degradation via ubiquitination (Santner and Estelle, 2010).

The identification and characterization of *CONSTITUTIVE PHOTOMORPHOGENIC*/*DE-ETIOLATED*/*FUSCA* (*COP*/*DET*/*FUS*) gene mutants revealed that ubiquitination mediated degradation also plays an important role in light mediated plant development. The proteins COP10, COP9 signalosome (CSN) 1-4, 7 and 8 along with COP1 were originally identified as repressors of seedling photomorphogenesis in Arabidopsis and were later identified in other eukaryotes including humans (Lau and Deng, 2012). All these molecules together form a complex and target proteins for ubiquitin-mediated proteasomal degradation (Hardtke et al., 2000). Several studies highlighted that distinct light signal transduction pathways often converge

at the level of COP1 protein that in turn interacts with other partners to regulate the photomorphogenic response (Schäfër and Bowler, 2002). COP1 being an E3 ubiquitin ligase with three distinct domains (ring finger, coiled coil and WD40) interacts with many molecules – Transcription Factors (TFs), transcriptional regulators, interacting proteins, photoreceptors etc, and targets them for degradation and therefore acts as the master regulator of light signaling pathways (Lau and Deng, 2012). COP1 and its interacting partners have been shown to be important for almost all facets of plant development (Ranjan et al., 2014). However, their role in regulating fruit ripening is not yet clearly defined.

Ripening is a complex process involving many physiological and biochemical events that lead to accumulation of pigments, metabolic changes related to flavor and nutrient composition, softening of cell wall, etc. and therefore aids in seed dispersal (Klee and Giovannoni 2011). Over last two decades, the molecular, biochemical and genetic investigations in tomato focused largely on the developmental regulation of fruit ripening including the elucidation of cell wall modifying proteins, ethylene biosynthesis and signaling, genes and pathways involved in pigment accumulation and other internal cues (like hormones). These studies also highlighted the role of light in regulation of fruit pigmentation. The tomato *hp1* and *hp2* mutants show exaggerated photoresponsiveness at seedling stage and enhanced fruit pigmentation. The molecular and genetic characterization of *hp1* (Lieberman et al., 2004) and *hp2* (Mustilli et al., 1999) mutants revealed that these loci are encoded by the homologs of *DDB1* and *DET* genes of Arabidopsis. Both DDB1 and DET proteins are auxiliary to COP1 and together act as negative regulators of photomorphogenesis (Yi and Deng, 2005). Apart from these two auxiliary molecules, the RNAi suppression of *Cullin4* (*CUL4*- encodes a scaffolding protein) and *COP1LIKE* (recently annotated as *SPA3LIKE*- whose protein is closely related to COP1) also exhibited altered photoresponses (Liu et al., 2004; Wang et al., 2008). Enhanced carotenoid content in tomato fruits of these RNAi suppression lines suggests that regulated proteolysis may be important for enhancing nutritional quality during fruit ripening.

Considering the importance of carotenoids for human health (Grune et al., 2010), and the demand for nutritionally enriched tomatoes, many genes in the carotenoid biosynthetic pathway and ethylene signalling pathway have been exploited. Also, from the foregoing it is evident that light signalling components influence the ripening of

climacteric fruits especially, formation of pigments and nutritional quality; however the underlying mechanisms are completely unknown. Though tomato has a large number of monogenic mutants, relatively few mutants are available which are compromised in light signaling. Considering this, our laboratory has adopted TILLING and Eco-TILLING, the reverse genetics strategies to isolate novel mutations in *COP1* and *SPA3LIKE* genes of tomato. Further, to understand the natural diversity of these two candidate genes, we checked their sequences in six wild relatives of tomato. The major objectives of this study are:

1. Isolation of tomato *COP1* mutants through TILLING
2. Identification of SNPs in *COP1* and *SPA3LIKE* genes of tomato by Eco-TILLING
3. Nucleotide diversity analysis of *COP1* and *SPA3LIKE* genes in six wild relatives of tomato

Chapter 2

Review of Literature

2.1 Regulated proteolysis: a crucial step for plant growth and development

Plants are sessile and photoautotrophic; therefore the entire life cycle is influenced by the ever changing environment surrounding them. In order to sense and respond to the fluctuating internal and external stimuli, plants need a rapid turnover of proteins. This higher order of proteome plasticity is achieved by the regulated proteolysis of proteins. Based on the experimental estimates in Arabidopsis, 10% of all intracellular proteins are short-lived with 30% of the nascent polypeptides degraded soon after translation (Vierstra, 2009). Most of the proteins are degraded by ubiquitin-mediated proteasome system and few are targeted via non-proteasome pathway. In plants, regulated protein degradation by the ubiquitin/26S proteasome contributes significantly to development by affecting a wide range of processes, including embryogenesis, hormone signaling and senescence etc (Moon et al., 2004).

2.2 Ubiquitin 26S Proteasome System (UPS)

The UPS involves a cascade of three enzymes- E1- ubiquitin-activating enzyme, which activates ubiquitin; E2- ubiquitin conjugating enzyme, which conjugates it and E3- ubiquitin protein ligase, which transfers ubiquitin to target protein and finally covalently binds to it. Often this cycle continues till a chain of ubiquitin molecules are attached to the target protein (Hartmann-Petersen et al., 2003). Depending on the substrate and E2/E3 complex, several types of conjugates are possible (Smalle and Vierstra, 2004). In some cases, only a single Ub is added; this monoubiquitination can direct proteolytic targets to the lysosomes/vacuole for turnover (Bach and Ostendorff, 2003) or modify transcription (Hicke, 2001). Target proteins can be in the cytoplasm, in the nucleus, on membrane surfaces that face these compartments, or even in the endoplasmic reticulum (Hershko and Ciechanover, 1998; Pickart, 2001). The resulting ubiquitinated (or ubiquitylated) proteins are then degraded by the 26S proteasome with the concomitant release of the Ub moieties for reuse (Weissman, 2001). Proteomic and genetic analyses, mainly using *Arabidopsis thaliana* as the model indicate that UPS is pervasive in plants, and it regulates almost all aspects of plant growth and development (Vierstra, 2009). Internally, the targets of UPS in plants include numerous intracellular regulators which have central roles in hormone signalling (Hellmann and Estelle, 2002; Vierstra, 2003), regulation of chromatin structure (Sridhar et al., 2007; Liu et al., 2007) and transcription (Hicke, 2001), self

recognition etc. Data accumulated also reveals that UPS is involved in host-pathogen interactions (Devoto et al., 2003; Rosebrock et al., 2007) and many other responses of plants to environmental challenges. Given the importance of light to a plant's survival, a number of UPS components have been implicated in photomorphogenesis too (Hoecker, 2005; Henriques et al., 2009).

2.3 Light signaling in plants: perception and transduction

Being the primary source of energy, light is one of the most important environmental factors for a plant (Hoecker, 2005). Apart from the crucial role in photosynthesis, it controls a large number of processes in the plant, such as seed germination, seedling photomorphogenesis, phototropism and gravitropism, chloroplast movement, shade avoidance, circadian rhythms and flower induction (Neff et al., 2000; Sullivan et al., 2003). The perception, interpretation and transduction of different wavelengths, intensities, direction and duration of light are accomplished by a sophisticated photosensory system which includes photoreceptors and signaling intermediates (Franklin et al., 2007; Josse et al., 2008).

2.3.1 Photoreceptors

Photoreceptors perceive the light signal which is then transduced via the downstream signaling intermediates to cause a physiological response (Chen et al., 2004; Quail, 2002; Somers and Fujiwara, 2009; Chen and Chory, 2011). The different photoreceptors known in plants are- Cryptochromes, Phototropins and ZTL/LKP2/FKF1 (Zeitlupe/LOV Kelch repeat Protein 2/Flavin-binding, Kelch repeat, F-box) are responsible for UV-A/blue region of spectrum, phytochromes perceive red and far-red light (Lin and Shalitin, 2003; Chen et al., 2004) and the most recently discovered UV-B receptor- UVR8 (Rizzini et al., 2011) is involved in UV-B signaling.

2.3.1.1 Cryptochromes

Cryptochromes are the blue/UV-A light sensing flavoproteins present in plants and animals (Cashmore et al., 1999; Li and Yang, 2007). In Arabidopsis, two members of this protein- CRY1 and CRY2 were identified (Lin et al., 1998). Three members of this family, CRY1a, CRY1b and CRY2 have been identified in tomato (Ninu et al., 1999; Perrotta et al., 2000). These are structurally similar to bacterial

DNA photolyases but do not possess any photolyase activity. Most of the plant cryptochromes are 70–80-kD proteins with two recognizable domains: amino-terminal photolyase homology region (PHR) carries two non-covalently bound chromophores: a primary or catalytic flavin adenine dinucleotide (FAD) and a second light-harvesting chromophore, a pterin or deazaflavin (Lin and Shalitin, 2003; Liscum et al., 2003), and a C-terminal extension. Cryptochromes mediate light dependent redox reactions. Accumulated data favor the nuclear localization of cryptochromes, where they interact with many downstream factors.

2.3.1.2 Phototropins

Phototropins are blue/UV-A light-absorbing flavoproteins that mediate phototropic responses to the direction of blue light (Briggs et al., 2001). Isolation of nonphototropic mutants in *Arabidopsis* like *nph1* led to the discovery two phototropins- phot1 and phot2 (Sakai et al., 2001). Tomato also has two phototropins that function in regulating many aspects of plant physiology and development. Majorly, they regulate phototropism, stomatal opening, chloroplast accumulation, leaf expansion and also show differential photosensitivity to blue light (Holland et al., 2009; Suetsugu and Wada, 2013; Sulabah et al., 2014). Structurally, these belong to a larger family of proteins known as LOV (**L**ight, **O**xygen, and **V**oltage) domain family (Briggs and Christie, 2002; Christie, 2007). The C-terminal portion of phototropin contains a classical Ser/Thr kinase domain, whereas the N-terminal half has two LOV domains, which are related to the PAS (**P**er- period circadian protein, **A**rnt- aryl hydrocarbon nuclear translocator protein, **S**im- single minded protein) domains (Kong et al., 2007; Inoue et al., 2008 a). Upon light absorption, the excited LOV domain causes in autophosphorylation of the receptor and thus initiates phototropin signaling involving the recruitment of multiple signaling partners (Christie and Murphy, 2013; Hohm et al., 2013).

2.3.1.3 Zeitlupelike/FKF1/LKP2

Recently new class of putative blue-light receptors related to phototropins were identified. This family of Zeitlupe/FKF1/LKP2 includes proteins that mediate targeted proteolysis of components associated with circadian clock function and flowering and have only one LOV domain (Banerjee and Batschauer, 2005). These proteins form niche of novel putative blue light receptors known as ZTL/ADO family and mainly

comprise of Zeitlupe (ZTL), Flavin binding, Kelch Repeat, F-box 1(FKF1) and LOV Kelch, Protein 2 (LKP2) (Jarillo et al., 2001, Kiyosue et al., 2000, Schultz et al., 2001, Somers et al., 2000). The role of these proteins was ascertained by loss and gain of function analyses, which predicts their need in circadian clock function and photoperiodic dependent flowering in Arabidopsis (Kim et al., 2007). To date only FKF1 (Flavin-binding, Kelch Repeat 1) has been demonstrated to function as a photoreceptor while no light regulated role was ascribed to others (Möglich et al., 2010).

2.3.1.4 Phytochromes

Phytochromes are red/far-red photoreceptors and, together with cryptochromes and phototropins, constitute one of the three major regulators of photomorphogenesis in plants. As in Arabidopsis, tomato also has five phytochromes- PhyA, PhyB1, PhyB2, PhyE and PhyF (Sharrock and Quail, 1989; Pratt et al., 1997). Based on their stability, phytochromes are divided into two types- the type I photo-labile phytochromes (PhyA) degrade rapidly upon light exposure whereas the type II photo-stable phytochromes (PhyB to PhyE) are relatively stable in light.

Phytochrome is a 125 kDa polypeptide with a linear tetrapyrrole chromophore, phytochromobilin. The protein has two structural domains- N- terminal photosensory domain and the C-terminal regulatory or signal output domain (Smith, 2000; Park et al., 2000; Rockwell et al., 2010). Structurally, phytochromes contain two carboxy-terminal histidine kinase domains (HKRD) and two PAS domains within the HKRD1 domain (Hughes, 2010; Nagatani, 2010). Functional phytochromes occur as two photoreversible forms, namely; Pr form-absorbing red light maximally at 660 nm, Pfr form- absorbing far red light maximally at 730 nm (Lagarias, 1985; Mancinelli, 1994). The photosensory properties of phytochromes depend on the reversible photoconversion of the two forms (Ulijasz et al., 2010). Phytochromes are synthesized in cytoplasm in their physiologically inactive red light absorbing Pr form. On excitation, the Pr form is converted to far red light absorbing Pfr (active) form and translocated to nucleus where it interacts with different transducers to initiate downstream transcriptional cascades (Yamaguchi et al., 1999; Kircher et al., 2002; Nagy and Schafer, 2002; Wang and Deng, 2003; Ulm et al., 2006; Klose et al., 2014).

2.3.1.5 UVB receptor: UVR8

UV-B light is not only a source of damage but also an informational source for plants. UV-B light responses are fluence rate dependent eliciting a variety of responses including cellular damage (Baker et al., 1997; Sullivan and Rozema, 1999), suppression of hypocotyl growth, expansion of cotyledons, phototropic curvature, biosynthesis of anthocyanins and flavonoids and opening of stomata (Beggs and Wellmann, 1994; Kim et al., 1998; Boccalandro et al., 2001; Eisinger et al., 2003; Suesslin and Frohnmeyer, 2003; Shinkle et al., 2004). The UV RESISTANCE LOCUS 8 (UVR8) in Arabidopsis has recently been established to be a UV-B photoreceptor (Rizzini et al., 2011; Heijde and Ulm, 2012). UVR8 is a seven-bladed β-propeller protein of 440 amino acids (Christie et al. 2012, Wu et al. 2012). UV-B perception by UVR8 is mediated by a chromophore made up of at least Trp-285 and Trp-233, which directly absorb UV-B light and are excited.

2.3.2 Light signaling cascade

After perception of light signal, photoreceptors initiate a complex, integrated molecular network that results in physiological responses. Different signaling cascades can be induced either by single or multiple photoreceptors and several components are involved in the process **(Fig. 2.1)**. Molecular genetic analysis of Arabidopsis mutants led to the identification of a large number of light signaling components. Two major classes of light signaling mutants have been identified through classical genetic screens in Arabidopsis. The first group includes the photoreceptor mutants and mutants defective in the biosynthesis of chromophores (Parks and Quail, 1991, 1993; Reed et al., 1993; Whitelam et al., 1993). These classes of mutants develop normally in darkness but exhibit reduced or enhanced responsiveness to various light inputs. The second class belongs to the *constitutive photomorphogenesis (cop)/de-etiolated (det)/fusca (fus)* group (Wei and Deng, 1996), which exhibit characteristic signs of photomorphogenesis in complete darkness. Members of this group act as negative regulators of photomorphogenesis.

Interestingly, the signal cascades involve overlapping set of genes, indicating that the different light signals converge at some points and share signaling components among them. The analysis of these signaling pathways reveals a number

of commonalities such as the importance of light-regulated protein–protein interaction, protein stability, and kinase activity.

2.3.3 Positive regulators of photomorphogenesis

2.3.3.1 HY5

As different light signaling pathways integrate at common points, few light-regulated transcription factors have been identified as key regulators during photomorphogenesis. HY5 (for ELONGATED HYPOCOTYL 5) acts downstream of multiple photoreceptor pathways and is the most extensively studied positive regulator. Being a transcription factor, it directly binds to promoters of many genes that possess G-Box elements and therefore, controls their expression levels (Lee et al., 2007). The regulation of HY5 protein levels occurs at transcriptional and post-transcriptional levels, the latter including phosphorylation and also proteasomal degradation (Quail, 2002). CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1), a member of the COP/DET/FUS proteins, interacts with HY5 and promotes its degradation through the Ubiquitin/26S proteasome system in darkness (Ang et al., 1998). On the other hand, light perception inactivates this process and thus stabilizes the HY5 protein (Osterlund et al., 2000). Additionally, phosphorylation of a serine residue in a consensus CASEIN KINASE II (CKII) site in the COP1-interacting domain weakens the COP1-HY5 interaction and thus stabilizes HY5. Also, the phosphorylated HY5 protein is less effective in binding to its target promoters (Hardtke et al., 2000).

2.3.4 Negative regulators of photomorphogenesis

2.3.4.1 PIFs

Phytochrome Interacting Factors (PIFs) function as negative regulators in the phytochrome mediated light signaling pathways (Martínez-García et al., 2000). Molecular genetic studies in *Arabidopsis* reveal a family of PIFs (PIF1, PIF3, PIF4, PIF5 and PIF6 etc) having distinct and overlapping roles (Ni et al., 1998; Oh et al., 2004; Lucas and Prat, 2014). Structurally, these proteins belong to basic helix-loop-helix (bHLH) super family of transcription factors containing a PAS domain (Ni et al., 1998; Halliday et al., 1999; Huq et al., 2004). The basic domain functions in DNA

binding, whereas the HLH domain is involved in protein-protein interaction by binding to the promoters of target genes (Leivar et al., 2008; de Lucas et al., 2008; Bai et al., 2012). The protein level of PIF3 has also been shown to be down regulated via phosphorylation, polyubiquitylation and subsequent 26S proteasome-mediated degradation upon either red or far-red light exposure (Park et al., 2004; Bauer et al., 2004; Al-Sady et al., 2006; Feng et al., 2008). The nuclear localization of PIF3 and its regulation through posttranslational protein degradation imply its key role in integrating photosignals perceived by photoreceptors and then regulating transcription of downstream genes that result in either inhibition or promotion of photomorphogenesis (Shen et al., 2007; Nozue et al., 2007; Lorrain et al., 2008; Al-Sady et al., 2006).

2.3.4.2 SPA quartet- Suppressor of Phytochrome A (SPA) family proteins

SUPPRESSOR OF PHYTOCHROME A is a family of four proteins that negatively regulates phytochrome-specific light responses (Hoecker, 2005; Yi and Deng, 2005). SPA1 (SUPPRESSOR OF PHYA-105) and SPA2 (SPA1-RELATED 2) are mainly involved in the suppression of photomorphogenesis in darkness. SPA3 (SPA1-RELATED 3) and SPA4 (SPA1-RELATED 4) are responsible for the prevention of light-overstimulation. SPA3 and SPA4 additionally have important roles in inhibiting the development of dwarfism (Laubinger et al., 2004). In light grown seedlings, SPA1, SPA3 (SPA1-RELATED 3) and SPA4 (SPA1- RELATED 4) are responsible for the prevention of light-overstimulation. These notions are in agreement with the SPA proteins having overlapping but distinct functions in the regulation of plant development. Consistently, single, double or triple *spa* mutants show no phenotype at all or exhibit only restricted phenotypes under certain light conditions (Laubinger and Hoecker, 2003; Laubinger et al., 2004). On the contrary, *spa* quadruple mutants develop severe *cop* phenotype almost indistinguishable from *cop1* mutants (Laubinger et al., 2004). These findings strongly indicate that the SPAs are required for the suppression of photomorphogenesis in darkness, and that they work in concert with COP1 to perform this regulatory task.

Structurally, SPA proteins contain an N-terminal kinase like domain, coiled-coil domain followed by seven WD40 repeats at the C-terminal region that is highly

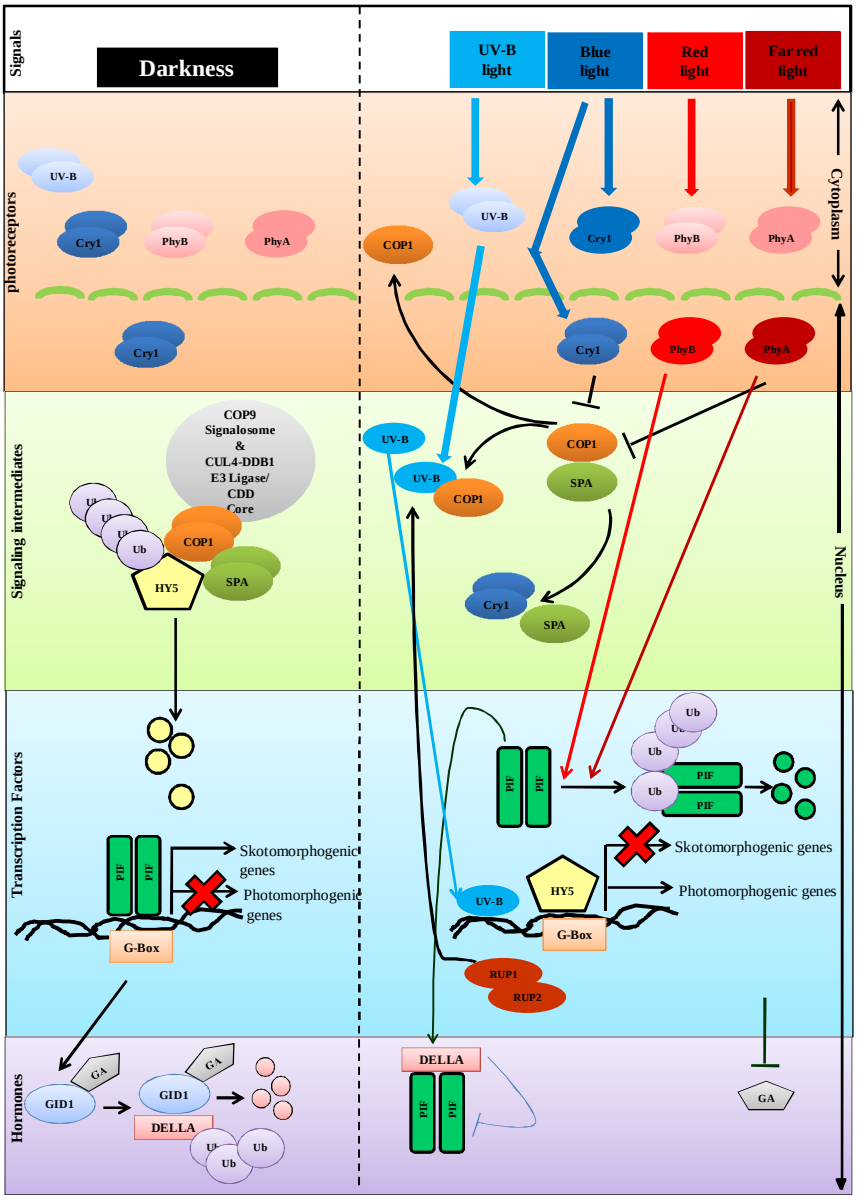


Figure 2.1 Light signal perception and signaling during Arabidopsis seedling photomorphogenesis. Left, The model shows light signaling molecules in darkness starting from the inactive photoreceptors, Ubiquitin/26S proteasome complex, in particular CUL4-DDB1$^{COP1-SPA1}$ E3 ligase complex degrades HY5 and other transcription factors that promote photomorphogenesis. High levels of gibberellins induce degradation of DELLA. **Right,** The mode of action of photoreceptors in the presence of light. In the nucleus, phytochromes (that migrate from the cytoplasm in their active form) and cryptochromes interact with COP1, reducing its activity and allowing the abundance of HY5 to increase. The activity of PIFs and gibberellin levels are reduced, DELLA accumulate and bind PIFs further reducing their activity. The above model is taken from Mazzella et al., (2014) and UV-B signaling is based on the model described in Heijde et al., (2012).

homologous to the COP1 WD40 domain (Laubinger and Hoecker, 2003; Laubinger et al., 2004). Biochemical studies demonstrated that SPA1 forms heterodimers with COP1. Similarly, all the SPA proteins interact with the coiled coil domain of COP1 (Hoecker and Quail, 2001; Laubinger and Hoecker, 2003; Saijo et al., 2003; Laubinger et al., 2004; Ranjan et al., 2014). Transient expression studies in Arabidopsis seedlings revealed that COP1 and SPA1 co-fractionate in high molecular weight complexes (Seo et al., 2003). In addition, along with LAF1 (**L**ong **A**fter **F**ar-red light1), COP1 and SPA1 form subnuclear speckles (Saijo et al., 2003; Seo et al., 2003). Furthermore, yeast two-hybrid assays showed that SPA1 binds to HY5 via WD40 domain. This is further supported by the accumulation of HY5 protein in light grown Arabidopsis *spa1* mutants (Saijo et al., 2003). Till date the accumulated data suggest that SPAs might enhance COP1 activity by helping to recruit its targets (Ranjan et al., 2014). Also, they might affect the stability of COP1 complexes by altering the nuclear accumulation. Though the exact mechanisms involved are not revealed and given the presence of kinase-like domains in the SPAs, one cannot rule out phosphorylation of COP1 and thus, its control by SPAs (Yi and Deng, 2005). Recent studies on Psychometrilla revealed that SPAs are functionally diversified and they control the COP1 activity in different species (Ranjan et al., 2014).

The identification of SPA like proteins- RUP1 (Repressor of UVR8 protein1) and RUP2 (Repressor of UVR8 protein2) involved in Arabidopsis UVB signalling (Gruber et al., 2010) also support the functional diversification of SPAs. Large scale sequencing and genomic studies in rice and Physcomitrella predicted the presence of orthologs of the Arabidopsis RUP genes (International rice genome sequencing project, 2005; Richardt et al., 2007). In tomato, the SPA3LIKE protein (present study) is predicted to be an ortholog of COP1 but is found to be highly similar to Arabidopsis RUP2 protein.

2.4 COP1: An E3 ligase

Many negative regulators of the light signaling cascades have been revealed over the years, mostly by genetic approaches. Genetic and biochemical studies in Arabidopsis have identified a group of pleotropic Constitutive Photomorphogeneic/De-etiolated/Fusca (COP/DET/FUS) proteins that act as central negative regulators of photomorphogenesis (Sullivan et al., 2003, Yi and Deng 2005).

Among them, COP1 (constitutive photomorphogenic protein 1) was one of the first cloned and is one of the most extensively studied. COP1 is an E3 ubiquitin ligase present in many organisms including plants and humans; however, it was first described in *Arabidopsis thaliana*.

At COP1 has three distinct domains – a RING finger motif, coiled coil region and seven WD40 repeats at its C-terminus. The RING finger motif mediates interaction of COP1 with the Ubiquitin conjugating enzyme (E2); coiled coil domain helps in the formation of COP1 homo-dimers or hetero-dimers with the SPA protein family (Hoecker and Quail 2001 and Yi and Deng 2005), the WD40 repeats region is the substrate binding domain. This is essential for the degradation of a number of bZIP-type transcription factors, such as HY5, HYH, LAF1 and HFR1 (Osterlund et al., 2000; Holm et al., 2002; Seo et al., 2003; Jang et al., 2005; Yang et al., 2005); transcriptional regulators and even photoreceptors like cryptochromes and phytochromes (Yang et al., 2005; Shalitin et al., 2002; Seo et al., 2004; Chen et al., 2010; Jang et al., 2010). Moreover, COP1 is responsible for light dependent gene regulation for the majority of light controlled genome expression (Ma et al., 2002). In the past few years, accumulated data show that the number of substrates of COP1 has more than doubled in plants which suggest the importance of COP1 in various plant developmental processes.

2.4.1 COP1 in other organisms

Although COP1 exists in almost all vertebrates including fish, amphibians, birds and mammals, unlike plants, these do not undergo photomorphogenesis. The function of COP1 is less understood among this group. Structurally mammalian COP1 has a high degree of sequence conservation and domain organization with AtCOP1 (Yi et al., 2002; Bianchi et al., 2003). Several initial studies have implicated that mammalian COP1 plays a role in tumorigenesis by targeting substrates like tumor suppressor p53, the proto-oncogene c-jun and more recently, metastasis associated protein 1 (MTA1) and the ETS transcription factor ERM (Dornan et al., 2004; Baert et al., 2010). In mammals, p53 is a potent tumor suppressor protein that controls the fate of cell cycle (Vousden and Prives 2009). *In-vitro* and *in-vivo* studies revealed that COP1 acts as an E3 ligase for p53 and thus negatively regulates its function (Dornan et al., 2004a). A protein kinase ATM (ataxia telangiectasia mutated) acts as a primary responder in

DNA damage signaling pathway and in response to DNA damage phosphorylates COP1 at Ser^{387}, leading to autodegradation of COP1 and therefore stabilization of p53. In addition to the above targets, MTA1, a component of nucleosme remodeling and histone deacetylation) complex involved in tumorigenesis has recently shown to be a target of COP1 (Li et al., 2009).

Recent studies also reveal the role of MmCOP1 in lipid and glucose metabolism. Acetyl-Coenzyme A carboxylase (ACC) is the rate limiting enzyme that catalyzes the formation of malonylCoA from acetylCoA. It is known that suppression of ACC promotes the utilization of lipids. During fasting, a pseudokinase Tribbles 3 (TRB3) interacts with COP1 and elevates the loss of fat by targeting ACC for ubiquitination and degradation (Qi et al., 2006). COP1 also functions as an E3 ligase for the key regulators TORC2/CRTC2 (an activator for cAMP responsive CREB) and FoxO1 (fork head transcription factor) which activate the expression of rate limiting enzymes, G6Pase and PEPCK in liver gluconeogenesis. Degradation of TORC2 and FoxO1 inhibits gluconeogenesis and therefore controls the blood glucose levels (Dentin et al., 2007 and Kato et al., 2008), a process usually regulated by insulin. As the hormone can induce COP1, it would be interesting to understand the role of COP1 in human diabetes and insulin resistance.

2.4.2 Multisubunit protein complexes

The pleotropic proteins that negatively regulate photomorphogenesis comprise of three distinct groups (**Fig. 2.2**) - COP1 and SPA complexes, COP9 signalosome (CSN) and the CDD complex (COP10, DDB1 and DET1). All of these proteins are involved in proteasome-mediated degradation of photomorphogenesis promoting factors (Saijo et al., 2003, Serino and Deng 2003, Yi and Deng 2005, Zhu et al., 2008; Huang et al., 2014). Recently, it had been shown that the COP1-SPA complexes and the CDD complex interact with CUL4 based E3ligases in vivo and therefore regulate the degradation of transcription factors involved in light mediated plant development (Chen et al., 2010).

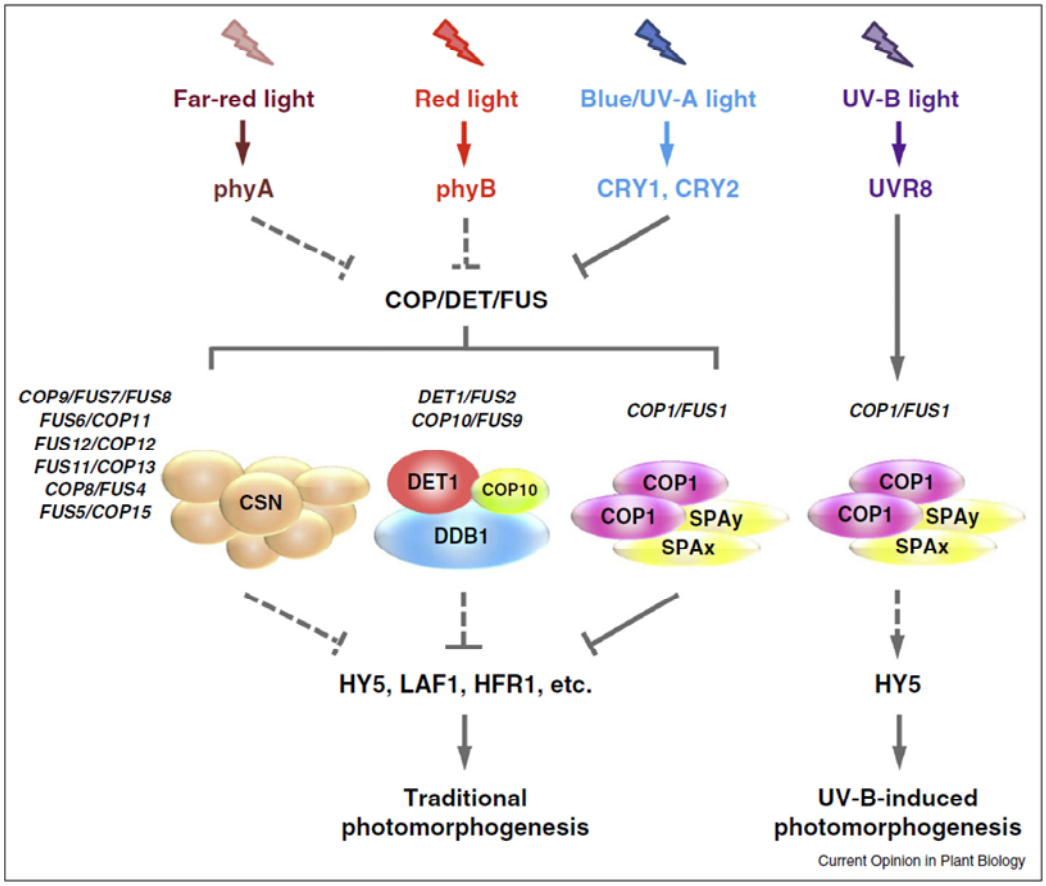


Figure 2.2 A simplified model representing the COP/DET/FUS complexes in photomorphogenesis. The COP/DET/ FUS proteins are components of the CSN complex, the CDD complex and the COP1–SPA complex. Photoreceptors are capable of sensing light ranging from UV-B to far-red light: phytochromes (phyA and phyB) for far-red and red light, cryptochromes (cry1 and cry2) for blue/UV-A light, and UVR8 for UV-B light. The degradation of photomorphogenesis-promoting transcription regulators like HY5, LAF1 and HFR1 under light conditions is depicted in the model. The activity of COP/DET/FUS complexes is inhibited by light signals. In UV-B-induced photomorphogenesis, the COP1-SPA complex accumulates and promotes light signaling by stabilizing HY5. The tetrameric COP1-SPA core complex consists of two COP1 proteins, and two SPA proteins. These two SPA proteins can be homogenous as indicated by SPAx, or heterogeneous as indicated by SPAx and SPAy. **Note**: solid lines represent direct effects, while dotted lines indicate indirect or unidentified direct regulation. This model is taken from Huang et al., (2014).

2.4.2.1 Cop9 signalosome (CSN)

Six of the subunits of COP9 signalosome were originally identified as repressors of photomorphogenesis in the original *cop*/*fus* mutant screens. CSN is a 550 kDa multimeric protein comprising of eight subunits that show structural and sequence homology to the 19S lid of the proteasome (Wei et al., 1998; Wei and Deng, 2003). *csn* mutants of Arabidopsis exhibit pleotropic responses and strong alleles are lethal. In addition, mutation in any single subunit can destabilize the entire complex (Schwechheimer et al., 2002) exemplifying the importance and role of CSN in plant development. Several biochemical studies indicated that the CSN interacts with a wide range of E3s, including COP1, SCF^{COI}, SCF^{TIR1}, and SCF^{UFO} (Schwechheimer et al., 2001; Feng et al., 2003) and modulates many plant responses.

2.4.2.2 COP10-DDB1-DET1 complex (CDD)

In-vivo, the pleotropic genes *cop/det/fus* are known to be involved in large protein complexes that help in recruiting the target protein and therefore helps in its degradation via UPS. Studies in Arabidopsis (Yanagawa et al., 2004) reveal that DET1 and DDB1 form a complex with COP10 called the CDD (for COP10-DDB1-DET1). It is well established that both the Arabidopsis *det1* and the *cop10* mutants have a phenotype similar to *cop1* (Suzuki et al., 2002) i.e. exhibit constitutive photomorphogenesis even in the absence of light. Extensive work by Deng et al., show that COP10 lacks a cys residue to function as an E2 and therefore is related to ubiquitin E2s. It was also shown that COP10 enhances the activity of E2s in vitro, either by itself or as part of the CDD complex (Lau and Deng, 2009). The complex interacts with both COP1 and the CSN. This emphasizes the central role of the COP-CDD complex in mediating physiological responses by promoting the degradation of positive regulators of photomorphogenesis like HY5 (Yanagawa et al., 2004).

2.4.2.3 Nucleocytoplasmic localization of *COP1*

The structural organization of COP1 protein was well defined for Arabidopsis. Many genetic, molecular and biochemical studies revealed that COP1 has three distinct domains along with a nuclear and cytoplasmic localization signals. While the core domains- ring finger, coiled coil and the WD40 domain were required for an efficient E3 ligase activity, a spatial and temporal regulation was mediated by its

nuclear and cytoplasmic localization signals. Studies reveal that COP1 was mainly localized to the nucleus of dark grown seedlings where it presumably ubiquitylates and thus promotes the degradation of the light development-promoting transcription factors and thereby hindering the early activation of photomorphogenesis genes. Light exposure results in the drastic reduction of nuclear COP1 levels, allowing these transcriptional regulators to re-accumulate and light induced transcription to occur (Ang et al., 1998; Holm et al., 2002; Seo et al., 2003, 2004; Jang et al., 2005; Yang et al., 2005). This contrasts with the fast change of expression of the early light-responsive genes (detectable already less than 1 hour after light exposure) suggesting that other mechanisms may also operate to rapidly repress COP1 function upon light stimulus (Tepperman et al., 2001). Nonetheless, its nuclear-cytoplasmic partitioning still plays a pivotal role in COP1 regulation.

The coiled-coil domain and NES signal of COP1 seem to partially overlap with signal sequence of sub-nuclear structures (Stacey et al., 2000). Additionally, the WD40 domain through which COP1 interacts with most of its partners has also been shown to be important for speckle formation. Three independent mutations in the WD40 motif abolish sub-nuclear speckles, and each of the corresponding mutant seedlings develop strong *cop* phenotypes and die before adult stage (McNellis et al., 1994b).These findings are consistent with the notion that incorporation of COP1 into sub-nuclear speckles is an essential aspect of its function. However, a more definite understanding of physiological significance of the sub-nuclear foci, as well as the elucidation whether they represent sites of active COP1 function or storage sites for inactivated COP1 proteins, requires further investigations including combined mutant and transgenic plant studies and detailed biochemical characterizations (Yi and Deng, 2005). In addition to its relatively slow light induced nuclear exclusion, another mechanism has been proposed for the inactivation of COP1 function, which may explain the fast responsiveness of the light activated genes (Yi and Deng, 2005). Cryptochromes, CRY1 and CRY2 have been shown to directly interact with of COP1 through their C-terminus and the COP1 WD40 domain in both light and darkness, and to negatively regulate COP1 activity in response to blue light (Wang et al., 2001). Conformational changes in the C-terminus of cryptochromes after light perception and activation of cryptochromes by auto-phosphorylation induces structural modifications of the receptor-bound COP1, that in turn releases the previously

attached HY5 from COP1 (Wang et al., 2001; Chen et al., 2004; Yi and Deng, 2005). Whether phytochromes also utilize a similar mechanism to shut off COP1 function in red and far-red light remains to be clarified. Notwithstanding this, rapid phosphorylation based cryptochrome activation might represent a fast acting COP1 deactivation mechanism allowing the dynamic onset of light induced transcription. The long-term inactivation of the COP1 function on light-responsive genes then in turn may occur by the subsequent nuclear depletion (Osterlund et al., 1999; Yi and Deng, 2005). Nuclear export-import mechanisms are not the sole regulatory steps that control COP1 function. In living cells, COP1 can be found in stable, high-molecular-mass protein complexes (Yi et al., 2002; Saijo et al., 2003). Genetic, yeast-two hybrid and protein complex purification approaches have identified numerous factors interacting with COP1, several of which are important physiological regulators of COP1 activity.

2.4.3 Known *cop1* and *spalike* mutants

The profound effect of light is demonstrated during the early development of seedlings. In the dark, seedlings undergo skotomorphogenic development wherein, the hypocotyl is elongated with closed cotyledons, weak root system and also the seedlings lack pigmentation. On the other hand when exposed to light, seedlings exhibit a characteristic de-etiolated growth. This process is termed as photomorphogenesis, where the seedlings have short hypocotyls, open and green cotyledons and a better root system. Therefore, the defective seedling phenotype in different light conditions is strong basis for isolating light signalling mutants. Arabidopsis *cop1* mutants were identified with great effort and most of the mutants exhibit strong phenotypes and do not survive. Only two mutants of Arabidopsis- *cop1-4* and *cop1-6* are weak and exhibit exaggerated photomorphogenic responses (Deng and Quail, 1992). COP1 functions have also been described in other flowering plant species. In rice the COP1 ortholog, PETER PAN SYNDROME1 (PPS) shortens the juvenile phase, a phenotype not reported for Arabidopsis, and delays flowering in short and long day (Tanaka et al., 2011). The COP1 ortholog of pea, LIGHT-INDEPENDENT PHOTOMORHOGENESIS1 (LIP1), regulates seedling growth by affecting gibberellic acid levels (Sullivan and Gray, 2000). In apple, *Md COP1* affects anthocyanin levels in the fruit peel (Li et al., 2012). COP1 also exists in non-plant lineages, e.g. humans,

where it acts as an E3 ubiquitin ligase to control the protein stability of a number of transcription factors, e.g. p53, cJun (Vousden and Prives 2009; Baert et al., 2010).

As discussed in the previous sections, COP1 and SPA work together in degrading target proteins. Mutations in *SPA* family of genes also exhibit *cop* like phenotype.

2.4.4 *COP1* and *SPA3LIKE* mutants in tomato

In tomato, many labs have tried to isolate *cop1* mutants. A recent report by Jones et al., (2012) showed that though *cop1* mutants were isolated, no observable phenotypes were detected. In the process of sequencing the tomato genome Liu et al., (2004) identified a negative regulator of photomorphogenesis- *COP1LIKE*, now denoted as *SPA3LIKE*. The RNAi lines of this gene also showed enhanced photomorphogenic responses at seedling level. Other than the RNAi lines, no known mutants of *SPA3LIKE* exist in tomato.

2.4.5 Physiological response controlled by *COP1* in plants

Light influences almost all aspects of plant development. The complex interactions involved in the perception and transduction of light signal are controlled by the master regulator COP1 to execute the downstream physiological responses. Therefore, COP1 is anticipated to play an important role in almost all the plant responses mediated by light. The major role of COP1 was elucidated by understanding major light responses like- photomorphogenesis, shade avoidance and flowering.

2.4.5.1 Photomorphogenesis

As discussed in the earlier section, in dark, seedlings are etiolated with elongated hypocotyls and unopened cotyledons. On the contrary, light grown seedlings are de-etiolated with short hypocotyl and opened cotyledons. This dramatic effect of light seen during the seedling development of higher plants requires the activity of the repressor CONSTITUTIVE PHOTOMORPHOGENIC1 (COP1). COP1 being an E3 ligase targets the light signaling components for ubiquitin mediated proteasomal degradation and suppresses photomorphogenesis in the dark (Hoecker, 2005). The presence of protein-protein interacting domains allows COP1 to interact with several target proteins. One of the first identified targets is HY5, a bZIP

transcription factor (Osterlund et al., 2000) abundant in light grown seedlings. Upon light to dark transition, HY5 is degraded. *cop1* mutants of Arabidopsis are impaired of this light-dark recognition and behave constitutively photomorphogenic suggesting that, functional COP1 is required for the HY5 destabilization in darkness. Further, it was shown that the two proteins directly interact through the WD40 domain of COP1, and mutations in both of these domains affected the association of the two proteins resulting in HY5 stabilization (Ang et al., 1998; Osterlund et al., 2000; Holm et al., 2001). Nuclear co-localization studies of COP1 and HY5 proteins support their physiological relevance in photomorphogenesis (Ang et al., 1998). And finally, on exposure of light, COP1 is excluded from nucleus thus stabilizing HY5 (von Arnim and Deng, 1994; Osterlund et al., 2000). HY5 forms functional heterodimers with another bZIP transcription factor known as HYH involved in PhyB signaling. It was also shown to be a target of COP1 in the dark (Holm et al., 2002). Many other transcription factors are known to be degraded in COP1-dependent manner. For example, LAF1, a MYC-type transcription factor, involved in positively regulating PhyA mediated far-red light signaling (Seo et al., 2003), HFR1, a bHLH transcription factor, another positively acting regulator of blue and far-red light signal transduction pathways, are targeted by COP1 mediated ubiquitylation for proteasomal degradation in darkness (Jang et al., 2005; Yang et al., 2005). Likewise, COP1 mediates ubiquitylation and rapid proteasomal degradation of the light labile PHYA and CRY2, therefore preventing the over activation of signaling cascades (Shalitin et al., 2002; Seo et al., 2004). Furthermore, the photoreceptors- PHYB and CRY1 have also been shown to directly interact with COP1 (Wang et al., 2001; Seo et al., 2003).

Interestingly, light-activated cryptochromes may directly turn off COP1 activity, thereby releasing genes suppressed by COP1. These findings support the hypothesis that in the dark COP1 mainly acts as a master switch that mediates the photomorphogenic responses. Additionally, a recent study demonstrated that COP1 is required for the nuclear accumulation but not for the red and far-red light triggered rapid degradation of the bHLH transcription factor, PIF3 that is a proposed negative regulator of phytochromes-mediated light responses (Bauer et al., 2004). Also, a positive regulatory function of COP1 in PhyB-mediated red light signalling has recently been proposed, suggesting that COP1 might play opposing roles depending on light conditions (Boccalandro et al., 2001).

2.4.5.2 Shade avoidance

Light is an important source of energy for proper development of plants. In natural vegetations, the amount and quality of light perceived by each plant depends on many factors, majorly the density of the canopy. Usually, plants grown in close proximity to each other compete for resources. Plants respond to this shade by avoidance mechanism that includes responses like elongation of stem and petiole, appearance of large and thinner leaves, reduction of photosynthate in the shoot at the expense of root growth, increased apical dominance etc. In severe conditions, the developmental decisions like flowering time, fruit and seed development are drastically affected (Smith and Whitelam, 1997). Therefore, shade is a major aspect that affects yield. Phytochromes are the major photoreceptors that mediate the shade avoidance response. Because shade reduces PHYA, PHYB, CRY1, and UVR8 activity, it is not clear whether shade should increase or decrease COP1 activity. COP1 is also required (likely indirectly) for the accumulation of PIF3 in the dark, suggesting that it could be required for the accumulation of PIFs in response to shade.

2.4.5.3 Flowering

Perception of day length mediates the transition from vegetative to flowering stages in plants, based on which they are classified either as long day (LD), short day (SD) or day neutral. The transcriptional regulator CONSTANS (CO) promotes flowering by triggering the expression of flowering locus FT that leads to flowering. Molecular investigations of photoperiodic flowering pathway reveal that circadian clock and light play a major role in regulating this pathway by transcriptional and translational regulation of CO (Hyama and Coupland, 2003). It has been established that the CO protein abundance is controlled by the E3 ubiquitin ligase- COP1, HOS1 (HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENE1) in both LD and SD, and SPA1 (which targets degradation of CO only during SD). In the dark, COP1 degrades CO via Ub/26S proteasome pathway (Jang et al., 2008). Mutants of *cop1* flower early under SDs and the phenotypes are largely suppressed in *co* mutant background. In presence of light, in particular blue and far red, decrease in COP1 activity leads to the accumulation of CO. Furthermore, the recently identified of BBX (B-box transcription factors) proteins are found to interact with COP1 and thus, act to

modulate flowering through CO-dependent or -independent pathways (Khanna, et al., 2009).

The availability of *cop1* mutants in Arabidopsis allowed for the detailed characterization and elucidation of *cop1* role in plant development. However, not all plant developmental processes are well established and still need extensive research. One such fascinating area is fruit ripening.

2.5 Tomato- a model for fruit ripening

Tomato is an excellent model species for investigating mechanisms for fruit development and ripening of fleshy fruits because of characteristics such as efficient cross and self pollination ability, short generation time and year round growth potential in greenhouse and dramatic color changes during fruit maturation. In addition, diploid genome (n=12; genome size around 950 Mb) with same haploid chromosome number and a high level of gene synteny with the other Solanaceae plants also makes it a model plant for Solanaceae (Tanksley 2004). Since the intensive cultivation of tomato 150 years ago, a large variety of resources have been developed such as genetic maps, advanced breeding populations, collections of wild and cultivated accessions and complete genome sequence. Since tomato is amenable to transformation with Agrobacterium, it can be genetically engineered too. For all these mentioned above physiology, genetics of tomato is extensively studied and recently it is also used as a model for relating fruit development and ripening to metabolite, proteome and trancriptome changes. In view of tomato being an excellent source of dietary carotenoids, we focused our study on manipulating levels of carotenoid in tomato fruit.

Carotenoids are organic pigments present in chloroplasts and chromoplasts of all photosynthetic organisms, rendering many key physiological roles. As accessory pigments of light harvesting complex, they participate in photosynthesis and protect plants from photo-oxidation. Further, they contribute to the different colors of flowers and fruits and also form the precursors for abscisic acid, strignolactone and several volatiles. Some carotenoids fulfill the dietary requirements of humans and animals. β-carotene serves as the main dietary source of vitamin A (Grune et al., 2010), whilst there is a considerable body of evidence that carotenoids like- lycopene, zeaxanthin and lutein act as antioxidants when consumed in necessary amounts. Recent studies

also link the intake of tomato to reduced cancer risks (Giovannucci et al., 2002) which has attracted considerable attention for research on tomato as it is the major source of dietary lycopene.

2.5.1 Carotenoid biosynthesis and regulation during fruit ripening

Carotenoid biosynthesis in higher plants is a well studied pathway (Hirschberg 2001; Fraser and Bramley 2004; Lu and Li 2008). Carotenoids are synthesized by nuclear encoded enzymes that are synthesized in the plastids. Like all other isoprenoids, carotenoids are also built from IPP (isopentenyl disphosphate). In plastids, a set of enzymes majorly- DXS (1-deoxyxylulose 5-phosphate synthase), DXR (DOXP reductoisomerase), IPS (IPP isomerase) and GGPPS (geranylgeranyl diphosphate (GGPP) synthase) act sequentially to produce the precursor GGPP via the plastid localized MEP pathway. Phytoene synthase (PSY) catalyzes the condensation of two GGPP molecules to produce phytoene, the first committed step in carotenoid pathway. PDS (phytoene desaturase) and ζ-carotene desaturase (ZDS) are structurally and functionally similar enzymes. Next in the pathway, these enzymes convert phytoene to lycopene via ζ-carotene. This step is followed by the cyclization of lycopene, a branch point in the pathway. One branch leads to β-carotene and its derivatives xanthophylls, while the other leads to α-carotene and lutein. From lycopene, β-carotene and δ-carotene are produced by the action of chromoplasts specific lycopene β-cyclase (CYCB) and lycopene ε-cyclase (LCYE) respectively. Consequently, β-carotene is converted to zeaxanthin via β-cryptoxanthin by β-carotene hydroxylases. There are two β-carotene hydroxylases in both Arabidopsis and tomato. Zeaxanthin epoxidase (ZEP) converts zeaxanthin to violaxanthin via antheraxanthin. In leaves, violaxanthin can be converted back to zeaxanthin by violaxanthin deepoxidase (VDE). The whole pathway can be divided into two major steps controlled by the rate limiting enzymes – PSY1 and CYCB.

Fruit ripening is classified into two types- climacteric and non-climacteric, based on the respiration and ethylene production. The ripening of climacteric/fleshy fruits, such as tomato, cucurbits, avocado, banana, peach, plum and apple requires ethylene to coordinate and complete the ripening process. On the other hand, non-climacteric fruits such as strawberry, grape, and citrus respond to ethylene but do not require ethylene to complete ripening. It has been reviewed that fleshy fruit ripening

is a complex process involving many physiological and biochemical events that lead to accumulation of pigments, metabolic changes related to flavor and nutrient composition, softening of cell wall etc and therefore results in seed dispersal (Bapat et al., 2010; Klee and Giovannoni, 2011). Over two decades the molecular, biochemical and genetic investigations in tomato focused largely on different aspects of developmental regulation of fruit ripening.

Ethylene is the most crucial hormone in fruit ripening. All the defined components of Arabidopsis ethylene biosynthesis and perception are present in tomato. Antisense suppression of the two ethylene biosynthetic enzymes – ACS2 (aminocy-clopropane-1-carboxylic acid (ACC) synthase) and ACO (ACC oxidase) prevented normal ripening, thereby, demonstrating the importance of ethylene biosynthesis in ripening (Oeller et al. 1991 and Hamilton et al. 1990). Unlike the biosynthesis, perception and signaling of ethylene involves receptors and many downstream components. There are seven ethylene receptors in tomato, (ETR1, ETR2, ETR3/Nr, ETR4, ETR5, ETR6 and ETR7) and five in Arabidopsis, all act as negative regulators of ethylene signaling. Silencing of either ETR4 or ETR6 with a fruit specific promoter caused an early ripening phenotype (Kevany et al. 2008). CTR1 (CONSTITUTIVE TRIPLE RESPONSE1) is downstream of ethylene receptors and its abundance is induced by exogenous ethylene application and ripening (Adams-Phillips et al., 2004). The later steps of signaling is carried by EIN's (Ethylene insensitive) and EIL's (Ethylene insensitive like) factors which further activate a large family of ETHYLENE RESPONSE FACTORS (ERFs). Among these *ERF6* act as a negative regulator of ethylene and carotenoid biosynthesis in maturing tomato fruit (Lee et al., 2012).

Molecular genetic studies on three ripening inhibited mutants of tomato-*ripening-inhibitor* (*rin*), *non-ripening* (*Nor*) and *Colourless non-ripening* (*Cnr*) mutations led to the identification of RIN-MADS and CNR-SPL proteins that act as core regulators of ripening. RIN-MADS has been shown to bind promoters of target genes at CArG-box recognition sequences (Martel et al., 2011). On this basis, many targets for RIN-MADS have been identified (Qin et al., 2012). Martel et al., (2011) showed that PSY1 is a direct target of RIN-MADS protein and hence is regulated at the transcriptional level. During ripening process, the levels of PSY1 are elevated which leads to the flux in the carotenoid pathway. However, it is terminated in

lycopene due to the ethylene mediated repression of lycopene β-cyclase thus blocking further metabolism. The targets of CNR-SPL are still unspecified, the levels of *FRUITFUL1* and *FRUITFUL2* genes are found to be suppressed in *Cnr* mutant. In addition to these, several other transcription regulators were identified by using reverse genetic techniques. The HD-zip homeobox protein (HB1) binds to the promoter of *ACO1* gene and therefore reduces the transcription (Lin et al., 2008). Similarly, additional factors like *TAGL1* and *AP2* were identified by trancriptome studies.

Apart from regulation of fruit ripening by ethylene, studies also revealed the role of other hormones like auxin and ABA in fruit development (Jones et al., 2002; Sagar et al., 2013, Sun et al., 2012). Characterization of jasmonate deficient tomato mutants (Liu et al., 2012) and treatment of *Nr* fruit pericarp discs with 2, 4-epibrassinolide (Liu et al., 2014) highlight the role of these hormones in fruit ripening. In addition to hormones, metabolites such as sugar also have a role in fruit development and ripening (Sagar et al., 2013). The complexity of fruit ripening is further elucidated by the regulation of cell wall metabolism and cuticle formation. The enzymes involved in these processes cause structural changes in the fruit making them more palatable to seed dispersing animals (Seymour et al., 2014). Recent studies (Tomato Genome Consortium 2012, Zhong et al. 2013) also suggest the role of epigenome modification on tomato fruit development.

As discussed in the above section, fleshy fruit ripening is a complex process regulated at many levels. Several studies on tomato fruit ripening established that this process is mainly regulated by a well defined hierarchy of regulatory genes accompanied with coordinated changes in hormonal levels especially ethylene. Apart from this internal regulation, several studies have indicated that light also modulates fruit ripening, particularly the accumulation of pigments (**Fig. 2.3**). Previous studies showed the involvement of phytochromes, a red/far red photoreceptor in tomato fruit ripening (Khudairi and Arboleda 1971; Thomas and Jen 1975). Further, Alba et al., (2000) showed that fruit-localized phytochromes stimulate carotenoid formation as tomato fruits harvested at mature green stage on red light exposure accumulated higher levels of lycopene during subsequent fruit development. A recent study on tomato photomorphogenic mutants has revealed that phytochromes act as a regulatory switch during the fruit development (Gupta et al., 2014). Cryptochromes were also

predicted to influence fruit ripening, as showed by over expressing CRY2 gene in tomato. In addition to these studies on photoreceptor mutants, the tomato *hp1* and *hp2* mutants show exaggerated photo responsiveness at seedling stage and enhanced pigmentation in the fruit (Kilambi et al., 2013). The molecular and genetic characterization of *hp1* and *hp2* (Lieberman et al., 2004; Mustilli, 1999) mutants revealed that these loci are encoded by the homologs of *DDB1* and *DET1*genes of Arabidopsis respectively. Both DDB1 and DET1 proteins are auxiliary to COP1 and together act as negative regulators of photomorphogenesis (Yi & Deng, 2005). Apart from these two auxiliary molecules, the RNAi repression of *Cullin4* (*CUL4*), a scaffolding protein subunit of E3 ligase and *COP1LIKE* (recently annotated as *SPA3LIKE*-whose protein is closely related to COP1) also exhibited altered photoresponses (Liu et al., 2004, Wang et al., 2008). Enhanced carotenoid content in these RNAi lines suggests that regulated proteolysis may be important for enhancing nutritional quality during tomato fruit ripening and light has a direct/indirect role.

2.6 Importance of mutagenesis and mutation screening in plant biology

Genetic variation is the basis for improving the traits of an organism and is valuable for production of new cultivars in plant breeding. Induced mutations have been a major resource to introduce these variations. Koernicke (1904) first attempted to induce mutations in *Vicia faba*. Tremendous progress has been made to improve the mutagenesis methods and many crop varieties have resulted due to mutation breeding. Fragrant aroma in certain rice cultivars and soya bean due to loss of betain aldehyde dehydrogenase gene (Bradbury et al., 2005) has high market value. Today, the mutant variety database of FAO/IAEA has approximately 3200 entries (http://mvgs.iaea.org/Search.aspx) and almost half are cereals. Also, in many other instances induced mutations proved to enrich the genetic variation in plants.

In nature, mutations occur spontaneously due to errors in DNA replication or damage to DNA by environmental factors like ultraviolet radiation and chemical exposure (van Harten, 1998). However, the rate of spontaneous mutations is very low and hence mutations were induced by different means (chemical or physical) to increase the frequency. Though random, the goal of mutagenesis breeding is to maximize the genetic variation in agronomic traits with minimum decrease in viability. Chemical agents like ethyl methanesulfonate (EMS) and N-ethyl-N-

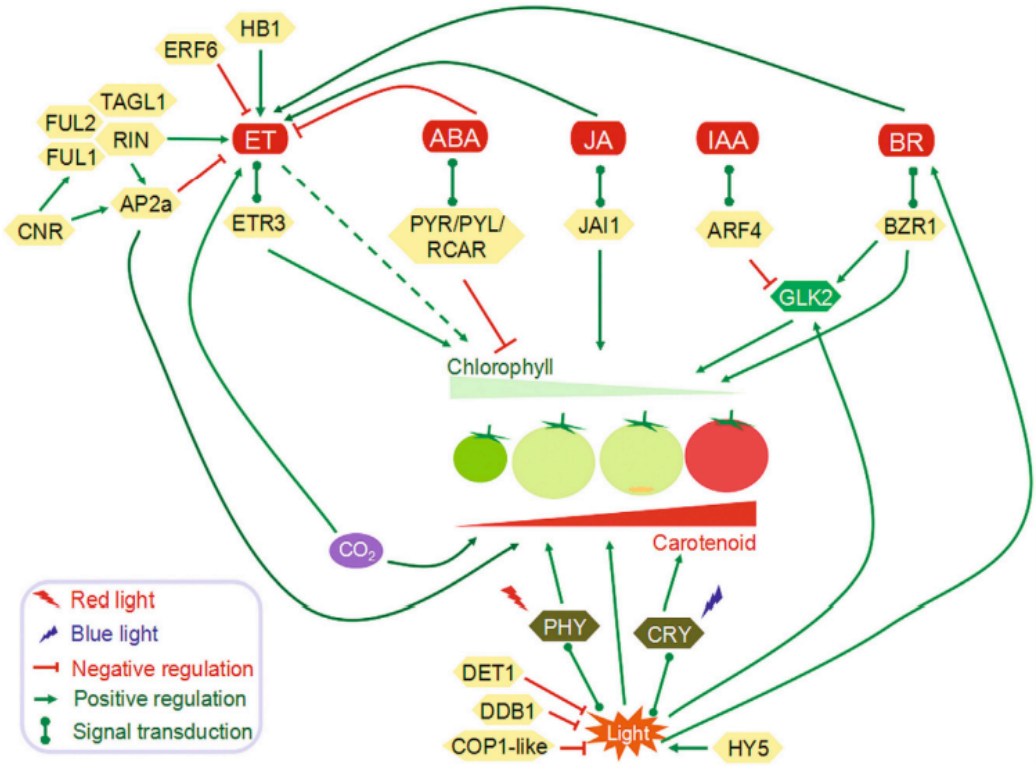


Figure 2.3 An overview of the interplay of plant hormones, CO2 and light in carotenoid accumulation during tomato fruit ripening. The hexagons represent the transcriptional factors and the signaling components. ABA, abscisic acid; APa2, apetala2; ARF4, auxin responsive factor4; BR, Brassinosteroids; BZR1, brassinazole resistant1; CNR, colorless non ripening; COP1, constitutively photomorphogenic1; CRY, cryptochrome; DDB1, UV-damaged DNA binding protein 1; DET1, de-etiolated 1; ERF6, ethylene responsive factor 6; ET, ethylene, ETR3, ethylene receptor 3; FUL1, fruitful 1; FUL2, fruitful 2; GLK2, golden2-like; HB1, an HD Zip homeobox protein; HY5, hypocotyl 5; IAA, indole-3-acetic acid; JA, jasmonic acid; JAI1, jasmonic acid- insensitive-1, PHY, phytochrome; PYR/PYL/RCAR, pyrabactin resistance 1/pyr1-like/regulatory components of ABA receptors; RIN, ripening inhibitor; TAGL1, tomato agamous-like. This picture is taken from Liu et al., (2014).

nitrosourea (ENU) are the most commonly used to generate point mutations (McCallum et al., 2000). These alkylating agents result in G/C to A/T transitions (Taylor et al., 2003) that cause amino acid substitutions affecting protein structure and function. Physical agents like gamma rays, X-rays, electrons and even fast neutrons are used to introduce deletions of different sizes ranging from few bases to even very large fragments (> 6 Mbp). Inversions, insertions and even translocations are also reported (Shikazono, 2005).

Once the mutagenized population is generated, mutations can be analyzed by different methods. Traditionally, forward genetics is used to screen for visible phenotypic changes. Early studies used PCR based techniques like- RFLPs (restriction fragment length polymorphisms), RAPD (random amplification of polymorphic DNA), SSRs (Simple sequence repeats) and AFLPs (amplified fragment length polymorphisms). However, these techniques are time consuming and laborious. Mutations can also be detected by using the genotypic information. This method is referred to as reverse genetics. A prior knowledge of the sequence is necessary in reverse genetic screens, which is normally gained from homology searches, expressed sequence tag (EST) databases, sequencing or transcript profiling. More sensitive and rapid mutation detection methods like TILLING, capillary electrophoresis, High Resolution DNA Melting analysis and very recently, next generation sequencing are being used currently.

2.7 Reverse genetics- Targeted mutagenesis

2.7.1 TILLING: a reverse genetic approach

Targeting Induced Local Lesions in Genomes is a high throughput reverse genetic approach based on mismatch enzyme cleavage. The modified bases in heteroduplexed DNA are identified by cleaving at the mismatch site and subsequently, the cleaved products can be visualized by gel electrophoresis (**Fig. 2.4**). It was first demonstrated in Arabidopsis (McCallum et al., 2000), later gained major attention in both plant and animal systems. After mutagenesis of seeds, pollen (Till et al., 2004) or ovaries, DNA is sampled from the population, extracted and pooled, most commonly with an 8-fold pooling depth. For high throughput screening of mutations, a two dimensional (2D) pooling strategy was introduced by Till et al., (2006a). A gene of interest was amplified by PCR, subjected to heteroduplex

formation, then digested by S1 endonucleases such as CEL1 (Colbert et al., 2001) or ENDO1 (Triques et al., 2007). Samples are then analyzed by polyacrylamide gel electrophoresis (PAGE) or on a capillary sequencer. TILLING has been adopted for many plant species, especially crops such as barley (Caldwell et al., 2004), wheat (Slade and Kanuf, 2005), maize (Till et al., 2004), sorghum (Xin et al., 2008), tomato (Minoia et al., 2010), beetroot (Kornienko, 2013) etc. **Table 2.1** provides an overview of all TILLING projects known to date. As shown in the **Table 2.1**, the mutation frequency varies in different plant species. This may be due to the factors- choice of mutagen, dose of mutagen, G/C content of the genome and tolerance of each genotype to mutagenesis. In spite of some technical bottlenecks, TILLING has led to the successful identification of mutants in many species.

2.7.2 Natural variations and Eco-TILLING

The domestication and cultivation practices resulted in the geographical distribution of a particular plant species. Genotypically these plants might have accumulated variations due to the existing abiotic and biotic conditions. Hence over time, they have either lost or acquired new traits or retained some traits. Unlike induced mutations, this genetic pool which includes the widely distributed ecotypes, is rich of information and is exploited to understand many aspects of plant development. TILLING was modified by Comai et al., (2004) to allow discovery of polymorphisms in such ecotypes and was hence named Eco-TILLING. The major difference in the protocol is the DNA is isolated from natural accessions instead of mutagenized population. Eco-TILLING was first performed in *Arabidopsis thaliana* to study the genetic variability among 150 accessions (Comai et al., 2004) and later it has been used routinely to study the single nucleotide polymorphisms (SNPs) in many species. For example, ecotilling has been carried *S. lycopersicum* (Rigola et al., 2009), *Brassica oleracea* (Wang et al., 2010a), Musa sp. (banana, Till et al., 2010), *Vigna radiata* (mung bean, Barkley et al., 2008) and in Cucumis sp. (Nieto et al., 2007) etc. Therefore, Eco-TILLING is a fast and accurate tool to discover the new alleles in the naturally available non-mutagenized populations. Also it can be used to identify the heterozygosity level and estimate phylogenetic diversity, selection pressures acting on a gene fragment.

2.7.3 Capillary electrophoresis (CE)

Capillary electrophoresis systems were developed to overcome the labor intensive procedure followed for detection on Li-COR DNA analyzers. CE systems are reported to be more sensitive and easy to handle than gel electrophoresis systems. Exploiting the physical properties of heteroduplex DNA different methods of CE were used (Triques et al., 2008). A major advantage of capillary instruments is the elimination of post-run lane re-tracking and it can be carried on different matrices (Mardis, 1999). Various modifications have been incorporated to capillary electrophoresis method to detect either with or without enzymatic cleavage of target amplicons, usually on ABI systems. Gady et al., 2009 have detected 5 mutations using confirmative sensitive capillary electrophoresis on ABI 3130 × L capillaries filled with CAP (semi denaturing polymer). Optimized protocol for detection of polymorphisms in pooled samples of Saccharum sp provides an alternative for high-throughput Eco-TILLING (Cordeiro et al., 2006).

2.7.4 High-Resolution DNA melting (HRM) analyses

HRM is one of the recent non-enzymatic mutation detection methods (Zhou et al., 2004; Gady et al., 2009; Botticella et al., 2011). In this method a PCR is performed in the presence of saturating fluorescent DNA dye and unlabeled primers followed by the detection of fluorescent melting curves of PCR amplicons (Wittwer et al., 2003). Optimized conditions for HRM analyses allows proper control of temperature and continuous monitoring of fluorescence in instruments such as the Lightscanner (Idaho Technology, Salt Lake City, UT, USA) or the Rotor-Gene (Qiagen, Hilden, Germany). The major advantage of this method is, firstly it allows the detection of single base mismatches in target fragments from the mutant pools even in the heterozygous state (Dong et al., 2009b). Secondly, as the detection sensitivity of HRM is optimized to small amplicon sizes (<450 bp), it is best suited for target genes with multiple small exons separated by large introns (Parry et al., 2009). Gady et al., (2009) have exploited this technique to identify mutations in tomato. They could successfully identify two mutations in the *phytoene synthase 1* gene of tomato leading to altered fruit color and carotenoid content (Gady et al., 2012). However, they observed that the ratio of false positives was high, which may need a lot of proof-reading and re-screening.

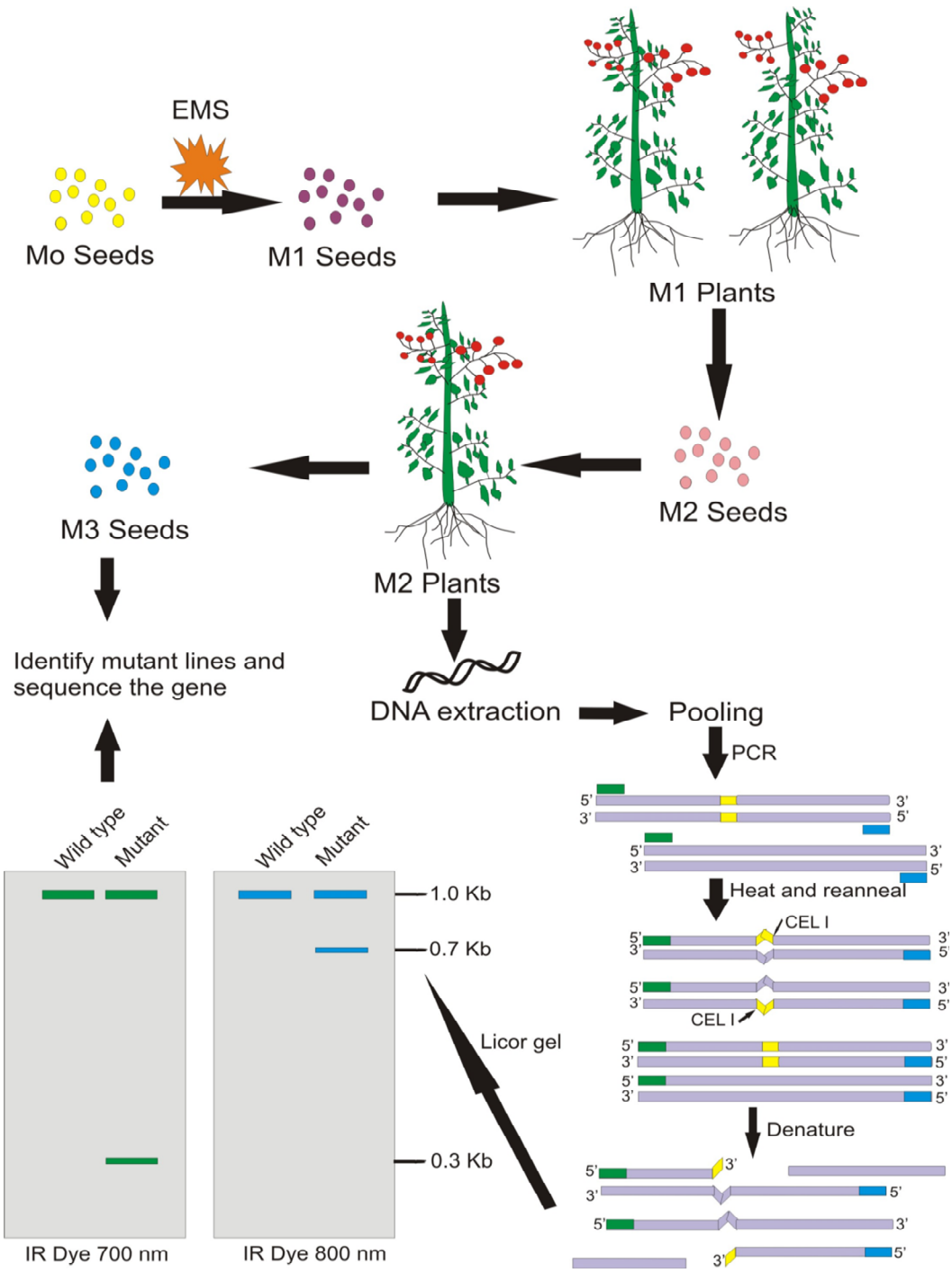


Figure 2.4 Overview of TILLING in Tomato.

2.7.5 Next Generation Sequencing (NGS)

Next generation sequencing (NGS) methods allow the rapid and high-throughput sequencing of large stretches of DNA base pairs, which can be used for direct detection of mutations without any pre-screening. This method is much reliable and considerably less expensive than Sanger sequencing (Deschamps et al., 2010). Rigola et al. (2009) used the 454 Roche technologies and detected two mutations for *eIF4E* gene in a single GS FLX sequencing run. Illumina GAII platform (http://tilling.ucdavis.edu/index.php/Main_Page) was used to detect mutations in rice and wheat using multidimensional pooling schemes (Tsai et al., 2013). Though NGS is gaining popularity due to its wide applications, very few studies are reported for detection of mutations in mutagenized populations. Cost of the required equipment and reagents is still a major drawback for NGS. Also the technical and analytical expertise to handle huge data sets is a limiting factor for widespread use of NGS.

2.7.6 Exome Capture

Recently, target exome capture or exome sequencing has been used for selective sequencing of only coding regions. It is a better approach than whole genome sequencing due to the selective sequencing of exons at considerably low costs when compared with the high costs involved for whole-genome sequencing. Exome capture can be effectively used to detect all changes in the coding region including very rare and spontaneous mutations. It has been used for the crop improvement (Singh et al., 2012) and has proved beneficial in case of allohexaploid wheat genome (Winfield et al., 2012). Henry et al. (2014) used multiplexed global exome capture and next generation sequencing to identify mutations in mutant populations of rice and wheat. In rice, they identified ~18,000 induced mutations from 72 independent M2 individuals. This technique would likely replace other technique for mutation detection in near future due to robustness of mutation detection (Till et al., 2012).

2.8 Wild relatives of Tomato

Solanaceae or Nightshade family is diverse, comprising of approximately ten thousand species. The domesticated tomato (*Solanum lycopersicum*) and its wild relatives belong to the *Lycopersicon* clade of this family. Members of this clade have

Table 2.1 List of TILLING platforms in different crop species.

Species (ploidy level)	Mutagen	Concentration	Screening Method	Size of M2 population	No.of genes analyzed	Total no of mutations detected	Mutation frequency	Reference
Arabidopsis thaliana (2x)	EMS	20-40 mM	Li-COR	3,712 6,912 3,072	14 1,090 192	450 1,774 1,890	1/89 Kb 1/170 Kb 1/300 Kb	Martin et al., 2009 Till et al., 2003 Greene et al., 2003
Arachis hypogea (2x)	EMS	32/80 mM	Li-COR	3,420	3	27	1/967 Kb	Knoll et al., 2011
Avena sativa (6x)	EMS	73 mM	MALDI TOF	2,600	2	16	1/33 Kb	Chawade et al., 2010
Brassica napus (4x)	EMS	48/80 mM 80 mM 16-96 mM	Li-COR	5,355 3,488 1,344	4 2 1	498 94 19	1/16 Kb 1/61 Kb 1/86 Kb	Harloff et al., 2011 Harloff et al., 2011 Wang et al., 2008
Brassica oleracea (2x)	EMS	32 mM	Li-COR	8,750	15	25	1/447 Kb	Himelbalu et al., 2009
Brassica rapa (2x)	EMS	16-40 mM	CE	9,216	6	617	1/60 Kb	Stephenson et al., 2010
Cucumis melo (2x)	EMS EMS	80-240 mM	Li-COR Li-Cor	4,023 2,368	11 7	134 33	1/573 Kb 1/1.5 Mb	Dahmani-Madras et al., 2010 Gonzalez et al., 2011
Glycine max (2x)	EMS MNU EMS	40 mM 2.5 mM 40 mM	Li-COR	529 768 768	7 7 7	32 47 12	1/140 Kb 1/140 Kb 1/550 Kb	Cooper et al., 2008a Cooper et al., 2008a Cooper et al., 2008a
Helianthus annus (2x)	EMS	68 mM	Li-COR	3,651	4	9	1/475 Kb	Sabetta et al., 2011
Hordeum vulgare (2x)	Az-MNU Az MNU EMS EMS NaN3 Gamma rays	N/A 10 mM N/A 20-60 mM 20/30 mM 1.5 mM N/A	ND Li-COR ND Li-COR dHPLC CE ND	10,000 4,906 1,372 10,279 9,216 9,573 1,753	12 4 2 6 2 2 1	174 22 9 81 10 5 1	1/235 Kb 1/374 Kb 1/486 Kb 1/500 Kb 1/1,00 Kb 1/2,500 Kb 1/3297 Kb	Kurowska et al., 2011 Talame et al., 2008 Kurowska et al., 2011 Gottwald et al., 2009 Caldwell et al., 2004 Lababidi et al., 2009 Kurowska et al., 2011
Lotus japonicus (2x)	EMS	N/A	Li-COR	4,904	61	576	1/502 Kb	Perry et al., 2009
Medicago truncatula (2x)	EMS	12/16 mM 12/16 mM	CE Li-COR	4,350 4,500	56 56	456 456	1/424 Kb 1/485 Kb	Le Signor et al., 2009 Le Signor et al., 2009
Oryza sativa (2x)	MNU Az-MNU EMS Az-MNU EMS	1 mM 1 mM Az 15 mM MNU 120 mM 1 mM Az 15 mM MNU 64/80 mM	CE Li-COR Li-COR NGS Li-COR	767 768 768 768 2,000	3 10 10 13 10	24 30 27 81 na	1/135 Kb 1/265 Kb 1/294 Kb N/A 1/1,000 Kb	Suzuki et al., 2008 Till et al., 2007 Till et al., 2007 Tsai et al., 2011 Wu et al., 2005
Pisum sativum (2x)	EMS	16/20/24 mM	Li-COR	4,717	1	50	1/193 Kb	Triques et al., 2008
Solanum lycopersicum (2x)	EMS	80 mM 68 mM 48 mM 48 mM 80 mM	Li-COR CE NGS Li-COR HRM	4,741 1,926 15,000 4,759 8,225	7 7 1 19 5	25 41 2 256 44	1/322 Kb 1/574 Kb N/A 1/574 Kb 1/737 Kb	Minoia et al., 2010 Minoia et al., 2010 Rigola et al., 2009 Piron et al., 2010 Gady et al., 2009
Solanum tuberosum (2x)	EMS	80 mM	SS	2,748	1	19	1/91 Kb	Muth et al., 2008
Sorghum bicolor (2x)	EMS	8-48 mM	Li-COR	1,600	4	5	1/526 Kb	Xin et al., 2008
Triticum aestivum (6x)	EMS HII	60/80 mM 73/80 mM 50 Gy	Li-COR Taq Man	10,000 1,536 4,500	2 3 1	196 186 294	1/24 Kb 1/38 Kb 1/84 Kb	Slade et al., 2005 Uauy et al., 2009 Fitzgerald et al., 2010
Triticum durum (6x)	EMS	60/80/96 mM 57/60 mM	Li-COR PAGE NGS	8,000 1,386 1,386	2 2 5	50 93 112	1/40 Kb 1/51 Kb N/A	Slade et al., 2005 Uauy et al., 2009 Tsai et al., 2011
Zea Mays (2x)	EMS	80 mM	Li-COR	750	11	17	1/485 Kb	Till et al., 2004
Beta vulgaris (2x)	EMS	77 mM 88 mM	Li-COR Li-COR	3,418 2,688	10 7	32 32	1/1062 Kb 1/1,108 Kb	Kiel et al., 2013

originated from South America near the Andean region. The micro habitats and extreme ecological conditions contribute to the diversity of wild species. This broad variation is reflected at various aspects *i.e.* morphological, physiological, sexual and molecular features of tomato wild relatives (Spooner and Peralta, 2006). The wild relative species can be discriminated based on their fruit color. Most of them have green colored fruits, with the exception of two species – *S. galapagense* (yellow and orange fruits) and *S. pimpinellifolium*, the only wild species with red colored fruits. *S. lycopersicum var cerasiforme* fruit is larger than *S. pimpinellifolium* and is commonly referred to as cherry tomato. This has been considered as the direct ancestor of domesticated tomato (Rick and Chetelat, 1995). The modern cultivated tomato is cosmopolitan and is grown under various environmental conditions all over the globe.

Wild relatives harbor high genetic diversity and are a primary source of genetic variation. Wild relatives of many crops have been extensively used in breeding programs to improve the agronomic traits. Several studies on tomato wild relatives identified many useful traits, such as tolerance to drought and salinity in *S. pennellii* (Dehan et al., 1978, Shalata et al., 1998), accumulation of health promoting phytochemicals in *S. habrochates* and resistance to multiple pathogens exhibited by *S. cheesmannii*, *S. lycopersicoides*, *S. neorickii* and *S. chilense* (Robert, 2001; Rick et al., 1994). Previously breeding studies were carried on mapping populations that are based on interspecific crosses between a cultivar and related wild species from the *Lycopersicon* group. Though this approach permitted the identification of QTL for a trait of interest, it is time consuming and laborious. With the emergence of new molecular techniques and bioinformatics tools, and the availability of genome sequences, genotyping by sequencing is possible (Hohenlohe et al., 2011; Elshire et al., 2011). Increasing community wide resources and germplasm banks have added value in understanding the genetic diversity of wild species. Tomato JBrowse, a recent resource developed by Wageningen University, Netherlands provides huge data on the SNPs of more than 150 tomato species that can be used for better understanding of natural diversity (Aflitos et al., 2014).

Chapter 3

Materials and Methods

3.1 Plant material and growth conditions

3.1.1 Plant material

The seeds of *Solanum lycopersicum* cultivar Arka Vikas, a local South Indian variety obtained from IIHR Bangalore and EMS mutagenized M_2 seed of M82 population (kindly provided by Dani Zamir, Hebrew University, Israel) were used for the TILLING study. For Eco-TILLING, a collection of 585 *S. lycopersicum* accessions obtained from NBPGR (National Bureau of Plant Genetic resources (http://www.nbpgr.ernet.in/), IIVR (Indian Institute of Vegetable Research, Varanasi, India (http://www.iivr.org.in/) Bejo Sheetal, Jalna, India and TGRC (Tomato Genetic Resource Center, University of California, Davis (http://tgrc.ucdavis.edu/) were used. The wild tomato TGRC accessions *S. pimpinellifolium* (LA1589), *S. galapagense* (LA0483), *S. chilense*, *S. pennellii* (LA0716), *S. neorickii* (LA2133) and *S. habrochaites* (LA1777) were obtained from Tomato Genomics Resource Centre (University of California, Davis).

3.1.2 Growth conditions

Seeds from all the tomato genotypes mentioned were surface sterilized with 4% (w/v) sodium hypochlorite for 10 min to soften the seed coat and then washed in running tap water till the traces of hypochlorite were removed. Thereafter, the seeds were spread on wet blotting paper placed in petriplates and incubated in darkness for germination. Emergence of radicle was scored as the first sign of germination. The germinated seeds were then transferred to coconut peat mixture (Sri Balaji Agroservices, Madanapalle, Andhra Pradesh, India) placed in a plastic germination box (9.5 cm length, 9.5 cm breadth, 5 cm height). The boxes were kept under continuous white light (100 µmol $m^{-2}s^{-1}$) in growth room for the establishment of seedlings. Unless specified all the plants were grown in an open field at University of Hyderabad under drip irrigation.

3.1.3 Physiological response of seedling to darkness and light

The germinating seeds were subjected to darkness or white light treatment after emergence of radical from seed coat. Unless otherwise mentioned, all experiments

were performed 3-4 times and for each set, 10-12 seedlings were used. The length of hypocotyls and roots were measured and recorded. The phenotypes were recorded using an Olympus camera.

3.2 Southern blot analysis for estimation of copy number of *COP1* gene

The copy number of *COP1* gene in tomato was determined by Southern blot analysis. The major steps include- genomic DNA isolation and restriction digestion, treatment and transfer and lastly, hybridization with the probe followed by autoradiography.

3.2.1. Isolation and restriction digestion of genomic DNA

DNA was isolated in bulk from leaves of AV by CTAB method (Murray and Thompson 1980). Approximately, 5 g of leaf material was homogenized to fine powder in liquid nitrogen using mortar and pestle. The powdered tissue was transferred to a 50 mL centrifuge tube and re-suspended in pre-warmed extraction buffer (100 mM Tris-HCl pH 8.0, 20 mM EDTA, 1.4 M NaCl, 2% (w/v) CTAB, 2% (w/v) PVP and 2% (v/v) β-ME). The suspension was incubated in water bath for half an hour at 65°C. The slurry was intermittently mixed by occasionally inverting the tubes. Thereafter it was allowed to cool to room temperature. Equal volume of chloroform: isoamyl alcohol (24:1) mixture was added to it and mixed gently by inverting the tubes. The tubes were then centrifuged at 8000 rpm, at room temperature for 10 min. The upper aqueous layer was transferred to a fresh tube and the chloroform-isoamyl alcohol step was repeated. To the clear supernatant obtained, two volumes of ice cold ethanol and 0.5 volume of 5 M NaCl was added and the tubes were incubated at 4°C for 15 min. The precipitated DNA was collected by centrifuging at 8000 rpm for 15 min at room temperature. The pellet was washed twice with 70% (v/v) ethanol, followed by drying of DNA in SpeedVac (Thermo, USA) and resuspended in TE buffer (10 mM Tris pH 7.5, 1 mM EDTA pH 8.0). The samples were treated with RNase-A (10 µg/mL) to remove the dissolved RNA. After 2 h of incubation at 37°C, the proteins were removed by phenol- chloroform treatment. This was followed by the chloroform- isoamyl treatment to the aqueous phase. The DNA was precipitated by adding 0.1 volume of 3 M sodium acetate pH 5.2 followed by 2.5 volumes of ice cold ethanol. The mixture was kept at -20°C for 2

h for maximum precipitation of DNA and after centrifugation at 12000 rpm for 10 min at 4°C, the DNA pellet was washed with 70% ethanol, dried and dissolved in TE buffer. The purity and integrity of the DNA was checked by Nanodrop readings and agarose gel electrophoresis.

A pilot digestion was set up with a small aliquot of the genomic DNA (0.5-1µg) with the restriction enzyme *BamHI*. The digestion quality of DNA was determined by agarose electrophoresis on a mini gel (1%). Thereafter, digestion of bulk DNA (approx. 10 µg) was carried out for Southern analysis. The genomic DNA was digested with the restriction enzymes *BamHI*, *EcoRV* and *HindIII* which either have a single site or no site in the *COP1* cDNA. Restriction digested (~10 µg) and undigested (~4 µg) DNA samples were quantified on a 1% (w/v) agarose gel containing 5 µg/mL EtBr along with standard size marker (λ DNA *HindIII* digest) and electrophoresed in 1 X TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.3) at 30 V for ~ 20 h.

3.2.2 Treatment of Gels and capillary transfer

Random digestion of genomic DNA releases fragments of more than 15 Kb. Therefore, to ensure efficient transfer from gel to membrane, the gel was sequentially treated with the following solutions:

1. Denaturation solution-: 0.15 N HCl for 30 min.
2. Depurination solution –1.5 M NaCl, 0.5 M NaOH, for 45 min.
3. Neutralization solution-: 1 M Tris-HCl pH 8.0, 1.5 M NaCl, for 20 min at room temperature.

The agarose gel was rinsed with distilled water every time before treating with the solutions. After neutralization, the gel was capillary blotted on Hybond-N membrane (Amersham International Inc., UK) using 20X SSC (saline sodium citrate, 20X stock contains 3 M NaCl and 0.3 M TriSodium Citrate).The capillary transfer was set up as follows: two sheets of Whatmann 3 MM paper longer than the gel were used as a wicks. The wick was saturated in transfer buffer and placed over the glass sheet supported by a baking dish filled with transfer buffer so that the wick gets

submerged. A Whatmann 3MM paper was placed over the wick on which the gel was inverted. The area surrounding the gel was covered with strips of cellophane to prevent the short-circuiting during transfer. The membrane was overlaid on top of the gel followed by a sheet of Whatmann 3MM paper. Air bubbles were eliminated after each layer by rolling the gel with a glass rod. Blotting sheets cut to the size of the gel were placed over the above setup. A 250 g or a heavier bottle was placed over the paper towels and transfer was carried out overnight. After the completion of the transfer, the position of wells was marked on the membrane with a ball-point pen. The membrane was then rinsed with 2X SSC, air-dried on Whatmann 3MM paper for 30 min. The DNA was cross-linked to the membrane with the help of a UV crosslinker at the rate of 254 nm for 1 min 45 seconds and 1.5 J/cm2.

3.2.3 Hybridization and autoradiography of a Southern blot

Usually, hybridization of a Southern blot is carried out in two steps- first, the blot was pre-hybridized to block the unused DNA sites on the membrane. In this process, the blot was incubated for 16 h at 42°C in a buffer containing 50% Formamide, 5X SSC, 10X Denhardt's solution (20 mg/ml each of PVP, Ficoll 400 and Bovine Serum Albumin), 50 mM sodium phosphate buffer, pH 6.5, and 250 µg/mL sonicated and denatured Herring sperm DNA. In the second step, the actual hybridization was carried out with the probe at 42°C for 24 h. A 414 bp of *COP1* cDNA fragment was radiolabeled with α P-32 dATP radioisotope (JONAKI, India) using Mega prime DNA labeling kit (Amersham Pharmacia., UK). The radiolabeled DNA probe was denatured in boiling water bath for 2 min and added to pre-hybridization buffer containing 10% (w/v) dextran sulphate.

At the end of hybridization the DNA blot was washed, in the following sequence, (i) 50% formamide, 5X SSC, 0.1% (w/v) SDS. (ii) 2X SSC, 0.1% SDS (w/v) (iii) 2X SSC, 0.1% SDS (w/v) and (iv) 1X SSC, 0.1% SDS (w/v) for 10 min each at RT. For providing high stringency conditions, the last washing of the blot was done at 37°C or 42°C depending on the requirement. The differential radioactive binding in the blots was monitored throughout the course of washing with the help of a radiation detection counter (GM counter). In case the counts were too high to discern regions containing genomic DNA from background, an additional wash for 30

min at 60 °C with a 0.1 X SSC buffer and 0.1% (w/v) SDS wash solution was performed (Sambrook et al., 2001). After air drying, the blots were wrapped in saran wrap and exposed to X-ray films (Kodak X-OMAT AR) enclosed in cassettes at –80°C. Depending on the radioactive counts detected at the last wash, the cassettes were incubated for 3-7 days. The exposed X-ray films were developed by dipping in developer solution followed by twice or thrice washing with water. After the appearance of band the film was fixed by soaking in fixer solution, followed by washing with water. The fixed films were then air dried and photographed or scanned for image analysis.

3.3 TILLING (Targeting Induced Local Lesions In Genomes)

3.3.1 EMS mutagenized tomato population

Seeds of Arka Vikas were mutagenized with EMS (Ethyl Methane Sulfonate) as per the standard mutagenesis protocol (Koornneef et al., 1990). In consideration with the previous reports on induced mutagenesis in tomato (Menda et al., 2004), a concentration of 60 mM EMS was used to mutagenize AV seeds. However, based on the phenotypic data and lethality rates (Sherinmol Thomas, PhD thesis, 2012), in order to increase the frequency of mutations, a higher dosage of EMS at 120 mM was used to develop another population. **Table 3.1** describes in detail the dosage of EMS and size of populations developed in our study.

The detailed process of generating EMS mutagenized tomato populations was described earlier (Sherinmol Thomas, PhD thesis, 2012). In short, the M_0 seeds were treated with the required concentration of EMS and germinated. Three week old M_1 seedlings were propagated in the field. Each M_1 plant was labelled with a unique number for subsequent collection of M_2 seeds. The plants were watered using drip irrigation and M_2 seeds were manually collected. Depending on the seed availability, between one and four M_2 plants per M_1 family were grown for M_3 seed production by using identical conditions as described above. The M_2 DNA and M_3 seeds from the mutated populations were arrayed and catalogued properly, thereby establishing a TILLING platform.

3.3.2 DNA extraction and pooling strategy

TILLING requires DNA with high yield and good quality. In the present study we used an in-house standardized DNA extraction protocol (Sreelakshmi et al., 2010) from tomato cotyledons/leaves. Briefly, the harvested tissue (100 mg) was mechanically disrupted using three steel balls in a Mini-Bead Beater for 2 min in the presence of 750 µL preheated (65°C) extraction buffer (0.1 M Tris-HCl, pH 7.5; 0.05 M EDTA, pH 8.0; 1.25% (w/v) SDS) containing 0.2 M β-mercaptoethanol and 20 mg of insoluble polyvinylpolypyrrolidone (PVPP).

Table 3.1 TILLING population structure. Number of M_2/M_3 lines contributing to the populations generated using different EMS concentrations.

S. No	Population	Tomato cultivar	EMS concentration	Population size
1	I	Arka Vikas	60 mM	2304
2	II	Arka Vikas	120 mM	4608
3	III	M82 M_2*	~60 mM	3840

* This population was originally obtained from Dani Zamir, Hebrew University, Israel

The samples were then incubated with 4 µL of RNase (10 mg/mL stock) for 30 min at 37°C in a water bath. This was followed by the addition of 400 µL of 6.0 M Ammonium acetate in order to precipitate the proteins. Thereafter, the samples were incubated at 4°C for 15 min followed by centrifugation at 4700 rpm for 30 min (for 96 well plates) in Sorvall RC evolution centrifuge or at 13,000 rpm (for individual microfuge tubes). The clear supernatant was transferred to a new microfuge tube and an equal volume of chilled iso-propanol was added. The samples were incubated for 2h at -20°C. Samples were then centrifuged at 4°C for 30 min followed by 70% ethanol wash. To remove the traces of salts from samples, one more round of ethanol wash was given. The DNA samples were air dried and the pellet was dissolved in 200 µL TE buffer supplemented with RNase (3.3 µg/mL) and kept for overnight dissolution. The DNA was then quantified and stored in -80°C till further use. For screening, the DNA isolated from the population was normalized and aliquots of 5 ng were made in 96 well format. The plates were, dried in a speedvac and stored in -20°C. The samples were resuspended in 10 mM Tris HCl [pH] 7.5 just before performing PCR.

In order to isolate DNA on a high throughput scale with low cost, we also adopted 8-fold pooling of tissue prior to DNA isolation (Sreelakshmi et al., 2010). DNA from all the three populations was arrayed into 16 plates (96 well formats) for screening mutations.

3.3.3 Primer design and Nested PCR strategy

Tomato *COP1* homolog has 13 exons and 12 introns covering 11 Kb of the total genome. *COP1* sequence was initially obtained from a BAC sequence (C12HBa0026C13.1) upon BLAST search in NCBI. Later with the availability of tomato genome sequence at SGN (http://solgenomics.net/), the gene details were verified and considered further. The gene sequence was fed in the Codons Optimized Detection of Deleterious Lesions (CODDLE) input utility page (http://www.proweb.org/input/). CODDLE generates a gene model (intron/exon position) and a protein conservation model by searching the Blocks Database (Henikoff et al., 2000). Once a gene model and a protein conservation model are assembled, CODDLE (http://www.proweb.org/coddle/) presents candidate regions

that are most likely to be susceptible to deleterious mutations after EMS treatment (McCallum et al., 2000b). **Figure 3.1** represents the conserved domain structure of COP1 protein and the CODDLE output file wherein the exons encoding the domains were aligned accordingly. According to the prediction, exon 3 and 4 appear to be more susceptible for mutagenesis by EMS, however, we have designed different sets of primers to cover the coding region for screening. The primers were designed using Primer 3 software with melting temperatures from 62 to 70°C and a length in the range of 25–32 nucleotides (**Table 3.2**). 100 µM solution of each primer was prepared in TE buffer and aliquots were stored at -80°C to avoid repeated freeze–thaw cycles.

In order to improve the sensitivity and specificity of PCR for robust mutant detection, we adopted nested PCR strategy (**Fig. 3.2**) with labeled M13 universal primers. In nested PCR, two separate amplifications are used. The first uses a set of primers that yields a large product, which was then used as a template for the second amplification. The second set of primers anneal to sequences within the initial product producing a second smaller product. Based on nested strategy, we designed primers for screening the remaining exons (**Table 3.3**). An advantage of nested PCR is the consistency of reactions with all samples so that any fraction of the pooled DNA which bears the mutation can be amplified and visualized. Moreover, the use of M13 universal primers is cost effective.

3.3.4 Development of fluorescent dye labeled size standards

TILLING is a high throughput technique widely used to detect a mutation or SNP in a large population. Among different detection methods, the most commonly used platform for mutation detection is separating the DNA fragments after endonuclease digestion on polyacrylamide gels on a Li-COR DNA Analyzer. This method is purely based on the detection of fluorescent dye labelled PCR products. The molecular size of any digested product can be easily estimated by comparing with IR dye labeled markers of known sizes. The commercially available size standards were expensive and are restricted up to only 700 bp which renders estimation of products of sizes greater than 700 bases inaccurate. In the present study, we developed a method that can be used to generate size standards for routine gel. The molecular size of any digested product can be easily estimated by comparing with

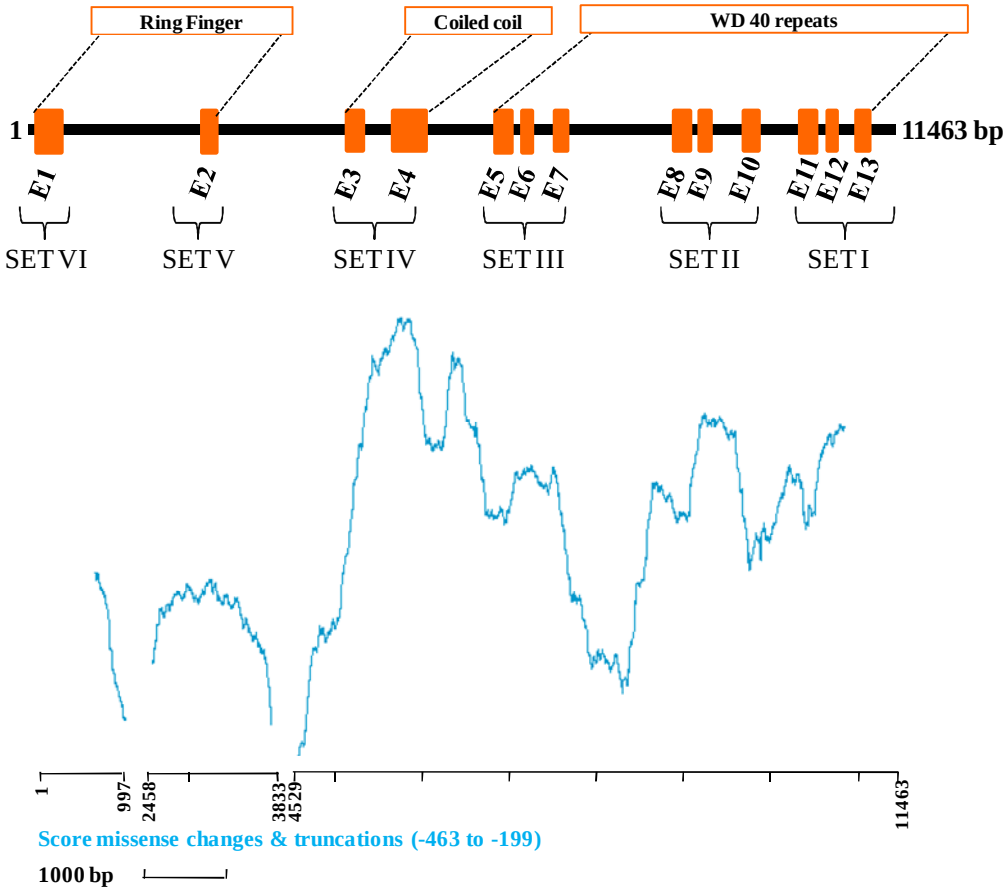


Figure 3.1 Output of the CODDLE program using *COP1* sequence. The gene represented by orange open boxes (exons) connected by black lines in between (introns). The probability curve traced in blue below the gene model represents those regions of the gene in which G: C to A: T transitions are most likely to result in deleterious effects on the encoded protein. The domains (Ring finger, coiled coil, WD-40 repeats) encoded by the corresponding exons are represented above the gene model. The tomato *COP1* gene is of 11.4 Kb and consists of 13 exons and 12 introns.

Table 3.2 Primer sequences designed for exon 5, 6 and 7 of tomato *COP1* gene.

Gene	Primer		Start-end position	Length	Tm (°C)	GC (%)	Sequence
COP1	SET III	F1	6441	30	64	40%	*TGGACTTGTGACTTATCGGGAAATTTGATG*
		R1	7635	30	64	43%	*AAACTTTTATACGCCGTGAAACTCCAGCAG*

IR dye labeled markers of known sizes. The commercially available size standards were expensive and are restricted up to only 700 bp which renders estimation of products of sizes greater than 700 bases inaccurate. In the present study, we developed a method that can be used to generate size standards for routine gel electrophoresis. We surpassed the requirement of restriction digestion and used commercially available size standards that were PCR amplified and fluorescently labelled. The pool of DNA size standards carrying A at their 3′ ends were ligated to pGEM®-T Easy Vector (3015 bp, Promega) carrying M13 forward and reverse sequencing primer binding sites. The ligated product was transformed into ultra-competent *E. coli DH5α* cells followed by plating on Luria Bertani agar plates that were supplemented with ampicillin (50 µg/mL) and incubated at 37°C overnight. Transformants were analyzed by colony PCR using M13F (TGTAAAAC GACGGCCAGT) and M13R (AGGAAACAGCTATGAC CAT) primers. Plasmid DNA was isolated from the positive clones by alkali lysis method and purified by PEG 8000. To determine the exact molecular size of the standards generated using M13 forward and reverse unlabeled primers, the PCR products were sequenced by ABI (Applied Bio systems) Big dye terminator chemistry version 3.0 (Gupta et al., 2011). **Figure 3.3** represents a part of the polyacrylamide gel with in-house generated size standards loaded adjacent to mutant DNA samples showing the complementary excised products in the 700 and 800 IRD channels.

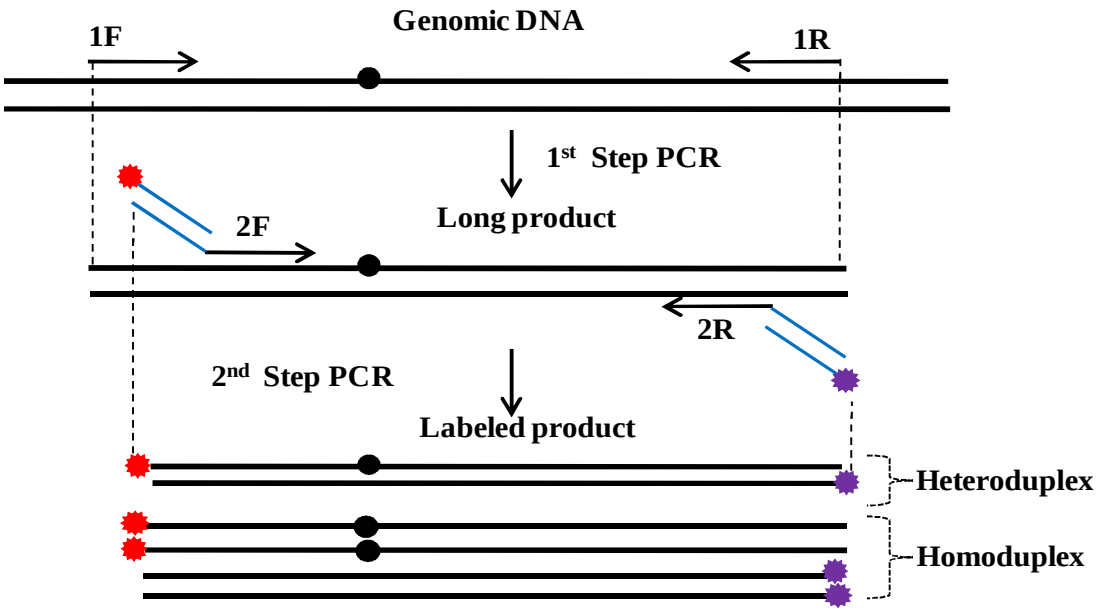


Figure 3.2 Nested PCR strategy with universal primers. Targets were amplified with sequence specific primers 1F and 1R (black arrows). The large product was taken as a template and a nested PCR with an internal gene specific primers 2F and 2R (black part of the arrow) containing universal M13 primers (blue part of the arrow) in combination with IR-dye labeled universal M13 primers (blue arrows with dye) was performed. The red and purple stars indicate the position of the IRD700 and IRD800 lablelling respectively. Black circle indicates the point mutation in the gene. After nested PCR, both homoduplex and heteroduplex DNA strands were obtained with the IRD labelling, which were further subjected to mismatch enzyme cleavage and detected on Li-COR.

3.3.5 PCR based mutation screening

The DNA from mutagenized populations was used for screening mutations in the gene of interest. We used PCR based mutation screening which involves three steps – firstly, amplification of fragment of interest followed by heteroduplex formation and mismatch cleavage by CELI enzyme and thirdly, detection on denaturing polyacrylamide gels. PCR was performed in Tetrad Machine (Bio-Rad) in 96-well microtitre plates. Two different PCR strategies were followed to amplify the gene. SET III primers of *COP1* gene were designed in such a way that the primer itself was labelled with fluorescent dyes, with the forward primer having the IR 700 dye and the reverse with IR 800 dye. In this case we used a single step touchdown PCR. In touchdown PCR, an annealing temperature is selected which is initially above the determined empirical Ta, and as the reaction progresses it gradually declines and falls to a lower level. This strategy helps ensure that the initial primer-template hybridization events occur with the greatest stringency; i.e., those yielding the target amplicon. A touchdown cycle of 94°C - 4 min, 5 cycles of 94°C - 20 sec, 72°C - 45 sec with a ramp of - 2.0°C per cycle, 72°C - 2 min followed by 25 cycles of 94°C - 20 sec, 62°C - 45 sec, 72°C - 1 min 30 sec, 72°C - 10 min was used for SET III primers.

The rest of the primers were designed based on nested strategy, where a two step PCR was programmed for amplifying the fragment of interest. First step PCR amplification was performed in a volume of 20 µL with 5.0 ng DNA. The reaction consisted of 5 µL of DNA template, 1X PCR buffer, 0.2 mM dNTPs, 0.18 µL *Taq* polymerase (In-house isolated) and 3 pmoles each of first step forward and reverse primers. In the second step PCR reaction, diluted first step PCR product was used as template and 1 pmoles of primers combined in a ratio of 1.2:0.7:1.5:0.5 (forward labeled (F_L): forward unlabeled (F_{UL}): reverse labeled (R_L): reverse unlabeled (R_{UL})) were used. The cycling conditions were 94°C - 4 min, 5 cycles of 94°C - 20 sec, 62°C - 45 sec with a ramp of - 2.0°C per cycle, 72°C - 2 min followed by 25 cycles of 94°C - 20 sec, 52°C - 45 sec, 72°C -1 min 30 sec, 72°C - 10 min. In either case of PCR amplification, the products were allowed to form heteroduplexes by subjecting to the following cycling conditions - 99°C-10 min, 80°C-20 sec, 70 cycles of 80°C-7 sec with a ramp of -0.3°C per cycle and held at 4°C.

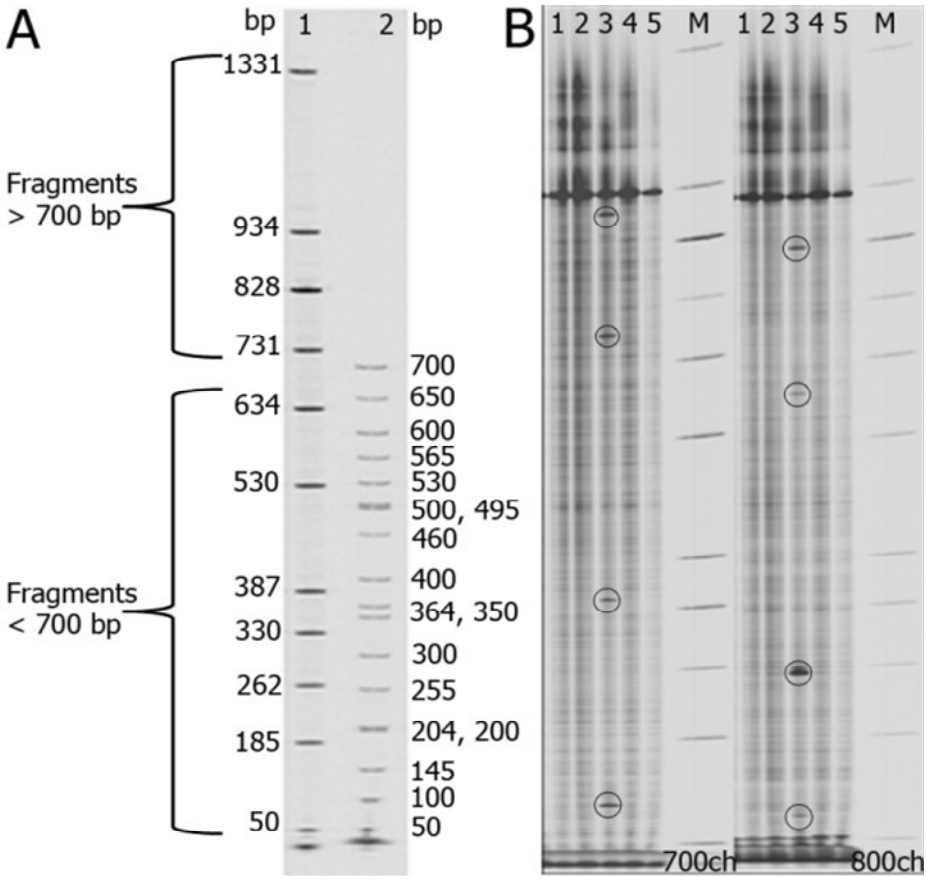


Figure 3.3 The dye labeled size standards for molecular size determination. A, IR dye labeled size standards for Li-COR Analyzers. Comparison of custom generated size standards with the commercially available size standard from Li-COR Biosciences. The numbers on left and right side of gel picture refer to size of fragments in bp respectively. Lane 1 - Custom generated size standards, Lane 2 - Li-COR size standards. B, Customized size standards used for SNP detection. Images of Li-COR gel showing resolved size standards in 700 and 800 channels. One of the samples having multiple SNPs (demarcated as black circles in the image) is shown. Lanes 1, 2, 3, 4, 5 - DNA samples from different tomato accessions mixed with reference DNA (Arka Vikas); Lane M - customized size standards.

Table 3.3 List of primer sequences designed for TILLING *COP1* gene by nested PCR strategy.

Gene	Primer		Start-end position	Length	Tm (°C)	GC (%)	Sequence
	M13 F			18	52		*TGTAAAACGACGGCCAGT*
	M13 R			19	52		*AGGAAACAGCTATGACCAT*
COP1	SET I	F1	9909	26	59	50	*CCTTTCTACTCATCTAGCCCATGGGT*
		R1	11424	26	59	50	*GGTGTGTAAATCCGATGTCCTTCAGC*
		F2	10284	25	56	44	*CTCAAGTTCTTTGTGCTTCAACCCA*
		R2	11355	24	55.7	46	*ACCAAGCCCTACAACTTGTGTTAC*
	SET II	F1	7976	24	57.4	50	*TGAACTTGGCTCAGCTTGTGGAAC*
		R1	9934	26	59.5	50	*ACCCATGGGCTAGATGAGTAGAAAGG*
		F2	8346	24	55.7	46	*TTACAGCTCGAACAATCACGACCTT*
		R2	9827	24	57.4	50	*TGCATTGACCAGCAGTTCAGAAGTG*
	SET IV	F1	4484	26	53.81	38.4	*TTGTCCTCCTGTTTTACTTCCTAACT*
		R1	6488	26	53.93	38.4	*GAAATGTCTGAGCAAACTGTAGTCAT*
		F2	4958	24	51.06	33.3	*TAACTATGAAAGCGATTGGTGAAA*
		R2	6164	24	51.80	37.5	*ACAAGTGGGAAGGGTAATATTGAA*
	SET V	F1	1719	27	58.2	44	*GAGGCACTTACTTTCCCACACATATGA*
		R1	3558	24	57.4	50	*CACTACAAGGATGGCCTTGTGAGA*
		F2	1949	24	57.4	50	*CTCCCTCTGTCCCAATTGATGTGA*
		R2	3224	24	57.4	50	*AAAGCAAACTCCGGAGACTAACGC*
	SET VI	F1	-456	26	59.5	50	*GGCTGCTTACGATGTGAAGATAGGAG*
		R1	1285	24	57.4	50	*CAGTTGGGAAGTTGACACGTCCAT*
		F2	-269	24	54	42	*AAGTTGACGCTTGATTCACTGATG*
		R2	1113	27	56.7	41	*TCTGTACCTAAGATAGTGCTCCTGTAA*

3.3.6 Mismatch cleavage reaction and precipitation of PCR products

Once the fragment of interest was amplified, a mismatch cleavage reaction was performed by an endonuclease, CELI which cleaves the mismatches in double stranded DNA. The mutated nucleotides form a loop during heteroduplex formation and become the targets of digestion for CELI enzyme. The digestion is carried out in a total volume of 45 µL containing 10 µL of labelled PCR product, 1X CELI digestion buffer (10 mM HEPES buffer pH 7.0, 10 mM KCl, 10 mM MgCl2, 0.002% (v/v) Triton X-100 and 10 µg/ml BSA) and CELI enzyme at 1: 300 dilution (1 µL/300 µL CELI digestion buffer). The mixture was incubated at 45°C for 15 min and then the reaction was stopped by adding 10 µL stop solution (2.5 M NaCl, 75 mM EDTA, pH 8.0 and 0.5 mg/ml blue dextran). The DNA was precipitated by the addition of 125 µL of cold absolute ethanol and incubated at -80°C for 15- 30 min. This is followed by centrifugation at 4500 rpm for 30 min in a SH-3000 plate rotor. The pellet was washed with 70% (v/v) ethanol, dried in a dry bath at 80°C and then resuspended in 8 µL of formamide loading buffer (37% (v/v) de-ionized formamide, 1 mM EDTA and 0.02% (w/v) bromophenol blue). The PCR products were denatured by heating the plates/tubes to 94°C for 2 min and then placed on ice.

3.3.7 Li-COR based detection

The labeled samples were then analyzed on Li-COR DNA analyzers which include two major steps– acrylamide gel preparation, sample loading and gel running.

3.3.7.1 Acrylamide gel preparation

In TILLING protocol, denaturing acrylamide gels were used for separation of CELI digested PCR products. Prior to casting of a gel, the glass plates were cleaned and assembled with spacers and rails following the protocol provided with the Li-COR DNA analyzer. 20 mL of gel matrix is sufficient to cast a 15X15X2 mm sized 6.5% gel. For the preparation of gel matrix, 4 mL of 20% acrylamide gel matrix, 8.8 g of Urea, 2 mL of 10X TBE (Tris Boric acid EDTA) buffer were mixed by gentle stirring. If necessary the magnetic stirrer was heated upto 35°C for proper dissolution of urea. The final volume was made up with MQ water. To this mixture, 15 µL of TEMED and 150 µL of 10% (w/v) APS was added and the matrix was mixed with the

help of syringe. Quickly the mix was poured carefully between the glass plates, avoiding air bubbles. Later, the shark tooth comb was inserted and the pressure plate was applied. The comb was inserted in such a way that the flat end of the comb creates a shallow trough after solidification. The excess acrylamide was poured onto the comb. The gel was allowed to set for 90 min before starting the run. After complete polymerization, the gel cast was thoroughly washed on both the front and back sides of the casting plate, the casting comb was removed and remaining acrylamide was washed off from the well. Before placing the plates in the Li-COR DNA analyzer, the plates were cleaned with Kim wipes, to ascertain that the laser detection region of plate is clean and also the comb was re-inserted such that small wells are formed between two shark teeth.

3.3.7.2 Loading of Sample and Running Gel

The plate assembly was inserted into Li-COR DNA analyzer and the wells were cleaned before beginning the electrophoresis run. The upper and lower tank buffer was filled with 1X TBE buffer. The pre-run of the gel was started for 15 min, ensuring that the lasers focus properly. About 0.5 µL of the sample was loaded per well on acrylamide gel (6.5%, w/v) and the electrophoresis was carried out for 4 h at 1,500 V, 40 W and 40 mA at 50°C in Li-COR 4300 DNA analyzer. The two TIFF images obtained for 700 and 800 channels of Li-COR analyzer were analyzed in Adobe Photoshop software (Adobe Systems Inc.) and the gel was visually assessed for mutations. Mutations, which were visually discovered, were later validated by analysis of target gene from mutant lines by PCR amplification and sequencing.

3.4 Eco-TILLING

In our study, Eco-TILLING protocol involves steps that are almost similar to TILLING procedure with few exceptions. Majorly, genomic DNA isolated from the collection of tomato natural accessions was used as a template. In order to facilitate heteroduplex formation, reference cultivar (AV) genomic DNA was used. Along with *COP1* gene, we also considered *SPA3LIKE* gene for Eco-TILLING. The details of primers used for PCR amplification of *COP1* and *SPA3LIKE* genes were listed in **Table 3.3** and **Table 3.4** respectively. Nested PCR was set up in 20 µL volume. The reaction mixture consisted of 5 ng of genomic DNA as template, 1X PCR buffer, 2.5

mM dNTPs, 0.18 µL *Taq* polymerase, 1 picomoles each of labeled and unlabeled primers in a ratio of 1.2:0.7:1.5:0.5 (F_L: F_{UL}: R_L: R_{UL}). The PCR conditions were 94°C-4 min, 5 cycles of 94°C-20sec, 60°C-45 sec with a decrement of 2°C per cycle, 72°C-1 min 30 sec, 30 cycles of 94°C-20 sec, 52°C-45 sec, 72°C-1 min 30 sec followed by 72°C-10 min and held at 4°C. The amplification was followed by further denaturation and SNP detection as described for TILLING (**section 3.3**)

3.5 Amplification of *COP1* and *SPA3LIKE* genes from six tomato wild relatives

To assess the natural distribution and diversity of single nucleotide polymorphisms existing in *COP1* and *SPA3LIKE* genes, the target genes were amplified from the six wild relatives of tomato considered in the study. Genomic DNA isolated from the young leaves of these plants was used as template. The primers and PCR conditions designed and optimized for TILLING and Eco-TILLING protocols were used for amplifying the target genes. For SNP analysis, the amplified products were purified and directly sent for sequencing.

Table 3.4 Details of *SPA3LIKE* gene primers used for Eco-TILLING.

Gene	Primer	Start-end position	Length	Tm (°C)	GC (%)	Sequence
	M13 F		18	52		*TGTAAAACGACGGCCAGT*
	M13 R		19	52		*AGGAAACAGCTATGACCAT*
SPA3LIKE	F1	-564	26	59.5	50	AGAGGGTGGGATGAGATGTACCTAGT
	R1	+465	24	57.4	50	*GGCACACAAATCGATCGGCCTTTA*
	F2	-18	25	54.4	40	ACGTAGTATGAATGGGAAATGACAG
	R2	+47	24	54	42	*CCGATACTTCGTCTTTTACTGGAT*

3.6 Silica based purification of PCR products

Apart from the target sequences amplified from tomato wild relatives, the gel based confirmed mutations were re-amplified and sequenced to know the exact position and nature of the nucleotide base changes. Before sequencing, we ensure the purity and integrity of the amplified products by purifying them by silica based purification. It is a simple and time saving protocol as described in Boyle and Lew (1995). The protocol involves two major steps- preparation of silica suspension and purification of PCR products.

3.6.1 Preparation of Silica suspension

5.0 g of Silica (Sigma, S-5631) was suspended in 50 mL of autoclaved milli Q water and allowed to settle down for 2 h. The supernatant containing the fine particulate was discarded and this step was repeated once followed by centrifugation at 2000 g for 2 min. The supernatant was discarded and the slurry was stored at 4°C. Usually, one mg of the silica suspension (= 10 µL) binds to 3-4.5 µg of DNA.

3.6.2 Purification of PCR products

The PCR product was mixed with two volumes of 6 M NaI in a microfuge tube and incubated for 5-10 min. For agarose gel excised DNA fragments, the agarose is melted at 55°C in 6 M NaI with occasional mixing. To this mixture, 10 µL (~1 mg of silica suspension) of silica suspension was added and incubated at room temperature for 5 min with intermittent mixing or gentle vortex. The contents of the microfuge tube were then centrifuged for 1 min at 13,000 rpm. The supernatant was discarded and the pellet was resuspended in 500 µL wash solution (50 mM NaCl, 10 mM Tris HCl pH 7.5, 2.5 mM EDTA and 50% (v/v) ethanol). The suspension was gently mixed and then centrifuged for 1 min. The wash step was repeated once again followed by a pulse spin to remove the residual liquid. The pellet was then air dried for 10 min and an appropriate volume of milliQ water or TE was added to dislodge the pellet. The contents were allowed to stand for 5 min at room temperature and then centrifuged at 13,000 rpm for 1 min. The supernatant containing the DNA of interest was then transferred to a fresh microfuge and thus, the samples were purified.

3.7 Sequencing and Sequence analysis

Silica purified amplicons were outsourced to Macrogen Inc. (Korea) for sequencing on both the strands. The chromatogram of the sequences was analyzed in Chromas v 2.4. Sequences were assembled and compared, and inspections of differences or similarities were analyzed using the Multialin Interface page (Corpet et al., 1988). The protein sequences were aligned using PRALINE (Bawono and Heringa, 2014) software (http://www.ibi.vu.nl/programs/pralinewww/). Other software used for further analysis of the nucleotide sequence was:

3.7.1 Project Aligned Related Sequences and Evaluate SNPs (PARSESNP)

PARSESNP is a web-based tool useful for the analysis of polymorphisms in target genes in TILLING and Eco-TILLING protocols. It determines the position of the altered nucleotide base from the reference sequence (genomic or cDNA) and the effect of the polymorphism on nature of the amino acid changed (if any). If a homology model was provided, PARSESNP also determines the severity of the mutation or polymorphisms based on PSSM and SIFT scores. Since all plant proteins do not have homology models, these scores cannot be considered accurate and therefore in the present study, we neglected the PSSM scores and relied on the SIFT scores that were calculated separately.

3.7.2 SIFT (www.sift.dna.org)

For missense/nonsynonymous polymorphisims, the deleterious effect of the mutation was estimated by assigning SIFT (Sorting Intolerant From Tolerant) value using version 4.0.5 of SIFT (Ng and Henikoff, 2003; Sim et al., 2012). The prediction is based on database sequence homology and physical properties of amino acids. The following parameters were used: median conservation was set to 2.75, sequences with 90% similarity or more were discarded, and the database used was the Uniprot-SWISS-PROT+ TrEMBL 2010_09. The probability of SIFT score less than 0.05 was predicted to be damaging to the function of the encoded protein, those greater than or equal to 0.05 were predicted to be tolerated.

3.7.3 BLAST (http://www.ncbi.nlm.nih.gov/BLAST)

The software is designed to take protein and nucleic acid sequences and ascertain homology with respective sequences present in NCBI databases. Protein sequence of tomato SPA3LIKE was used as query in BLAST. The query sequence was compared (aligned) with the database sequence of NCBI and the amount of similarity with related proteins was inspected.

3.7.4 Phylogenetic analysis

We used MEGA 4 (Tamura et al., 2007) software for analyzing phylogenetic relationship of nucleotide or protein sequences. The sequences were aligned using Multialin so that the substitutions can be accurately enumerated. During alignment, gaps were introduced in sequences that have undergone deletions or insertions. Phylogenetic relationships were performed using neighbor joining method. Neighbor joining is a bottom up clustering method which uses distance measures to correct for multiple hits at the same sites and chooses a topology showing the smallest value of the sum of all branches as an estimate of a correct tree. This method computes the length of the branches of the tree and in each stage; two nearest nodes of the tree are chosen and defined as neighbors. This process is carried out until all the nodes are paired together. The bootstrap values are denoted on the branch which is a measure of the evolutionary changes existing within the genotypes.

3.7.5 Codon evaluation

The codon usage frequencies for the synonymous or silent changes identified for each gene were evaluated by comparing the frequencies obtained for the reference codon and the modified codon from the database- http://www.kazusa.or.jp/codon/ maintained by the Kazusa DNA Research Institute (Chiba, Japan). The calculated codon frequency for the cultivated tomato *S. lycopersicum* was given in **Table 3.5**.

3.7.6 Tests of Neutrality

To test for deviations from the neutral equilibrium model of evolution, we performed Tajima's D (Tajima 1989) analysis using the program DnaSP. In each locus, Tajima's D was calculated at all sites and at silent sites separately. In Tajima D

analysis, negative values indicate an excess of low-frequency polymorphisms, whereas positive values indicate an excess of intermediate polymorphisms. Therefore it signifies the mode of selection. DnaSP version 5.10 (Librado and Rozas, 2009) was used for carrying out statistical neutrality tests, and Tajima's D test on each sequence. In addition, the Tajima D value for synonymous and nonsynonymous SNPs reveals the selection pressure acting on the loci.

3.8 In silico characterization of tomato SPA3LIKE

Homology modeling of SPA3LIKE of WT, NB518 and NB182 was performed using SWISS-MODEL software (Arnold et al., 2006; Kiefer et al., 2009; Guex et al., 2009; Biasini et al., 2014), with the template protein sequence of SPA3LIKE from WT tomato and *SPA3LIKE* variants. The predicted models were evaluated by ploting Ramachandran plots and calculating the energy values. A score of more than 20% or more is predicted to be a good model. For the SPA3LIKE protein of WT, NB518 and NB182 the scores were above 90% indicating a very good model. The structural difference between mutant and WT were analyzed by using the binding site prediction software FT site (http://ftsite.bu.edu/cite).

Table 3.5 Codon usage frequency table for *Solanum lycopersicum*.

***Solanum lycopersicum* [gbpln]: 1452 CDS's (634390 codons)**
fields: [triplet] [amino acid] [fraction] [frequency: **per thousand**] ([number])

UUU	F	0.6	26	16504	UCU	S	0.26	21.2	13464	UAU	Y	0.6	18.6	11799	UGU	C	0.62	10.8	6883
UUC	F	0.4	17.5	11123	UCC	S	0.12	9.9	6268	UAC	Y	0.4	12.4	7855	UGC	C	0.38	6.7	4223
UUA	L	0.15	14.4	9111	UCA	S	0.25	20.7	13102	UAA	*	0.4	0.9	598	UGA	*	0.37	0.9	553
UUG	L	0.25	24.2	15336	UCG	S	0.07	5.6	3521	UAG	*	0.23	0.5	340	UGG	W	13.5	13.5	8563
CUU	L	0.26	24.9	15814	CCU	P	0.39	19.2	12192	CAU	H	0.67	15.5	9827	CGU	R	6.9	6.9	4404
CUC	L	0.12	11.2	7127	CCC	P	0.12	5.7	3645	CAC	H	0.33	7.8	4926	CGC	R	3.1	3.1	1946
CUA	L	0.11	10	6348	CCA	P	0.39	19.2	12205	CAA	Q	0.6	21	13315	CGA	R	5.4	5.4	3414
CUG	L	0.11	10.5	6665	CCG	P	0.09	4.6	2920	CAG	Q	0.4	14	8861	CGG	R	3.1	3.1	1981
AUU	I	0.5	28.2	17914	ACU	T	0.39	19.9	12612	AAU	N	0.64	30.5	19339	AGU	S	15.2	15.2	9669
AUC	I	0.25	14	8906	ACC	T	0.17	8.6	5464	AAC	N	0.36	17.3	11002	AGC	S	9.3	9.3	5897
AUA	I	0.25	14	8893	ACA	T	0.35	17.9	11333	AAA	K	0.5	31	19724	AGA	R	16.4	16.4	10398
AUG	M	1	24.7	15651	ACG	T	0.09	4.6	2939	AAG	K	0.5	31	19651	AGG	R	11.9	11.9	7541
GUU	V	0.43	28	17762	GCU	A	0.45	30.7	19474	GAU	D	0.72	39.3	24917	GGU	G	23.9	23.9	15192
GUC	V	0.15	10.1	6382	GCC	A	0.15	10.1	6426	GAC	D	0.28	15	9509	GGC	G	9.7	9.7	6265
GUA	V	0.17	11.2	7126	GCA	A	0.33	22.2	14107	GAA	E	0.57	34.8	22053	GGA	G	25.6	25.6	16258
GUG	V	0.25	16	10181	GCG	A	0.08	5.2	3327	GAG	E	0.43	26.6	16906	GGG	G	10.8	10.8	6839

Coding GC 42.52% 1st letter GC 50.16% 2nd letter GC 39.87% 3rd letter GC 37.53%

3.9 Biochemical analysis

3.9.1 Estimation of Leaf Chlorophyll content

The chlorophyll from mature green fruits was estimated in 80% (v/v) acetone as described in Arnon et al. (1949). Briefly, a known amount of leaf tissue (50 mg) in 2 mL 80% (v/v) acetone was ground using three steel balls in a Mini-Bead Beater for 1 min and it was stored in darkness overnight. After centrifugation at 5000 rpm at 4°C for 5 min, the supernatants were collected and absorbance was recorded at 663 nm and 645 nm using Uvikon spectrophotometer. The chlorophyll a and b were calculated by the following formulae:

$$\text{Chl a } (\mu g/mL) = (12.7 \text{ X } A_{663} - 2.79 \text{ X } A_{645})$$

$$\text{Chl b } (\mu g/mL) = (22.9 \text{ X } A_{645} - 4.68 \text{ X } A_{663})$$

where A is the absorbance at 663 nm and 645 nm.

3.9.2 Estimation of Anthocyanin content in seedlings

Ten seedlings of uniform size/height were harvested and extracted with 1.2 mL of acidified methanol (1.0% HCl (v/v)). The solution was incubated at 4°C for 24 h with constant shaking. A folch partitioning was performed after adding 0.9 mL water and 2.4 mL chloroform to the extract. After centrifugation at 1600 *g* for 30 min, the absorbance of the upper phase containing anthocyanin was determined at A_{535nm} (Harborne, 1967).

3.9.3 Estimation of carotenoids

Due to the number of samples and also the availability of fruit tissue for carotenoid analysis, different samples were processed differently.

3.9.3.1 Carotenoid content in fruits of *COP1* variants

About 1.0 g of red ripe fruit tissue was homogenized with mortar and pestle using liquid nitrogen. The fine powder was mixed with 25.0 mL of organic mixture consisting of hexane: acetone: ethanol (2:1:1 (v/v/v)). After mixing for 5 min, 3.9 mL of water was added and was centrifuged at 4°C, 12,000 rpm for 10 min. After phase

separation, 1.0 mL of upper hexane layer was collected and quantified spectrophotometrically at A_{503} for lycopene and A_{470} for β– carotene.

3.9.3.2 Carotenoid content in fruits of TILLING mutant line

Red ripe fruits were flash frozen in liquid nitrogen and stored at -80°C. Lycopene and β-carotene were extracted from fruits and estimated using a previously described procedure (Kilambi et al., 2013). In brief, the pericarp tissue was homogenized with mortar and pestle and 1 g of homogenate sample was mixed with 25 mL extraction solvent consisting of hexane/acetone/ethanol (50:25:25 v/v/v). After 5 min vortexing, 3.9 mL water was added to above mixture. After mixing, the mixtures were left for 2-3 min to allow phase separation. A 1 mL aliquot was withdrawn from the upper phase consisting of hexane and was dried in a Speed-Vac concentrator. The dried extract was dissolved in tetrahydrofuran/acetonitrile/methanol (15:30:35 v/v/v) and was incubated overnight in dark at -80°C. The extract was then centrifuged at 13,225 x *g* for 10 min at 4°C, and supernatant was filtered through 0.2 µm membrane filters (Ultipor® N_{66}® Nylon 6, 6, membrane; PALL Life Sciences, Bengaluru, India). A 20 µL filtered sample was analyzed by HPLC (Shimadzu) for separation of carotenoids on a C18 column (250 X 4.6 mm; 5 micron, Phenomenex, Torrance, CA) using a 60 min isocratic gradient of methanol/acetonitrile (90:10 v/v) + triethylamine (TEA) 9 µM as mobile phase. The absorbance for carotenoids was read at 470-503 nm with a UV/VIS detector (model SPD-20A). Lycopene and β-carotene content of sample was calculated by comparing the area of standard lycopene (L-9879 from tomato, Sigma-Aldrich) and β-carotene (C-4582, Sigma-Aldrich).

3.9.3.3 Carotenoid content in fruits of *SPA3LIKE* Eco-TILLING lines

The carotenoid profiling in the fruits of *SPA3LIKE* variants was performed using Accela U-HPLC system (Thermo Fisher Scientific, Bremen, Germany). 15 different carotenoid standards (violaxanthin, neoxanthin, antheraxanthin, lutein, zeaxanthin, phytoene, β-cryptoxanthin, phytofluene, α-carotene, β-carotene, Zeta-carotene, δ-carotene, γ-carotene, neurosporene and lycopene) were purchased from CaroteNature, (Lupsingen, Switzerland) and used in this study. For chromatographic separation by HPLC, a YMC polymeric C30 column (250 mm × 4.6 125 mm I.D., 3 µm particle size) and YMC guard cartridge (23 mm × 4.0 mm I.D., 3 µm particle 126

size) from YMC Co. (Kyoto, Japan) was used. The extraction procedure was carried in a room shielded from light. 150±10 mg of the freeze-dried tomato fruit tissue was homogenized to a fine powder with a mortar and pestle. The homogenate was transferred to 2 mL micro-centrifuge tubes on ice followed by the addition of 1000 µL of chloroform and 500 µL of dichloromethane. The resultant suspension was mixed on a thermomixer at 1000 rpm for 20 min at 4°C. To the tube, 500 µL NaCl-saturated water was added and contents were mixed by inversion. The tubes were centrifuged at 5000g for 10 min at 4°C and the organic phase was collected with a pipette. The aqueous phase was re-extracted with 500 µL chloroform and 250 µL dichloromethane. The chloroform and dichloromethane extracts were pooled and dried by centrifugal evaporation. The residue was stored at −80°C under nitrogen until U-HPLC analysis. The dry residue was dissolved in the 20 µL of injection solvent MeOH:MTBE (1:3) immediately prior to analysis. The analysis of the chromatographic data was carried out by Xcalibur software v 2.2 (Thermo Fisher Scientific, Bremen, Germany).

3.10 Gene expression analysis

3.10.1 RNA extraction

Total RNA was extracted from the pericarp tissue of fruits in three biological replicates at different developmental stages using hot phenol method (Verwoerd et al., 1989). Approximately, 2 g of fruit pericarp tissue was ground into fine powder in liquid nitrogen with pre-chilled mortar and pestle. The powder was added to a falcon tube containing 4 mL of buffer A (0.1M LiCl, 0.01 M EDTA, 1% SDS, 0.1 M Tris [pH 9.0])/phenol mixture at 80°C. Two ml chloroform was added to it and kept in a shaker for 30 min at room temperature. The samples were transferred to 1.5 mL Eppendorf tubes and then were centrifuged at 10,000 rpm for 10 min at 4°C. The upper phase was transferred to a new Eppendorf and 1/3 volume of 8 M LiCl was added to it and incubated overnight at 4°C. Next day, samples were centrifuged at 10,000 rpm for 30 min at 4°C. Supernatant was discarded and pellet was dissolved in 2 M LiCl followed by a short spin at 10,000 rpm for 5 min. Supernatant was discarded and pellet was washed three times with 70% ethanol by centrifuging at 10,000 rpm for

5 min. The pellet was air dried and dissolved in 30 µL of DEPC treated MQ and kept at -70°C till further use.

RNA from leaves was isolated using TRIZOL (Sigma) method. Approximately 100 mg leaf tissue was homogenized to a fine powder in liquid nitrogen using mortar and pestle. 1.0 mL of TRI reagent was added to this and the homogenate was incubated for 5 min. Later, the homogenate was transferred into a microfuge tube and 200 µL of chloroform was added. The mixture was gently mixed and incubated at RT for 2-3 min followed by centrifugation at 12,000 rpm for 15 min at 4°C. The aqueous upper phase was transferred to a fresh microfuge. RNA was precipitated by adding 500 µL of isopropanol. After 10 min incubation at RT, the samples were centrifuged at 12,000 rpm for 15 min at 4 °C. The supernatant was discarded and the pellet was washed twice with 75% ethanol (v/v). The pellet was then air dried and dissolved in RNase free water.

3.10.2 Determination of RNA quality

The RNA was incubated with RNAse free DNAse (Promega) as per manufacturer's protocol to eliminate the genomic DNA contamination. RNA concentration and quality was determined using an ND0100 Nanodrop UV-Vis spectrophotometer. To check the integrity of isolated RNA, 1-2 µg of RNA sample was mixed with 5 µL MiiliQ water and 2 µL 6X loading dye containing EtBr and loaded on agarose gel (1.2%, w/v) containing 4% (v/v) formaldehyde and 1X MOPS. All the solvents and the gel apparatus used were pretreated with 0.01% (v/v) diethyl pyrocarbonate (DEPC). Electrophoresis of RNA was performed in 1X FA (formaldehyde) buffer at constant voltage (50 V) for 1-2 hr. Thereafter, gel was visualized using UV filter in a gel documentation unit (Alpha Imager, India) and integrity was determined by checking the presence of two intact bands of 18S and 28S rRNA as a marker for RNA quality (Sambrook et al., 1989).

3.10.3 cDNA synthesis and Quantitative Real-time PCR

cDNA was prepared from 2 µg of RNA using cDNA synthesis kit (SuperScript III; Invitrogen, USA). Quantitative Real time PCR (qRT PCR) was carried out in EP realplex (Eppendorf). The transcript abundance was measured in 10 µL volume of the

SYBR Green PCR Master Mix (Takara, Japan) containing cDNA corresponding to 5 ng of total RNA with gene specific primers (Kilambi et al., 2013). The ΔΔCt value was calculated by normalizing Ct values of each gene to the mean expression of the two internal control genes (*β-actin* and *ubiquitin3*). [ΔΔCt value = (ΔCt value of gene from mutant)-(ΔCt value of gene from WT/reference)]. Primers for *SPA3LIKE*, *COP1*, *SPA1*, *HY5*, *β-actin* and *ubiquitin3* were designed from Primer 3 (**Table 3.6**).

3.11 Western blotting of tomato COP1 protein

3.11.1 Protein extraction and estimation

Three day old dark grown seedlings were used for protein extraction. The extraction procedure was conducted in complete darkness with a very low intensity of safe green light. Seedlings weighing around 150 mg were snap frozen using liquid nitrogen and placed in a pre-frozen 1.5 mL microfuge tube. The tissue was ground using a homogenizer for few seconds. To maintain the low temperature, the pestle of homogenizer was also pre-frozen in liquid nitrogen. 300 µL of extraction buffer (0.5M Tris HCl, pH 7.5; 0.5 M sucrose; 0.05 M EDTA; 1.0 mM PMSF; protease inhibitor cocktail) was added to the ground tissue and immediately centrifuged at 5000 rpm for 5 minutes in Sorvall microfuge. The rotor was also pre-chilled to avoid any protein degradation. The supernatant was transferred to a fresh microfuge tube. An aliquot of the extract was taken aside for protein estimation using Bradford's assay and to the remaining extract 75 µL of 4x sample loading dye was added and boiled for 5 minutes at 100°C in a water bath. Before the samples were subjected to gel electrophoresis, protein estimation was performed using Bradford's micro essay (Bradford, 1976).

3.11.2 SDS gel electrophoresis

The proteins extracted from the seedlings were separated using sodium dodecyl sulfate gel electrophoresis (SDS-PAGE) according to Laemmli (1970). One millimeter (mm) thick separating gel of 12% (w/v) was polymerized in 1.5 M Tris-HCl (pH 8.8) containing 0.1% (w/v) SDS, 0.1% (w/v) APS and 0.01% (v/v) TEMED. 5% (w/v) stacking gel was made in 1 M Tris-HCl (pH 6.8) containing 0.1% (w/v)

Table 3.6 Primers used for Real time PCR analysis.

Gene	Forward primer (5′- 3′)	Reverse primer (5′- 3′)	Product size (bp)
COP1	GTGATTGCCCCTGTTGTTCT	AACAGGGGATGCAGTTTTTG	118
SPA3LIKE	GGTTTTAGGGTCGGGTGATT	ACCGGATCCGACTGACAATA	132
HY5	CTAGTTTCGGGTGGATTGTT	AGCCACAATCCTTCACTCTC	101
SPA1	GTGATCTGCTCCCTCTGTTT	CGCAGATTCATTCAAAAGTG	118
PSY1	CAGCCTTAGATAGGTGGGAAAAT	CTGAATGGCTGAATATCAACTGG	121
CYC B	GGGTAATGAGCCATATTTAAGGG	TCCAACGACTCTCTGAGGTA	92
β-ACTIN	GTCCCTATTTACGAGGGTTATGC	CAGTTAAATCACGACCAGCAAGATT	108
UBIQUITIN 3	GCCGACTACAACATCCAGAAGG	TGCAACACAGCGAGCTTAACC	110

Phytoene synthase1 (*PSY1*, Solyc03g031860.2.1), *Lycopene-β-cyclase* (*CYCB*, Solyc06g074240.1.1), *Constitutive photomorphogenic1* (*COP1*, *Solyc12g005950*), *β-actin* (FJ532351.1) and *ubiquitin3* (X58253.1), *Suppressor of PhytochromeA 3 Like* (*SPA3LIKE*, Solyc11g005190.1), *Long hypocotyl5* (*HY5*, Solyc08g061130.1).

SDS, 0.1% (w/v) APS and 0.01% (v/v) TEMED. Gel slabs were prepared in an AE-6530 mPAGE electrophoresis unit (ATTO, Japan) according to the manufacturer's instructions. Samples were prepared in 4X sample buffer containing 1 M Tris-HCl (pH 6.8), 4% (w/v) SDS, 20% (v/v) glycerol, 0.1% (w/v) bromophenol blue and 10% (v/v) β-mercaptoethanol. The samples were boiled for 5 minutes for uniform coating of the detergent. After cooling to room temperature, the samples were loaded and electrophoresed at room temperature at 100 V constant power for approximately 90 minutes. The gel running buffer was made of 250 mM Tris, 200 mM glycine and 0.1% (w/v) SDS.

3.11.3 Electroblotting of SDS-PAGE gel

Proteins were electroblotted onto PVDF (Immobilon-P, No P-2938; Sigma) membrane using semi-dry blotting method. The PVDF membrane was cut to the size of the resolving gel and wetted with methanol for 2 min, washed with distilled water and then soaked in transfer buffer (25 mM Tris base, 192 mM glycine, 20% (v/v) methanol). Whatman No. 3 chromatographic paper was cut to the size of the gel and soaked in transfer buffer. After electrophoresis, the stacking gel was excised and the bottom of the left hand corner of the resolving gel was marked with a cut. The gel was first washed with distilled water, followed by transfer buffer. Semi-dry blotting was carried out in a custom built apparatus. Both anode and cathode graphite plates were washed with distilled water. On the anode plate, three sheets of Whatmann papers were layered carefully so that no air bubbles were trapped between the sheets. The membrane was layered over the sheets and a small cut was made at the left hand bottom corner of the membrane to coincide with that of the gel. The gel was carefully layered on top of the membrane on which three more sheets of Whatman papers were layered. Finally, on the top cathode plate was placed. The blotting unit and the whole sandwich consisting of the plates, papers and gel were tightened using clamps. The blotting was done at a constant current of 0.8 mA/sq.cm area for 1.5 h.

3.11.4 Immunoblotting of COP1 protein

Immunoblotting was done following the procedure of Towbin et al., (1979) at room temperature. After electroblotting, the non-specific binding sites on the membrane were blocked by incubating the membrane in 10 mL of blocking buffer

containing 1% BSA (w/v) in 20 mM Tris –HCl, pH 8.0. 150 mM NaCl, 0.1% Tween-20 (v/v) (Tris buffered saline + Tween-20, TBST). After blocking, the membranes were washed thrice with 20 mL of TBST for 10 min each. Next, the membrane was incubated in 10 mL of primary antibody (AtCOP1 at 1:2,000 dilution, Agrisera) in TBST for 1 h with gentle shaking. The unbound primary antibody from the membrane was removed as the gel was washed three times with TBST. It was then incubated for 1 h in secondary antibody, which is horseradish peroxidase, coupled anti-rabbit IgG (1: 80,000 dilution in TBST, Sigma). Thereafter the membrane was again washed thrice for 10 min each with 10 mL of TBST to remove the unbound secondary antibody.

Antigen-antibody complexes on the membrane were incubated in Amersham ECL primer western blotting reagent as per the manufacturer's guidelines. Horseradish peroxidase (HRP)-catalyzed oxidation of luminol generates chemiluminescence at a wavelength of 425 nm which was visualized using Versa Doc Gel Imaging System (Bio-Rad). Image lab software was used to real time monitor the chemiluminescence and when the bands were clearly visual, the reaction was stopped by rinsing the membrane in distilled water for few min.

Chapter 4

Isolation of tomato *COP1* mutants through TILLING

4.1 Introduction

TILLING (Targeting Induced Local Lesions in Genomes) is an alternative reverse genetic technique developed for the detection of point mutation in Arabidopsis (McCallum et al., 2000a). Four major steps define this technique- development of a mutagenized population, pooling of DNA and amplification of target genes, heteroduplex formation and mismatch enzyme cleavage by endonuclease like CELI or EndoI followed by subsequent fragment length analysis. Detection of mutations was initially carried out by HPLC (Caldwell et al., 2004) and the later advancements included detection by denaturing acrylamide gels on Li-COR analyzers (Till et al., 2004), capillary sequencers (Stephenson et al., 2010), by MALDI-TOF (Chawade et al., 2010) or by high-resolution DNA melting analysis (HRM) (Bush and Krysan, 2010). In combination with chemical mutagenesis and the availability of genome sequence information, TIILING has been widely accepted and applied to a broad range of plant species irrespective of the ploidy level. A major advantage of this technique is based on the random mutagenesis that creates an array of mutations including missense, truncations and splice junction mutations. This possibility of detecting multiple alleles has made TILLING a more preferable technique. In addition, TILLING is a nontransgenic approach and hence do not have any ethical issues as encountered for transgenic plants. In consideration with all the above discussed advantages TILLING has been widely used in many crop species including tomato as mentioned in **Table 2.1** (**Chapter 2**) to develop new crop varieties.

Tomato is a diploid organism with a relatively small genome size of 950 Mb. Being a very good source of health promoting nutrients like carotenoids, flavonoids, ascorbic acid, etc and a characteristic ripening process, tomato has gained both economic and scientific importance. Tomato fruit ripening is a complex process regulated by many internal and external factors. In the past two decades, a handful of ripening mutants of tomato were exploited to understand the biochemical and molecular aspects of ripening. Light is one of the external cues that influences many aspects of plant development including fruit ripening (Alba et al., 2000; Gupta et al., 2014). The tomato *hp1* and *hp2* mutants exhibit exaggerated photomorphogenesis and enhanced carotenoid fruit content (Kilambi et al., 2013). In addition, the recent characterization of tomato *COP1LIKE RNAi* (*SPA3LIKE*) and *CUL4 RNAi* lines

further support that light signaling components can be manipulated for enhancing the nutritional quality of tomato. Interestingly, the genes encoding for the *hp1* and *hp2* mutations- Arabidopsis homolog of *DDB1* and *DET1* along with *CUL4*, *SPA3LIKE* were considered as negative regulators of photomorphogenesis and work in synchrony with *COP1* and therefore degrade the target proteins via the Ub- mediated 26S proteasome pathway.

Constitutive photomorphogenic1 (COP1) is an E3 ubiquitin protein ligase which acts downstream of photoreceptors (Deng et al., 1992). It was established in Arabidopsis that COP1 functions as an essential negative regulator of light mediated plant development, evidenced by *cop1* mutant seedlings undergoing photomorphogenic development even in the absence of light, and *cop1* null alleles never surviving to adult stage. COPl has a novel combination of the following recognizable motifs: a ring-finger zinc-binding domain at the amino-terminal region, followed by a coiled coil helix structure, and several WD40 repeats at the carboxyl-terminus (Deng et al., 1992; McNellis et al., 1994a). The coiled-coil and WD40 repeats are likely surfaces for mediating protein-protein interactions, whereas the zinc finger domain has the potential to interact with either or both proteins and nucleic acids. The multiple interactive motifs of COPl provide a physical basis for being a central player in interacting with multiple upstream and downstream components in the light regulatory cascades. Therefore, *COP1* is involved in many plant developmental processes- seedling photomorphogenesis, shade avoidance, anthocyanin accumulation, flowering etc. However, its role in fruit ripening is not yet elucidated and many obvious questions remain unanswered in this regard such as- What are the interacting partners of the *COP1* gene product in tomato? What are their roles during fruit ripening? How do they regulate their downstream targets? How are light signals from multiple receptors integrated to modulate fruit ripening?

In order to answer these questions one has to obtain mutations in tomato *COP1* gene. Several attempts were made in our lab to isolate *COP1* mutant through forward genetic screens in tomato, however, these attempts were unsuccessful. We adopted TILLING; a reverse genetic strategy to isolate *COP1* mutants from an EMS mutagenized population.

4.2 Results

4.2.1 TILLING platform

EMS mutagenized populations (I, II and III) were used for TILLING in our study (**Table 3.1,** Chapter 3). The size of each population varied due to the percentage of lethality observed. 2304 M_2 individuals of 60 mM population (I), 4608 M_2 individuals of 120 mM (II) and 3840 individuals of M82 (III) populations formed the core genetic resource for TILLING. Use of in-house purified *Taq* polymerase and CELI enzyme aided in reducing the cost of TILLING in our study. Further, eight fold tissue pooling and two dimensional strategy of arraying samples had increased the efficiency of high throughput screening by reducing the number of reactions to be performed per primer pair. The upper limit of CODDLE based primer design was 1500 bp, which defined the fragment length to be chosen for screening. In almost all TILLING platforms, the primers were designed to amplify a fragment of 600 bp-1000 bp length. The amplicon length within the above range was considered as optimum due to the limitation with commercially available DNA size markers which can detect fragments up to 700 bp only. For genes like tomato *COP1* that are lengthy with small exons and long introns, more primer sets are required for screening the entire gene. This would indirectly increase the total number of reactions per gene and therefore, the cost of TILLING. In order to circumvent this problem, we used in-house developed labelled DNA size marker for efficient detection of digested fragments beyond 700 bp (Gupta et al., 2011). This helped us to design primers that could amplify a fragment more than 700 bp, thereby, reducing the number of reactions to be performed per gene. Overall, an efficient TILLING platform was set that could be screened for mutations in the gene of interest.

4.2.2 Tomato COP1 is encoded by a single copy gene

Once the TILLING platform was set, the next step was to select a target gene. Usually, single copy genes with known sequence information were considered as ideal targets for TILLING because in this case any drastic mutation could be scored at the phenotype level. If the gene is very important for plant development, this may result in lethality. In cases, where precise sequence information of the gene in the target species was not available, the gene information was facilitated by using the homologous sequences found in other species or by Southern analysis in the target

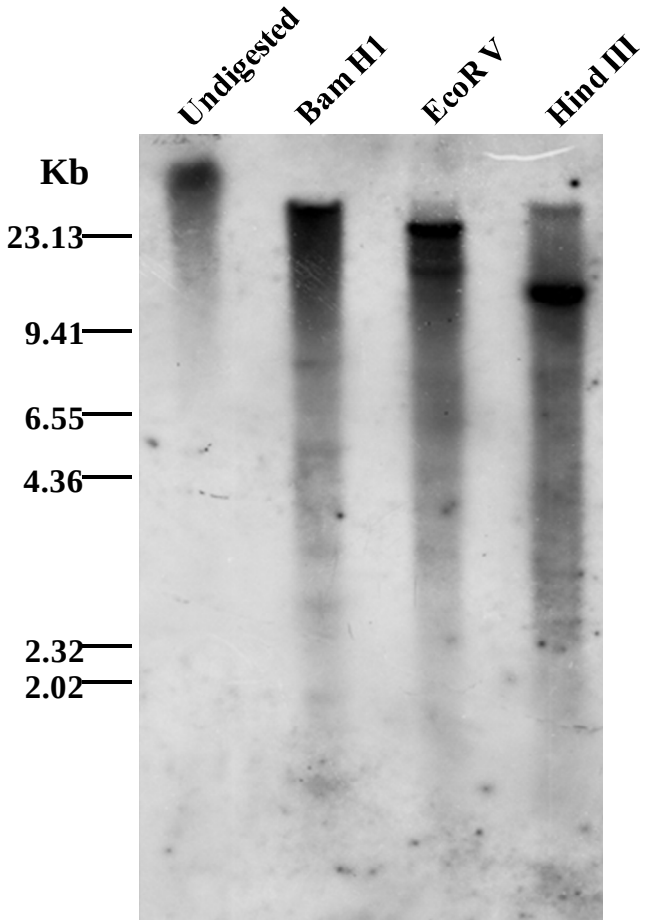


Figure 4.1 Southern analyses to determine the copy number of *COP1* gene in tomato. Total genomic DNA in 10 µg aliquots was digested with *BamHI*, *EcoRV* and *HindIII*. Southern blotting and hybridization was done as described in the methods. The blot was probed with a 414 bp of *COP1* coding sequence. DNA marker sizes are indicated on the left. A prominent band in each lane with digested DNA confirms the presence of a single copy of *COP1* gene in tomato.

species. In the present study we chose tomato *COP1* as our target and when we started TILLING, the tomato genome sequencing was still under way. We could retrieve only the BAC sequence information based on the *COP1* gene bank accession number AF029984.1 submitted in the NCBI. In order to confirm the copy number of *COP1* in tomato, we performed Southern analysis. A Southern blot of *S. lycopersicum* cv. Arka Vikas genomic DNA, individually digested with three restriction enzymes - *BamH1*, *EcoRV* and *HindIII* (**Fig. 4.1**), showed a single prominent hybridized band in each lane indicating that tomato genome has a single copy of *COP1* gene. The sequence retrieved from BAC sequence was used to design primers and was later crosschecked with the tomato genome submitted in the Solanaceae genomics network (SGN) website (http://solgenomics.net).

4.2.3 Optimization of PCR for TILLING *COP1* gene

The major bottleneck in high throughput PCR based screening is efficient standardization of PCR conditions. When developing a new protocol for PCR amplification, it is necessary to optimize all parameters involving- reagent concentrations (template, dNTPs, buffer, Magnesium chloride, primers and *Taq*. polymerase); cycling temperatures and cycle number. Since we used an in-house isolated *Taq* polymerase, we optimized the PCR conditions for stringent amplification of the target gene. The optimal conditions for PCR with the gene specific primers were standardized both on agarose and Li-COR gels. In the current study, different sets of primers (six sets for *COP1*) were optimized (**Table 3.3, Chapter 3**). Initial standardization was carried out using SET III primers for *COP1* gene. The established conditions were further used with slight modifications to expedite the screening with other primers. The following parameters were optimized:

4.2.3.1 Annealing Temperature (T_a)

To improve the specificity of the PCR reaction, the optimal annealing temperature is empirically determined by performing a gradient PCR at different temperatures (Tm ± 5°C). The *COP1* region encompassing the exons 5, 6 and 7 was amplified with SET III forward and reverse primers. A temperature gradient from 54°C to 72°C was set using following cycling conditions - denaturation at 94°C (20 s), annealing 45s, elongation at 72°C (90s) x 35 cycles. The calculated primer

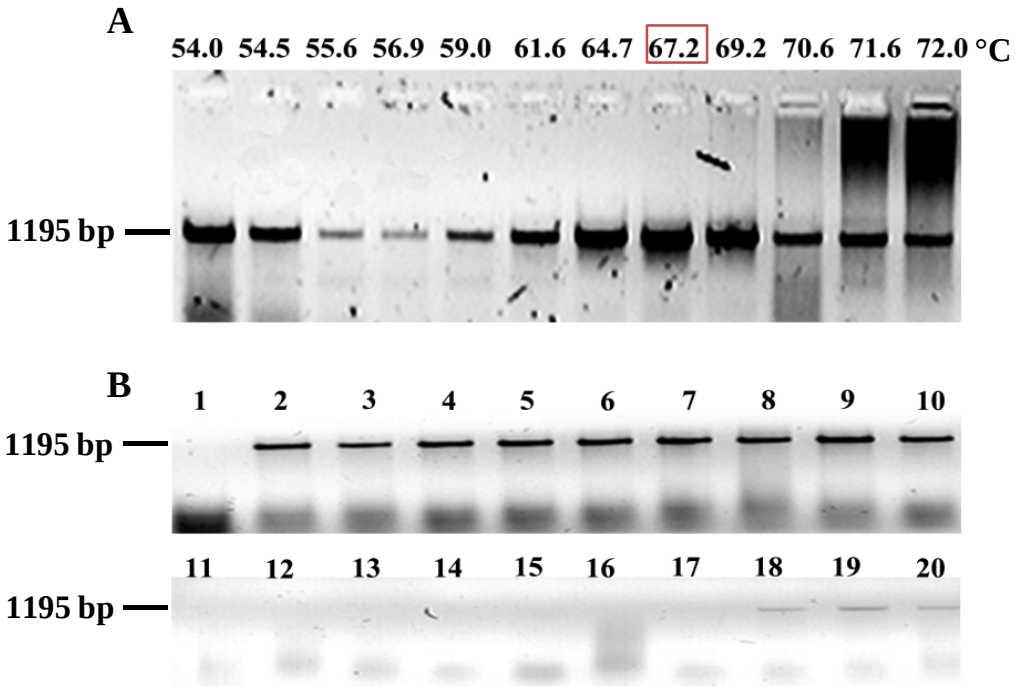


Figure 4.2 Optimization of TILLING PCR conditions on agarose gels for TILLING. A, AV DNA was amplified with SET III primers at a temperature gradient of 54°C to 72°C. The calculated primer annealing temperature was 64.0°C and the best annealing temperature obtained was 67.2°C (marked in red open box). **B**, Amplification of 8X pooled DNA in comparison with AV DNA was checked with optimized annealing temperature (67.2°C). Lane number 1-10 represent the amplified product of AV DNA and from 11-20 represent the 8X pooled DNA. SET III primers with the following cycling conditions were used: denaturation at 94°C (20 s), annealing for 45 s, elongation at 72°C (90 s) x 35 cycles. Final reactions were analyzed on 1% agarose gel stained with ethidium bromide.

annealing temperature was 64°C but the actual annealing temperature determined was 67.2°C (**Fig. 4.2**). Optimized T_a was used to amplify the pooled DNA (used for TILLING) and wild type DNA (Arka Vikas). The Li-COR gel image in **Figure 4.3** clearly indicates less product yield with pooled DNA, suggesting that the reaction needs further optimization.

4.2.3.2 Cycling conditions

Though the standard PCR (straight cycle) cycling conditions were suitable, we changed the cycling conditions to increase the yield of PCR product. We tried both step-down and touchdown cycles. A touchdown program was used where the T_a was progressively lowered from 72°C to 62°C by 2°C every cycle followed by additional 25 cycles at 62°C. With this modified PCR conditions, we obtained a threefold better amplification with SET III primers (**Fig. 4.4**).

4.2.3.3 Reagent concentrations and PCR additives

A second level of standardization was done on polyacrylamide gels in Li-COR DNA Analyzer 4300. The units of *Taq* required per reaction, concentration of Mg^{2+} ions and the amount of 10X buffer added into the PCR reaction (0.8X-1.5X) were standardized **(Fig. 4.5**). The effect of PCR additives like glycerol, BSA (for more yield and specific product) in the PCR reaction was also tested. For the SET III primers, a slight increase in the Mg^{2+} concentration to 2.5 mM and a decrease in monovalent ions in the PCR buffer (0.8X) showed profound effect on Li-COR gel. However, the overall reaction profile on the gel after digestion with CELI enzyme was faint and in spite of a prominent main product, the intensity of the digested fragments was weak. Hence, we standardized the next parameter, i.e. the ratio of labeled primers to unlabeled primers and also the template concentrations. Theoretically, even a single molecule of the template DNA is sufficient for the polymerase to amplify the target sequence. However, the amount of template depends on the type of DNA (plant, human, bacterial, plasmid etc.) and also the quality of the DNA isolated. Sometimes the amplicon size might also affect the amount of template needed. For Li-COR based detection, a multiplex reaction with labeled and unlabeled primers was used. Therefore, a ratio of the primers and the starting material has to be standardized for a better amplification of the target. We optimized the template and primer concentration simultaneously. A gradient of genomic DNA ranging from 1 ng

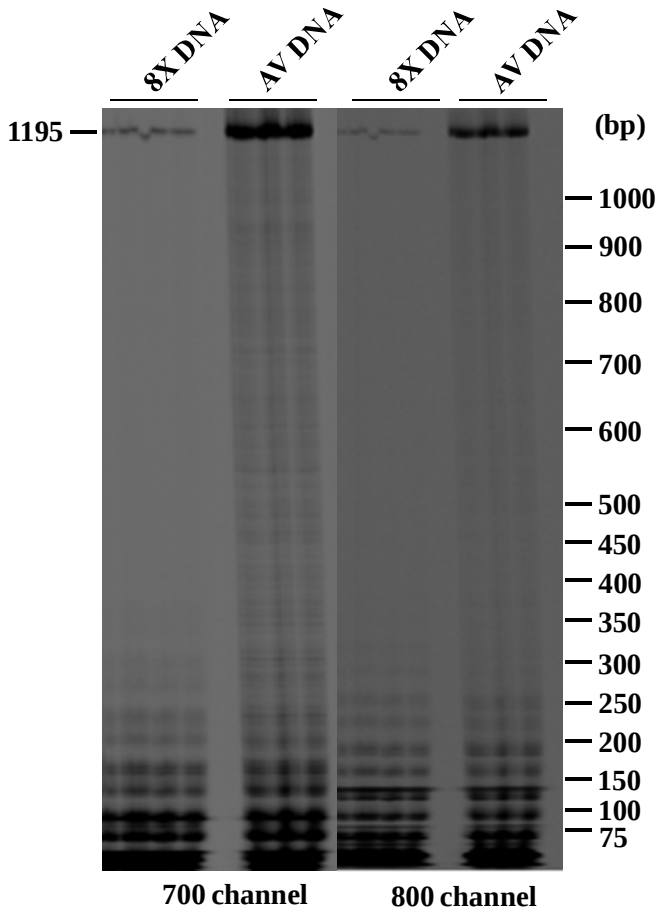


Figure 4.3 Analysis of labeled PCR products from 8X Pooled DNA on Li-COR. Eight fold pooled DNA and AV DNA was amplified with a primer cocktail ration of 2:3:1:4 (F:F_L:R:R_L) at an annealing temperature of 67.2°C. The 1195 bp amplicon is marked at the top of the gel and the labeled DNA size marker on right side. F- Forward, R- Reverse, F_L- Forward labeled with IRD 700, R_L- Reverse labeled with IRD 800.

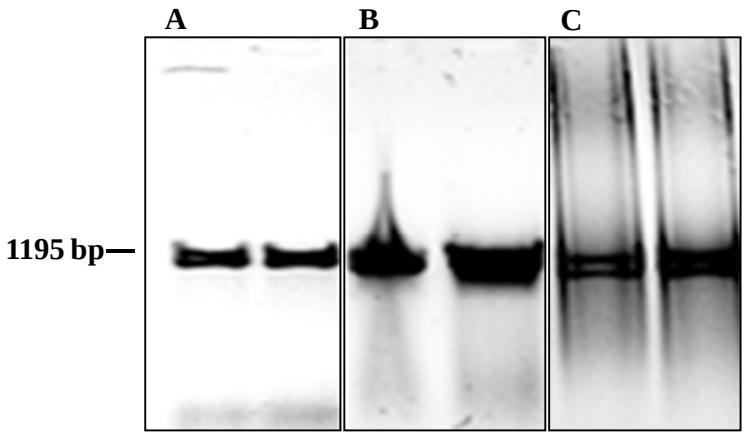


Figure 4.4 Optimization of PCR cycles for TILLING. Three different PCR cycling conditions were checked to improve amplification. **A-** Straight, **B-** touchdown and **C-** step down cycles. SET III primers were used to amplify 1195 bp product from AV DNA. Touchdown cycles were found to be more efficient.

to 15 ng was amplified with 1 pmole and 2 pmole primers. The ratio of unlabeled to labeled primers used was - F_{UL}: F_L: R_{UL}: R_L = 1:1:0.5:1.5. With SET III primers, a template concentration of 5 ng and 2 pmole primers (**Fig. 4.6**) were finalized for further screening. The rest of the primers were designed as per the nested strategy and the basic optimization parameters remained the same. Therefore, similar standardization was carried with other primer pairs used for both TILLING and Eco-TILLING.

4.2.4 Detection of mutations in *COP1* gene in EMS mutagenized M_2 populations

We screened the three mutagenized populations for identifying mutations in *COP1*. The 16 plates of genomic DNA (8 fold pooled), isolated from the EMS mutagenized tomato M_2 lines of the three populations were screened for the presence of *COP1* mutations in the coiled coil and WD40 region from 3rd to 13th exon. On screening 9216 individuals with 4 sets of primers in the mutagenized populations, 3 putative mutations were identified in population II (120 mM EMS treated). One mutation was identified in well number 51, plate no. 2 with SET III primers. The size of the CELI digested fragments observed were 550 and 645 bp in 700 and 800 channels respectively (**Fig. 4.7**). The combined size of the cleaved fragments was equal to the size of the corresponding main product, indicating the presence of a

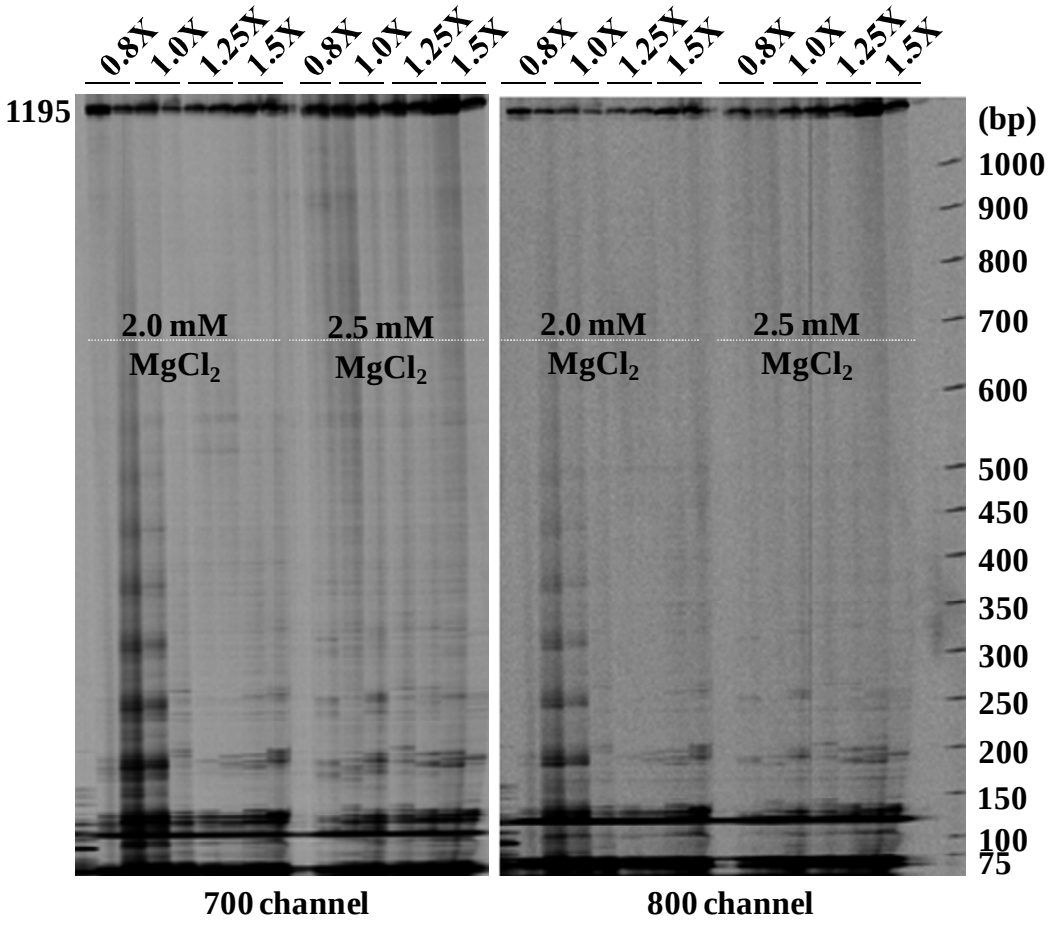


Figure 4.5 Effect of salt and PCR buffer composition on *COP1* amplification Different concentrations of 10X PCR buffer (0.8X, 1.0X, 1.5X) in combination with $MgCl_2$ (2.0 mM and 2.5 mM) were tested with SET III primers to amplify a 1195 bp amplicon from AV DNA. Images of denaturing gels representing both fluorescent channels (700 & 800) were shown. The main product is represented by a bright band at the top of the gel. A combination of 0.8X buffer with 2.5 mM $MgCl_2$ was found to yield optimally. The labeled DNA size marker is indicated on right side.

mutation. Upon screening with SET I primers, two other mutations were detected in well no. 82 in row plate number 6 and well number 55 in the row plate 3. These positive pools were de-convoluted to detect and identify the line number harboring the mutation (using 2D pooled column plates). Out of the three mutations, only two could be confirmed by screening the complementary plates (**Fig. 4.8 and 4.9**). The line number that was common in both row and column plate was considered to bear mutation. Based on the 2-D pooling, the mutations were present in the line numbers AV 120 mM 1894 and AV 120 mM 2346.

4.2.5 Confirmation and identification of the mutated plant

Unlike animals, the germ line specification in plants is not predetermined and also occurs in the later developmental stages (Wilson and Yang, 2004). As a result the seeds (heritable material) are multicellular structures. EMS mutagenesis of seeds causes random changes in the genome leading to different mutations in different cells of the same seed (Henikoff and Comai, 2003). Therefore, the M_1 plants would be chimeric and the segregating M_2 population, where the somatic mutations disappear and germ line mutations get fixed is used for mutant screening. High throughput screening makes it easier to identify the mutations. Once a mutation is observed, the individual line harboring the mutation must be identified. Ten seeds from each of line AV 120 mM 1894 and AV 120 mM 2346 were grown for further analyses. The DNA from these lines was isolated, amplified with gene specific primers followed by CELI digestion and subjected to gel based analysis to detect the cleaved products. Out of the two mutations detected, we could identify and confirm only one mutation in AV 120 mM 1894. We screened approximately 25 M_2 seeds from the line number AV 120 mM 2346 but we could not identify the mutated plant. Due to limited seeds in both M_2 and M_3 stocks of AV 120 mM 2346, we could not extend the screen size for this particular line. The mutation frequency for the gene of interest in the population studied was calculated using the formula:

$$\text{Mutation frequency} = \frac{\text{total number of mutations identified}}{\text{No. of individuals screened} * (\text{No. of base pairs screened} - 50)}$$

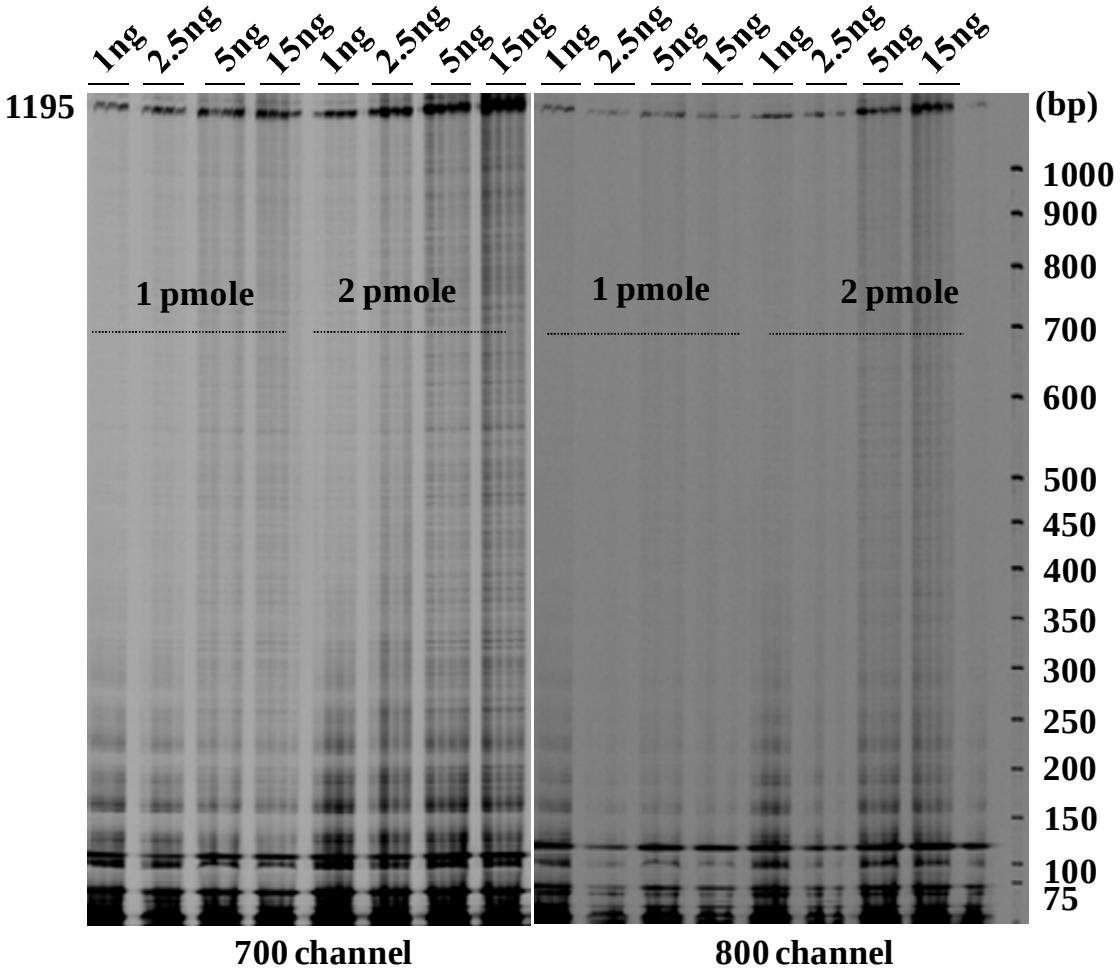


Figure 4.6 Effect of primer and template concentrations on PCR for TILLING and detection on Li-COR. Different concentrations of template DNA (1-15 ng) with varying labeled primer concentration (2- 4 pmoles) were tested. Images of LiCOR gel were shown for both fluorescent channels (IRD 700 & IRD 800). The labeled DNA size marker is marked on right side.

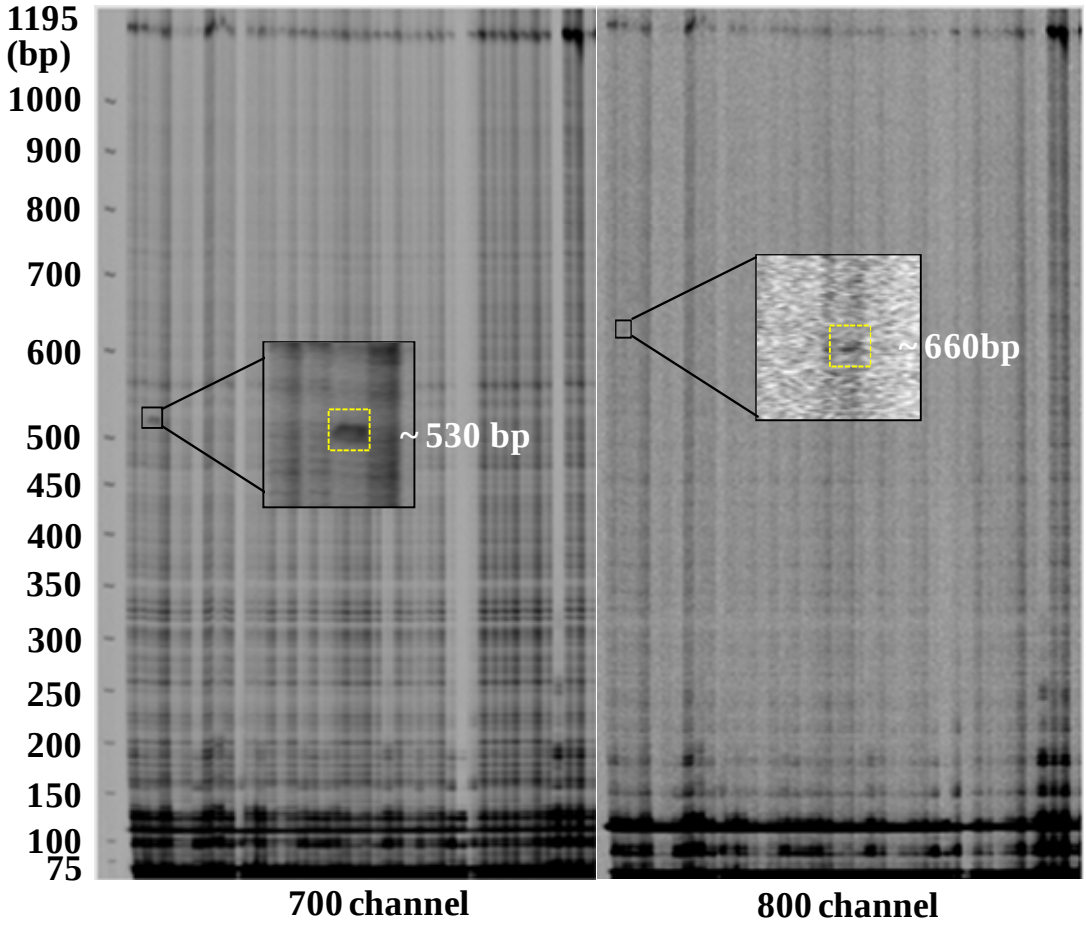


Figure 4.7 Detection of mutations in population II with SET III primers of *COP1* gene. Mutations were scored based on the complementary fragments observed on the gel. Enzymatic cleavage of 1195 bp product, amplified by SET III primers of *COP1* gene was detected in both the IRD700 and IRD800 channels. The unique digested fragments of approx. 530 bp and 660 bp were marked by black open boxes in both the IRD700 and IRD800 channels respectively. The cleaved fragments were highlighted in an inset and marked by yellow open boxes. The labelled molecular size marker is loaded on the left side of the gel.

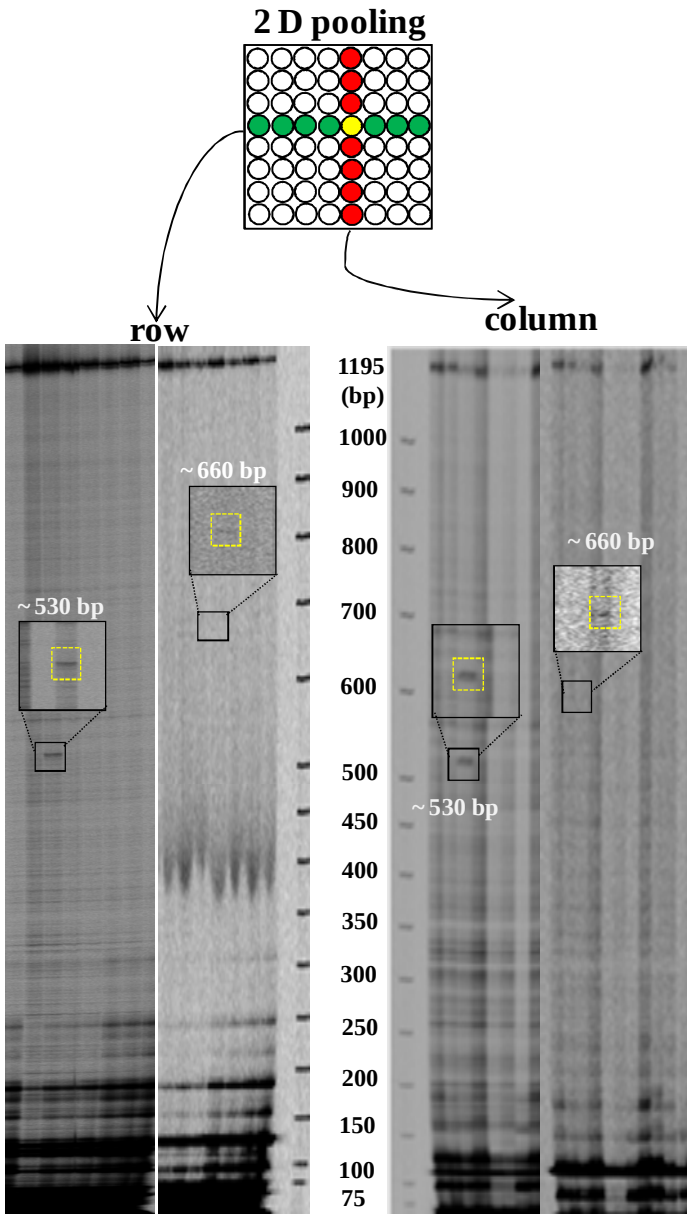


Figure 4.8 Confirmation of the putative mutations in the 2-D pooled row and column plates with SET III primers. With reference to the mutations detected in population II, the corresponding complementary plates were screened with SET III primers respectively to de-convolute the pool containing mutation. A portion of the Li-COR image showing the digested complementary products are highlighted in boxes and the inset shows the enlarged picture with the approximate sizes (in bp).

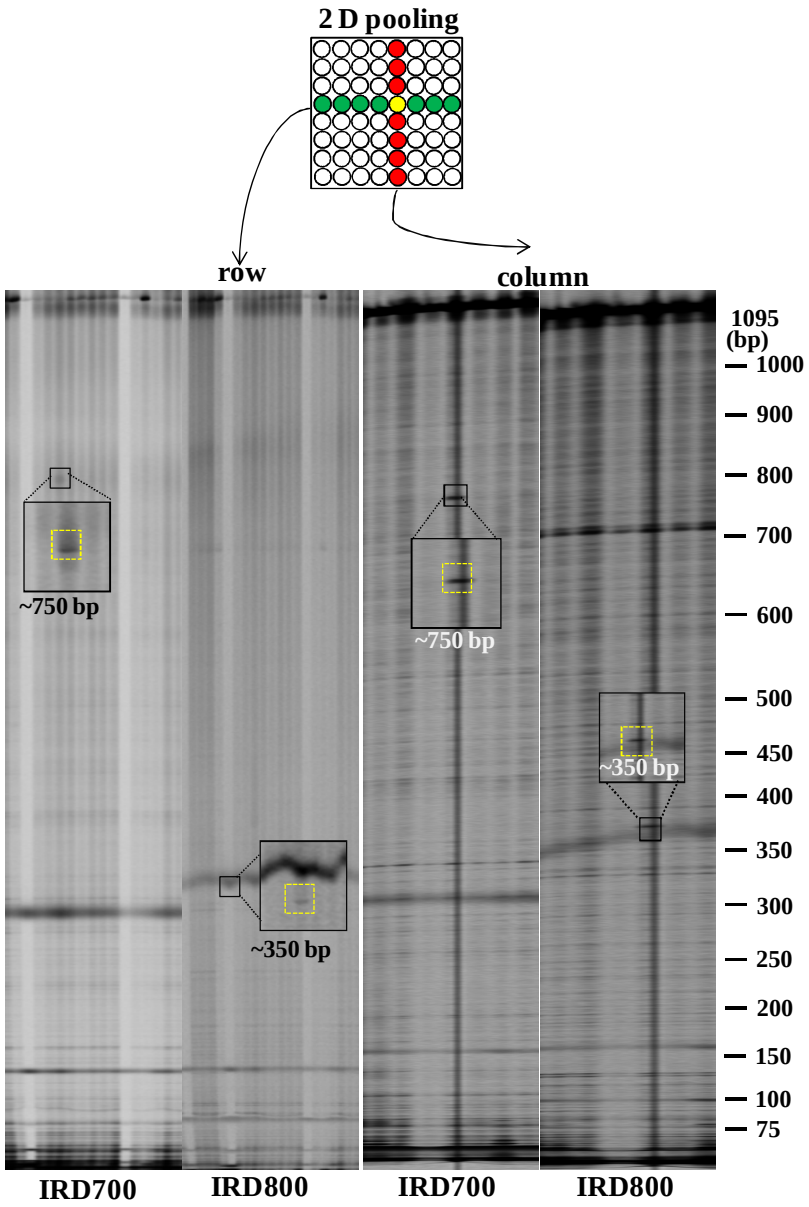


Figure 4.9 Confirmation of the putative mutations in the 2-D pooled row and column plates with SET I primers. With reference to the mutations identified in population II, the corresponding complementary plates were screened with SET I primers respectively to de-convolute the pool containing mutation. A portion of the Li-COR image showing the digested complementary products of 1095 bp amplicon are highlighted in boxes and the inset shows the enlarged picture with the approximate sizes (in bp). The labeled molecular size marker is marked on the right side of the gel.

Upon screening approx. 10,000 mutated lines, only 3 mutations were obtained for *COP1* gene in tomato. The mutation frequency for *COP1* gene calculated based on our data was very low i.e. $6.54x10^{-8}$, a frequency very close to the spontaneous mutation rate ($1x10^{-8}$).

4.2.6 Genotypic and phenotypic characterization of AV M_2 120 mM 1894

To examine the presence of mutation at the individual plant level, twenty five seeds from the M_2 seed packet for the mutant line were sown. After surface sterilization, the seeds were germinated in dark on wet tissue paper. Once the radicle emerged, they were transferred to coconut peat and seedlings were grown under white fluorescent light (100 μmol $m^{-2}sec^{-1}$) with a 16/8 hr light cycle in growth room at 22 ± 2°C. Out of 25 seeds, only 14 germinated. The seedlings were designated as P1, P2…P14. A variation in the growth of putative mutant seedlings was observed when compared with AV. Two of the seedlings died at an early stage. A clear phenotypic difference was observed in the height of one month-old seedlings: five of them were very short, three of them were intermediate in length and the rest were similar to AV (**Fig. 4.10**). Moreover, the phenotype of 3-day old dark grown seedlings of AV M_2 1894 was similar to that of AV. They characteristically lacked the *'cop'* phenotype (**Fig. 4.10**). Nine of the lines (P1, P2, P5, P6, P7, P8, P9, P10 and P11) that survived were then transferred to green house and were kept under observation. DNA was isolated from these plants and a second round of screening was done to identify the plant which is actually harboring the mutation. AV DNA was used as a reference. Labeled PCR was performed with the putative mutant DNA alone and that mixed with AV DNA followed by digestion with CELI enzyme. The products were loaded on Li-COR gel to observe the excised fragments (**Fig. 4.11**). The complementary cleaved products were observed only in four samples. In order to confirm the hetero/homozygosity, the four samples were reloaded on Li-COR gel. Out of 9 plants, two of the plants, P1 and P3 were found to be homozygous for the mutation as the excised products were observed only when mixed with AV DNA and not with individual DNA. P5 and P6 plants were heterozygous as the excised products were observed in both mixed and individual samples after CELI digestion (**Fig. 4.11**) and there were no specific fragments observed in AV.

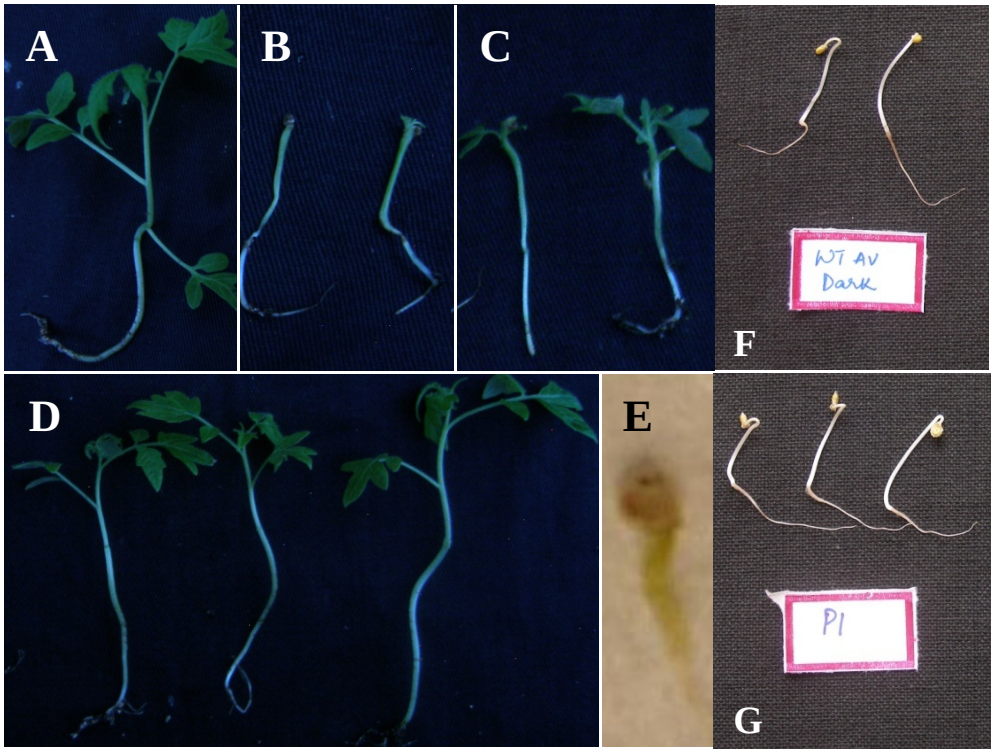


Figure 4.10 Phenotype of the AV and the 120 mM AV M_2 individuals. The phenotype of the mutant line at early stages of development (1 month old), A, Wild type AV; B, represents a dwarf phenotype; C, medium length; D, phenotype that was comparable with wild type; E, seedlings that died. F and G represent phenotype of 3 day old dark grown seedlings of WT and P1 (homozygous for mutation in exon 5,6 and 7 region) respectively.

In order to identify the molecular nature of the mutation, the PCR product amplified from putative homozygous mutant plants, P1 and P3 was column purified using Bioserve kit and sequenced. The sequence alignment of mutant with AV showed a single base change from C6964 to T on the genomic co-ordinates (**Fig. 4.12A**). The sequence was analyzed using web based tool PARSESNP (http://www.proweb.org/parsesnp) (Taylor & Greene 2003). This software predicts the mutation type and also gives a SIFT (Sorting Intolerant from Tolerant) score. The PARSESNP output resulted in a change from C6964 to T that corresponds to the 5th exon (WD40 domain) of *COP1* (**Fig. 4.12B**). The altered codon from CUA (CTA in gene sequence) to UUG (TTA) also codes for Leucine in tomato, thereby causing no change in the amino acid. The three month old plants showed an observable phenotypic difference that may be correlated with the zygosity of the mutant plants identified (**Fig. 4.13**). The homozygous plants, P1 and P3 showed delayed growth and development, whereas the heterozygous plants, P5 and P6 grew normally as AV. Only single fruit was obtained from P1 that was advanced to the next generation, hence, we could not quantify the carotenoid content in the fruit in this generation. Since the M_2 seeds were obtained from chimeric M_1, it is expected that the segregation of the mutation in M_2 is random and the plants have to be scored for any phenotypic differences in the next generation.

4.2.7 Phenotypic segregation of the mutant in M_4 generation

The M_3 lines homozygous for the mutations were allowed to self fertilize to obtain the M_4 seeds. The M_4 generation was raised to follow the segregation pattern of the mutation and also to study the phenotype. The M_4 seeds from M_3 P1 plant were further considered to examine the genotype to phenotype correlation. After germination, the seedlings were transferred to coconut peat till the arrival of first leaves and were acclimatized in greenhouse. Later they were transferred into pots in greenhouse. The M_4 seeds from P1 segregated into two phenotypic classes – short statured plants (P1-S) and normal plants (P1-L) which resemble AV (**Fig. 4.14**) and the phenotypes of P1 were monitored. The short plants bore only 5 internodes after which the internode differentiation was not clear and the inflorescence was observed. On the other hand the shoot architecture of the P1-L plant was similar to AV. They showed proper internode differentiation and the inflorescence appeared on the 8th node.

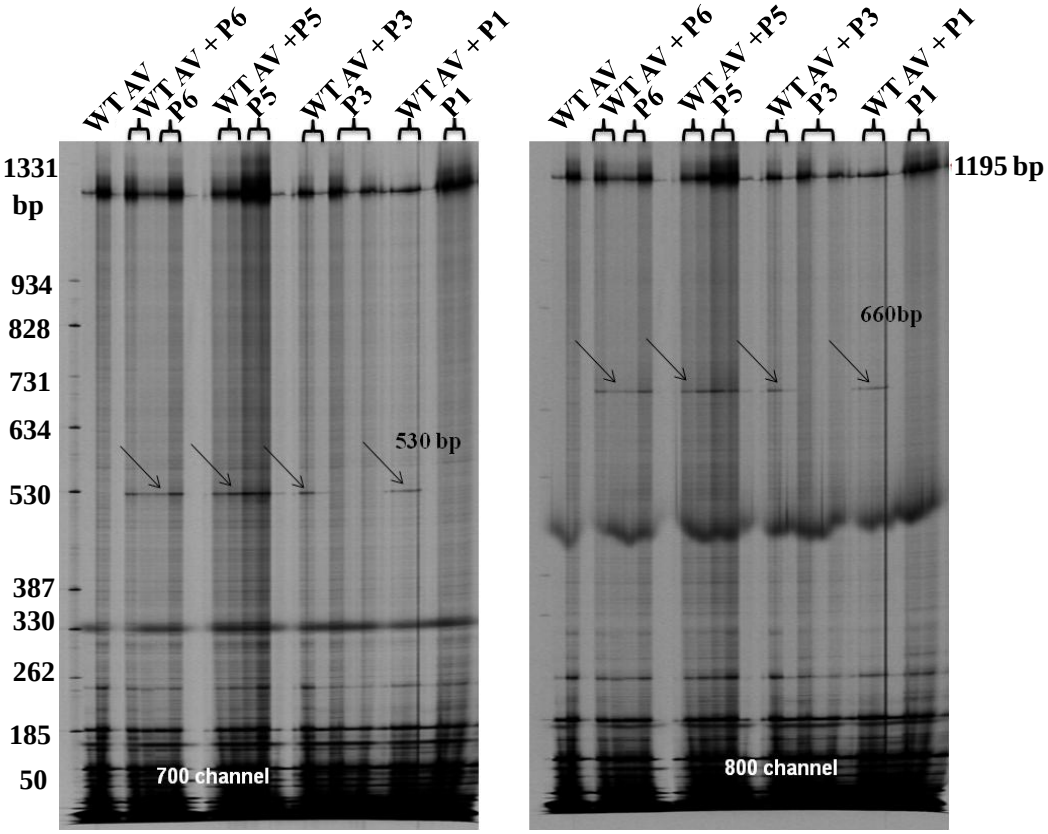
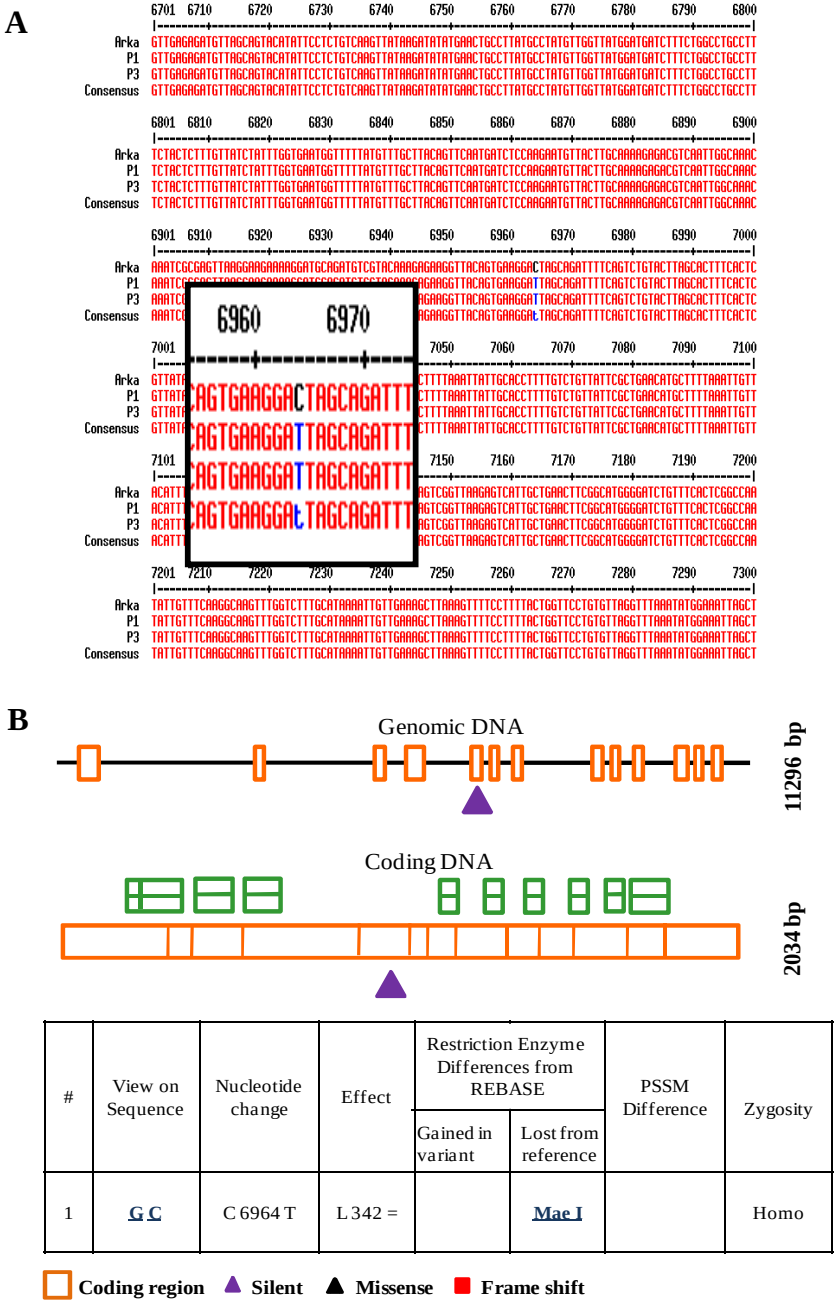


Figure 4.11 Confirmation of mutation in individual plants of population II-1894 line. The 1195 bp amplicon size is represented by thick bands at the top of the gel and the digested products by black arrows in both 700 channel (530 bp) & 800 channel (660 bp). P1 and P3 are homozygous for the mutation identified while P5, P6 are heterozygous. AV was used as a control. The DNA size strand marker is marked on the left side of the gel.



#	View on Sequence	Nucleotide change	Effect	Restriction Enzyme Differences from REBASE		PSSM Difference	Zygosity
				Gained in variant	Lost from reference		
1	G C	C 6964 T	L 342 =		Mae I		Homo

Coding region ▲ Silent ▲ Missense ■ Frame shift

Figure 4.12 A, **The sequence alignment of homozygous mutant with AV** reveals a single base change from C 6964 to T on the genomic co-ordinates of tomato *COP1* gene sequence. **B**, Schematic representation of the mutation identified in *COP1* gene. The figure was designed from PARSESNP output files (Taylor & Greene 2003). The mutation is denoted both on genomic sequence (11296 bp) and coding sequence (2034 bp). The table below shows the position of nucleotide change C6964T on genomic sequence, the effect on amino acid in protein is also given. An L342 = symbolizes that there is no change in amino acid at 342 position in COP1 protein.

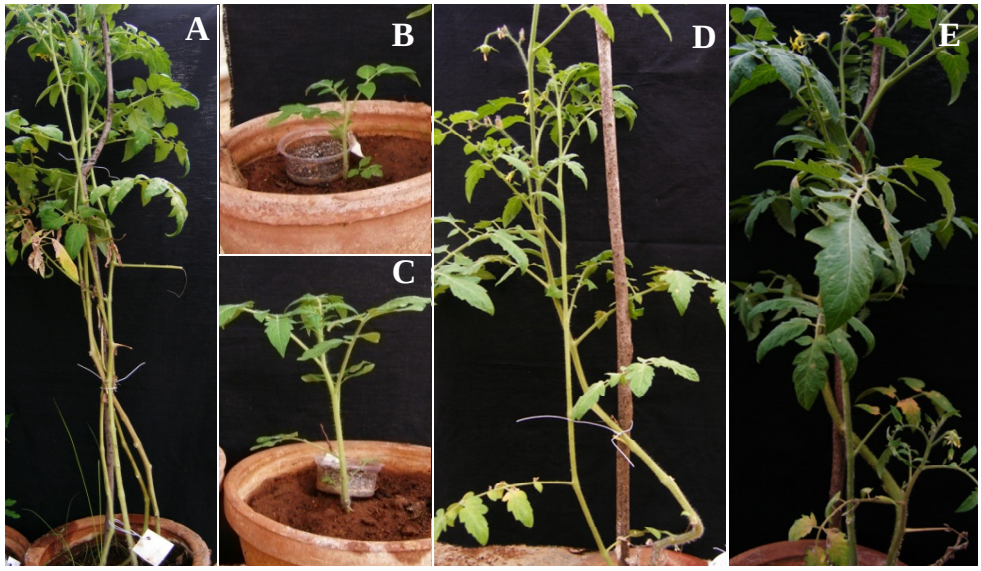


Figure 4.13 Phenotype of adult plants of 120 mM AV 1894. The homozygous P1 and P3 individuals of the mutant line are short and slow growing. The heterozygous P5 and P6 individuals are similar to wild type. A, Wild Type (AV), B & C, top view of the homozygous plants, D, the homozygous P1 on left side and P3 on right side, E, the heterozygous P5 and F, the heterozygous P6 plant.

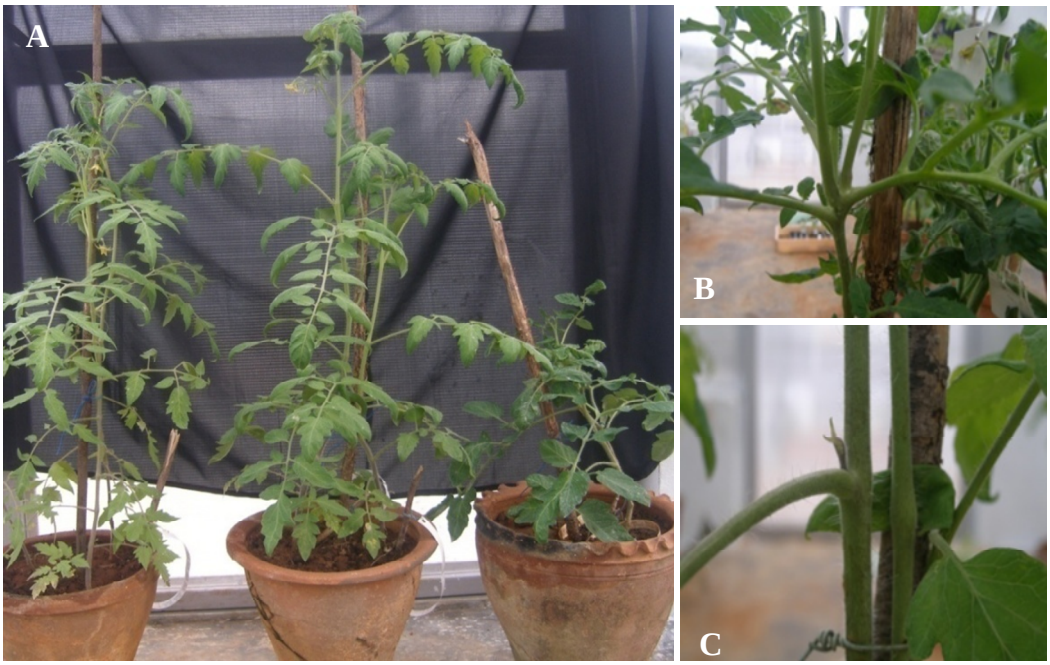


Figure 4.14 Phenotypic segregation of the 120 mM AV 1894 homozygous individual P1 in M_4 generation. The progeny of homozygous P1 plant segregated as short and long plants. A-shows the comparison of height of the segregating M_4 with WT. B & C show the architectural differences of the internode at the 5^{th} node in P1-S and P1-L respectively.

4.2.8 Carotenoid content is enhanced in AV 120 mM 1894

The carotenoid content of the red ripe fruits of the mutant P1-S and P1-L was quantified by HPLC. We estimated the levels of lycopene and β-carotene, as these compounds make up the bulk of the total carotenoid content at red ripe stage. In spite of phenotypic differences in the mutant plants, we observed almost similar levels of lycopene in both P1-S and P1-L fruits (**Fig. 4.15A**) with almost 50% increase in the lycopene content compared to AV. However, β-carotene was found to be approximately four fold high in fruits of P1-L than AV, and in P1-S fruits the level was almost similar to that in AV with no significant difference (**Fig. 4.15B**). Based on the observed phenotypes, we expected a segregation of the mutation as the phenotype of M_3 was neither uniform nor followed a monogenic ratio. It is also plausible that any secondary mutations might have caused the above phenotypic differences.

4.2.9 Silent mutations in tomato *COP1* cause the altered phenotype

As both the segregating homozygous P1-S and P1-L plants of AV 120 mM 1894 line exhibited enhanced levels of carotenoids than AV, we tried to investigate the actual genetic cause. Initially, we confirmed the genotype of these plants, i.e. the presence and position of the mutation in the *COP1* gene. DNA from the M_4 plants was isolated and a labeled PCR was performed, followed by mismatch cleavage analysis and the point mutation was detected on Li-COR system. To our surprise, the position of the original silent mutation in the exon 5 was re-confirmed in both P1-S and P1-L plants without any genotypic segregation (**Fig. 4.16**) i.e. the two plants, P1-S and P1-L harbored the same mutation. Therefore, we ruled out the possibility of segregation of mutation and considered the fact that random mutagenesis by EMS would have caused any secondary mutations in *COP1* gene or in any other gene. In order to rule out this possibility, we sequenced the entire *COP1* gene from P1 plants (both P1-S and P1-L) and AV for identifying any secondary mutations within the gene (**Fig. 4.17 to 20**). Interestingly, there were no secondary mutations in the entire gene except for another base change from T126C in exon 1 of the P1-S (**Fig. 4.21**). PARSESNP analysis of this base change revealed another synonymous change, i.e., there was a change in the codon from ACT to ACC; however the amino acid remained the same i.e., threonine. Though it was presumed that a change in the wobble base would not affect the amino acid and hence the protein function, the

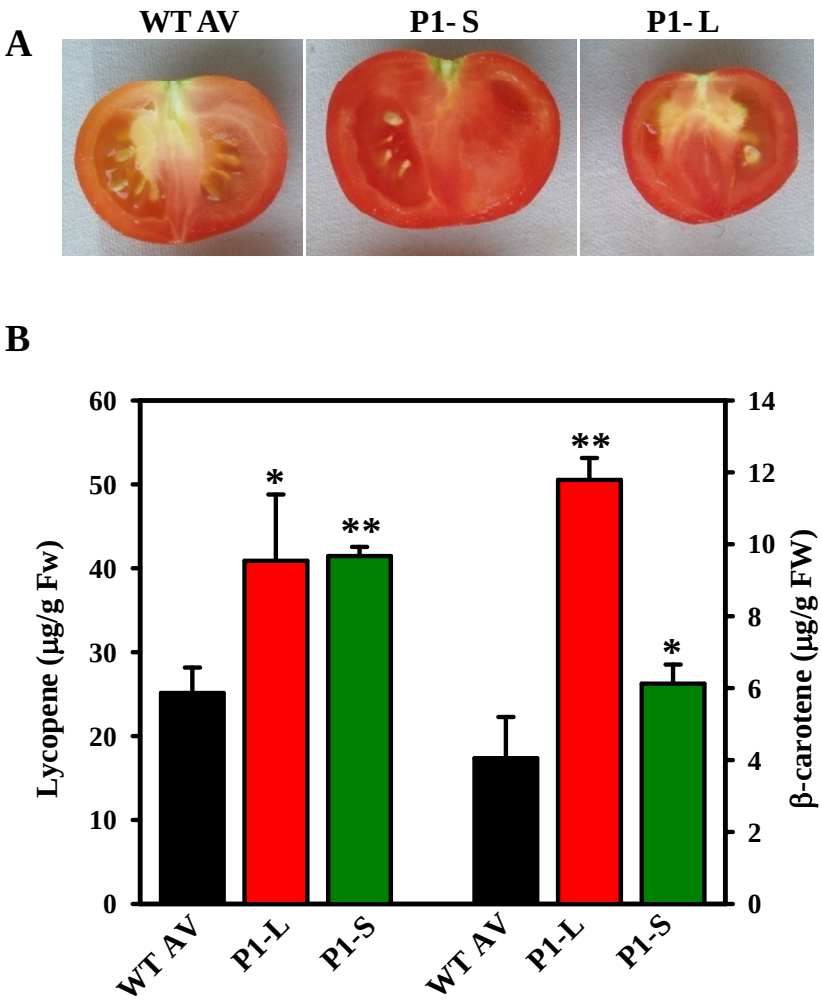


Figure 4.15 The fruit phenotype of WT and the 120 mM AV 1894 M_4 individuals. A, The fruit section of the homozygous P1 short and long individuals with WT. B, Estimation of lycopene and β- carotene content in these fruits by HPLC. $n \geq 3 \pm SD$, where n= no. of replicates $p \leq 0.1$ (*) and $p \leq 0.05$ (**) were considered for the analysis.

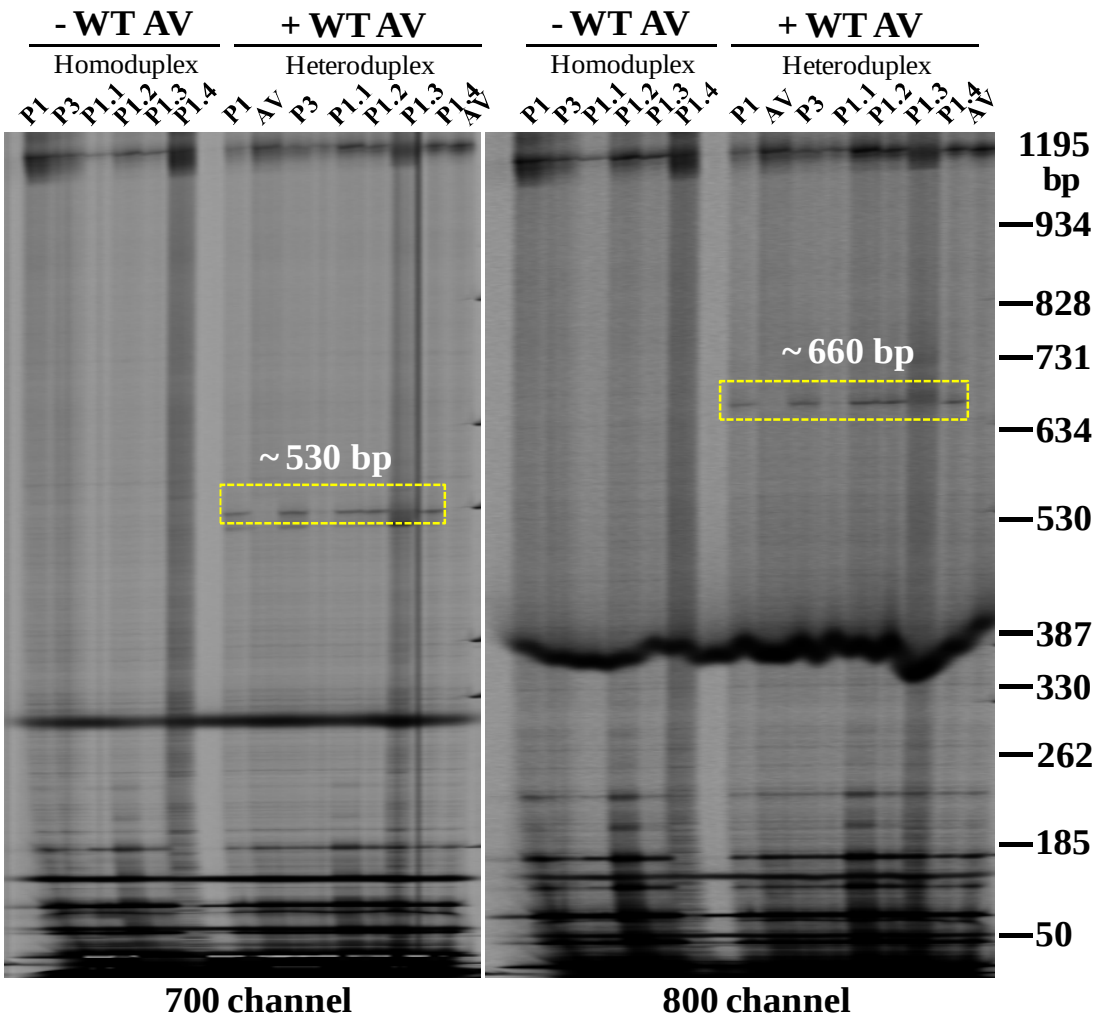


Figure 4.16 Genotypic confirmation of the silent mutation in 1894 M_4 individuals. The individual DNA isolated from M_4 individuals was amplified using SET III primers. The 1195 bp amplicon was subjected to enzymatic cleavage and the digested fragments were resolved on Li-COR gel. No complementary products were observed in homo duplex samples. The yellow dashed line highlights the cleaved products obtained in the heteroduplex samples. The DNA ladder is marked on extreme right of the 700 channel.

Exon 2 (3107 bp – 3184 bp)

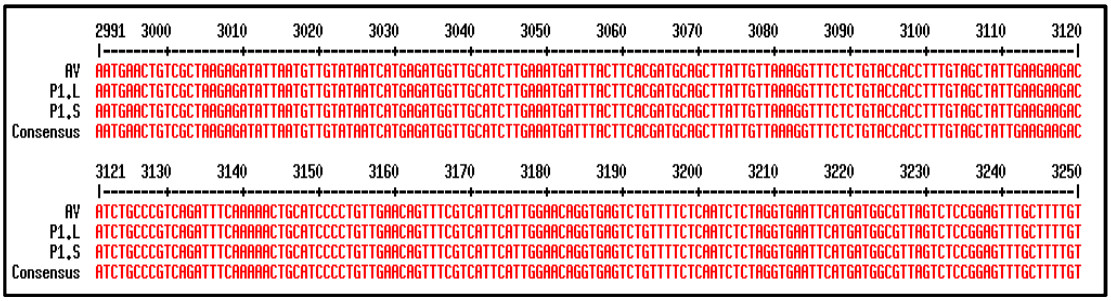


Figure 4.17 Multiple alignment of the exon 2 region amplified by SET V primers from the DNA of WT, P1-Short (P1-S) and P1-Long (P1-L).

Exon 3 and 4 (4958 bp – 6187 bp)

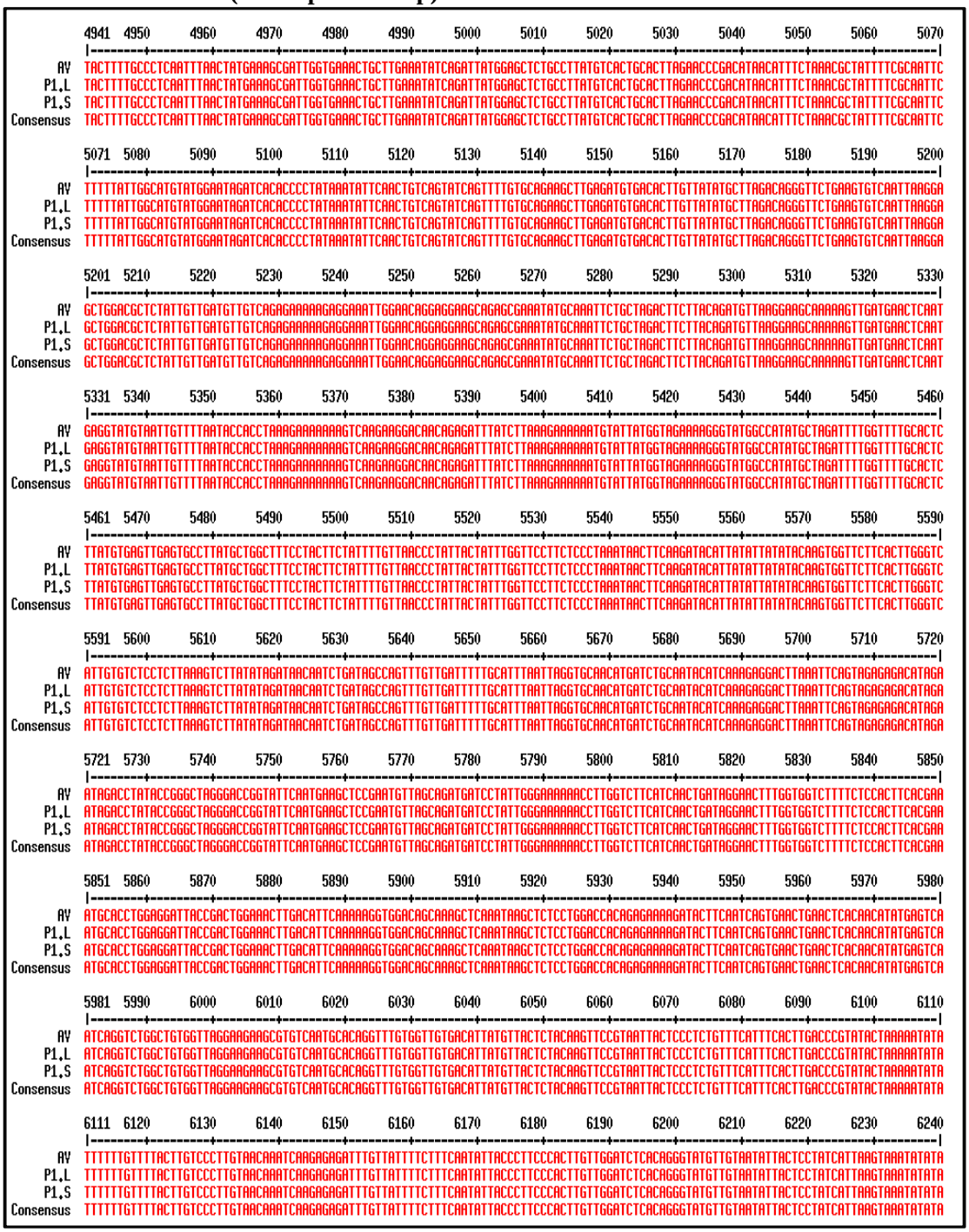


Figure 4.18 Multiple alignment of the exons 3 and 4 region amplified by SET IV primers from the DNA of WT, P1-Short (P1-S) and P1-Long (P1-L).

Exon 8, 9 and 10 (8346 bp – 9827 bp)

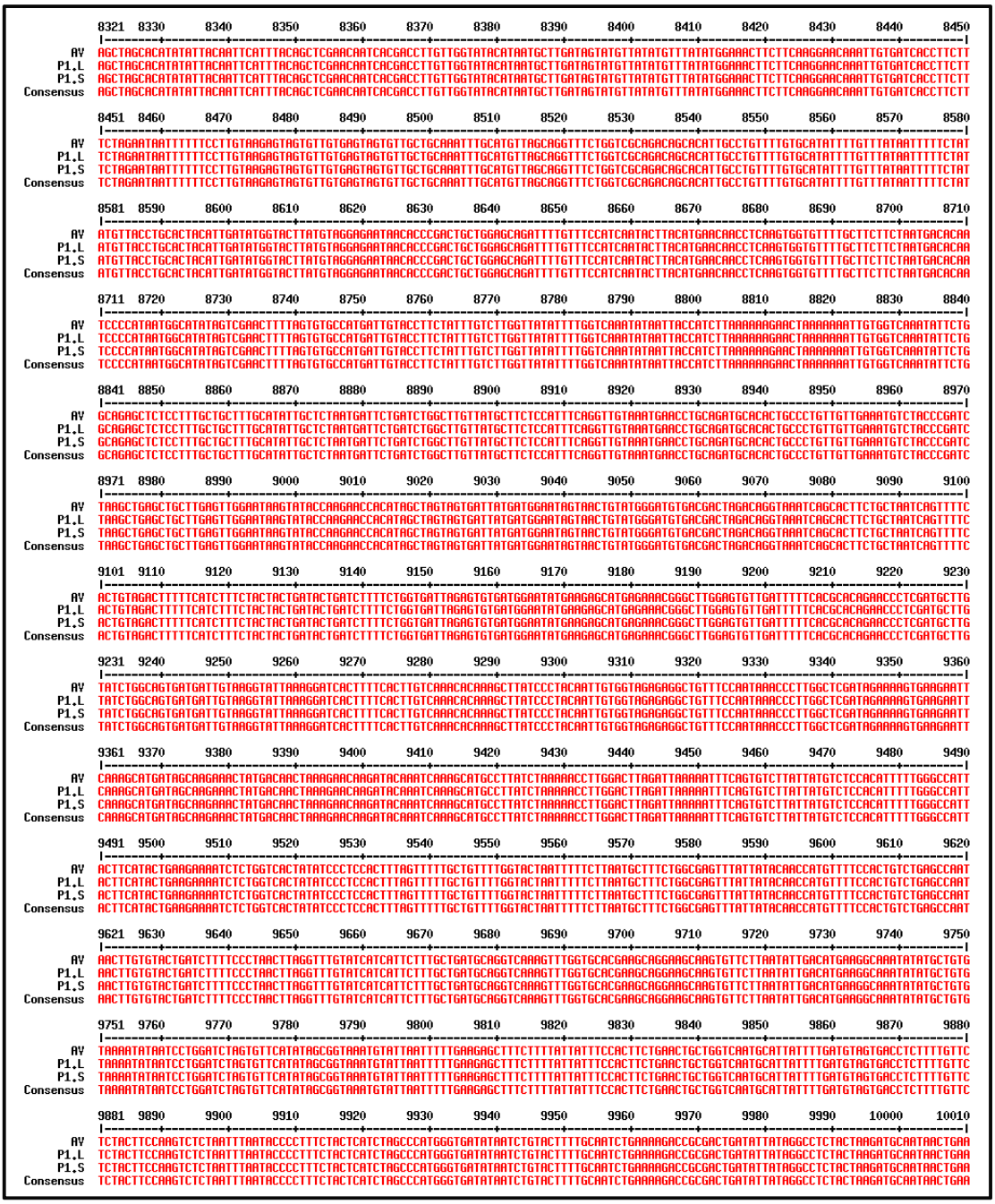


Figure 4.19 Multiple alignment of the Exons 8, 9 and 10 region amplified by SET II primers from the DNA of WT, P1-Short (P1-S) and P1-Long (P1-L).

Exon 11, 12 and 13 (10284 bp –11355 bp)

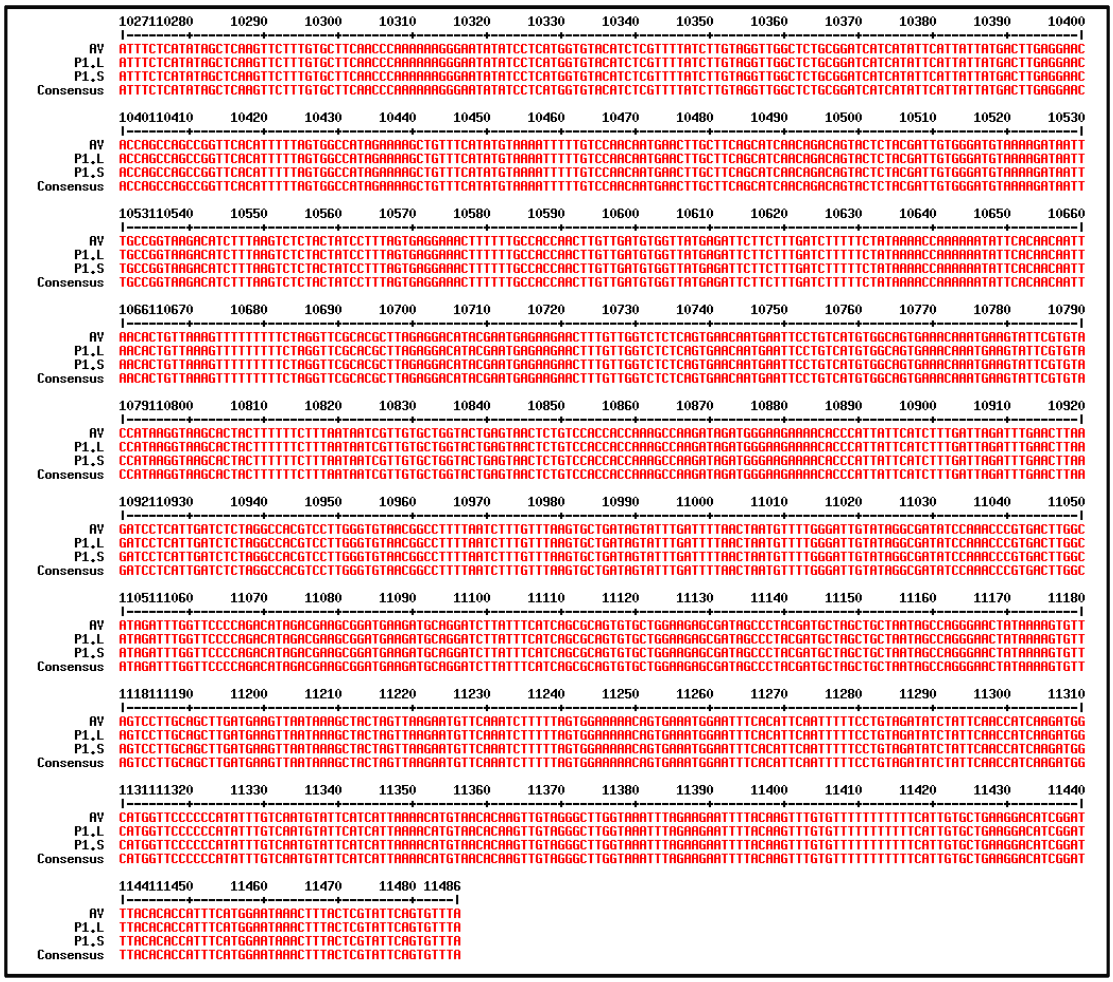
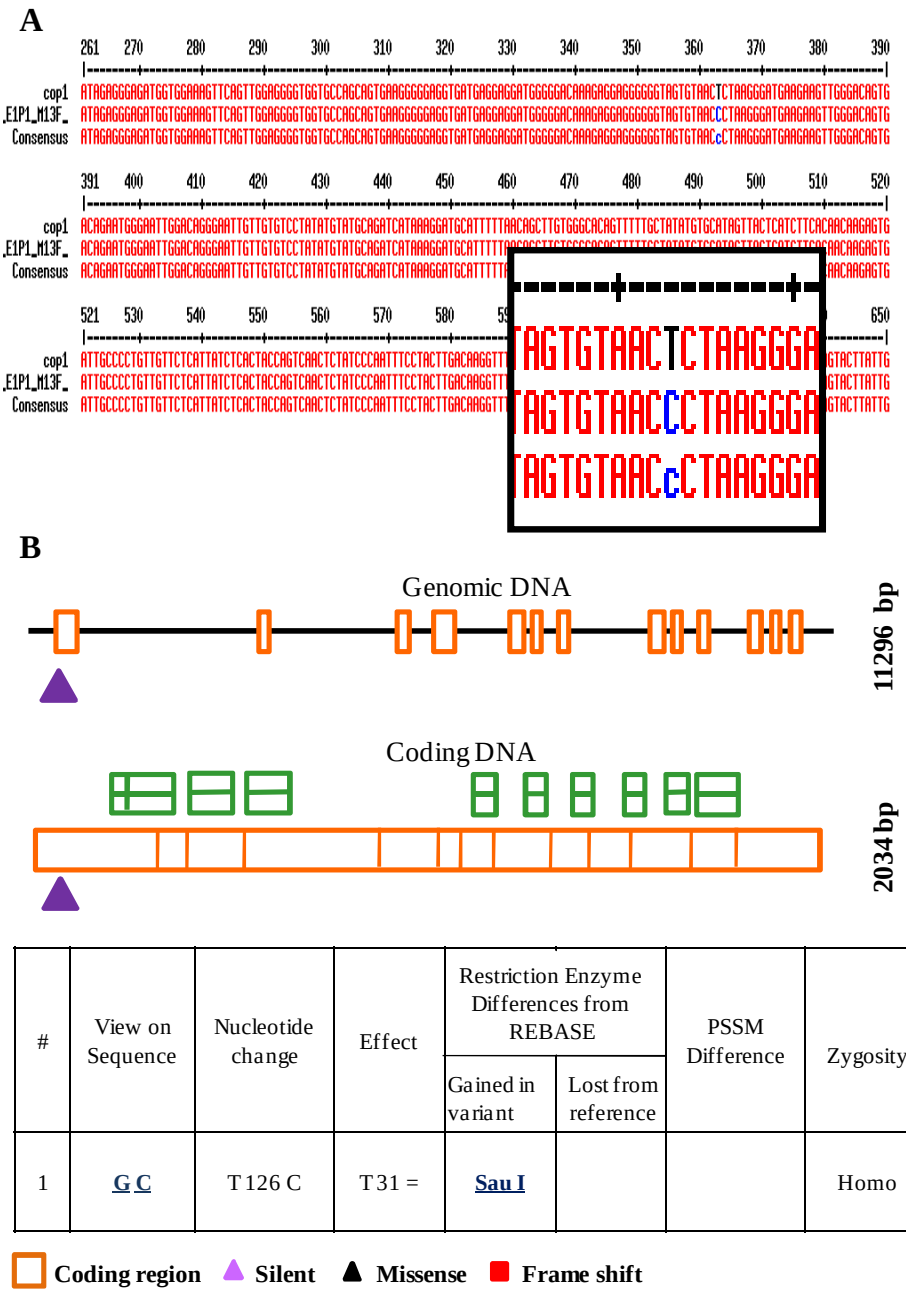


Figure 4.20 Multiple alignment of the Exons 11, 12 and 13 region amplified by SET II primers from the DNA of WT, P1-Short (P1-S) and P1-Long (P1-L).



#	View on Sequence	Nucleotide change	Effect	Restriction Enzyme Differences from REBASE		PSSM Difference	Zygosity
				Gained in variant	Lost from reference		
1	G C	T 126 C	T 31 =	Sau I			Homo

Figure 4.21 A, **The sequence alignment of homozygous mutant with WT** reveals a single base change from T126 to C on the genomic co-ordinates of tomato *COP1* gene sequence. **B**, PARSESNP output files (Taylor & Greene 2003) denoting the mutation both on genomic sequence (11296 bp) and coding sequence (2034 bp). The table below shows the position of nucleotide change T126C on genomic sequence, the effect on amino acid in protein is also given. An amino acid T31 = symbolizes that there is no change in amino acid at 31 position in COP1 protein**.**

accumulating evidence on synonymous changes and their effects on protein function tempted us to ask the next question - does the silent change affects the phenotype? If so what could be the possible mechanism? While there are no detailed studies on synonymous changes affecting the phenotype in plant system, in animal system, several studies provide evidence for the affects of synonymous changes. As reviewed in Hunt et al., (2014), synonymous changes can affect the gene function by three major mechanisms- altering the cis regulatory elements like splice sites, miRNA and exonic transcription factor binding sites; modifying the mRNA structure, thereby affecting the turnover of mRNA or by lowering the translational speed, thereby altering the protein abundance. In order to answer the above question, we made an attempt to check the possibilities sequentially.

Firstly, we tried to identify the presence of any miRNA binding sites in our target sequence of *COP1,* harboring the point mutations. From the study on miRNA targets in tomato fruit development by Karlova et al., (2013), we could find seven miRNAs for *COP1* gene. The corresponding SGN-Unigene sequences which were considered as the targets of the above mentioned miRNA are listed in the **Table 4.1**. Only one Unigene- SGN-U579490 was associated with *COP1* locus in the SGN data base. Considering the fact that miRNAs have multiple target sites, we checked the sequence similarity of all the listed miRNAs with tomato *COP1* gene. The probable target sites and the corresponding exon –intron position has been mentioned in **Table 4.1**. Of the seven miRNAs listed, only one (miRNA 5303) corresponded to the exon 5, 6 and 7 region of *COP1* gene. Also, the site for miRNA 5303 is present few bases (617 bases) ahead of the site of mutation. We do not know the exact mechanism by which these miRNA regulate the levels of *COP1* and further detailed study has to be taken up for a better understanding for their role.

We crosschecked the next possibility, and examined the codon usage bias for the amino acids leucine and threonine in tomato. From the codon usage table for *S. lycopersicum* (http://kazusa.org.in), we examined the corresponding codon usage frequencies for leucine and threonine (**Table 4.2)**. It is very clear from this table that both the synonymous mutations i.e. leucine at 345 and threonine at 31 of COP1 protein led to non-preferred codons in the mutant line as the frequency usage of both codons was 50% lower than the wild type codons. The codon CTA is the most frequent (24%) codon for leucine and TTA is the least preferred one (with a frequency

Table 4.1 Probable miRNA sequences detected for *COP1* gene in tomato as taken from Karlova et al., (2013).

miRNA	Tomato Unigene	Sequence of the miRNA	Target Site in COP1 genomic DNA	Corresponding exon/intron
miR396	SGN-U564425	CAATACAATAAAGCTGTGGGA	10496-10516	E11
miR414	SGN-U579490	GGGAGGTGATGAGGAGGATGG	77-97	E1
miR5303	SGN-U564425	GATGGTGAATTTTTTAAAAC	7581-7606	E7
miR6023	SGN-U579490	ATTTACACACCATTTCATGGAA	11439-11460	E13
miR167	SGN-U585184	AAGGTCGTGCTGGCACTTCA	10823-10843	I 12
miR6026	SGN-U564423	GCAATAGAACTTAGAGAAGAA	4393-4412	I 2
miR472	SGN-U576848	GGTATGGGCGGAGTAGGAAAAA	21-42	E1

Table 4.2 Silent mutations in AV 1894 mutant line with the altered codons and their relative frequencies highlighted in red.

***Solanum lycopersicum* [gbpln]: 1452 CDS's (634390 codons)**
fields: [triplet] [amino acid] [fraction] [frequency: **per thousand**] ([number])

UUU	F	0.6	26	16504	UCU	S	0.26	21.2	13464	UAU	Y	0.6	18.6	11799	UGU	C	0.62	10.8	6883
UUC	F	0.4	17.5	11123	UCC	S	0.12	9.9	6268	UAC	Y	0.4	12.4	7855	UGC	C	0.38	6.7	4223
UUA	L	0.15	14.4	9111	UCA	S	0.25	20.7	13102	UAA	*	0.4	0.9	598	UGA	*	0.37	0.9	553
UUG	L	0.25	24.2	15336	UCG	S	0.07	5.6	3521	UAG	*	0.23	0.5	340	UGG	W	13.5	13.5	8563
CUU	L	0.26	24.9	15814	CCU	P	0.39	19.2	12192	CAU	H	0.67	15.5	9827	CGU	R	6.9	6.9	4404
CUC	L	0.12	11.2	7127	CCC	P	0.12	5.7	3645	CAC	H	0.33	7.8	4926	CGC	R	3.1	3.1	1946
CUA	L	0.11	10	6348	CCA	P	0.39	19.2	12205	CAA	Q	0.6	21	13315	CGA	R	5.4	5.4	3414
CUG	L	0.11	10.5	6665	CCG	P	0.09	4.6	2920	CAG	Q	0.4	14	8861	CGG	R	3.1	3.1	1981
AUU	I	0.5	28.2	17914	ACU	T	0.39	19.9	12612	AAU	N	0.64	30.5	19339	AGU	S	15.2	15.2	9669
AUC	I	0.25	14	8906	ACC	T	0.17	8.6	5464	AAC	N	0.36	17.3	11002	AGC	S	9.3	9.3	5897
AUA	I	0.25	14	8893	ACA	T	0.35	17.9	11333	AAA	K	0.5	31	19724	AGA	R	16.4	16.4	10398
AUG	M	1	24.7	15651	ACG	T	0.09	4.6	2939	AAG	K	0.5	31	19651	AGG	R	11.9	11.9	7541
GUU	V	0.43	28	17762	GCU	A	0.45	30.7	19474	GAU	D	0.72	39.3	24917	GGU	G	23.9	23.9	15192
GUC	V	0.15	10.1	6382	GCC	A	0.15	10.1	6426	GAC	D	0.28	15	9509	GGC	G	9.7	9.7	6265
GUA	V	0.17	11.2	7126	GCA	A	0.33	22.2	14107	GAA	E	0.57	34.8	22053	GGA	G	25.6	25.6	16258
GUG	V	0.25	16	10181	GCG	A	0.08	5.2	3327	GAG	E	0.43	26.6	16906	GGG	G	10.8	10.8	6839

Coding GC 42.52% 1st letter GC 50.16% 2nd letter GC 39.87% 3rd letter GC 37.53%

of 14.4%). The second site of silent mutation also conferred a reduction in the codon usage frequency as the codon ACT has 19.9% usage frequency and on the other hand, ACC is the least preferred codon (8.6%). As evident from the literature, change of codon affects the gene function by regulating the mRNA secondary structural stability and transcript levels. These in turn may affect the protein abundance, structure and therefore function.

4.2.10 Silent mutation affects *COP1* transcript level

Our analysis of miRNA targets and codon usage bias were predictive. Assuming that either of these mechanisms could have caused the observed phenotypes due to *COP1* mutation, we examined the transcript levels of *COP1* in the mutant lines, P1-S and P1-L compared to AV. We isolated total RNA from young leaves and red ripe fruit tissue, synthesized cDNA and then used it for real time PCR analysis of the transcripts- *COP1*, *COP1LIKE*, *SPA1* and *HY5*. The Ct values of these transcripts were normalized with the Ct values of internal control genes *β-actin* and *ubiquitin 3*. **Figure 4.22** represents the relative expression of transcripts in leaf tissue and **Figure 4.23** relates to red ripe tissue. As expected, the levels of *COP1* were significantly reduced in the leaf (**Fig. 4.22**) and red ripe fruit (**Fig. 4.23**) of both P1-S and P1-L plants. Consistent with this, in the leaf the *SPA1* levels were reduced and *HY5* levels were elevated in both P1-S and P1-L. The levels of *SPA1* and *HY5* were also reduced in the red ripe fruit tissue of P1-S with an exception of P1-L. This could be due to the additive effect of the two silent mutations in P1-S and not in P1-L.

As we observed an enhanced level of lycopene, we made an attempt to understand the role of *COP1* on carotenoid pathway. Therefore, we checked the transcript levels of *PSY1* and *CYCB* in the red ripe fruit tissue. Interestingly, the transcript levels of *PSY1* increased with a reduction in *COP1* levels. Also, the transcript data could be correlated to the metabolite accumulation in ripe fruits of P1-S and P1-L suggesting that the phenotype might be caused due to a repression in the levels of *COP1* which might have resulted due to the silent changes.

4.2.11 Analysis of COP1 protein in tomato fruit

The weak alleles of Arabidopsis *cop1* mutants exhibited strong phenotypes at the seedling level due to the loss of protein function at the WD40 domain (Deng et al.,

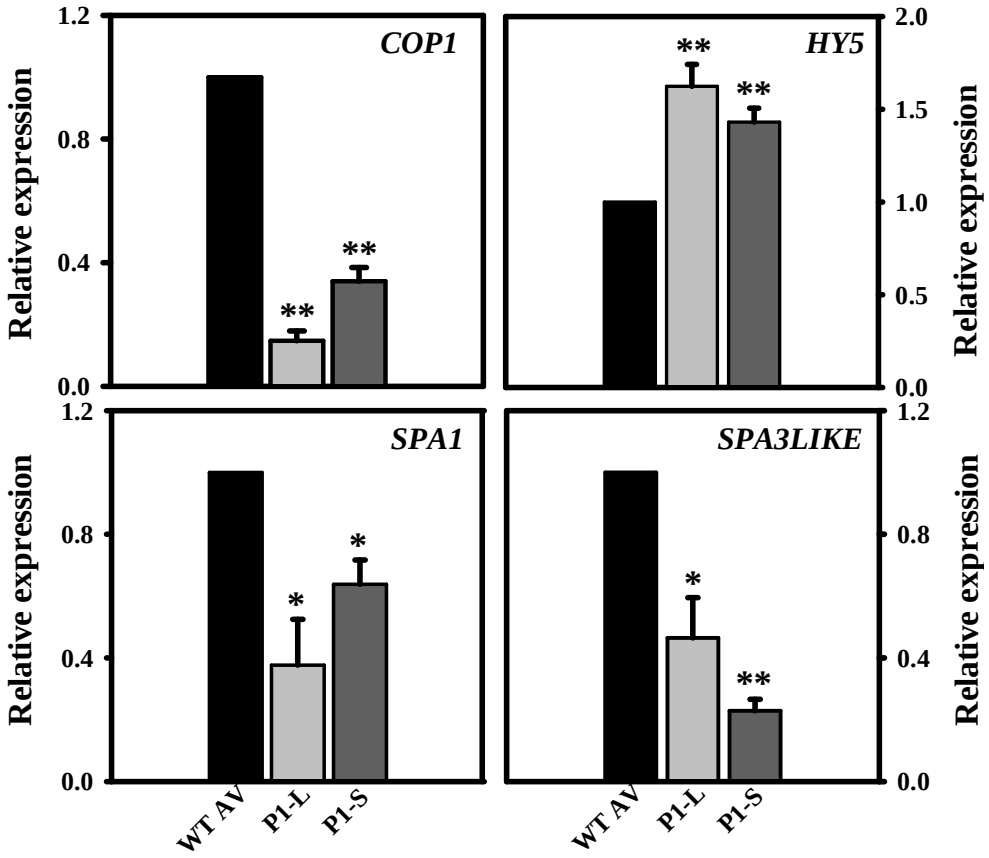


Figure 4.22 Gene expression analysis in young leaf tissue from wild type (AV) and mutant plants. The relative fold expression of the transcripts analyzed was plotted after normalization to WT sample. Leaf samples ($n \geq 3 \pm$ SE, n= no. of samples) of WT, P1-S and P1-L from green house grown plants were used. ** denotes the significant differences between WT and mutant plants (P1-S and P1-L) as determined by statistical tests with $P \leq 0.05$. * represents the statistical significance with $P \leq 0.1$.

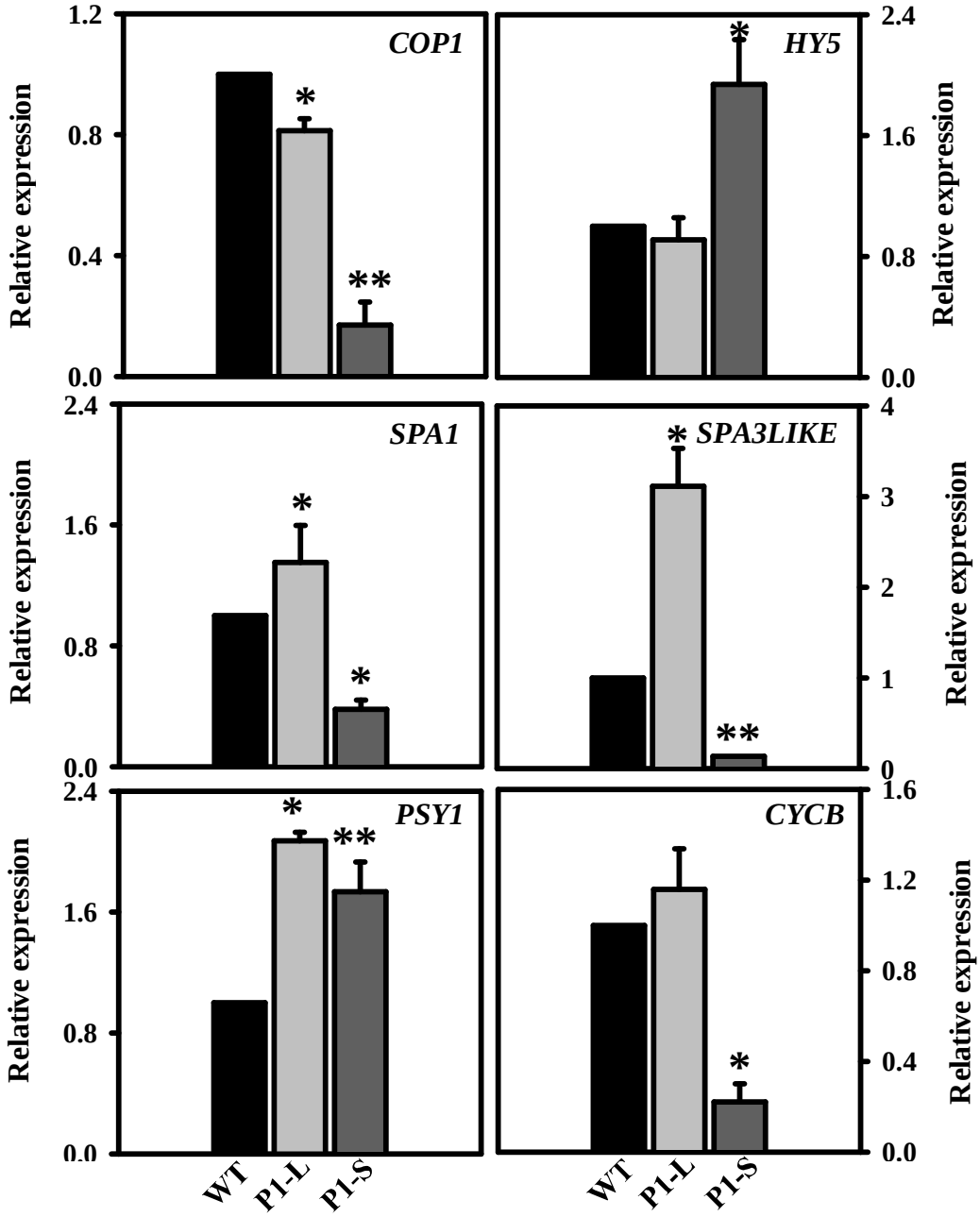


Figure 4.23 Expression analysis of *COP1, COP1LIKE, SPA1, HY5, PSY1 and CYCB1* in red ripe fruit tissue from wild type (AV) and P1-L and P1-S. The relative fold expression of the transcripts analyzed was plotted after normalization to WT sample. Red ripe fruit samples ($n \geq 3 \pm SE$, n= no. of samples) of WT, P1-S and P1-L from green house grown plants were used. ** denotes the significant differences between WT and mutant plants (P1-S and P1-L) as determined by statistical tests with $P \leq 0.05$. * represents the significant difference with $P \leq 0.1$.

1992). Though a silent mutation in tomato *cop1* gene did not cause drastic phenotype at seedling level (**Fig. 4.10**), but altered carotenoid content in the fruits, the mutant could possibly have either an insufficient amount or a nonfunctional tomato COP1. Considering the above results, we presumed that codon bias would alter the protein abundance and therefore checked the protein levels in fruit tissue. Immunoblotting of COP1 protein from WT tomato seedlings (grown in darkness) probed with At-anti COP1 antibodies detected a weak band of 76 kD protein (**Fig. 4.24**). Prominent low molecular weight products were also observed. As there were no reports on COP1 in fruit, we made an attempt. A very faint band of 76 kD was observed in green fruits of WT and no detectable band was observed in ripe fruit. In view of these results and the limitation of fruit tissue, we could not determine the COP1 protein levels in the mutant.

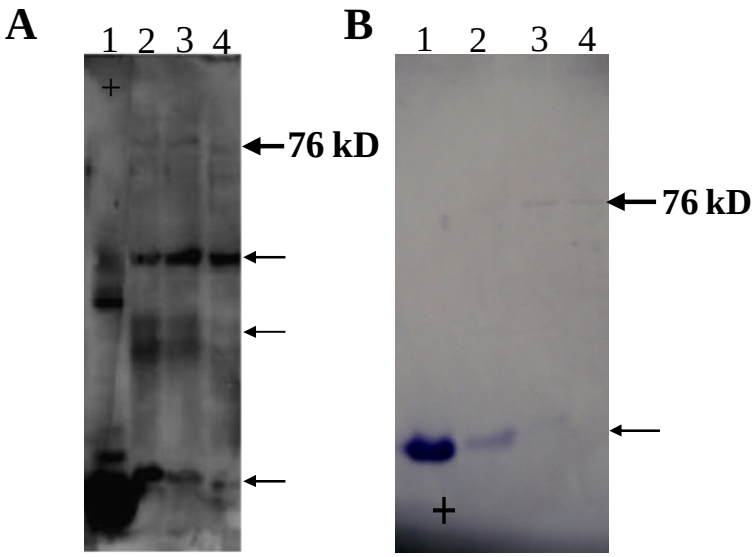


Figure 4.24 Immunodetection of tomato COP1. Arabidopsis anti-cop1 antibody recognizes protein of approximately 76 kD. A-faint band was detected in the dark grown WT seedlings (Lane 2, 3 and 4). B- COP1 protein levels in tomato fruit tissue. Very faint band were detected in green fruit (Lane 3 and 4), negligible or not detectable levels in ripe fruit (Lane 2). In both the blots Lane 1 represents positive control (denoted by +) and the thin arrows represent the degradation products.

4.2.12 Genetic analysis of the *COP1* mutant

As we observed a segregation in the plant height (P1-S and P1-L) but no segregation in the fruit carotenoid content, it was not clear whether the observed phenotype is due to the existence of background mutations or the incomplete penetrance of the mutation. To examine this, we performed a back cross of the homozygous P1-S plant (♂) which harbored two mutations and also exhibited drastic phenotypes (short stature and enhanced lycopene) with wild type Arka Vikas (♀). The F1 plants had normal development as wild type parent; however, the fruit carotenoid content was almost similar to the mutated parent. The genetic analysis of 53 F_2 plants was carried out based on Li-COR analysis of the SET III primer amplified products subjected to mismatch enzyme cleavage assay (**Fig. 4.25 A&B**). A ratio of 8:1:1 (homozygous wild type: heterozygous plants: homozygous mutant) F2 segregation (**Table 4.3**) with respect to the silent change observed in exon 5,6 and 7 region did not fall in accordance to Mendelian ratio. These results suggest a case of expressivity or penetrance of the gene. At this point, one cannot rule out the size of F_2 considered for analysis and also the segregation of the second silent mutation present in exon 1. Due to limitation of seeds, and persistent PCR problems with SET VI primer (amplifying exon 1, where the AT content is high), we could not do a thorough statistics for calculating the Mendelian segregation ratio for a monogenic trait. These results can only be speculated to be the pleotropic affects of *COP1* gene, owing to its importance in regulating many plant developmental pathways. Due to the inability to conclude from the genetic analysis, we plan to analyze the heterozygous plant obtained while screening the M_2 population and subsequently consider the homozygous mutant that segregates from this line. In addition, a reciprocal backcross with P1-S (♀) and AV (♂) also needs to be performed to understand the segregation of the mutation and this is currently underway. A further detailed study is needed to dissect the nature of the mutation, by using the homozygous plant and to determine the phenotype of *COP1* mutant in tomato.

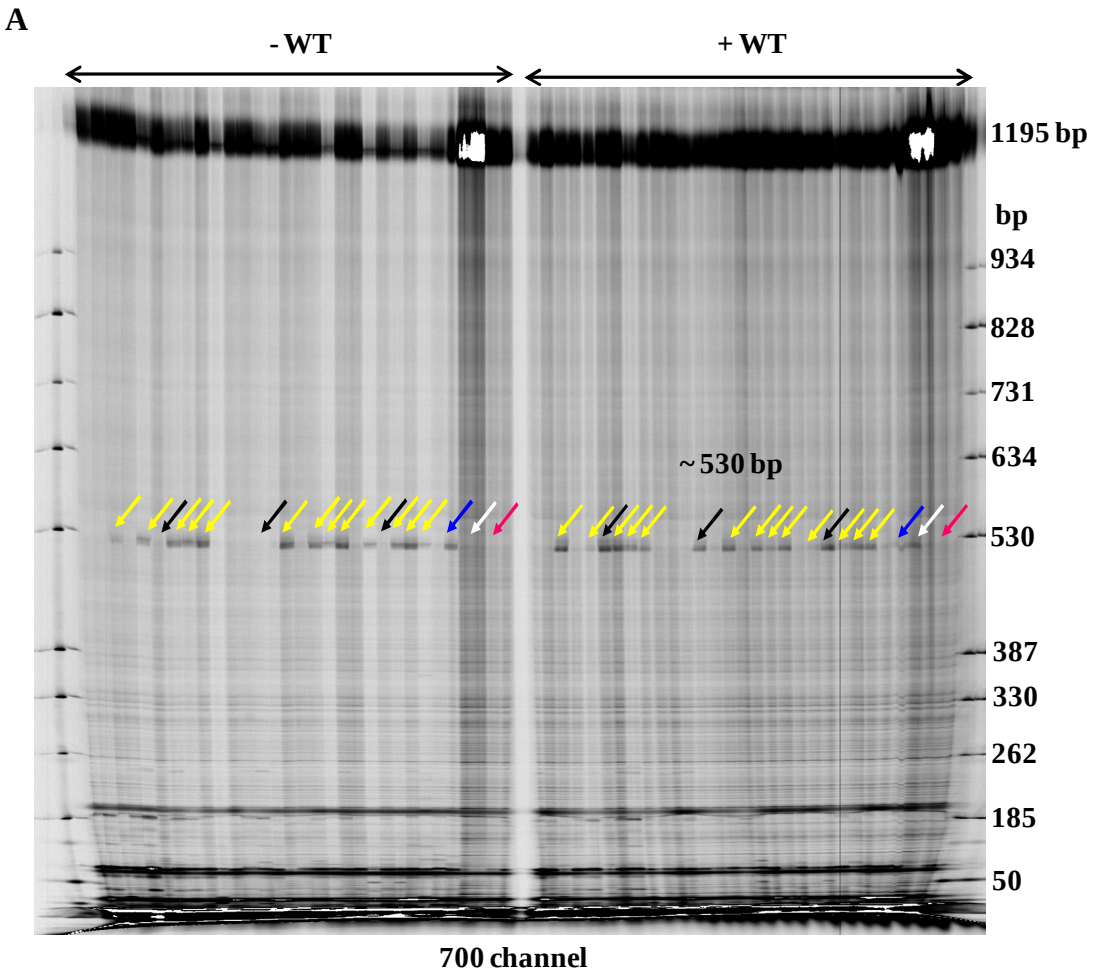


Figure 4.25 Genotype confirmation of the silent mutation (L342=) in F_2 segregating population of the cross WT AV (♀) x P1-S (♂). **A**, represents the IRD 700 channel of the Li-COR image with the excised product at ~ 530 bp. The solid yellow arrows represent the heterozygous plants, where the digested fragment is observed in both homoduplex (-WT) and hetero duplex (+ WT).The solid black arrows represent the homozygous segregants for mutant, where the digested fragment is observed only in hetero duplex (+ WT) condition. The solid white arrow represents the homozygous mutant parent, solid red arrow represents the homozygous WT parent and the solid blue arrow represents the F_1 plant. The DNA ladder is marked on right of the 700 channel. **Note**: The individual DNA isolated from F_2 individuals was amplified using SET III primers. The 1195 bp amplicon was subjected to enzymatic cleavage and the digested fragments were resolved on Li-COR gel.

B

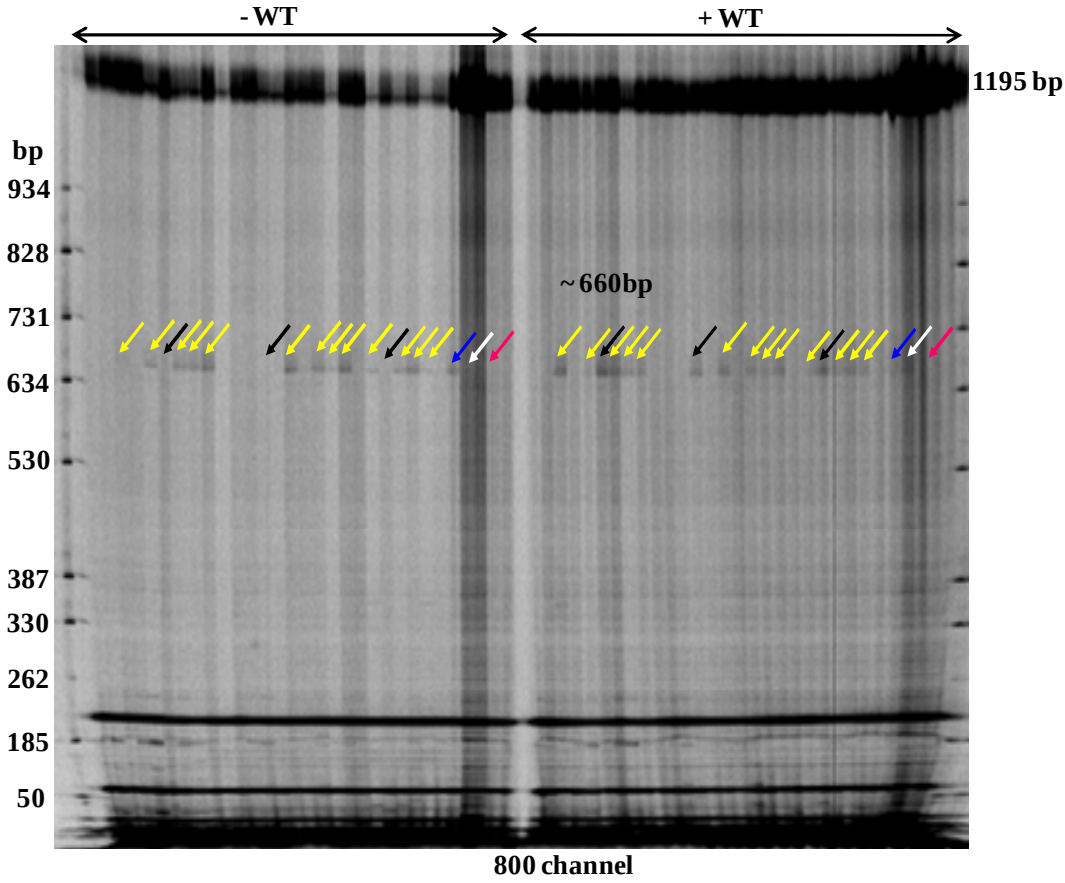


Figure 4.25 Genotype confirmation of the silent mutation (L342=) in F_2 segregating population of the cross WT AV (♀) x P1-S (♂). B, represents the IRD 800 channel of the Li-COR image with the excised product at ~ 660 bp. The solid yellow arrows represent the heterozygous plants, where the digested fragment is observed in both homoduplex (-WT) and hetero duplex (+ WT).The solid black arrows represent the homozygous segregants for mutant, where the digested fragment is observed only in hetero duplex (+ WT) condition. The solid white arrow represents the homozygous mutant parent, solid red arrow represents the homozygous WT parent and the solid blue arrow represents the F_1 plant. The DNA ladder is marked on left side of the 800 channel. **Note**: The individual DNA isolated from F_2 individuals was amplified using SET III primers. The 1195 bp amplicon was subjected to enzymatic cleavage and the digested fragments were resolved on Li-COR gel.

Table 4.3 Segregation Analyses of F_2 population of the cross WT (♀) x P1-S (♂).

Total F_2 plants	No. of Homozygous (WT)	No. of Heterozygous (WT + Mutant)	No. of Homozygous (Mutant)
53	43	5	5
Ratio	8	1	1

4.3 Discussion

Tomato is a rich source of dietary nutrients and also the most highly consumed vegetable next to potato. Ripe tomato is rich in carotenoids like lycopene and beta carotene and many other micronutrients that have potential health benefits. For a developing country like India, improving the quality and meeting the demands of food supply would be a formidable challenge. Enhancing the nutritional value of tomato, a poor man's apple would be of great contribution to the society. Though plant breeding has led to the selection of best breeding varieties, molecular interventions and technical inventions had great impact on research and development. Considerable number of studies has been conducted in understanding the mechanisms and minimizing the post-harvest losses of tomato by targeting the genes involved in cell wall biosynthesis, carotenoid biosynthesis, ethylene biosynthesis and regulation etc. Recent studies on manipulating the light signal for fruit ripening (Liu et al., 2004) suggest that many gene isolation efforts and metabolic engineering studies have primarily targeted genes involved in carotenoid biosynthesis. However, factors that regulate flux through the carotenoid pathway remain largely unknown. In the present study we did TILLING of *COP1* (a master regulatory gene in photoreceptor pathway) to examine its effects on tomato fruit quality.

Pre- requisites for TILLING

The Targeting Induced Local Lesions IN Genomes (TILLING) is a widely accepted and practiced reverse genetic tool. Because of its high throughput screening, cost effectiveness, specificity of mismatch-specific endonucleases and detection sensitivity, many TILLING platforms have been developed both in plants including rice (Suzuki et al., 2008, Till et al., 2007), wheat (Slade et al., 2005), maize (Till et al., 2004), melon (Dahmani-Mardas et al., 2010, González et al., 2011), barley (Caldwell et al., 2004, Talamè et al., 2008), pea (Triques et al., 2007), ground nut (Knoll et al., 2011), beetroot (Hohmann et al., 2005 and Frerichmann et.al., 2013), Medicago (Carelli et al., 2013), tomato etc; and in animal species like the worm - *Caenorhabditis elegans* (Gilchrist et al., 2006), the fruit fly - *Drosophila melanogaster* (Winkler et al., 2005) and Zebra fish - *Danio rerio* (Moens et al., 2008). In tomato, four independent TILLING platforms employing a variety of genetic backgrounds have been developed by different groups (Okaba et al., 2013). All these

studies report that TILLING is an efficient tool for mutation discovery in crop species as in Arabidopsis. In the current study we used a collection of approximately 9126 mutagenized M_2 tomato plants (of cv Arka Vikas and M82) for mutation detection and frequency determination. The efficiency of a TILLING platform relies on many factors - (1) a well-mutagenized population (2) availability of gene sequences and optimized primer parameters, (3) optimization of PCR conditions and (4) mainly the gene selected and the mutation frequency.

Fluorescent labeled marker for a high throughput screening

Though TILLING is a high throughput and reliable technique, one has to consider the overall cost evaluation to perform this technique on a large scale. In order to make it cost effective without compromising the quality, many in house developed protocols (well optimized) were adopted. NEATTILL, a time-saving, less labour intensive and reasonably cost-effective DNA isolation method (Sreelakshmi et al., 2010) has been developed in the lab. The 8X pooling of the isolated mutant DNA and 2-D array strategies have further reduced the cost of Li-COR based mutation detection. A major advantage of Li-COR based TILLING over competing methods is that the approximate position of each detected mutation is inferred from the size of the fragments obtained after CELI digestion, which greatly facilitates subsequent analysis (Henikoff et al., 2004). In spite of the many in-house developed protocols, the number of reactions per gene is the deciding factor for high throughput screening of large populations. The commercially available labelled markers could detect only up to 700 bp and this made many of the TILLING labs to design primers that result in only $\leq$ 1000 bp amplicon. However, CODDLE can facilitate the prediction of hot spots in the target gene upto 1200- 1500 bp (max. window length is 1500 bp). In consideration with this fact, one can amplify fragments $\leq$ 1500 bp. On large scale TILLING, this would definitely reduce the number of reactions to be carried out per gene. The fluorescent labelled marker developed in our lab which can detect DNA fragments upto 1500 bp, alleviated this problem encountered with the commercial markers (low molecular weight detection) thus, making our TILLING platform more efficient and economical.

Optimization of primer parameters and PCR conditions

Once the target genes are selected, primer designing would be a major task in performing efficient TILLING. In our study, we used Primer3, a web-based primer design tool to generate gene-specific primers. Many other TILLING platforms (Greene et al., 2003; Perry et al., 2003; Talamè et al., 2008), also used the same web based tool for primer designing. Several factors have to be considered to design a successful primer pair, like- nucleotide composition at the 3′-end being essential for the control of mispriming (Kwok et al., 1990) and complementation of a primer pair (Dieffenbach et al., 1993). The tendency to form primer dimers, which will compete with the template in the PCR reaction, can be minimized by checking the homology at the 3′-end. Furthermore, the melting temperature (Tm) between the forward and reverse primers should not differ more than 5°C, as this will have a negative effect on primer binding to the template, as the optimal annealing temperature is shifted (Singh and Kumar, 2001). In consideration of the above discussed factors, each primer would have different success rate, because of the region chosen for amplification and the resulting nucleotide compositions. In the present study, different procedures and conditions were evaluated for PCR amplification of six different sets of primers used for *COP1* TILLING. A systematic standardization was done with the SET III primers and an optimal PCR protocol was developed for TILLING of *COP1* gene.

The main goal of optimizing PCR conditions was to achieve proper and clean amplification of the target region i.e. a minimum of 10 ng/µL of PCR product in a total reaction volume of 10 µL as discussed in Till et al., (2006a). Various stages of optimization were conducted to attain the proper amplification. Initially, the primers were tested on agarose and the parameters like - annealing temperature, cycling conditions, PCR components and additives were optimized. Optimization of cycling conditions was the most critical step that has improved the quality of PCR on agarose gels. Touchdown PCR, a variant of normal PCR, employs an annealing temperature that is 5 to 10°C higher than the original Tm of the primer. It has been reported that this improves the specificity of PCR amplification (Don et al., 1991) by avoiding the generation of contaminating PCR products (Wu et al., 2005). However, the inclusion of labeled primers in PCR mixture required additional optimization than unlabelled primers, therefore PCR optimization was reinitiated. The next optimization step included the variation in the amount of reaction components.

The concentrations of monovalent and divalent ions in the reaction mixture affect the activity of the *Taq* polymerase. As we used an in-house *Taq* polymerase, we tested the different concentrations of $MgCl_2$ and 10X PCR buffer (this provides K^+ and Cl^- as the monovalent ions). To improve the profile of PCR product observed on Li-COR DNA analyzer, varying concentrations of DNA template and primer concentrations were tested. In order to further improve the efficiency, consistency of PCR reactions and also to cut down the cost of TILLING, we followed nested PCR. Nested PCR increases the efficiency and specificity of PCR amplification by re-amplifying the product of the first PCR step in a second step PCR. To deal with major challenges related to PCR amplification, several groups used nested PCR as a part of their optimization (Wienholds et al., 2003b, Winkler et al., 2005). The main advantage of this method is that, it reduces the influence of the variation in quality and quantity of the genomic DNA that is used as a template for TILLING PCR (Wienholds et al., 2003b).

TILLING of *COP1* gene

TILLING allows for the identification of an allelic series of mutants with a range of modified functions for desired genes (McCallum et al., 2000a, McCallum et al., 2000b, Colbert et al., 2001, Comai and Henikoff 2006). In this study, we targeted *COP1* for identifying multiple alleles. The 8-fold pooled DNA from three different populations (**Table 3.1)** was used for the screening of mutations on Li-COR DNA analyzer. With the optimized PCR conditions, amplification of the desired fragment was performed and the products were denatured and reannealed to form heteroduplexes between the mutated sequence and the wild type counterpart. We used the classical analysis method of TILLING that was based on CELI endonuclease enzyme which cleaves at the point of mutation by recognizing the mismatches in the DNA heteroduplexes, followed by the resolution of PCR products on denatured polyacrylamide gels.

The COP1 (constitutive photomorphogenic 1) protein is conserved in both higher plants and vertebrates (Yi and Deng, 2005). It is one of the best characterized proteins among the pleotropic proteins - COP/DET/FUS. It acts as an E3 ubiquitin ligase and is a 673 amino acid protein in *S. lycopersicum* (675 a.a in Arabidopsis and 733 a.a in Humans) with a molecular weight of 76 kDa. It comprises of three

domains- RING finger, coiled-coil and WD40 domains. The ring finger motif mediates interaction with the Ub - conjugating enzyme (E2), while coiled coil region allows the formation of COP1 homo-dimers or hetero-dimers between COP1 and members of SUPPRESOR of PHYA (SPA) protein family in Arabidopsis (Hocker and Quail 2001). The WD40 repeat region is the substrate interacting domain (Holm et al., 2001, Yi and Deng et al., 2005). Many of the COP1 substrates are transcription factors or transcriptional regulators. COP1 has a major role in seedling photomorphogenesis with *HY5* (a bZIP transcription factor of *PHYA* gene) being its prime target.

In view of the importance of WD40 domain in mediating COP1 function, the primers were designed for screening mutations in this region. Three sets of primers (SET I, SET II and SET III) were designed to cover the entire WD40 region. The CODDLE output predicted the exons 3 and 4 (encoding coiled coil domain) to be most susceptible for EMS treatment. Therefore we designed the SET IV primers to screen this region. Studies in Arabidopsis reveal that any mutations in the N-terminal region would be lethal (Deng et al., 1992). Therefore, this region was not considered for TILLING. Moreover, this domain encoded by the exons 1 and 2 was also not predicted by the CODDLE software. This supports the fact that the N-terminus of the protein is very important for COP1 function (Mc Nellis et al., 1996). On screening 9126 individuals with four sets of primers, only three putative mutations were identified. Upon de-convoluting the positive pools by screening the corresponding column plates, only two mutations were confirmed. The third putative mutation might have been a false positive as we could not confirm it in the corresponding column plate. Out of the two confirmed, the individual plant was identified only for one mutation detected in population II when amplified with SET III primers. For one of the putative mutations identified upon screening with SET I primers in population II, we could not trace the individual plant carrying the mutation. This may be because the mutant allele could be lost from M_2 to M_3. This could happen if the M_2 plant was heterozygous for the mutation and during M_3 seed production, some form of gametic selection against the mutant allele may have occurred due to lower fitness. Various degrees of gametic selection have been reported for soybean (Kopisch-Obuch and Diers, 2006) rice (Guiderdoni, 1991), mustard (Cheung et al., 1997) and sugar beet (Pillen et al., 1992).

The overall frequency of mutations for *COP1* gene in our population was found to be 1/15.1 Mb. In other words, the rate of occurrence of a mutation in *COP1* gene is in the order of 6.54 x 10^{-8}. This value is almost equal to the estimated rate of a spontaneous mutation. Hence, we could conclude that the mutation frequency for the targeted gene *COP1* is very low. The reason for such a low frequency of mutations for *COP1* in tomato M_2 lines was not known and it can only be speculated. One probable explanation could be *COP1* gene of tomato is recalcitrant to mutagenesis. To date there are no known spontaneous mutations reported for *COP1* in tomato. Though we have covered 85% of the coding region (except for exon 1 and 2) for TILLING, the very low frequency of mutations in *COP1* suggests the conserved nature of this gene and its importance in plant growth and development (Ranjan et al., 2014). Though we could not isolate multiple alleles for *COP1* by TILLING, Jones et al., (2012) reported seven independent mutations in tomato *COP1*. However, none of them exhibited any of the expected characteristic phenotype, including enhanced nutrient quality of fruits. One more reason for the low frequency might be the detection method opted. Though Li-COR based TILLING is sensitive, for genes like *COP1* which have very small exons and large introns, the frequency of mutations in coding region might be low. The next generation mutation detection methods like HRM or NGS would be more apt to screen small exon regions. Another possible reason for the low mutation frequency observed could be the type of cultivar used for mutagenesis. As described in **Table 2.1,** the different genetic backgrounds of tomato cultivars used for mutagenesis have different mutation frequencies. In Micro-Tom using 0.5% EMS mutagenesis, a frequency of 1 mutation/2180 bp was reported, which increased to 1 mutation in 737 bp when 1% EMS was used as mutagen (Okabe et al., 2013). The cultivar specific differences exist in tomato with respect to frequency of mutations, where, Red setter variety of tomato TILLING population showed 1 mutation/574 bp with 0.7% EMS mutagenesis (Minoia et al., 2010).

The mutant line (Accession number AV- M2- 1894) harboring mutation in *COP1* gene was further analyzed to detect the molecular nature of the mutation. Sequence analysis of *COP1* gene from the above mutant line revealed the presence of a nucleotide change from cytosine (C) to thiamine (T) at 6964^{th} position on genomic co-ordinates. The conversion of amino acid showed that this change causes a silent mutation as it alters the codon CTA to TTA, but do not lead to any change in the

amino acid at 342 (L342 =) in the mutant COP1 protein. However, the codon usage efficiency of altered codon (TTA) is approx. 50% less than the wild type codon (CTA). Detailed phenotype analysis of the plant revealed no characteristic change at the seedling stage (open cotyledons and short hypocotyls- as observed in Arabidopsis *cop1* mutants), the adult mutant plants phenotypically segregated as short (P1-S) and long (P1-L) with some architectural differences of internode differentiation. The genotype of the two phenotypic classes remained the same, i.e. silent change from L342=. Sequencing of the entire coding region from the segregating plants revealed another silent mutation at T31= in P1-S. Considering the fact that EMS causes random mutations, the two silent mutations were taken as two independent events. Moreover, a reduction in transcript levels of *COP1* in both the phenotypic classes supports the effect of mutation. The observation that the mutant plant exhibited altered phenotype establishes a functional consequence of the silent mutations.

With the growing complexities in relating genotype and phenotype, one cannot stick on to the presumption that only missense changes in a gene would have an impact on the phenotype. Recent studies by Kimchy-Sarfaty et al., (2007) on *Multidrug Resistance 1 (MDR1)* gene product -P-glycoprotein, or P-gp, the attenuation of polio virus by introducing 600 silent mutations (Coleman et al., 2008) show that the presence of rare codons encoded by silent mutations affect the function of the gene. Almost 50 genetic disorders have now been attributed to the silent mutations, many involving altered splicing of mRNAs. A recent summary by Chamary and Hurst (2009) outline a number of mechanisms that can account for phenotypic consequences of silent mutations. These mechanisms include: effects on the translational rate which in turn affects the protein folding, inactivation of the native splicing donor site, which may result in premature stop codon or exon skipping yielding a shorter mRNA and also synthesis-rate-dependent effects on protein folding, leading to altered protein function (this is because of the effect on co-translational rates).

Concordant with the above discussion, the reduction in the transcript levels of *COP1* in both the phenotypic classes (80% in short and 60% in long) compared to AV might be a consequence of silent mutations. As described in Hoecker and Quail (2001), the *SPA1* acts together with *COP1* and *HY5* is an antagonist. Further, the lower transcript levels of *SPA1* and on the contrary, increased transcript levels of *HY5*

support the functional role of *COP1*. In spite of architectural variations of the adult mutant plants, the red ripe fruits have similar two fold increased lycopene levels in comparison to AV. Two important genes, *PSY1* and *CYCB* (encoding rate limiting enzymes) of the carotenoid biosynthetic pathway are considered for transcript analysis. The increase in lycopene and β-carotene contents in the fruits of both P1-S and P1-L could be correlated to the increased expression of *PSY1* (in P1-S and P1-L) and not statistically significant increase in P1-L for *CYCB*. As mutations in the light signaling pathway positively influence the pigmentation of ripe fruit, targeting the light signaling pathway might be an effective means of engineering fruit nutritional quality (Liu et al., 2004). Moreover, *PSY* transcript abundance is upregulated during photomorphogenesis via a phytochrome-mediated pathway (Welsch et al., 2000). Taking into account the carotenoid content and the transcript data, P1-S harboring two silent mutations appears to exhibit strong phenotype probably due to lack of *COP1* function. These results also raise the possibility of incomplete penetrance or expressivity of the gene.

In order to understand the phenotype of mutant plants, we performed a genetic cross between WT AV (♀) and P1-S (♂) parents to check segregation of genotype and phenotype. Segregation analysis of F_2 population with SET III primer for the mutation in exon 5, 6 and 7 (**Table 4.3**) did not correspond to the segregation pattern of a monogenic trait. Moreover, we failed in our attempts to segregate the silent mutation observed in exon1 (Thr a.a at 31=) in this F_2 population. This was due to the difficulties faced during TILLING PCR reactions and sequencing of exon 1 amplicon as the AT content of the region was found to be high (63%) and this region also had nucleotide repeats which may have interfered during sequencing. It is also not clear at this point whether the phenotypic changes in P1-S are a result of either both the silent mutations or one of them. Despite the tremendous improvement in sequencing technology, sequencing of the templates having various nucleotide repeats, high GC content, long homopolymer stretches and also those which tend to form hairpin structures is challenging (Kieleczawa, 2006). Taking into account these bottle necks during PCR and hypothesizing that the two silent mutations co-segregate, we still do not have a clear segregation pattern. Also we cannot rule out the possibility of the two independent events to segregate, and for this we need to further extend our back

crosses to segregate the pure line. In addition, we cannot ignore the possibility of other background mutations co-segregating with the observed mutations.

In conclusion, in the present study we identified a novel allele for *COP1* in tomato. A preliminary characterization of the mutant line established a putative role of *COP1* in fruit ripening. However, identification of more alleles would be helpful in fully understanding its role in tomato development and fruit ripening.

Chapter 5

Eco-TILLING of tomato *COP1* and *SPA3LIKE* genes

5.1 Introduction

Eco-TILLING is a modified TILLING strategy that is widely used for the detection of polymorphisms in target genes in natural populations (Comai et al., 2004). Similar to the pooled DNA from mutagenized population used for TILLING, the DNA from different natural accessions are pooled and used as the starting material in Eco-TILLING. The DNA from natural accessions can be either mixed with reference cultivar DNA or it can be mixed from different lines.

Similar to TILLING, heteroduplex formation at the point of polymorphism is detected by the action of mismatch specific endonucleases like CELI and the digested fragments are detected by any standard method such as agarose gel based or Li-COR gel methods; capillary electrophoresis; HRM etc. Eco-TILLING is a cost effective technique as pooling and later de-convolution of samples is less critical and also the number of individuals to be sequenced for each representative haplotype can be minimized. Like TILLING, Eco-TILLING is a general tool and has been widely employed in various plant species from Arabidopsis to the orphan crops for which genome sequences are not available like cassava, ground nut, millets etc (Till et al., 2007). Eco-TILLING has been extensively used to identify the genetic variability in *Arabidopsis thaliana* (Comai et al., 2004), *Populus trichocarpa* (Gilchrist et al., 2008), *Vigna radiata* (Barkley et al., 2008), *Phaseolus vulgaris* (Galeano et al., 2009), *Musa* spp. (Till et al., 2010) etc. In addition to the identification of genetic variation, this technique has also been applied to detect new alleles that conferred biotic and abiotic stress in *Hordeum vulgare* (Mejlhede et al., 2006), *Oryza sativa* (Kadaru et al., 2006; Negrao et al., 2011), *Cucumis* spp (Nieto et al., 2007), *Solanum tuberosum* (Elias et al., 2009). In tomato, Eco-TILLING has been used to identify novel alleles for *eIF4E* gene that confers resistance to viruses (Rigola et al., 2009) and also for improving the fruit quality (Bauchet et al., 2009).

Tomato is a dicotyledonous plant belonging to the most diverse family - Solanaceae (with more than 10,000 sps). Though it is a new world plant, it is presently grown and consumed in almost all parts of the world. Domestication of tomato has led to the development of large number of tomato cultivars and varieties differing in many morphological aspects from leaf to fruit. Since tomato is s rich source of health promoting nutrients like lycopene and β-carotene, the major

carotenoids in ripe fruits (Paiva et al., 1999; Omoni et al., 2005) considerable efforts have been directed towards development of nutrient rich tomato cultivars. Tomato fruit ripening is a typical climacteric process with a rise in the levels of ethylene accompanied by sudden respiratory burst during ripening. Molecular genetic studies on non-ripening mutants of tomato- *Nor* (non ripening), *Nr* (never ripening), *Rin* (ripening inhibitor) and *cnr* (colorless non ripening) (Osorio et al., 2011) and also with the transgenic suppression of genes involved in ethylene biosynthesis and signalling (Karlova et al., 2011; Lee et al., 2012), and cell wall softening (Qin et al., 2012; Fujisawa et al., 2013) elucidated the genetic hierarchy in regulation of fruit ripening. In addition to ethylene, hormones like auxin, ABA, polyamines etc. are also involved in regulation of fruit ripening (Trainotti et al., 2007).

Similar to elucidation of genetic hierarchy regulating tomato ripening, the usage of photomorphogenic mutants independently showed the regulation of fruit ripening and quality. Molecular genetic studies of *hp1* (*high pigment1*, homolog of *AtDDB1*) and *hp2* (*high pigment2*, homolog of *AtDET1*) mutants of tomato indicated an important role of light signal transduction components for the accumulation of carotenoids in fruits (Lieberman et al., 2004; Mustilli et al., 1999 and Kilambi et al., 2013). The role of light signaling is supported by enhancement of carotenoid levels in RNAi silenced lines of tomato *CUL4* (a scaffolding protein of E3 ligases) and *COP1LIKE* (recently annotated as *SPA3LIKE* and is a negative regulator of photomorphogenesis) genes. These reports highlight that manipulating light signaling components is also an alternate approach for improvement of the nutrient quality of tomato (Wang et al., 2008 and Liu et al., 2004). Furthermore, recent analysis of ripening in tomato phytochrome mutants implicated a complex interaction between the phytochrome species to regulate different phases of fruit ripening (Gupta et al., 2014).

Notwithstanding the wealth of information generated using tomato mutants to date, majority of studies were conducted using either spontaneous mutants (that can be introgressed into local cultivars by repeated back crosses to generate a pure line) or using transgenic lines for a set of genes of interest (which requires regulatory body approval and also has ethical issues). To avoid issues with genetically modified organisms (GMOs), high throughput techniques like Eco-TILLING can be applied to natural accessions to identify alleles that can lead to the production of nutrient rich

tomatoes. The identification of a natural accession with high soluble sugar content (brix value) in tomato fruit (Bauchet et al., 2009) has set an example for improving the fruit quality using Eco-TILLING**.**

In the present study, we examined a large collection of natural accessions of tomato to identify allelic variants of candidate genes for enhanced fruit lycopene and β-carotene content. As candidate genes we chose (i) *COP1*, as its homolog plays master regulatory role in light mediated responses in Arabidopsis (Lau and Deng, 2012) (and also due to the low frequency of induced mutations observed in our TILLING study) and (ii) *SPA3LIKE* because of its functional similarity with tomato *COP1*, *AtRUP1* (repressor of UV-B photomorphogenesis) and also its reported negative regulatory role in photomorphogenesis and carotenoid accumulation (Liu et al., 2004). We found that allelic variation at the *SPA3LIKE* locus was associated with the enhanced lycopene and β-carotene levels in red ripe tomato fruits. Our data provide the evidence for the functional relevance of *SPA3LIKE* in tomato fruit ripening, and is also relevant for tomato breeding. This study also furthers our understanding of the role of light signaling components in regulating fruit ripening and carotenoid accumulation in tomato.

5.2 Results

Taking in account the functional similarity of *COP1* and *SPA3LIKE* genes, the nucleotide diversity for these two genes was evaluated in a population comprising of 300 accessions from NBPGR, 140 accessions from IIVR, 60 accessions from BSS, India and around 85 accessions from TGRC.

5.2.1 Eco-TILLING of tomato *COP1* gene

A total of six primer sets were used to assess the natural variation in *COP1* gene. All the primer sets were optimized for touchdown PCR. PCR products were analyzed on polyacrylamide gel on Li-COR based system. Based on the presence or absence of complementary bands in 700 nm and 800 nm channels in LiCOR obtained after CELI digestion, the polymorphisms were manually scored in accessions. **Figure 5.1** shows a representative Li-COR gel with the cleaved PCR products. Some of the accessions showed similar cleavage pattern while others were different. Multiple cleavage bands were detected in some lanes representing several polymorphic sites in an individual accession. Each SNP is defined as a unique fragment visible on the polyacrylamide gel and the prominent bands were considered for haplotype classification.

5.2.1.1 Genotyping and sequencing of variants

Based on the similarities and differences observed on Li-COR gels, the *COP1* variants were categorized into 11 haplotypes with the reference accession (WT AV) designated as HT0. The accessions showing similar SNP pattern were grouped into one haplotype. Each haplotype included variable number of accessions exhibiting identical SNP pattern for *COP1* gene. As expected, majority of the accessions were similar to the reference cultivar and thus were included in the haplotype HT0. The accessions included in each haplotype are listed in **Table 5.1**. Haplotype 3 was represented by the maximum number (38) of accessions, while the rest of the haplotypes were represented by one or more accessions. One representative genotype for each unique haplotype was sequenced on both the strands. The raw sequences were analyzed using the Chromas 2.4 software. Alignment of the variant sequences with reference cultivar helped in the identification of the exact position of the SNPs.

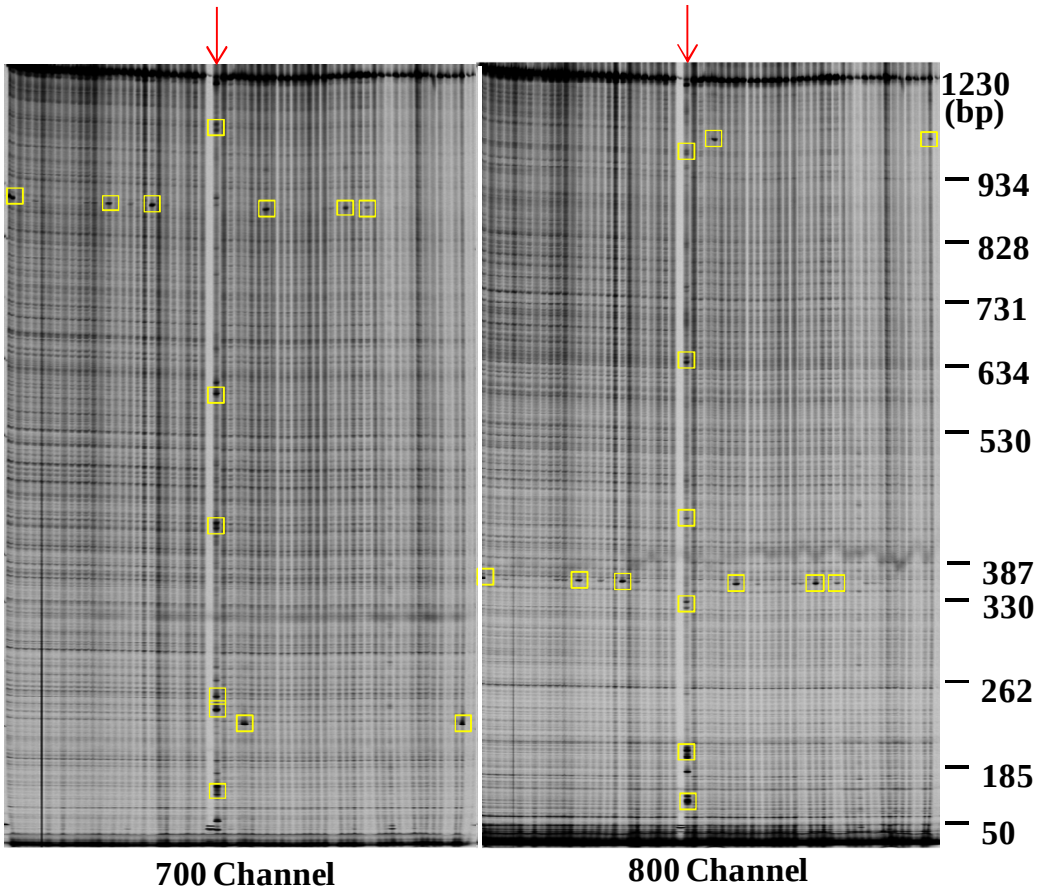


Figure 5.1 Detection of SNPs in tomato accessions using Eco-TILLING. The Li-COR gel analyzer images in IRD700 nm and IRD800 nm channel are shown for a 1230 bp region of *COP1* gene. Each lane in a 64 well gel was loaded with digested DNA from a unique accession. The polymorphisms are boxed. The accession harboring multiple SNPs is marked by a red arrow. Complementary bands, one in IRD700 and another in IRD800 were observed for each polymorphism. The right side of image shows DNA markers with respective sizes.

Table 5.1 *COP1* haplotypes in tomato accessions. The SNPs identified by Eco-TILLING were confirmed by sequencing the amplicons of *COP1* gene. Based on the gene sequences, the polymorphisms were sorted into haplotypes.

Haplotypes	Accessions
HT0	Arka vikas (reference cultivar) and genotypes similar to reference
HT1	NBPGR 45
HT2	NBPGR 597, NBPGR 514
HT3	IIVR 99, IIVR 85, IIVR 64, IIVR 33, IIVR 25, IIVR 23, IIVR 116, TGRC 85, NBPGR63, NBPGR67, NBPGR84, NBPGR413, NBPGR 424, NBPGR 433, NBPGR60, NBPGR27, NBPGR101, NBPGR104, NBPGR110, NBPGR117, NBPGR353, NBPGR 589, NBPGR 600, NBPGR 480, NBPGR388, NBPGR 515, NBPGR200, NBPGR 504, NBPGR 469, NBPGR 471, NBPGR 484, NBPGR 434, NBPGR 578,NBPGR 226, NBPGR 485, NBPGR 588, NBPGR321, NBPGR 477
HT4	NBPGR85, IIVR 77
HT5	NBPGR24
HT6	NBPGR25
HT7	NBPGR32
HT8	NBPGR343, NBPGR349, NBPGR401, NBPGR347
HT9	IIVR 87, IIVR 53, IIVR 102, IIVR 104, IIVR 37, IIVR 74, IIVR 120, IIVR 114, IIVR 116
HT10	NBPGR4, NBPGR 485
HT11	NBPGR32, NBPGR86, NBPGR89, NBPGR 479, NBPGR352, NBPGR 515, NBPGR200, NBPGR 504, NBPGR 469, NBPGR 417, NBPGR371, NBPGR347,NBPGR 514, NBPGR 591, NBPGR 222

5.2.1.2 SNP data analysis

Approximately 4.5 million bp (no. of individuals screened x total amplicon size) were screened via Eco-TILLING technique. A total of 11 unique haplotypes were obtained for *COP1* gene for 6 amplicons including the reference sequence. The PARSESNP analysis of the 11 haplotypes file represented the position of single nucleotide polymorphisms, insertions and deletions. Most of the variations identified were intronic (**Fig. 5.2**). All the polymorphisms in coding region were categorized into two types- synonymous SNPs (S) that do not cause any change in the protein encoded by the target gene and nonsynonymous SNPs (NS) that cause a change in the protein sequence. The haplotypes 3, 4 and 6 share a nonsynonymous base change- G5848A, which changes the amino acid arginine (R) at 245th position to glutamine (Q) in the protein sequence. Two more nonsynonymous SNPs were identified- A5242C and G11077A in HT3 and HT10 each separately. The corresponding amino acid changes were lysine (K) at 153 to threonine (T) and glutamic acid (E) at 639 to lysine (K) respectively. The SIFT scores for all the three nonsynonymous changes were predicted to be tolerated, therefore, the function of COP1 protein may not be affected drastically by these natural polymorphisms.

Apart from these observed nonsynonymous changes, we could identify only one synonymous SNP in HT2, HT5 and HT8 each respectively. The details of the position of the base change and the unaltered amino acids were listed in **Table 5.2**. Since recent reports including our TILLING results suggest that some of the silent mutations also affect the phenotype, we checked the codon usage bias for these three synonymous changes. Interestingly, none of the synonymous changes led to drastic variations in the codon usage frequencies. Though most of the SNPs were intronic, some of the changes were common among few haplotypes. For example, the three intronic base changes - T10358G, G10826A and C10925CACC were shared among the HT8, HT9 and HT10. The significance of intronic SNPs is still unknown. In total, 28 SNPs and 3 INDELS were detected with a mean of 1 SNP/160 Kb and 1 INDEL/1497 Kb. A total number of 3 nonsynonymous and 3 synonymous SNPs were identified upon Eco-TILLING of *COP1*. The number of SNPs observed for each possible base change was calculated including the INDELS (**Fig. 5.3A**). The observed SNPs were categorized according to nucleotide substitution as either transitions (C/T

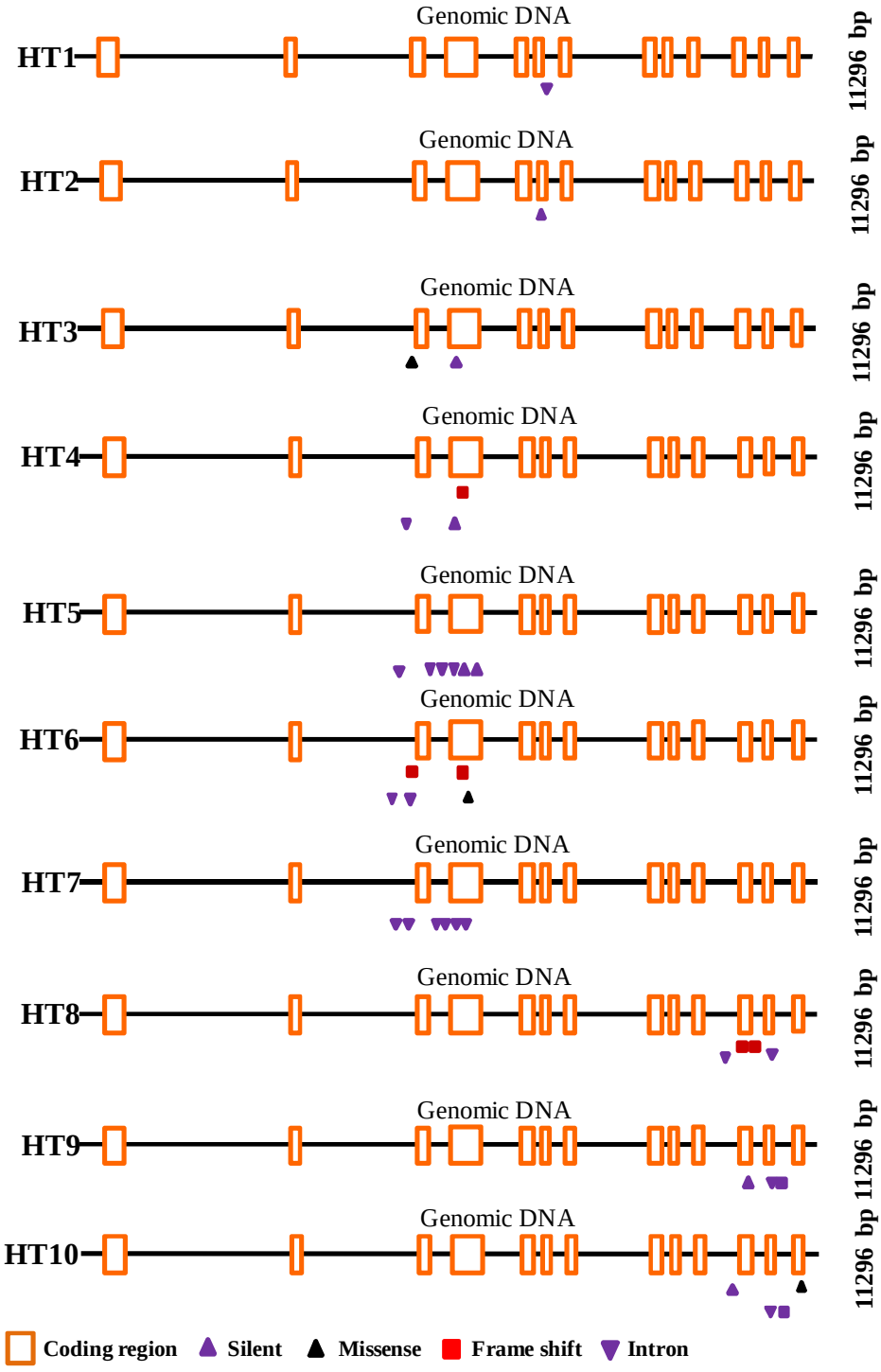


Figure 5.2 PARSESNP output files of the *COP1* haplotypes. Open orange boxes denote exons interconnected with introns (solid black line). Black upright triangle indicates NS changes in the DNA sequence that affect the amino acid. Purple upright triangle indicates S changes that do not affect the amino acid. Purple downward triangles indicate nucleotide changes in introns. Red boxes represent the insertions or deletions. INDELs when present in coding region lead to change in the reading frame.

Table 5.2 The *COP1* haplotypes identified with the respective nucleotide and amino acid changes.

Haplotype	No. of SNPs	Base Change	Amino acid change	Sift Prediction	Accessions Harbouring SNP	Homo/Heterozygous	Nature of the SNP
HT0	-	-	-	-	-	Homo	
HT1	1	G7280A	Intron	-	1	Homo	
HT2	1	A7210G	S377=	-	1	Homo	
HT3	2	G 5848A A5242C	R245Q K153T	-0.214	37	Homo Hetero	Ns Ns
HT4	Multiple	A5112T G5848A T5858TT T6096:	Intron R245Q Frameshift Intron		1	Hetero Homo Hetero Hetero	Ns
HT5	Multiple	A5082G A5382G T5397G G5526A C5636T A5720T T5769C G6031GG	Intron Intron Intron Intron Intron R202= Intron		1	Hetero Hetero Homo Homo Hetero Homo Homo Hetero	S
HT6	Multiple	T5008 C5070T T5215TT G5848A T5858TT	Intron Intron Frameshift R245Q Frameshift	-0.214	1	Hetero Hetero Hetero Hetero Hetero	Ns
HT7	Multiple	A5026C A5051T C5115T T5337A T5343G A5352T A5369AA	Intron Intron Intron Intron Intron Intron		4	Hetero Hetero Hetero Hetero Hetero Hetero Hetero	
HT8	Multiple	C10339T G10826A T10319A G10331GG T10332G T10358G A10381AA A10408AA C10925CACC	Intron Intron Intron Intron V525= Frameshift Frameshift Intron		9	Hetero Homo Homo Hetero Hetero Hetero Hetero Hetero Hetero	S
HT9	Multiple	T10358G G10826A C10925CACC	V525= Intron Intron		2	Hetero Homo Homo	S
HT10	Multiple	T10358G G10826A C10925CACC G11077A	V525= Intron Intron E639K		15	Hetero Homo Homo Hetero	S Ns

or G/A) or transversions (C/A or T/G) and we found that the percent of transitions was more than that of transversion (**Fig. 5.3B**).

The overall distribution of allelic variations was evaluated on the basis of the exons encompassing the domains of COP1 (**Table 5.3**). Exons 3 and 4 coding for coiled coil domain harbored maximum polymorphisms followed by exon 11, 12 and 13 which code for a part of WD40 domain. The exons coding for ring finger domain-exon 1 and 2 did not show any polymorphism. These results are in agreement with the earlier *in-silico* CODDLE predicted analysis i.e. the exons 3 and 4 of *COP1* were more prone to mutations/the most mutable region.

5.2.1.3 Genotype to phenotype co-relation in *COP1* variants

Based on the phenotypic variation observed in the carotenoid content in our Eco-TILLING collection, we identified few accessions (NB515, NB104, NB471 and NB110) that exhibit enhanced carotenoid content when compared to the reference (**Fig. 5.4**). It has been also hypothesized that down-regulation of *COP1* might enhance the carotenoid content of tomato (Liu et al., 2004). In order to link the genotype and the phenotype of the *COP1* variants, we examined the SNPs identified in these accessions. Interestingly, the accessions NB515, NB104, NB471 and NB110 (belonging to HT3) harbored SNPs in the 3rd and 4th exons of *COP1* gene. NB515 (HT11) also harbored polymorphisms in the last three exons-11, 12 and 13 (**Table 5.2**). Though the observed genetic variation in *COP1* gene was low in our collection, our preliminary screening revealed few accessions with high carotenoid content.

In Arabidopsis, *cop1* mutants exhibit constitutively photomorphogenic phenotype in darkness with open cotyledons, short hypocotyls and also accumulate anthocyanins (Deng and Quail, 1992). To examine whether the SNPs observed in tomato COP1 affect its function, we monitored the seedling phenotype of these accessions in both light and dark conditions. The accessions NB24 and NB32 were also included as they harbor INDELs that might cause frame shift. However, for some unknown reasons, we could not germinate the seeds of NB32. The seedlings of accession IIVR 53 (which harbored intronic and synonymous SNPs in *COP1*) and *hp1* (with a lesion in *DDB1* gene) mutant were also considered as controls along with the reference WT AV. The length of hypocotyls and roots was measured after 7 days. No significant difference was observed in the length of hypocotyls or roots in the

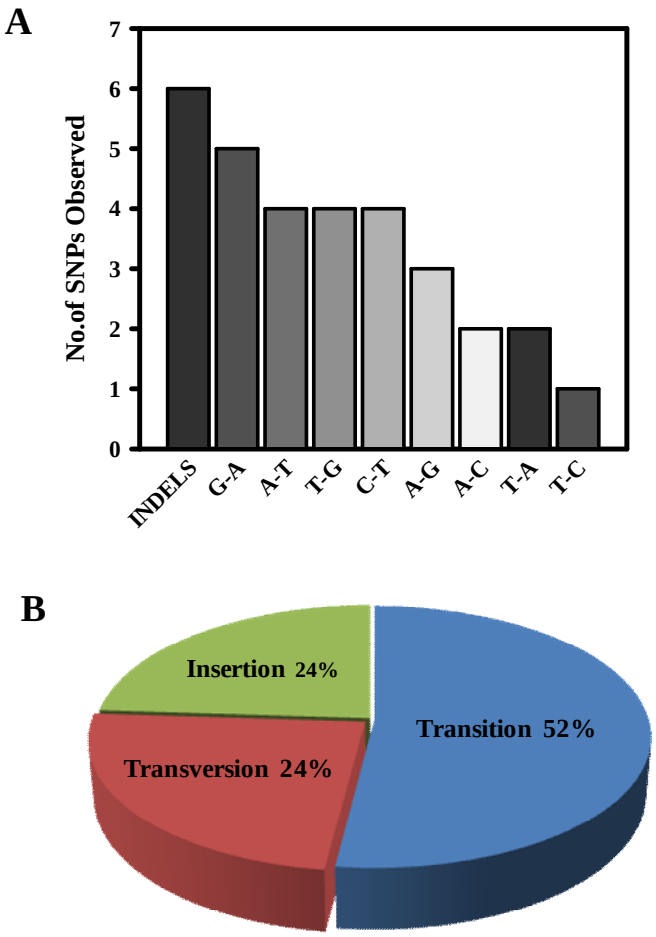


Figure 5.3 The distribution of SNPs in *COP1* gene. A, the type of nucleotide base changes observed and **B**, the frequency of different types of base changes.

etiolated seedlings of these accessions **(Fig. 5.5**) however; in some of the accessions- NB110, NB104 and NB471, the de-etiolated seedlings exhibited slight reduction in hypocotyl length (**Fig. 5.6**) as compared to the reference. The seedling anthocyanin levels were not significantly different when compared with the reference.

In our study, in spite of low genetic diversity in tomato *COP1*, we have identified few accessions which exhibited enhanced carotenoid content. These accessions can be potentially used as a starting material to develop new varieties.

Table 5.3 The distribution of SNPs in *COP1* gene and the corresponding domains of the protein.

Candidate gene	Region screened	Genomic co-ordinates (bp)	Primer Used	SNPs identified
COP1	Ring finger (Exon 1)	-269 to 1141	SETVI	-
	Ring finger (Exon 2)	1972 to 3224	SETV	-
	Coiled coil (Exon 3 & 4)	4958 to 6187	SETIV	22
	WD40 (Exon 5, 6 & 7)	6441 to7606	SETIII	2
	WD40 (Exon 8, 9 & 10)	8346 to 9850	SETII	-
	WD40 (Exon 11, 12 &13)	10284 to 11378	SETI	9

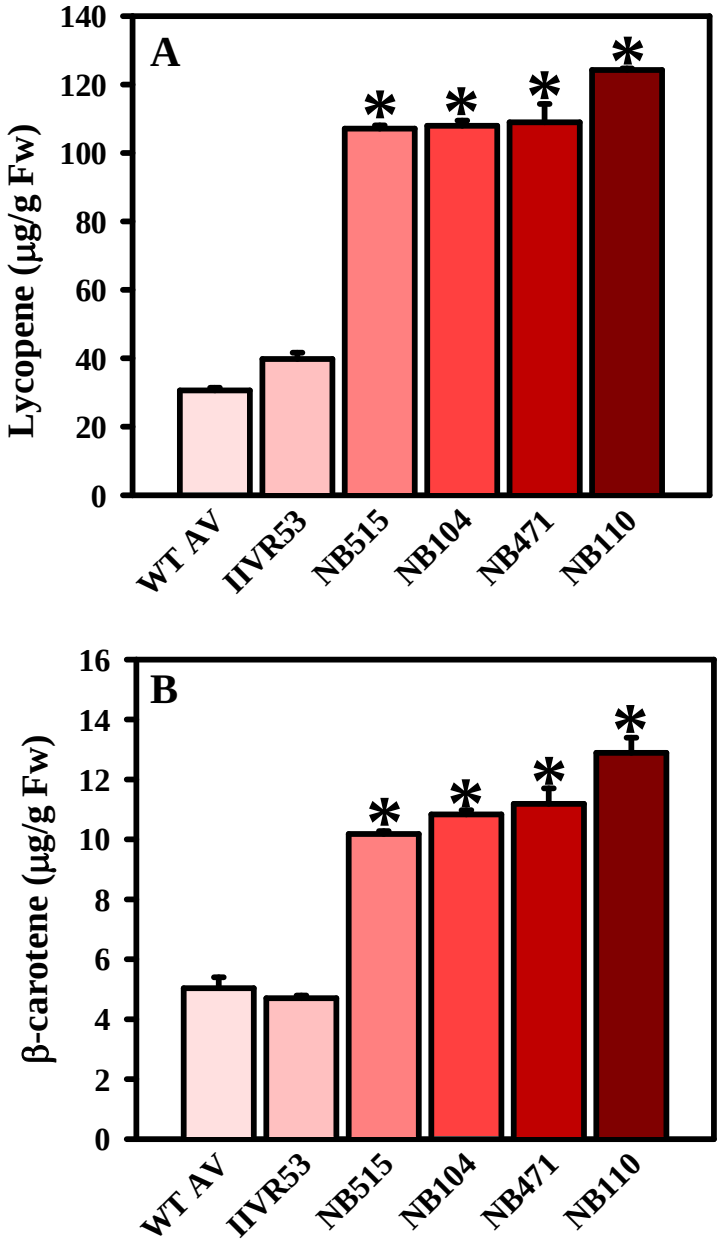


Figure 5.4 Carotenoid content in *COP1* variants. **A**, Lycopene and **B**, β-carotene content measured at red ripe stage in *COP1* variants. Carotenoid levels in the accessions NB515, NB104, NB 471 and NB110 are enhanced. n ≥ 3 ± S.E where n= fruit number, * = statistical significance with p≤ 0.05.

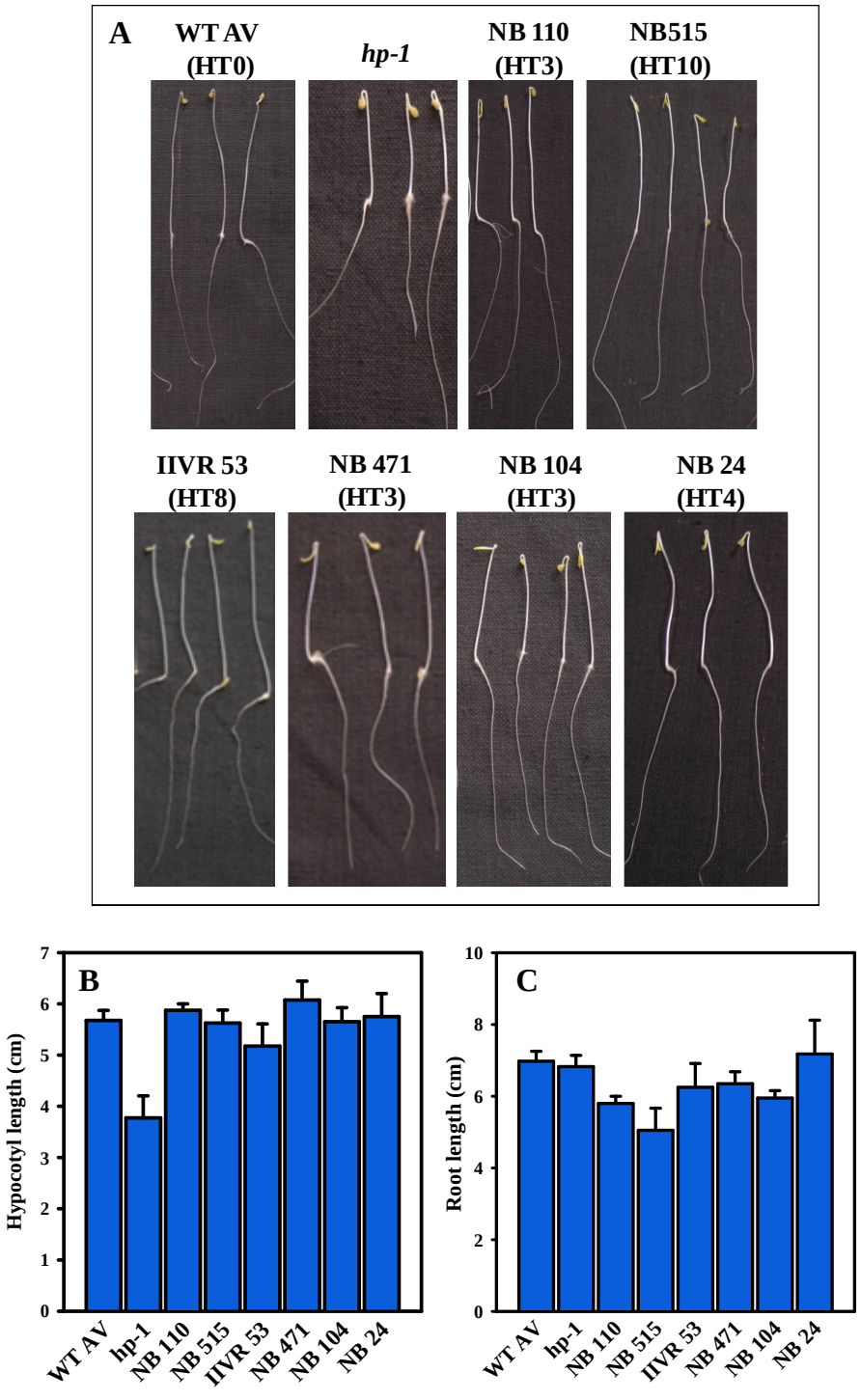


Figure 5.5 Phenotype of etiolated seedlings of *COP1* variants. **A**, dark grown seedlings of *COP1* variants. **B** and **C**, Comparison of hypocotyl and root lengths of reference (AV) and *COP1* variant seedlings grown under darkness. n ≥ 30 ± S.E using 3 x (10 seedlings) each time.

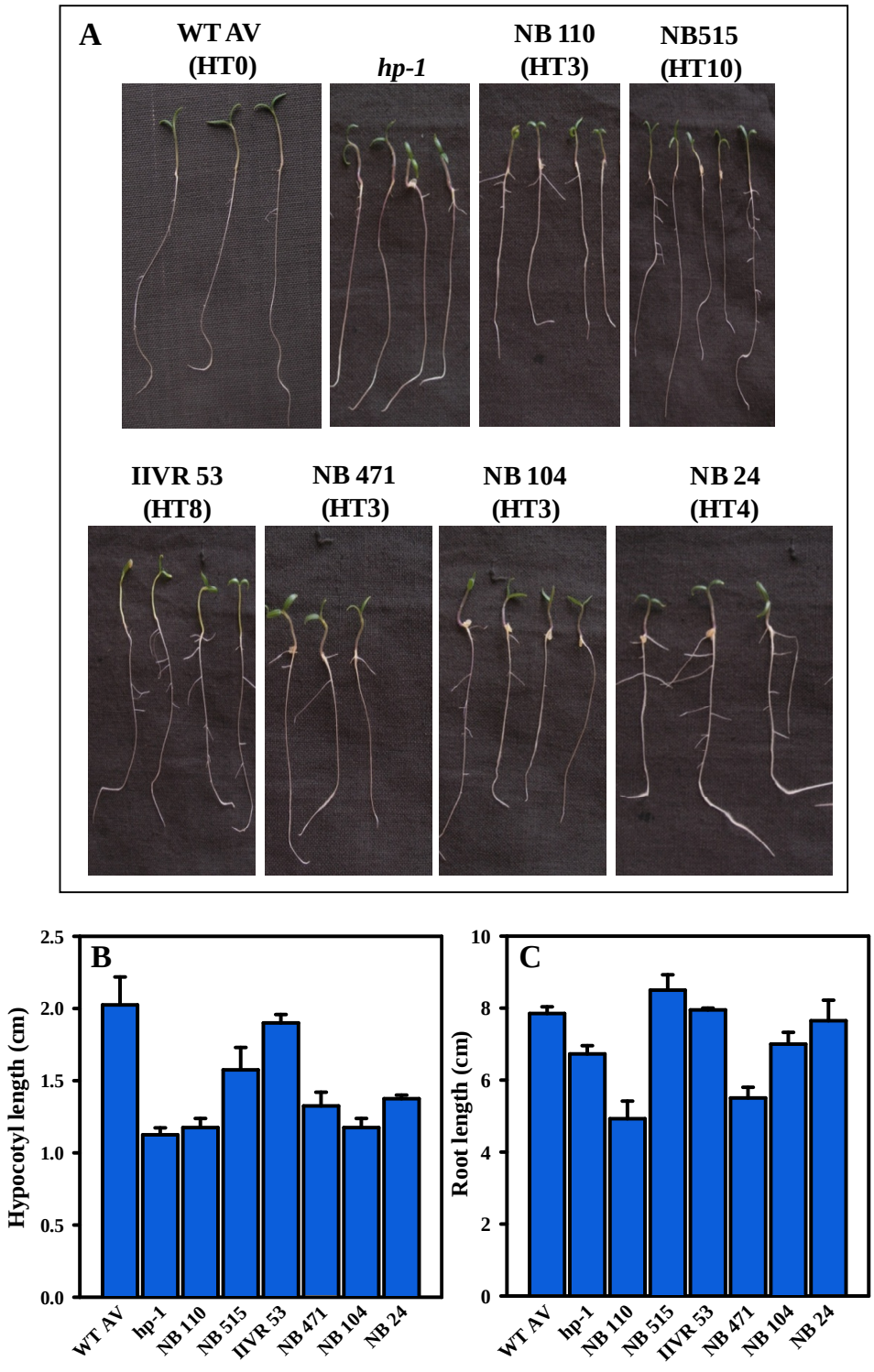


Figure 5.6 Phenotype of light grown seedlings of *COP1* variants. A, light grown seedlings of *COP1* variants. **B** and **C**, Comparison of hypocotyl and root lengths of reference (AV) and *COP1* variant seedlings grown under light. n ≥ 30 ± S.E using 3 x (10 seedlings) each time.

5.2.2 Eco-TILLING of tomato *SPA3LIKE* gene

With reference to the work done by Liu et al., 2004 and the evidence from tomato *hp1* and *hp2* mutants that light signaling components regulate the accumulation of carotenoids in tomato, we screened our Eco-TILLING population for natural variations in *SPA3LIKE* gene. *SPA3LIKE* in tomato is a 1028 bp gene with no defined introns. Owing to its negative regulatory role in photomorphogenesis (Liu et al., 2004) and the presence of WD40 domains, *SPA3LIKE* was initially considered as an ortholog of *COP1* in tomato. Previously it has been designated as *COP1LIKE* as it has similar role like *COP1*. In the recent update, tomato genome sequence was reannotated and the nomenclature of *COP1LIKE* was changed to *SPA3LIKE*. Consistent with Ranjan et al., (2014) the usage of NCBI Blast program (**Fig. 5.7**) revealed that the SPA3LIKE protein is closely related to proteins that consist of COP1/SPA like WD40 domains. The SPA3LIKE proteins also shared 61-81% amino acid sequence identity with the Arabidopsis SPA like protein called RUP2 (Repressor of UVR8 protein 2). The interpro scan results predicted that tomato SPA3LIKE and Arabidopsis RUP1 and RUP2 proteins consist of WD40 repeats (**Fig. 5.8**) with apparently no other domains. Taking into account the blast results and the structural predictions, it is plausible that tomato *SPA3LIKE* could be a functional ortholog of *AtRUP2*. Nevertheless, further studies to functionally complement the role of *AtRUP2* by *SPA3LIKE* would add wealth of information to understand the UV-B mediated responses in tomato.

Upon Eco-TILLING 0.6 million bp for *SPA3LIKE* gene, four accessions- NB182, NB518, NB8 and NB189 were identified as allelic variants on denaturing gels using Li-COR system. To determine the base changes in the polymorphisms observed in *SPA3LIKE* gene, six individuals from these accessions were grown and screened. Among these, homozygous individuals were selected by mismatch cleavage (**Fig. 5.9**). The cleavage pattern observed in these lines after CELI digestion was almost similar on the Li-COR gel. However, we observed differences in the number and the position of polymorphisms after sequencing. This discrepancy between Li-COR gel based screening and sequencing of individual accessions could be due to the presence of mutations in nearby sites, and also due to the presence of INDELs.

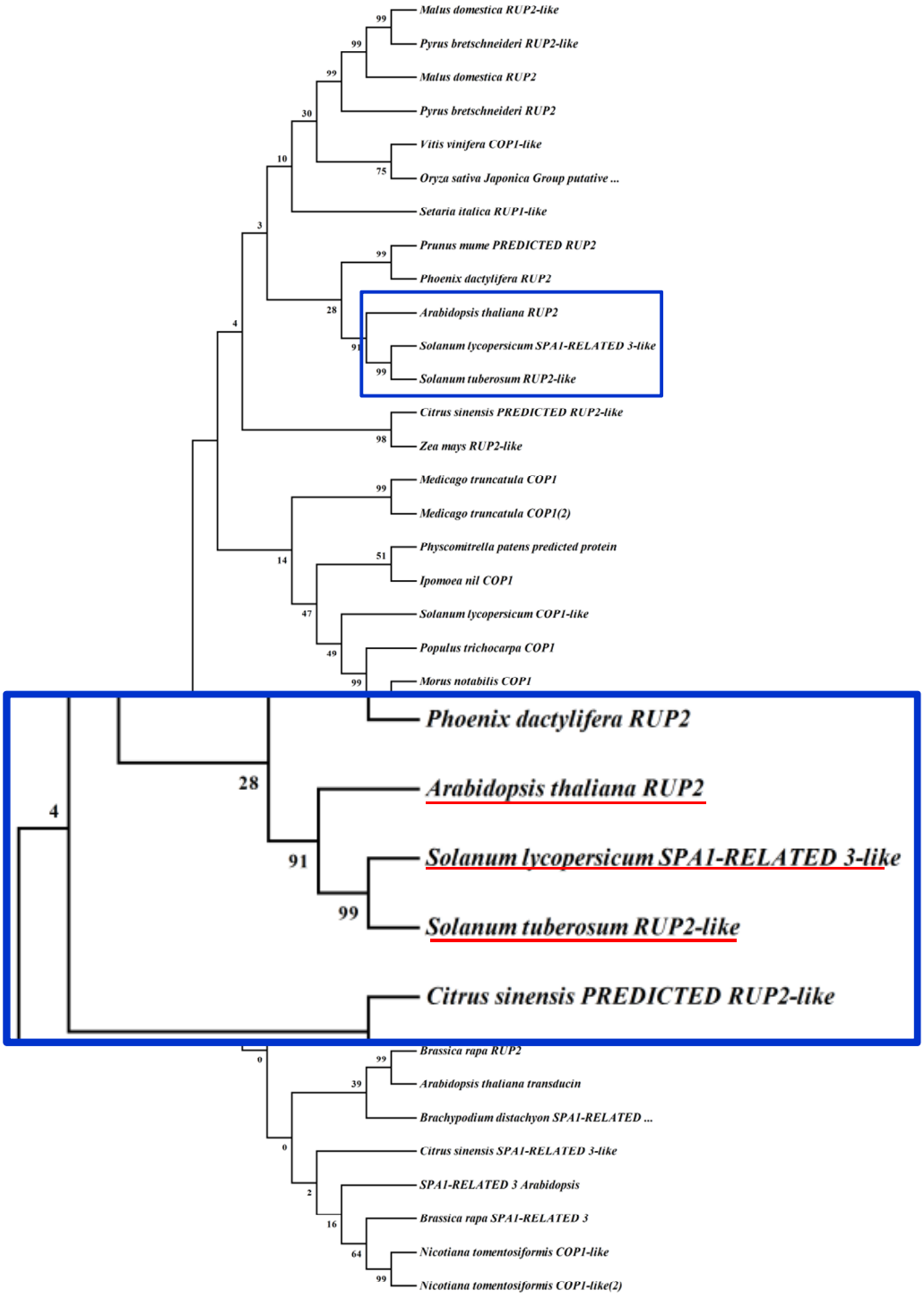


Figure 5.7 Phylogram representing the NCBI protein Blast results of tomato SPA3LIKE protein. The phylogram combines the phylogenetic relationships between the species analyzed which were obtained by maximum likelihood. Branch lengths are not in proportion to evolutionary times. All the proteins have WD40 repeats and inset shows that tomato SPA3LIKE protein is closely related to Arabidopsis RUP2 and Potato RUP2.

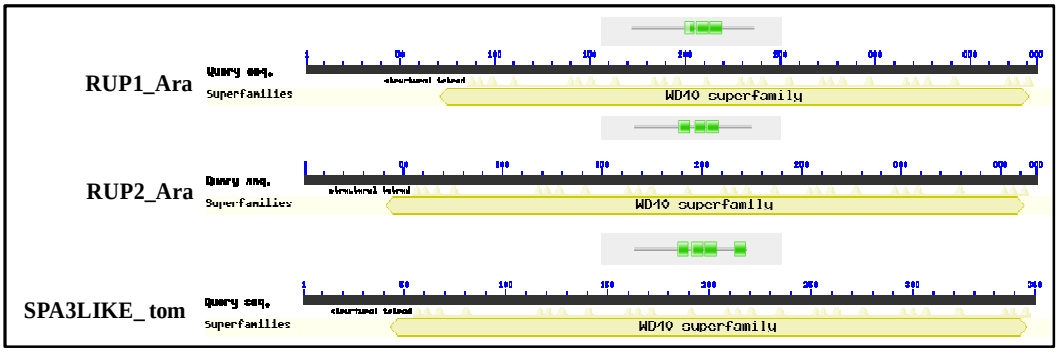


Figure 5.8 Domain structure of tomato SPA3LIKE and Arabidopsis RUP1 and RUP2 proteins. The conserved domains structure of the three proteins was taken from the data base at NCBI. The predictions of Pfam suggest the presence of three WD40 repeats represented by green blocks in Arabidopsis RUP1 and RUP2, whereas in tomato SPA3LIKE there are four WD40 repeats.

5.2.2.1 Distribution and frequency of SNPs in *SPA3LIKE* gene

The analysis of nucleotide sequences obtained for *SPA3LIKE* variants showed 17 changes. The majorities of the changes were specific and were represented by one accession only. Therefore a total of 5 haplotypes – HT1, HT2, HT3, HT4 and HT0 were designated for *SPA3LIKE* gene. The accessions which did not show any cleavage products on gel, including the reference were designated as HT0. PARSESNP analysis of the four different haplotypes (represented by an accession) revealed the distribution of the identified SNPs (**Fig. 5.10**). Six of the SNPs were found to be synonymous (purple triangles) and in addition, four frame shift changes (red boxes) were also observed. Ten SNPs were nonsynonymous and caused a change in the amino acid (black triangles). With an overall distribution of 1 SNP/24 Kb and 1 INDEL/124 Kb in *SPA3LIKE* gene, the frequency of nonsynonymous changes was double than the synonymous changes (**Fig. 5.11A**). Also the number of transversions was high when compared to transitions or INDELs (**Fig. 5.11B**).

Unlike COP1, SPA3LIKE is a single exon product and any drastic variation in the protein sequence would hamper the protein function. In order to predict the effect of changed amino acids, we used PARSESNP and SIFT tools. An amino acid change

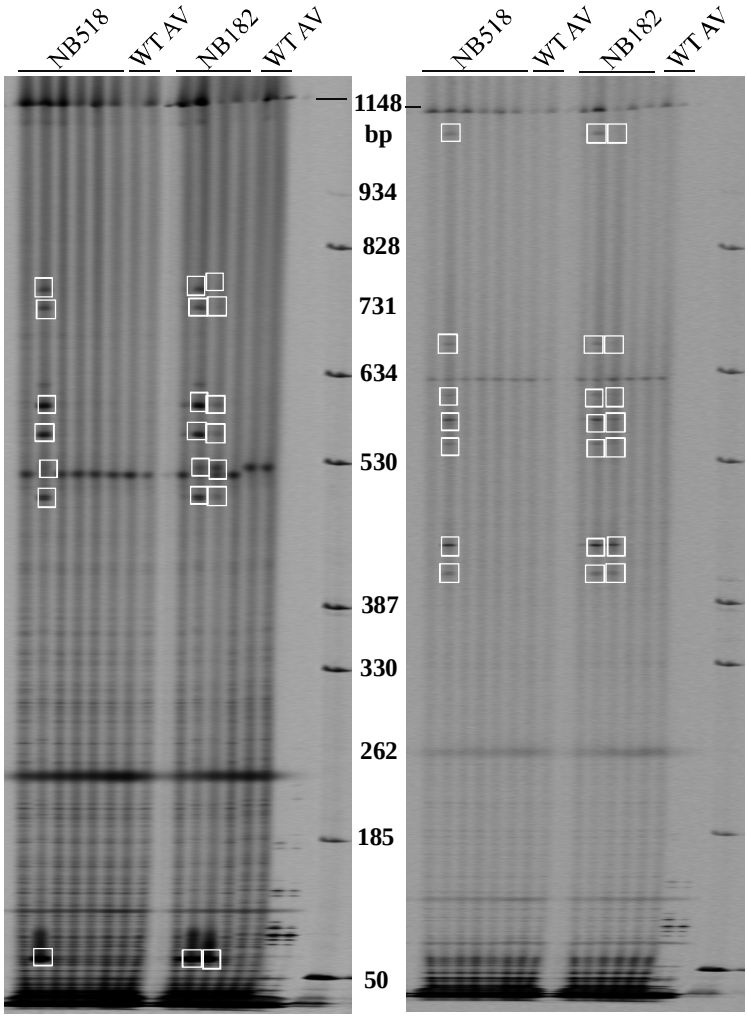


Figure 5.9 Confirmation of SNPs in individual plants of tomato *SPA3LIKE* variants. DNA from six individuals of the Eco-TILLING accessions was used for confirmation of SNPs. The complementary fragments are marked by boxes in both the channels. The open white boxes highlight the lane with SNPs. The amplicon (1148 bp) is detected as thick band at the top of the gel in both the IR 700 and 800 dye channels.

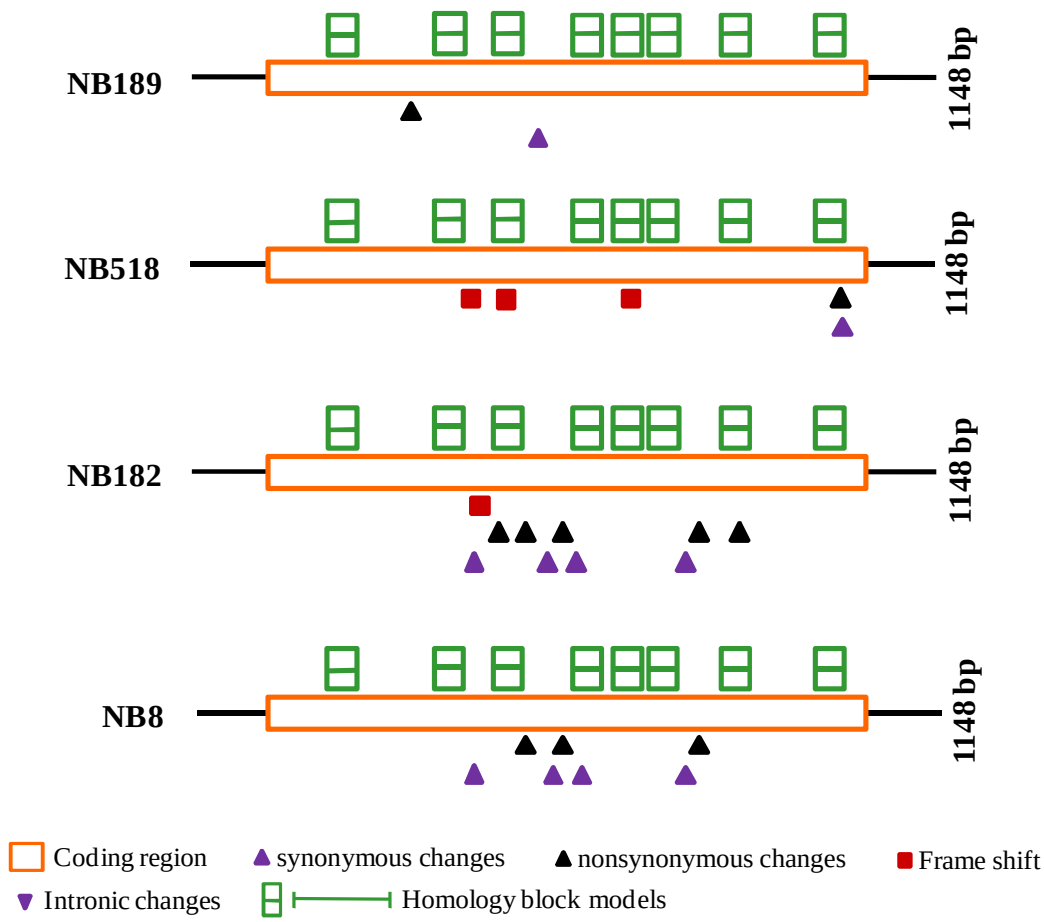


Figure 5.10 PARSESNP output file representing the distribution of SNPs in allelic variants obtained for *SPA3LIKE* gene. Open orange boxes denote exons interconnected with solid orange line introns. The green boxes and lines above gene model represent the homology block models. Black upright triangle indicates changes in the DNA sequence that affect the amino acid. Purple upright triangle indicates changes in the DNA sequence that does not affect the amino acid. Purple downward triangles indicate changes in introns. Red boxes represent the frame shift changes.

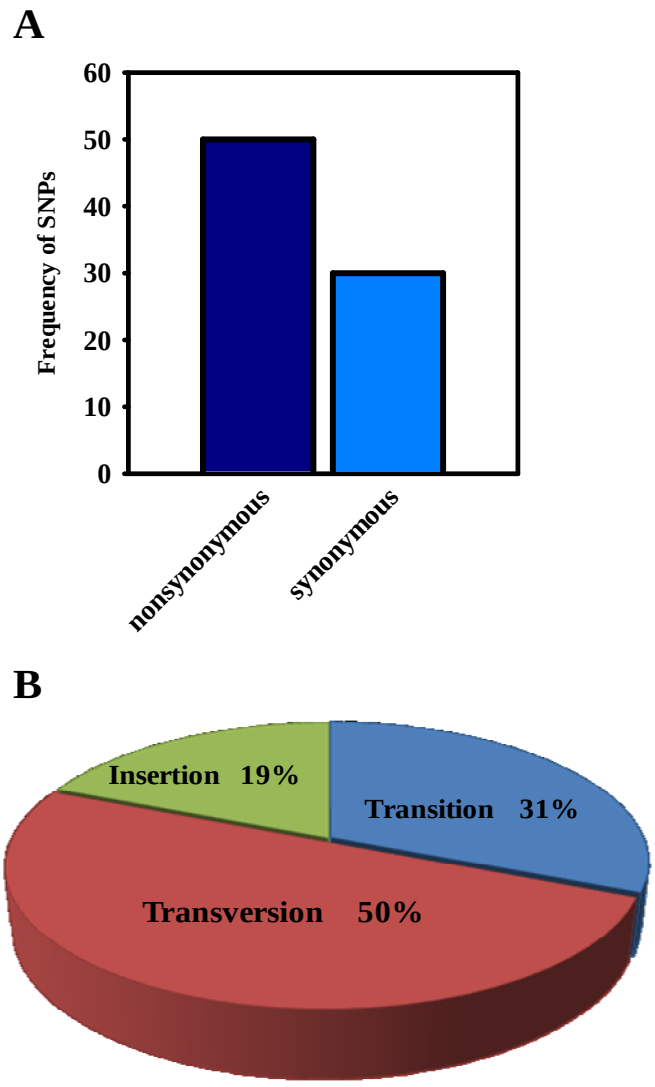


Figure 5.11 A, Frequency of nonsynonymous and synonymous mutations. **B**, Pie diagram illustrating the types of base change and their percent occurrence.

Table 5.4 Sequence based detection and genotyping of *SPA3LIKE* allelic variants.

Plant ID	Nucleotide change	Aminoacid change	SIFT Prediction
NBPGR8	G1460C	S154=	
	G1531A	S178N	-
	C1562A	T188=	
	G1567A	G190E	-
	C1592T	V198=	
	T1703C	H235=	
	G1731C	V245L	0.31
NBPGR182	A1513C	D172A	0.68
	G1531A	S178N	0.96
	G1567A	G190E	1.00
	G1731C	V245L	0.89
	T1787A	N263K	1.00
	C1487CC	Frame shift	-
NBPGR518	T1454TT	Frame shift	
	G1512GG	Frame shift	
	T1688TT	Frame shift	
	C2023G	T342R	0.00
	A2024T	T342=	
NBPGR189	T1326A	L110I	-
	G1568A	G190=	

from tyrosine at 342 to arginine was observed in NB518 and this change was predicted to be deleterious with a SIFT score of ≤0.05. The rest of the nonsynonymous changes does not appear to cause drastic changes in protein function as the SIFT scores obtained were tolerable (**Table 5.4**).

5.2.2.2 Analysis of the predicted SPA3LIKE protein sequence of the identified haplotypes

Apart from the nonsynonymous SNPs, INDELs in the coding region cause frame shifts while the protein is translated, thereby affecting the structure and function of the protein. To examine the effect of frame shifts on SPA3LIKE function, the coding DNA sequence (CDS) from all the haplotypes of *SPA3LIKE* was translated using EXPASY Translate, a bioinformatic tool. The reference SPA3LIKE protein is comprised of 361 amino acids with a predicted molecular weight of approximately 39.71 KDa. It has four WD40 domains spanning from amino acids 140-172, 185-221, 226-263 and 322-356.

When the protein sequences from all the haplotypes were multi-aligned, the N-terminal region was found to be highly conserved in all the haplotypes and most of the SNPs affected the C-terminal part of the protein (**Fig. 5.12**). The three insertions observed in NB518 resulted in a frame shift which prematurely terminated the protein after 169th amino acid. Interestingly, a single insertion in NB182 also resulted in a truncated protein of 164 amino acids due to a premature stop codon (**Fig 5.13**). To determine how SNPs would have altered the function of SPA3LIKE, the impact of frameshift mutation on the secondary structure of protein was examined. As no crystal structure of tomato SPA3LIKE was available, the secondary structure of the protein was predicted using the protein modeling (Swiss Prot) software. The energy calculation and validation of the predicted protein models was done by Ramachandran Plot with an R factor of > 90%. The R factor determines the quality of the model and the greater the percent, better is the prediction. The sturcture of WT, NB518 and NB182 do not look identical. WT has a cage like structure with two α-helices and due to mutation, there is a loss of 1 α-helix in NB518 and in NB182, both the α-helices were lost (**Fig. 5.14** Left Panel **A-C**). The binding activity of WT and variant proteins was also predicted using FT site software. Both WT and variant proteins have 3 active binding sites. Due to the distortions in the secondary structure of NB518 and NB182,

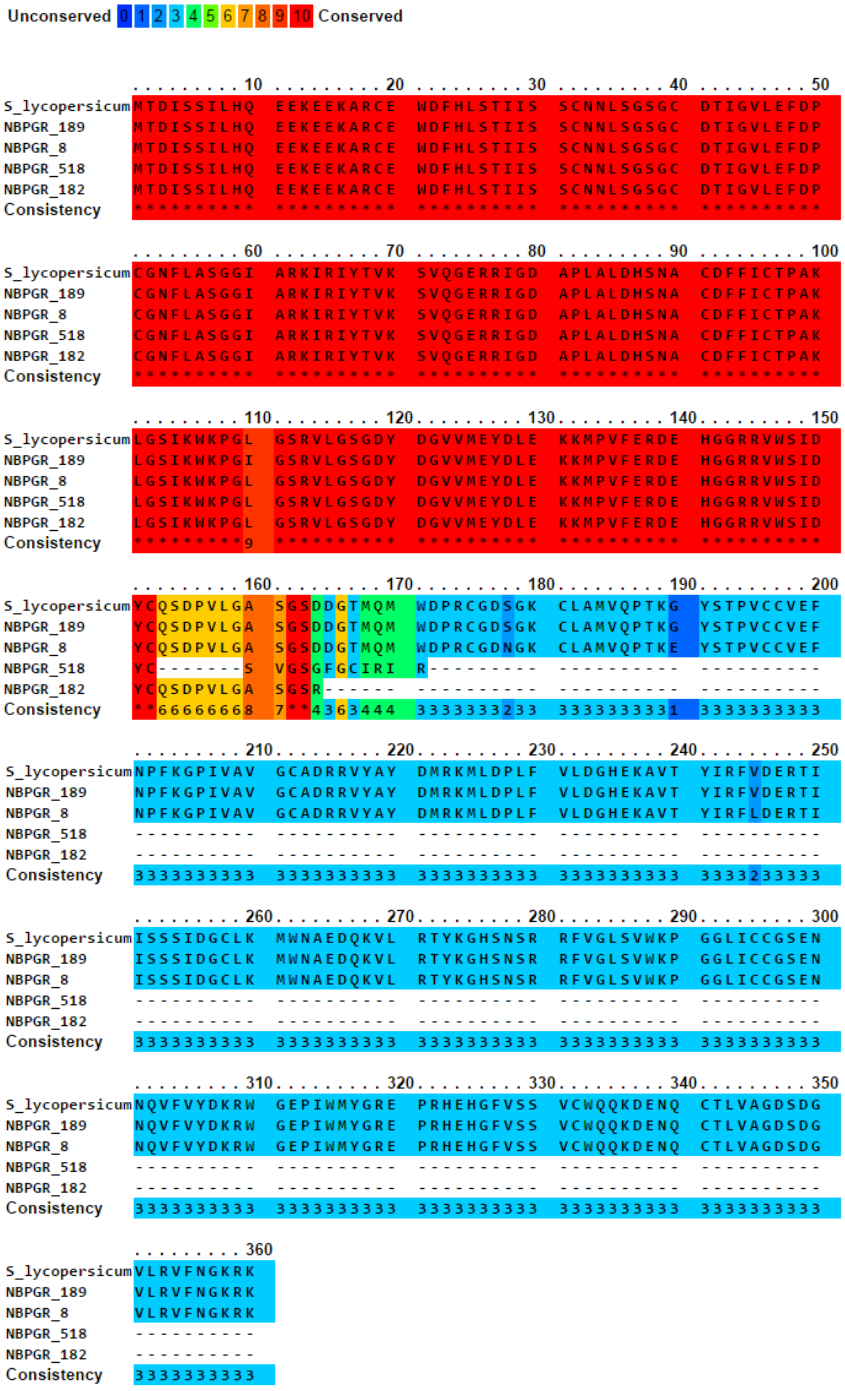


Figure 5.12 PRALINE multiple alignment output file of the derived SPA3LIKE protein sequences of haplotypes. The conservedness of the amino acids is denoted by a color code designed by the software. The highly unconserved moieties are denoted by dark blue color and on the contrary, conserved moieties by deep red color.

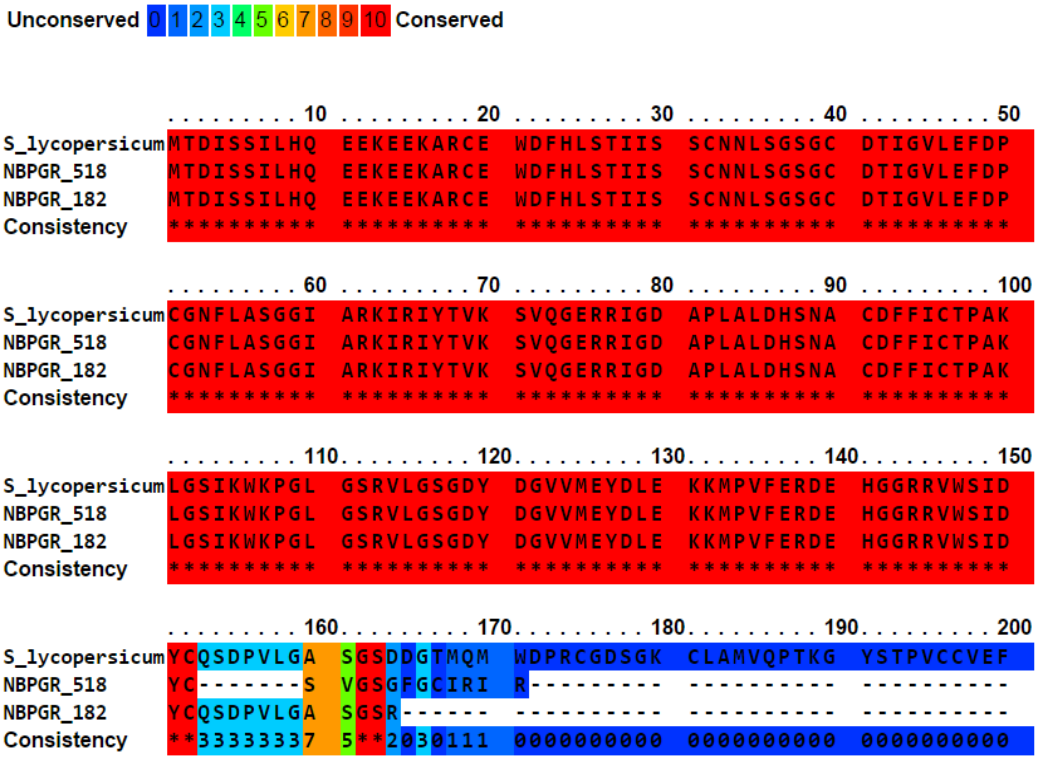


Figure 5.13 Multiple alignment of the predicted proteins of NB518 and NB182 which are truncated after 169th and 164th amino acid respectively.

the binding sites were disrupted as their orientations changed (**Fig. 5.14** Right panel **D-F**). It is likely that this structural modification may have reduced the affinity for binding of the interacting proteins with SPA3LIKE from NB518 and NB182, resulting in loss of function.

5.2.2.3 Marker based validation of genetic background of *SPA3LIKE* haplotypes

To check the genetic background of the two accessions NB518 and NB182 for further characterization and in order to validate the variants at genotype level, a preliminary marker based screening was attempted using SSR markers that are polymorphic to the reference- WT (AV), tomato cultivar Ailsa Craig (AC) and the tomato wild relative *S. pimpinellifolium*. The two SSR markers - L100042i that are polymorphic between AC and AV and U21085 (polymorphic between AV and *S. pimpinellifolium*) were amplified using the respective primers on the genomic DNA of the variants, the cultivars and *S. pimpinellifolium*. As evident from the electrophoretic separation of PCR products (**Fig. 5.15A**), the amplified products of the marker L100042i from the variants were similar to AV and *S. pimpinellifolium* but polymorphic to AC. The marker U21085 was polymorphic between AV and *S. pimpinellifolium* (**Fig. 5.15B**). Above analysis suggested that the genetic background of the variants NB518 and NB182 is likely similar to AV than AC and *S. pimpinellifolium*. Considering the above results, AV was used as a reference cultivar along the *SPA3LIKE* variants for further studies.

5.2.2.4 Morphological analysis of *SPA3LIKE* haplotypes

The homozygous individuals of NB518 and NB182 accessions harboring the SNPs were identified. The molecular nature of these two haplotypes was revealed by sequencing of 1028 bp of *SPA3LIKE* gene with specific primers. Since an earlier study showed that *COP1LIKE RNAi* line exhibited defects in photoresponses and increased lycopene levels (Liu et al., 2004), we also examined the phenotype of these lines both at the seedling and the adult stages.

5.2.2.4.1 Seedling morphology

The hypocotyl and root lengths of reference (AV) and variants seedlings were compared by growing germinated seedlings on vertically oriented petriplates containing 0.8% (w/v) agar. Seedlings were grown in the darkness and white light

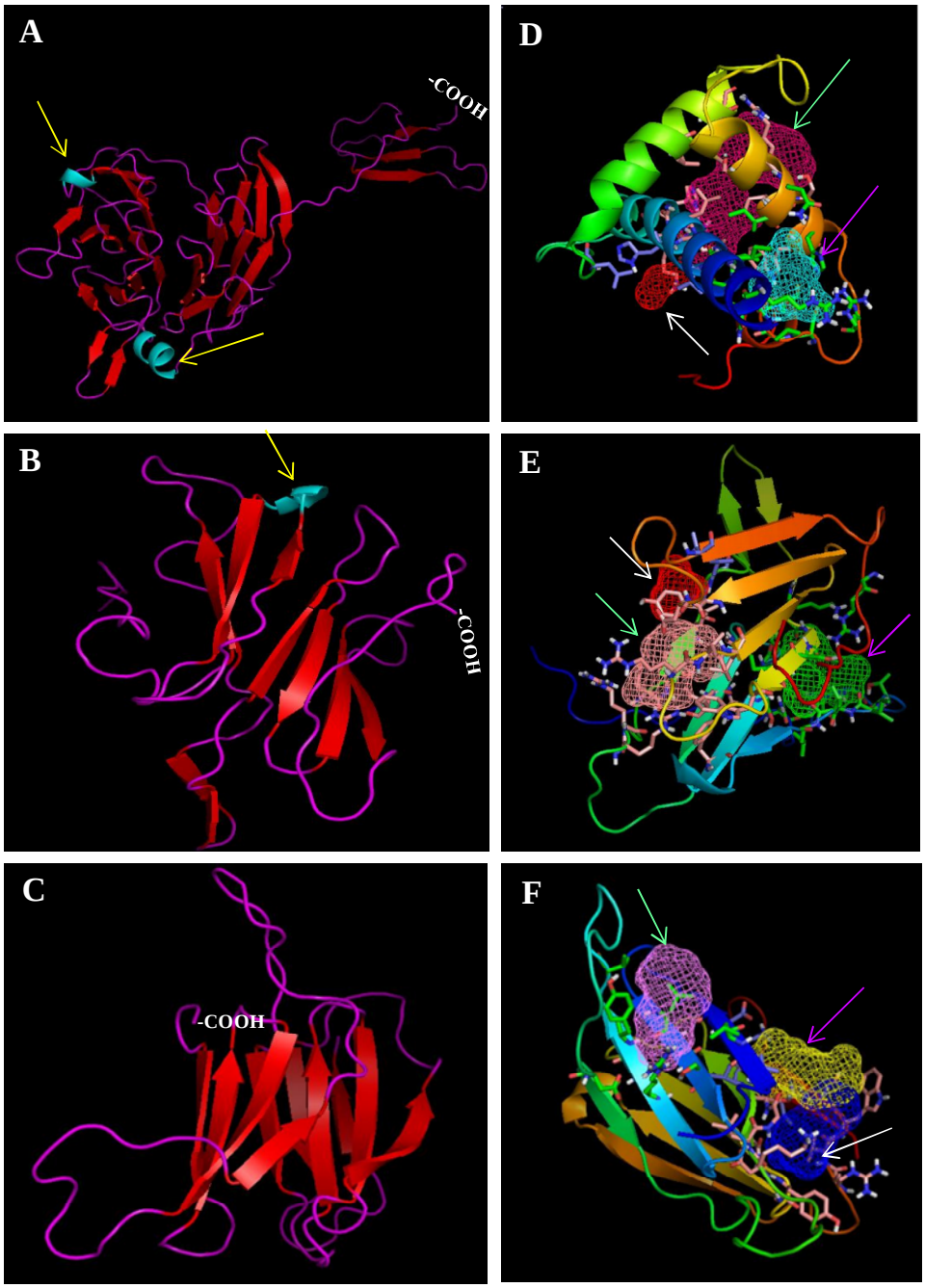


Figure 5.14 Predicted models of WT, NB518 and NB182 SPA3LIKE proteins. The left panel (A, B and C) show the secondary structure of WT, NB518 and NB182 SPA3LIKE protein. **Note,** in NB518, the polymorphisms caused loss of one α-helix and changed the orientation of other helix in the protein (marked with yellow arrow); in NB182, both the α-helices are lost. The right panel (D, E and F) show the secondary structure prediction with binding site of protein. The model predicts the presence of three binding sites in WT & the variants. However, the binding sites of variant proteins were disrupted, due to the change in orientation. The binding sites are represented by green, white and pink arrows.

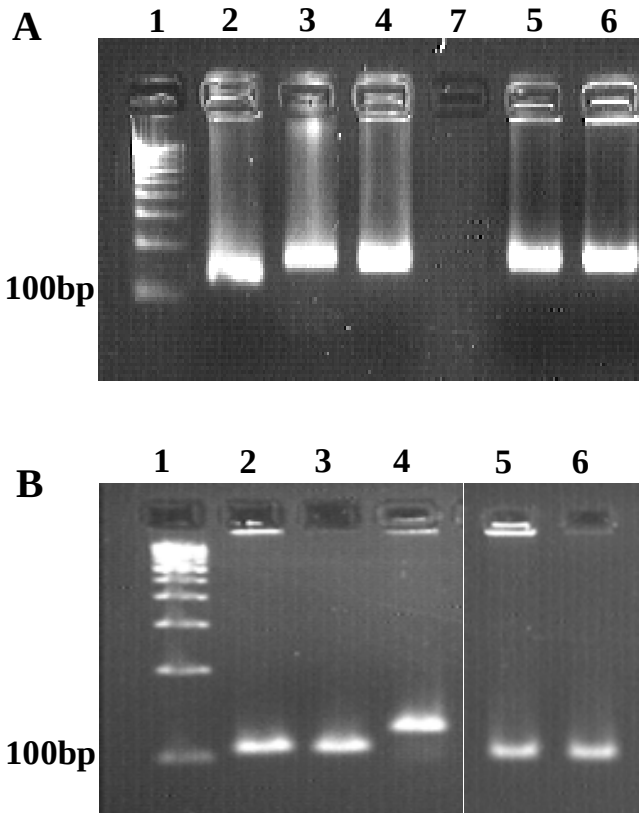


Figure 5.15 SSR marker analysis. A-SSRL100042i marker is polymorphic between AV and AC but not with *S. pimpinellifolium*, NB182 and NB518. B-SSRu21085 marker is polymorphic between AV and *S. pimpinellifolium* and not with others. A and B state that the *SPA3LIKE* variants NB182 and NB518 are also similar to AV. Lane 1- DNA Marker, Lane 2- AC, Lane 3- AV, Lane 4- *S. pimpinellifolium*, Lane 5- NB182, Lane 6- NB518 and Lane 7- empty.

(16/8 h day and night cycle) for three days and thereafter the hypocotyl and root lengths were measured. As expected, the hypocotyls were more elongated in the etiolated seedlings of both reference and variants. However, the 3-day old etiolated seedlings of the variant NB182 showed significant difference in the hypocotyl and root length as compared to the reference seedlings (**Fig. 5.16**). In case of NB518, no significant difference in etiolated seedlings was observed in the hypocotyl length as compared to reference. Interestingly, when grown under white light, seedlings of NB182 exhibited a 2-day delay in the opening of apical hook. The roots of both de-etiolated seedlings of NB518 and NB182 showed little or no difference as compared to the reference, whereas, hypocotyl of de-etiolated NB182 seedlings was shorter than the reference (**Fig. 5.17**). However, unlike the *cop1* mutants of Arabidopsis, the above accessions did not show any increased anthocyanin accumulation (**Table 5.5**).

5.2.2.4.2 Adult plant morphology

When grown under greenhouse conditions, the morphology of the allelic variants of *SPA3LIKE* gene was almost similar to the reference (**Fig. 5.18**). The chlorophyll a content measured spectrophotometrically was slightly high in NB182 as compared to reference (**Fig 5.18**). The flowering time in the variants was not altered as compared to the reference cultivar. 5-6 numbers of flowers were borne per inflorescence in the variants and reference. Though the presence of exerted stigma, a characteristic feature of Arka Vikas was not observed in the variants, the fruit set was similar in the variants to that of AV.

Table 5.5 Anthocyanin content in 7-day old light grown seedlings.

Genotype	Anthocyanin (A535/10 seedlings)
WT (AV)	0.3356 ± 0.003
NB518	0.2985 ± 0.098
NB182	0.4009 ± 0.006

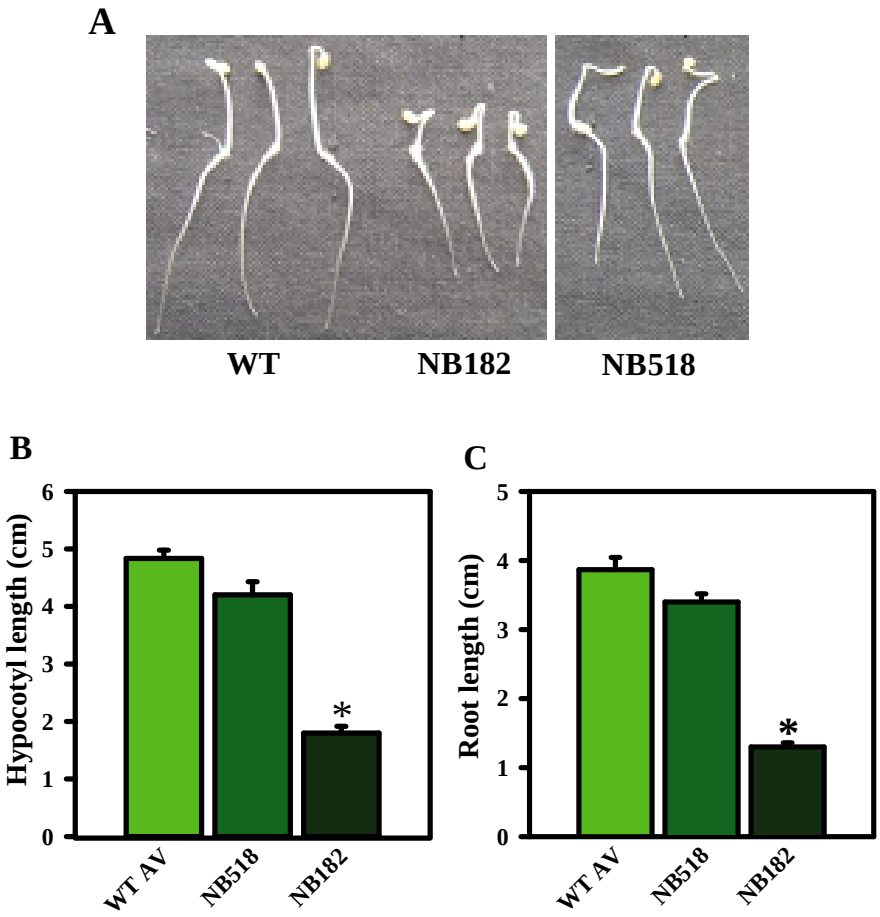


Figure 5.16 Phenotype of etiolated seedlings of *SPA3LIKE* variants. **A**, 3-day old dark grown seedlings of WT and *SPA3LIKE* variants. **B** and **C**, Comparison of hypocotyl and root lengths of reference (AV) and *SPA3LIKE* variant seedlings grown in darkness. Statistical significance is represented by * for P value ≤ 0.05.

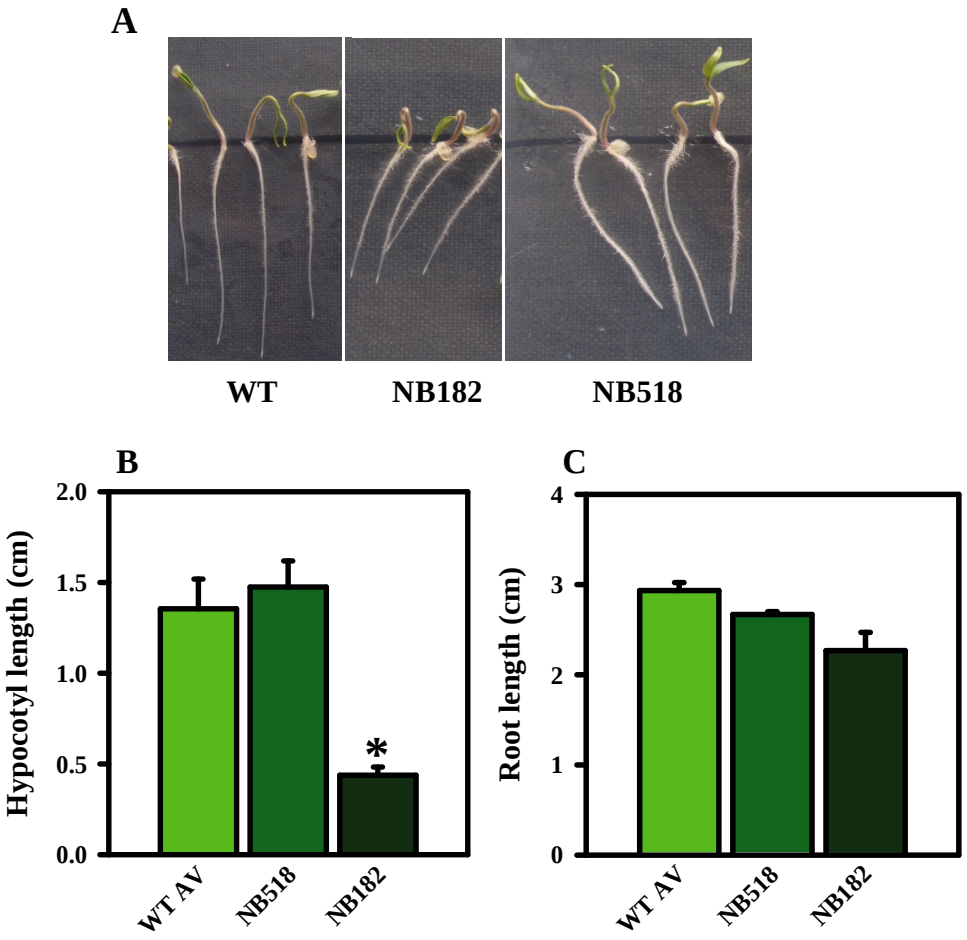


Figure 5.17 Phenotype of de-etiolated seedlings of *SPA3LIKE* variants. **A**, 3-day old light grown seedlings of WT and *SPA3LIKE* variants. **B** and **C**, Comparison of hypocotyl and root lengths of reference (AV) and *SPA3LIKE* variant seedlings grown under light. Each histogram is a representation of the mean value ±S.E of three independent experiments using 10 seedlings each. Statistical significance is represented by * for P value ≤ 0.05.

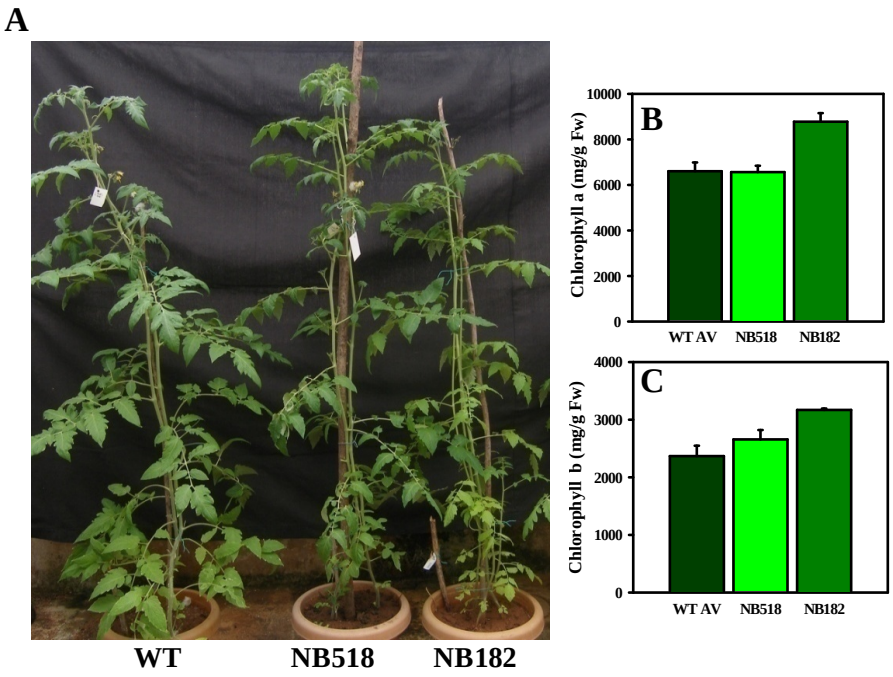


Figure 5.18 Phenotype of adult plants of *SPA3LIKE* variants. A, Two month old plants of NB518 and NB182 in comparison with AV. **B** and **C** represent the content of leaf chlorophyll a and b respectively. NB182 has slightly higher amount of chlorophyll a and b.

5.2.2.5 Fruit morphology

The molecular and genetic studies in tomato suggest that light has a role in fruit ripening (Alba et al., 2000). Studies have also shown that manipulating tomato light signal transduction genes can result in altered fruit carotenoid content (Liu et al., 2004). To examine whether *SPA3LIKE* gene affects fruit ripening, the chronology of the fruit development of *SPA3LIKE* variants was examined under greenhouse conditions and was found to be similar to that of the reference. In the course of fruit development, the days required for post anthesis to attain mature green stage (MG) was 26, and it took another three days to reach breaker stage (BR) (29 days) and then, the red ripe (RR) stage was achieved at 39 days. However, the size of the fruits varied drastically (**Fig. 5.19**). The perimeter of the fruits at different ripening stages was calculated using Tomato analyzer. **Figure 5.20** showed that the size of fruit in NB518 and NB182 was almost half in comparison with AV. Carotenoid profiling at different ripening stages of the variants was carried by UHPLC. Our analysis showed increased levels of phytoene, trans-lycopene, β-carotene and lutein in the variants when compared to the reference (**Fig. 5.21**). Interestingly we also identified gamma carotene in the red ripe fruits of the variants. In concordance with the published results of *COP1LKE* gene silenced lines (Liu et al., 2004), our results also show that *SPA3LIKE* variants exhibit enhanced carotenoid content in red ripe fruits of tomato. These promising lines can be exploited further to develop a carotenoid rich cultivar.

5.2.2.6 Gene expression analysis in *SPA3LIKE* haplotypes

To ascertain that the observed phenotypes correlated with the frame shift mutations, total RNA was isolated from three different fruit developmental stages- MG, BR and RR. Transcript levels of *SPA3LIKE* were analyzed by RT-PCR. In general, consequences of frame shift mutations will be observed at the translation rather than transcription of a coding sequence (Plotkin and Kudla, 2011). As a result, altered or truncated protein would be formed that may not function normally. In congruence with the frame shift polymorphisms identified in NB518 and NB182, we did not observe any significant difference in the transcript levels of *SPA3LIKE* as compared to the reference AV (**Fig. 5.22**). Since *COP1* is a close ortholog of *SPA3LIKE*, we also checked the transcript levels of *COP1*. The *COP1* transcript

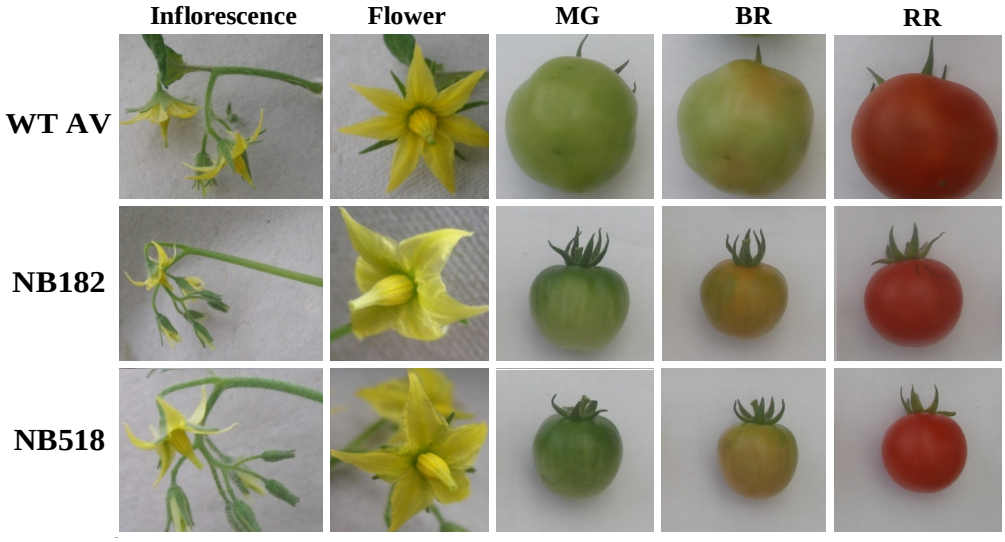


Figure 5.19 Flower and fruit morphology of NB518 and NB182 in comparison with AV. Fruits of the variants were small in size in all the three stages of development; mature green (MG), breaker (BR) and red ripe (RR) when compared with AV.

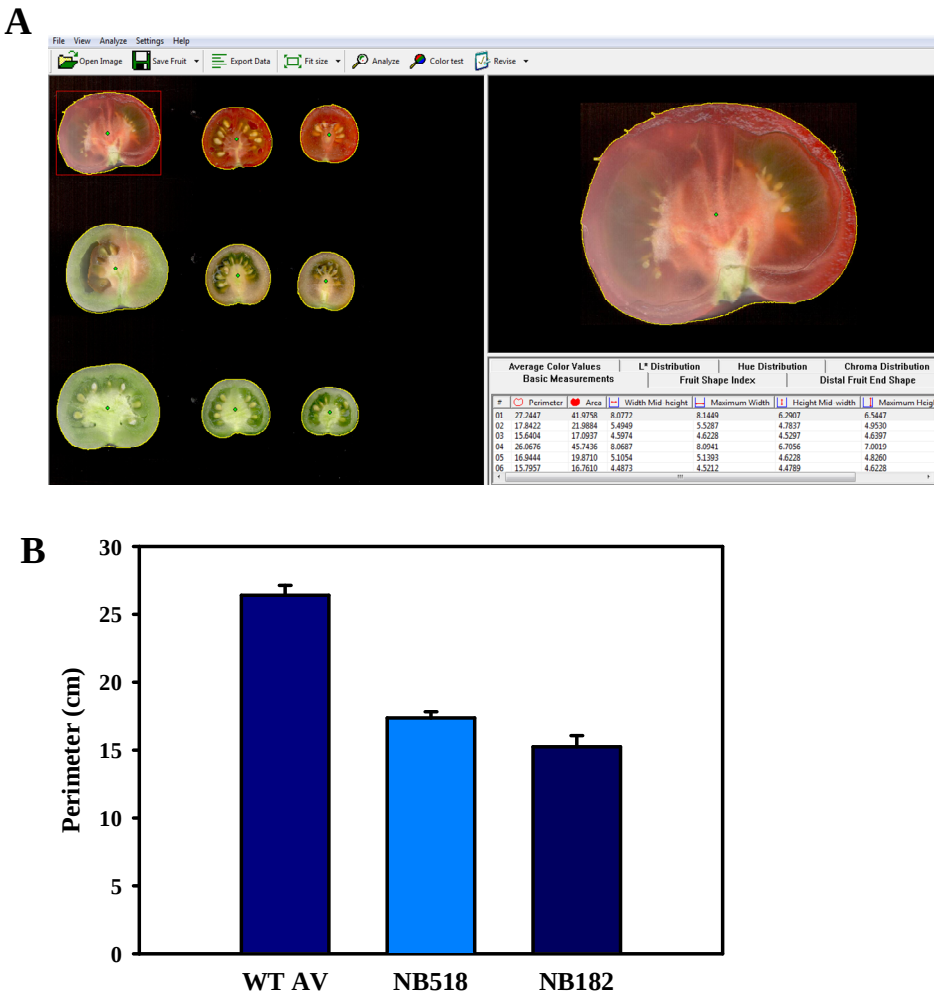


Figure 5.20 A, Screen shot of tomato analyzer software program to generate automated and standardized measurements of fruit shape. **B**, Perimeter of *SPA3LIKE* fruits at three stage- MG, BR and RR compared with the fruits of the reference cultivar AV. Statistical significance is represented by * for P value ≤ 0.05.

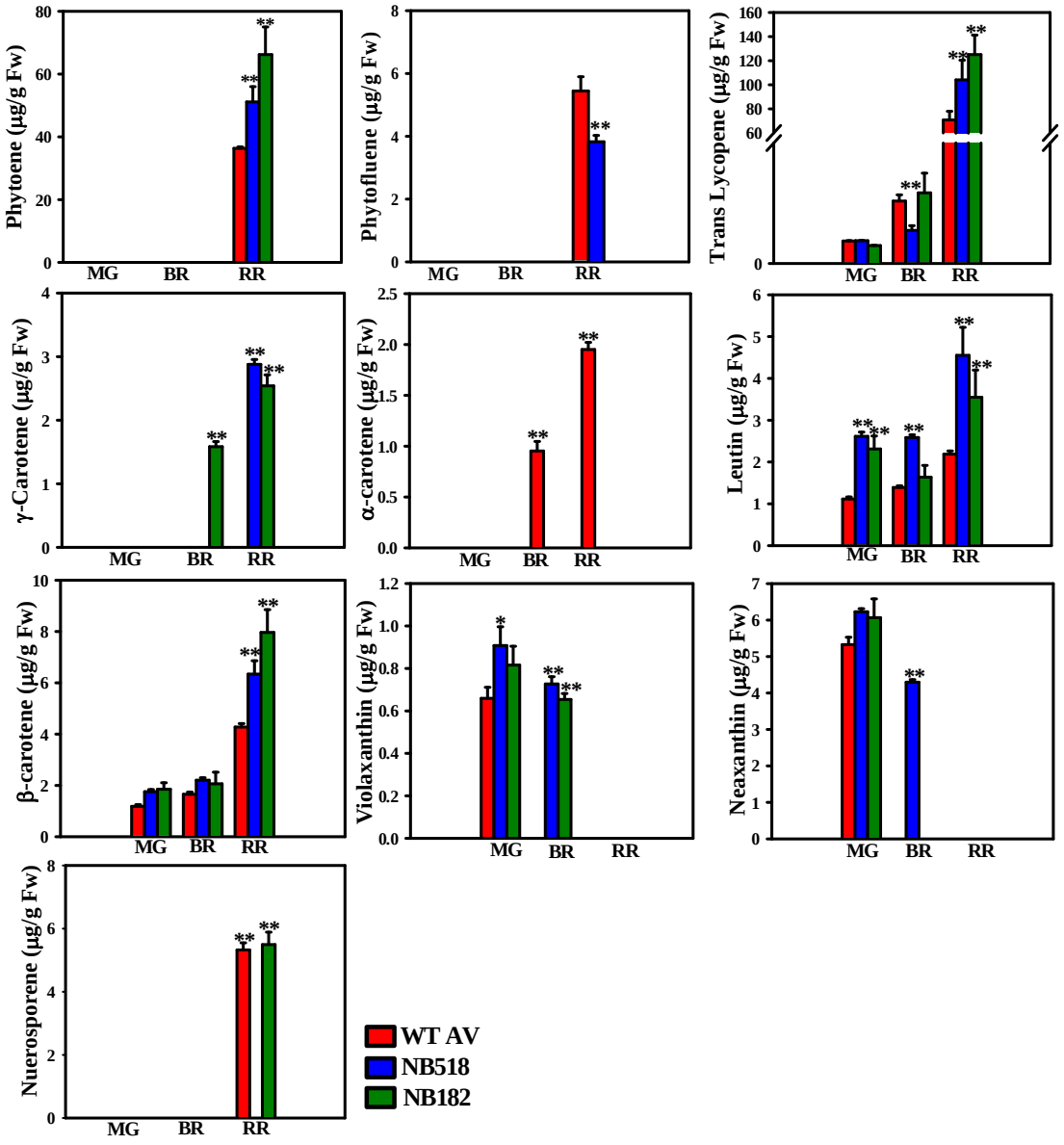


Figure 5.21 U-HPLC based carotenoid profiling in fruits of *SPA3LIKE* variants and AV at three stages – MG, BR and RR. Each carotenoid is represented as µg/g Fw of the fruit. Each histogram is a representation of the mean value ±S.E of three independent experiments. Statistical significance is represented by * for P value ≤ 0.05.

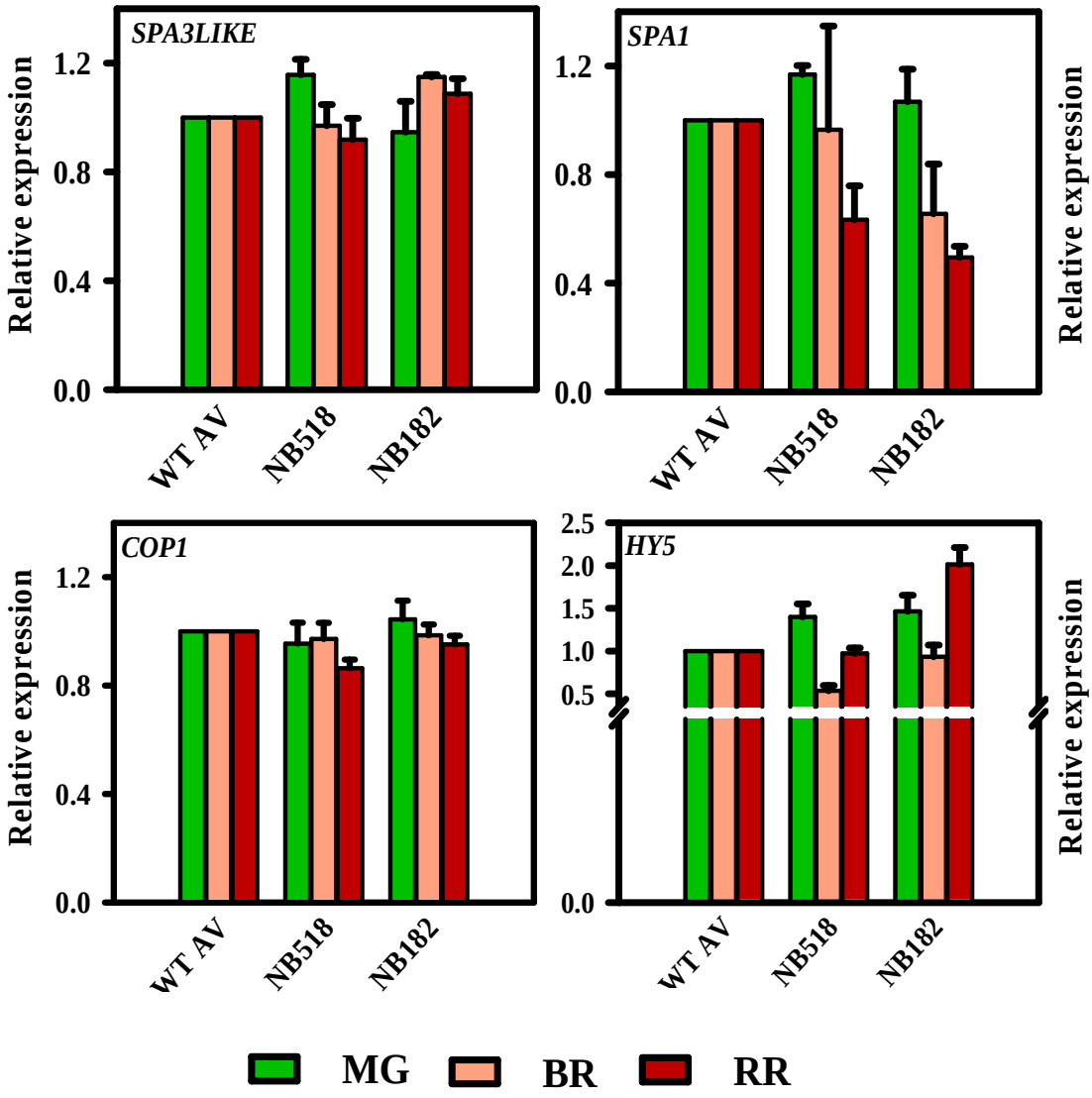


Figure 5.22 Gene expression analyses at three different fruit ripening stages in *SPA3LIKE* variants. The transcripts were expressed after normalization with two internal controls, *β-actin* and *ubiquitin 3* as described in methods. The transcript levels of *SPA3LIKE*, *COP1*, *SPA1* and *HY5* in NB518, NB182 and AV respectively.

levels were unaffected in the *SPA3LIKE* variants and were almost similar to the reference. Though there are no detailed reports of the interacting partners of *SPA3LIKE*, it has been well established that SPA1 interacts with COP1 in targeting the substrates for protein degradation and enhances COP1 function (Laubinger et al., 2004; Ranjan et al., 2014). Studies also state that the diversity in SPA family of proteins dictates the targets to be degraded by COP1 complex (Ranjan et al., 2014). Since HY5 is a major target for COP1 (Ang et al., 1998; Hardtke et al., 2000; Xu et al., 2014) mediated protein degradation, we also checked the transcript levels of these genes in the fruit tissue of the *SPA3LIKE* variants (**Fig. 5.22**). As expected, no significant differences were observed in the transcript levels of these genes compared to the reference, supporting that the frame shift mutations may have altered the protein function due to premature truncation.

5.3 Discussion

We established Eco-TILLING using a collection of 600 tomato accessions representing the natural gene pool of *S. lycopersicum*. In this study, we screened the allelic variation in two genes- *COP1* and *SPA3LIKE* that are components of light signaling pathway, negatively regulating photomorphogenesis. Further, we identified haplotypes that are associated with altered seedling response and enhanced fruit carotenoid content, which in turn might be useful for development of nutritionally enriched tomato cultivar.

Detection of polymorphisms by Eco-TILLING

For the Eco-TILLING protocols, optimized primers and in-house enzymes were used for SNP detection. By screening a large population, we were able to show that Eco-TILLING provides a high throughput technique for the analysis of polymorphisms in tomato. First target region was amplified, followed by digestion of the heteroduplexed DNA products with CELI endonuclease and the CELI cleaved PCR products were evaluated on Li-COR gels. The distinct fragments with complementary fragments in other channel (considered as a SNP) were grouped according to size and pattern. While pattern of fragments could be used for grouping of haplotype, a single accession/few accessions per haplotype is needed to be sequenced to know the exact number and position of the SNPs. However, in some cases, due to mispriming, weak fluorescence towards top of each lane and sometimes due to increased noise at the bottom of the gel, false positives may be detected in Eco-TILLING (Greene et al., 2003; Till et al., 2006a; Till et al., 2010). With a 1:1 (reference DNA: DNA from the natural accession) pooling strategy for Eco-TILLING, we could minimize the errors and also improve the stringency of detecting SNPs. For several accessions, we identified multiple cleavage patterns which could have arisen by the presence of nearby polymorphic sites, insertions and deletions. Though Li-COR analysis of some amplicons showed only one cleaved product, however, sequencing of such accessions revealed the presence of more than one SNP. This may be due to the possibility that if bands were cleaved at both ends, they would not fluoresce and therefore cannot be detected on Li-COR system (Liang, 2011). In spite of these minor drawbacks in Li-COR based detection of SNPs, our results

showed that Eco-TILLING is a suitable and cost effective method for identification of SNPs in tomato.

Single Nucleotide Polymorphisms in tomato *COP1* gene

Nearly 72 tomato accessions showed polymorphisms for *COP1* gene. Based on the cleavage pattern of PCR products on the gel, these accessions were classified into 11 haplotypes. Interestingly, most of the SNPs were observed in introns and in the coding region, only three nonsynonymous and three synonymous SNPs were identified. This result is consistent with that obtained from sequencing of large number cultivars, where only 10 % SNPs are located in coding region (Aflitos et al., 2014). The SIFT scores for the nonsynonymous SNPs were not appreciable, signifying that these changes do not alter the protein drastically. Also, the codon usage efficiencies of the synonymous SNPs were not drastically affected, conferring that induced mutations may have more impact on codon bias than the naturally occurring polymorphisms (Plotkin and Kudla, 2011). Moreover, consistent with the *in-silico* CODDLE analysis, the region of *COP1* encompassing the ring finger domain exhibited a very low nucleotide diversity as no SNP was identified in this region. The exons 3 and 4 (encoding for the coiled coil domain) exhibited highest number of SNPs, whereas the exon 5 to 13 (encoding for WD40 domain) showed low polymorphism with respect to the former region.

Two haplotypes of *COP1* gene (HT3 and HT10) had a significant increase in the fruit carotenoid content at the red ripe stage. Although mutations in *COP1* gene are known to affect the light mediated responses, majorly photomorphogenesis in Arabidopsis (Lau and Deng, 2012), the role of *COP1* in light regulated fruit ripening is not well understood. The genetic and molecular studies on *cop1* mutants of *Arabidopsis* revealed that most of the induced mutations were strong (Deng et al., 1998), causing lethality, and only a couple of weak mutants exhibited the characteristic de-etiolated phenotype. Consistent with these studies, our studies on TILLING for *COP1* gene resulted in a single mutant, with a probable phenotype. Compared to very low induced mutation frequency in EMS mutagenized population, Eco-TILLING for *COP1* yielded more alleles. Except for the two accessions NB110 and 104, all the other accessions of HT3 (NB471) and HT10 (NB515) that exhibited fruit phenotype did not show any drastic phenotype at the seedling stage. Unlike the

strong phenotypes observed in *cop*/*det*/*fus* mutants of Arabidopsis, the tomato mutant *hp1* (with a lesion in *DDB1* gene, Mustilli et al., 1999) and *hp2* (with a lesion in *DET1* gene, Liberman et al., 2004) do not show any open cotyledon phenotype in darkness. However, these mutants exhibit enhanced fruit carotenoid content and show exaggerated photomorphogenic response in terms of reduced hypocotyl length and anthocyanin accumulation (Lieberman et al., 2004; Mustilli et al., 1999). The complete absence of a strong constitutive photomorphogenic response i.e. open cotyledon phenotype in darkness in tomato mutant seedlings may emphasize the necessity of other interacting partners to initiate the response and this in turn may highlight the species-specific function of the pleotropic *COP*/*DET*/*FUS* genes.

Single Nucleotide Polymorphisms in tomato *SPA3LIKE* gene

The Arabidopsis SPA (Suppressor of Phytochrome A) protein family forms a part of COP1 (E3 ligase) complex and negatively regulates the transcription factors involved in light mediated responses (Laubinger et al., 2004). A recent study on COP1 and SPA proteins from Physcomitrella revealed that COP1 is highly conserved in plants and animals. Unlike COP1, SPA proteins were considered to have functional divergence, thereby, modifying the core function of COP1 (Ranjan et al., 2014). Notwithstanding the functional similarity of *SPA3LIKE* to *COP1* and also *RUP2*, we isolated allelic variants in *SPA3LIKE*. Only four accessions were identified to harbor SNPs, therefore, each one was considered as an individual haplotype (HT1, HT2, HT3 and HT4) with the reference as HT0. Interestingly, the INDELs in two of the accessions NB518 (HT3) and NB182 (HT4) caused frame shift mutation that might alter the protein function.

Despite screening a large population, *COP1* showed very low nucleotide diversity. This may be due in part to an underrepresentation of *COP1* and *SPA3LIKE* alleles in our collection. It was observed in our study that most of the accessions collected from NBPGR harbored SNPs rather than those collected from IIVR and TGRC. Most of the deposited accessions at TGRC are monogenic mutants or double /triple mutants of known genes, with a very few uncharacterized mutants. While other reasons cannot be excluded, one of the reasons for low SNP frequency could be because the amplicons selected for the current analysis were limited to the coding sequence of the targeted genes (highly conserved), and did not include the promoter.

Altered seedling phenotype of *SPA3LIKE* variants NB518 and NB182

The COP1/SPA complex is well characterized in Arabidopsis and acts as a negative regulator of photomorphogenesis (Ranjan et al., 2014; Laubinger et al., 2004). Therefore, in darkness the *cop1*/*spa* mutants of Arabidopsis exhibit constitutively photomorphogenic seedling phenotype characterized by open cotyledons, reduced hypocotyl length and anthocyanin accumulation (Deng et al., 1998). An earlier study revealed that RNAi suppression of tomato *COP1LIKE* (now annotated as *SPA3LIKE*) resulted in slight reduction in hypocotyl length of seedlings under white light and increased lycopene levels in the red fruit (Liu et al., 2004). The study also revealed that these seedlings did not exhibit any drastic photomorphogenic responses either in darkness or under white light conditions, as observed for defective pleotropic genes *cop*/*det*/*fus* (Wei and Deng, 1996; Hardtke, 2000; Huang et al., 2014) or the genes that negatively regulate photomorphogenesis in Arabidopsis (Leviar et al., 2012; Ranjan et al., 2014). Consistent with these studies, the seedlings of *SPA3LIKE* variant NB182 when grown in darkness or white light exhibited a significant reduction in hypocotyl length. On the other hand, the accession NB518 did not exhibit any significant seedling phenotypes both in darkness and light grown conditions. In addition to this data, there was no significant change in the anthocyanin levels as observed in these variants compared to the reference cultivar. The studies on Arabidopsis *rup* mutants reveal close but not completely similar phenotypes wherein, the *rup1* mutant behaves similar to the wild type and do not show any drastic photoresponses. However, it was established that the *rup2* mutant is slightly hypersensitive and the *rup1rup2* double mutant exhibits strong hypersensitivity at the level of hypocotyl inhibition, flavonoid and anthocyanin accumulation (Gruber et al., 2010). This reflects on the functional redundancy of *RUP1* and *RUP2* genes in Arabidopsis. Based on our blast results, it has to be noted that the Arabidopsis *RUP2* is the closest homolog of tomato *SPA3LIKE* and it may be possible that, tomato might contain another homolog of *SPA3LIKE* which might function as *RUP1*. Currently, no specific hit was observed for *AtRUP1* upon a blast analysis on tomato genome. This could be primarily because the genome is still only 95% annotated and this gene is either not annotated or tomato may not have a second functional gene. The weak seedling phenotypes observed for RNAi lines of tomato *COP1LIKE* can only be speculated as a result of leaky expression of the gene in the transgenic plant due to

incomplete silencing or may be due to the presence of functionally redundant genes. The seedling phenotypes of the Eco-TILLING variants can also be explained in the same manner or may be attributed to the species specific functional divergence of the genes. This can be further explained on the basis of functional diversity of COP1/SPA complex in different plant species like pea, rice and apple (Ranjan et al., 2014). Nevertheless, further genetic studies have to be done to screen for polymorphisms in any other genes in these haplotypes.

A two day delay in opening of the hook was another significant phenotype observed in the seedlings of NB182 when grown under white light. It is quite interesting that both light signal and hormones play an important role in regulating the hook development in Arabidopsis seedlings. Over a decade or two, several studies were conducted in Arabidopsis to understand the role of hormones in apical hook development. As reviewed by Mazzella et al (2014), the coordinated function of gibberellins (via the DELLA-PIF (PHYTOCHROME INTERACTING FACTORs) pathway), ethylene (via the EIN3/EIL1 (ETHYLENE INSENSITIVE 3/EIN3 like 1)-HLS1 (HOOKLESS 1) pathway) and brassinosteroids control the auxin distribution and therefore the development of hook in darkness. On the contrary, the repressors of photomorphogenesis such as PIFs and COP1 were also known to be important for hook development in darkness (Alabadí e t al., 2008, Leiv ar et al., 2008, Shin et al., 2009). It is evident from the recent studies that COP1 regulates the amount of EIN3/EIL proteins, thus playing a crucial role in ethylene development (Zhong et al., 2009, Liang et al., 2012). Considering these reports and the growing complexity of the factors, especially, light signaling intermediates involved in apical hook development, we cannot rule out the possibility that ethylene may be responsible for the hook development in the *SPA3LIKE* variant NB182. Though we could not critically examine this aspect of seedling development in NB182, further studies in this line would be required to understand the complex network of hormones in response to light signal.

SPA3LIKE variants – a functional relevance for fruit ripening

Tomato fruit ripening is associated with the accumulation of carotenoids-which are health promoting compounds. Many transgenic approaches were used to characterize the biosynthetic genes of carotenoid pathway (Just et al., 2007; Zhu et al.,

2008). The availability of natural ripening mutants in tomato had facilitated the molecular understanding of fruit ripening process (Osorio et al., 2013; Gapper et al., 2014). It is a complex process regulated by many regulators. Ethylene biosynthesis and signaling confer a major regulation apart from the genetic hierarchy that has been elucidated to be controlled by transcription factors like RIN/MADS or SBP/CNR.

In spite of these internal cues, light is also known to affect ripening of tomato. The characterization of spontaneous mutants *hp1* and *hp2* strongly supported the role of light in improving the nutritional quality of tomato. Further, studies on the transgenic lines - *COP1LIKE RNAi, DET1 RNAi* and also *CUL4 RNAi* lines (Liu et al., 2004; 2012; Wang et al., 2008) support the hypothesis that suppression of negative regulators of light signaling pathway enhances the carotenoid accumulation. With an aim to isolate new alleles, the two genes, *COP1* and *SPA3LIKE* were subjected for Eco-TILLING. Consistent with the studies of Liu et al., (2004) and the growing evidence for light regulation of fruit ripening and quality (Gupta et al., 2014) our data also support that alteration of light signaling components may be a good choice for enhancing nutrient quality of tomato. Though the fruit size of *SPA3LIKE* variants is small, the yield of the fruits was almost similar to the reference.

The transcript levels of *SPA3LIKE, COP1, SPA1* and *HY5* were almost unaltered and similar to AV in *SPA3LIKE* variants, supporting that the frame shift mutation may have altered the protein function due to premature truncation. As SPA3LIKE protein consist of WD40 region only, a truncated protein might cause a loss in the target recognition sites, thus modulating the downstream responses. As of now, the targets of SPA3LIKE were not clearly studied. The exact interaction of SPA3LIKE has to be established to understand the molecular mechanisms by which SPA3LIKE alters the fruit carotenoid levels. In light of the present data, *SPA3LIKE* variants, NB182 and NB518 appear to be probable candidates to improve the tomato fruit quality.

In spite of low nucleotide diversity of 1 SNP/160 Kb for *COP1* gene in Eco-TILLING, we could identify few alleles with promising phenotypes. Eco-TILLING in tomato has been carried by only few groups. Gady et al., 2012 studied 258 tomato lines for the genes involved in sugar metabolism. The frequency of polymorphism observed in this study was low, i.e. out of 20 targeted genes; only 6 of them were

found to harbor SNPs. Another study by Rigola et al., (2009) also found a low level of polymorphism in tomato lines from EU-SOL collection. In consistent with the low frequency of polymorphisms observed in tomato, our study also showed that the natural diversity for the two genes, *COP1* and *SPA3LIKE* is low. This low frequency might be due to the recalcitrance of tomato genome to polymorphisms or due to the importance of the genes that were considered for the study. More particularly, in our case, the two genes were involved in light signaling pathways and are important for plant development. Nevertheless, we identified novel alleles for *SPA3LIKE* that can be useful for the development of an elite tomato cultivar with enhanced carotenoid content.

Chapter 6

Nucleotide diversity analysis of *COP1* and *SPA3LIKE* genes in wild relatives of tomato

6.1 Introduction

Wild relatives of a crop species are the ancestors or closely related taxa that originated at distinct geographic locations but share a common gene pool. These wild species grow in different habitats which endows them with special features to tolerate a myriad of environmental and climatic changes. Latest estimates report that there are between 50,000-60,000 crop wild relatives in the wild (Maxted and Kell, 2009) and their genetic potential remains to be utilized for crop improvement.

A significant increase in crop production was brought about by the green revolution in the 1960s and 1970s. It was largely based on the discovery of new genes in the traditionally grown land races. Presently, with increasing concerns about climate change and its effects on agriculture, there is a high demand for ensuring food security. It has been argued that the development of new crop varieties or cultivars is required to achieve this goal (Lobell et al., 2008). This demand has promoted the quest for search of novel genes and gene systems present in the wild relatives. The studies on wild relatives also provide us information about the evolutionary relationships of the genes and the changes which have taken place as a part of natural selection. Traditionally, wild relatives contributed significantly to improve food production through conventional breeding programs. Development of Asian rice with resistance against grassy stunt virus is one of such examples of unlocking the potential of wild relatives coupled with biotechnology (Prischmann et al., 2009). The gene responsible for resistance from *Oryza nivara*, a wild rice species was introgressed into many rice cultivars (Prischmann et al., 2009). A similar survey for disease resistance against western corn rootworm in maize and fusarium wilt in banana identified wild relatives that possessed the resistance trait that were successfully introgressed (Hajjar and Hodgkin, 2007). Apart from disease resistance, wild relatives have been used to enhance the tolerance to drought and salinity in barley, wheat and sunflower (Nevo and Chen, 2010; Miller and Seiler, 2003). In broccoli and pearl millet, the quality and yield were enhanced by crossing with wild relatives that harbored these specific traits (Sarikamis et al., 2006 and Hajjar and Hodgkin, 2007). However, conventional breeding is not easy in all crop species due to self-incompatibility problems. Therefore, developing novel cultivars needs a good crop model system that can be exploited for more benefits.

In the recent years tomato has emerged as crop model because of advantages like- small diploid genome, a well defined ripening biology, and availability of genome sequence, monogenic mutants and wild species (Tomato Consortium, 2012). It is known that tomato consists of 9 closely related species or sub species: the wild or weedy forms of cherry tomato (*S. lycopersicum cerasiformae*) (Peralta et al., 2008), and the wild species- *S. chilense*, *S. chmielewskii*, *S. galapaganse*, *S. habrochaites*, *S. neorikii*, *S. pennellii*, *S. peruvianum* and *S. pimpinellifolium* (**Table 6.1**). Many useful traits found in wild tomato species such as tolerance to drought and salinity (*S. pennellii*) (Dehan et al., 1978, Shalata et al., 1998), accumulation of health promoting phytochemicals (*S. habrochates*) and resistance to multiple pathogens (*S. galapaganse, S. lycopersicoides, S. neorickii* and *S. chilense*) have been introgressed into cultivated tomato (Venema, 1999; Robert, 2001; Rick et al., 1994). It is estimated that cultivated tomato contains less than 5% of the genetic variation of the wild relatives (Miller et.al., 1990) and this genetic potential can be exploited to improve many agronomic traits in tomato.

Tomato fruit is rich in health promoting phytochemicals. The fruit of modern tomato or the domesticated tomato is large in size and red in color due to the accumulation of lycopene and β-carotene. On the contrary, the fruits of all the wild relatives are round in shape and much smaller in size than tomato. On the basis of accumulation of carotenoids during development and ripening of fruits, two categories are distinguished. Green fruited species are characterized by purplish green fruit color during ripening (Alexander and Grierson, 2002). The second group is characterized by the accumulation of carotenoids that impart orange or red color to the fruit (Bauchet and Causse, 2012) in the later stages of ripening. The biosynthesis of carotenoids begins with the conversion of IPP (isopentenyl disphosphate) to produce the precursor GGPP (geranylgeranyl diphosphate) via the plastid localized MEP pathway (Giuliano, 2014). Further, a series of enzymes catalyze the conversion of GGPP to end product lycopene. In tomato, most of the carotenoid biosynthetic enzymes are well characterized and are under tight transcriptional and translational regulation. In addition to genetic regulation, internal (hormonal regulation) and external stimuli (e.g., light, wounding, pathogen attack and temperature shift) also regulate fruit ripening. Among these factors light is known to play an important role to regulate carotenoid biosynthesis in tomato fruits via a post-transcriptional

Table 6.1 Principal features of the *Lycopersicon* subsection *(Solanum sect. Lycopersicon).* Data are compiled from Peralta et al., (2007), Moyle et al., (2008), Grandillo et al., (2011). SI- Self-incompatible, SC- Self-compatible, AI- Allogamous, AT- Autogamous, UI- Unilateral incompatibility, EL- Embryo lethality.

Species new nomenclature/(old nomenclature)	TGRC Accession Number	Geographic distribution and habitat	Mating system	Fruit color	Genetic polymorphism
S. neorickii/ (*L. parviflorum*)	LA2133	South Equador, South central Peru native, 1500 - 3000 m, rocky humid and well drained areas	SC, highly AT/reciprocal	Green with dark green stripes	Low
S. pennellii/ (*L. pennellii*)	LA0716	Peruvian coast native, 0-2000 m, dry and rocky hill side	SI usually, SC populations in southern parts/reciprocal	Green	High
S. habrochaites/ (*L. hirsutum*)	LA1777	South west Equador to South central Peru native, 500 – 3300 m, forest regions	SI/SC populations in southern parts/UI	Green with darker green stripes	High
S. chilense/ (*L. chilense*)	LA1969	South Peru to north Chile, 0 – 3000 m, dry river bed	SI, AL/UI, EL	Green to whitish green with purple	High
S. galapagense/ (*L. cheesmanii f. minor*)	LA0483	Galapagos islands endemic species, 0 - 1300 m from sea shore to volcanic areas	SC, AT/reciprocal	Yellow orange	Low
S. pimpinellifolium/ (*L. pimpinellifolium*)	LA1589	South west Equador – northern Peru native, under 1000 m, South valleys of the pacific coast	SC, AT facultative AL/reciprocal	Red	Intermediate
S. lycopersicum/ (*L. esculentum*)	Arka Vikas (reference cultivar)	South India	SC	Red	Very low

modification of phytoene synthase enzyme (Thomas and Jen, 1975; Gupta et al., 2014). Consequently fruits of *phytochrome* mutants of tomato ripened under darkness showed differences in the accumulation of lycopene (Gupta et al., 2014).

Unlike Arabidopsis, tomato lacks mutants defective in *COP1* gene. In order to gain better insight in light regulation of tomato fruit ripening and improvement of ripening quality attempts were made to isolate *COP1* and *SPA3LIKE* mutants by reverse genetics techniques- TILLING and Eco-TILLING. However, due to a low mutation frequency observed for *COP1* and *SPA3LIKE* genes, we examined the sequence diversity of these genes in six wild relatives of tomato. In the present study, we sequenced the coding regions of the two genes and performed *in-silico* analysis of the single nucleotide polymorphisms to understand the frequency of natural diversity of the two genes at both nucleotide and protein level.

6.2 Results

In earlier chapters, isolation of *COP1* and *SPA3LIKE* mutants through reverse genetic strategies was described. In this chapter, we examined the structural genomic data of these two candidate genes. More specifically, a search for single nucleotide polymorphisms in the selected wild species of tomato was undertaken. The overall number and nature of SNPs such as synonymous and nonsynonymous and their distribution can reflect on the selection pressures for the functional diversity of these genes and provide an idea on the evolutionary conserveness of the loci. The genes selected for this study include primarily *COP1* and *SPA3LIKE*.

6.2.1 Sequence diversity of *COP1* gene

The aligned length, including coding and non-coding regions of the *COP1* gene was 11478 nucleotides. A total of 17 INDELs were detected with an INDEL length ranging from 1 to 57 nucleotides (**Table 6.2**). The distribution of nucleotide variations found among the sequenced regions of *COP1* gene in different wild relatives considered were represented by PARSESNP output format (**Fig. 6.1**). It is noteworthy that the frequency of nucleotide variations in *COP1* is almost three times higher in introns than in coding sequences (**Fig. 6.2A**). The number of SNPs was higher than the number of INDELs, with a total of 67 SNPs in the coding region resulting in average of 1 SNP for 45 Kb of genomic DNA. Relatively fewer SNPs (22) and INDELs (5) were found in red colored fruit species than in the green colored fruit species (45 SNPs and 12 INDELs). On analyzing the polymorphism data per species, it was observed that the number of nucleotide polymorphisms observed for *COP1* was highest in *S. habrochaites* (green fruited) and least in the red fruited *S. pimpinellifolium* (**Fig. 6.2B**).

The nucleotide diversity analyses showed that the nucleotide diversity of green fruited varieties was higher ($\pi = 0.04965$) than the colored fruit species ($\pi = 0.00389$). The DnaSP analysis of the *COP1* gene revealed high nucleotide diversity at the region located at 5,000 to 7500 bp of the gene in green fruited species (**Fig. 6.3**). According to this plot, low nucleotide diversity was also observed at the regions 2400 to 2600, 8500 to 10,000 and 10,500 to 11,500 bp of the gene. This analysis further revealed that polymorphic sites were mostly located at the middle and at the end of the gene, with most being in the introns (**Fig. 6.3**).

Table 6.2 Summary of polymorphism analysis for *COP1* and *SPA3LIKE* genes as calculated by DnasP software.

Gene	*COP1*	*SPA3LIKE*
Total length including Indels	11552	3101
Total no. of sites excluding Indels	11478	3100
No. of Indels	74	1
No. of Indel polymorphisms	17	1
Length of coding region excluding Indels	2015	1081
Length of coding region including Indels	2021	1082
Length of noncoding region excluding Indels	9446	2019
Length of noncoding region including Indels	9457	2019
SNPs	67	32

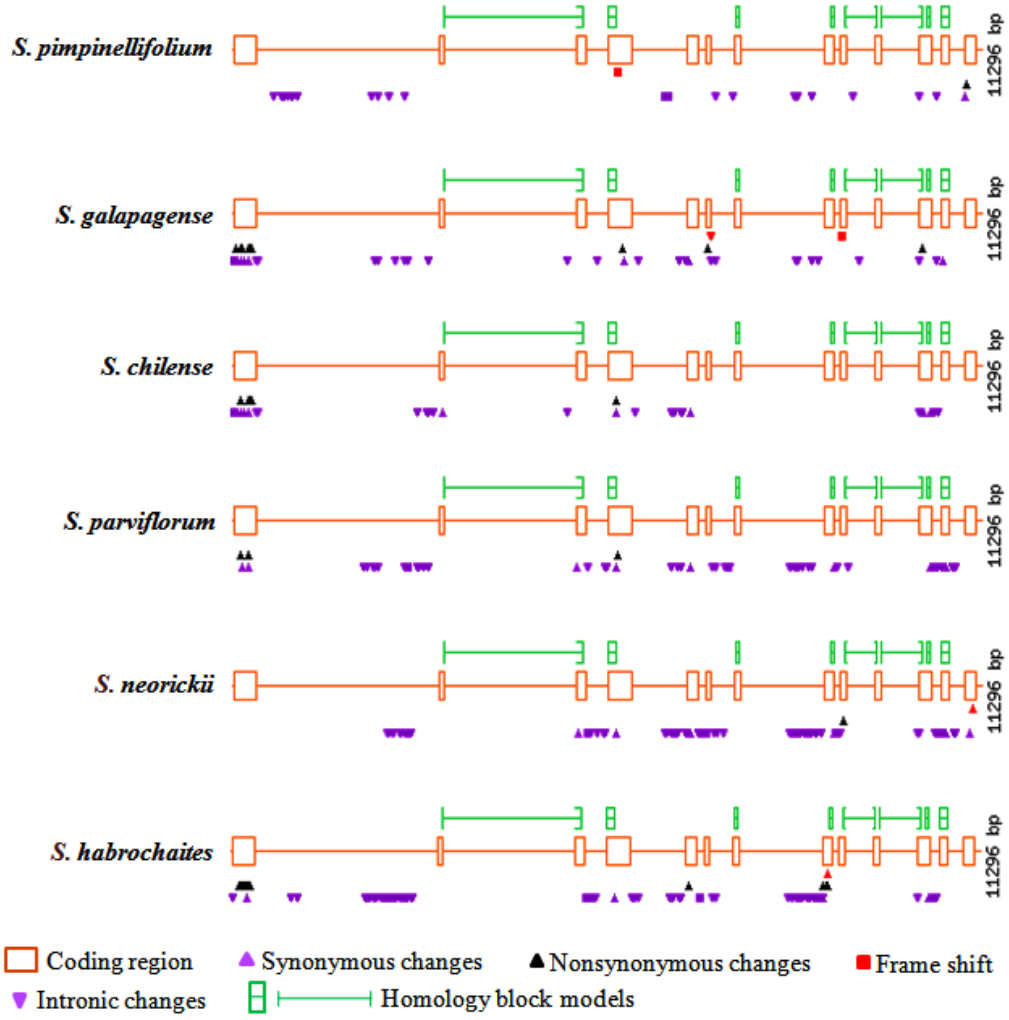


Figure 6.1 PARSESNP output file representing the distribution of SNPs for *COP1* gene in wild relatives of tomato. Open orange boxes denote exons interconnected with solid orange line indicating introns. Black upright triangle indicates changes in the DNA sequence that affect the amino acid. Purple upright triangle indicates changes in the DNA sequence that does not affect the amino acid. Purple downward triangles indicate changes in introns. Red boxes represent the frame shift changes.

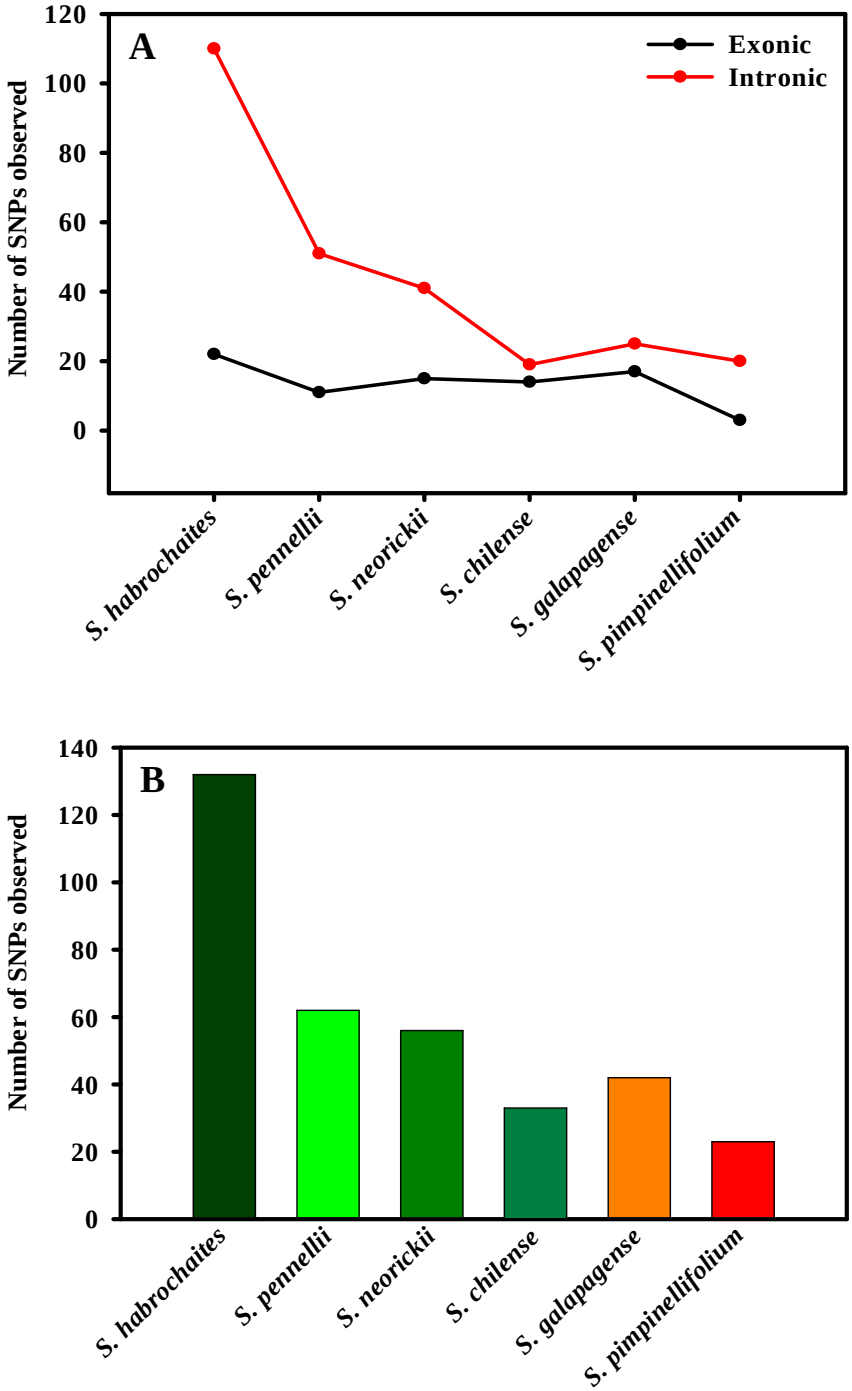


Figure 6.2 A, Comparison of exonic and intronic SNPs identified in various and **B** represents the total number of SNPs observed in various wild relatives of tomato for *COP1* gene.

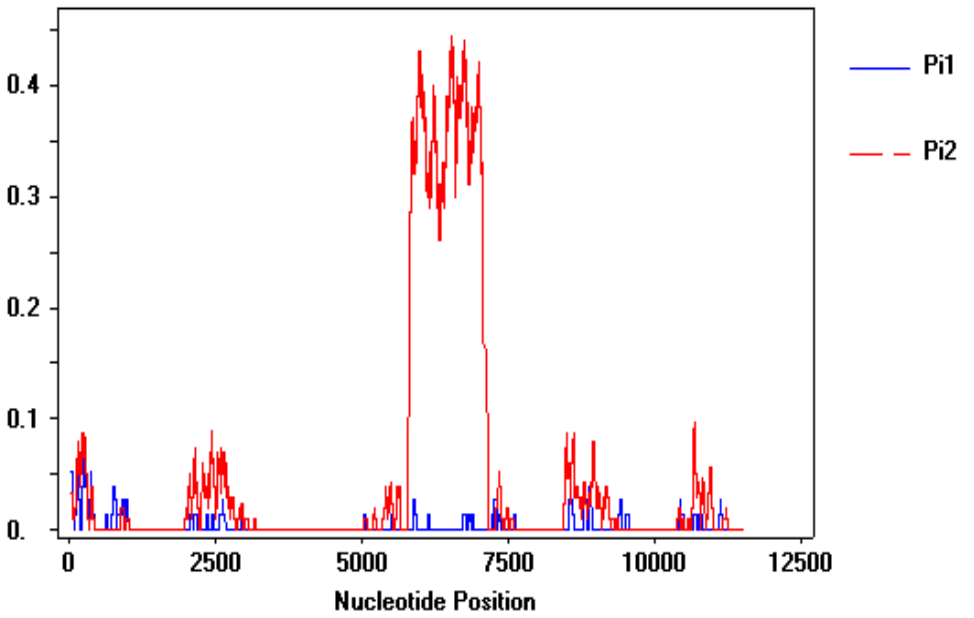


Figure 6.3 Nucleotide diversity (π) patterns along *COP1* gene in sliding window among color fruited and green fruited wild species. Analysis was performed using a window length of 100 bp in DNASP v 5.0. Pi1 (blue color graph) and Pi2 (red colored graph) are the nucleotide diversities representing the colored fruited and green fruited species respectively. Note: in the overlaid graph, Pi2 is high compared to Pi1 on the entire length of the *COP1* gene indicating the high nucleotide diversity in green fruited species.

6.2.1.1 Analysis of Nucleotide Base Changes

The observed SNPs were categorized according to the type of nucleotide substitution- either as transitions (C/T or G/A) or transversions (A/T, G/C, C/A, or T/G). The point substitution analysis showed that the C/T transition is the most frequent substitution in *COP1* gene with 14% frequency. We observed that G/A substitution is the next prevalent base change with a frequency of 13% (**Fig. 6.4A**).

Among 67 polymorphisms observed in the coding region, 22 were silent or synonymous and the rest were nonsynonymous (**Fig. 6.4B**). The details of the SNPs, position on the gene and the corresponding amino acid change are listed in **Table 6.3**. Some of the polymorphisms were common in the green fruited species. PARSESNP results revealed a nucleotide change from C7214 to T at the 3′ splice acceptor site of the *COP1* intron 6, present in *S. galapagense* species. This SNP was predicted to cause alternative splicing and this may alter the WD40 motif of COP1 protein. The sequencing results also revealed a base change from G8992 to A in *S. habrochaites* and C11161 to T in *S. pennellii*, resulted in a premature stop codon.

6.2.1.2 Variations in COP1 protein sequence

The effects of genetic variation on amino acid sequence were evaluated by SIFT values. All the six species used in the study showed different amino acid changes (**Table 6.3**). The nonsynonymous changes observed in *COP1* had no drastic affect on the protein function as the predicted SIFT values were higher than 0.5. However, the nucleotide changes G8992 to A in *S. habrochaites* and C11161 to T in *S. pennellii* resulted in a premature stop codon at position W430 and Q667 in *S. habrochaites* and in *S. pennellii* respectively. These changes might cause a truncation in the COP1 protein thereby affecting the protein function in these two species. These possibilities have to be validated further to understand the role of *COP1* in fruit ripening of green fruited species.

Synonymous SNPs do not cause any change in amino acid and therefore presumably have no effect on the protein function. However, several studies in animal systems revealed that synonymous or silent changes can affect the protein function by altering the translation rates or altering the secondary structure of the mRNA (Kimchi-Sarfaty et al., 2007). On the contrary, in plants, very few studies have been conducted to understand the role of synonymous SNPs. Taking into consideration of

Table 6.3 List of nucleotide polymorphisms in *COP1* gene and their effect on amino acids.

S. No	Nucleotide change	Amino acid change	*S. habrochaites*	*S. pennellii*	*S. neorickii*	*S. chilense*	*S. galapagense*	*S. pimpinellifolium*	SIFT Score
1	T38G	V2G					+		0.57
2	A66C	P11=				+			
3	G49A	V6I					+		0.98
4	G111A	E26=				+	+		
5	T122G	V30G				+			0.57
6	C127G	L32V	+		+	+	+		1
7	T128G	L32R				+			0.05
8	A138C	E35D	+						0.78
9	G139A	E36K					+		0
10	A141G	E36=			+				
11	A156C	T41=				+	+		
12	A159D	E42D	+						0
13	A172T	R47W	+						0.13
14	G183T	L50F	+						0.49
15	T184C	C51R	+						0.77
16	G198C	M55I	+						1
17	G211A	D60N	+						1
18	T217G	F62V	+						1
19	T231G	C66W			+	+	+		0.26
20	C237T	H68=			+	+	+		
21	G239A	S69N				+			0.14
22	G239C	S69T			+				1
23	T241G	F70V			+				1
24	A250C	M73L	+						1
25	G259A	V76I					+		0.98
26	T269C	L79P	+						0.21
27	C273G	H80Q				+	+		1
28	G281A	S83N			+		+		0.96
29	G281C	S83T	+						1
30	C289T	P86S	+						1
31	T294G	C87W				+			0.26
32	T297G	C88W					+		0.26
33	G3163A	Q124=					+		

Contd...

S. No	Nucleotide change	Amino acid change	*S. habrochaites*	*S. pennellii*	*S. neorickii*	*S. chilense*	*S. galapagense*	*S. pimpinellifolium*	SIFT Score
34	T5195C	I137=			+				
35	G5216A	L144=		+					
36	T5769C	L219=	+	+	+	+			
37	C5782A	P223H				+			0.53
38	G5801C	W229C			+				0.33
39	A5878C	N255T					+		1
40	A5891G	K259=					+		
41	T6874C	L312=					+		
42	A6897G	A319=		+		+			
43	A6898G	N320D	+						0.99
44	A6903G	K321=		+					
45	G7157C	V360L					+		0.9
46	T8917G	V405=	+						
47	T8923G	N407K	+						1
48	G8967A	R422Q	+						0.81
49	G8988T	S429I	+						0.8
50	G8992A	W430*	+						
51	G8998C	K432N	+						0.74
52	G9058A	V452=			+				
53	G9061C	T453=		+					
54	A9212C	T476P		+					2.8
55	A10395T	R538W					+		0.13
56	G10412T	P543=				+			
57	T10502A	T573=	+		+	+			
58	G10710A	T592=					+		
59	T10725C	F597=		+	+				
60	A10737T	S601=			+				
61	G10755C	L607=		+	+				
62	A10758G	S608=		+	+				
63	C11034T	S624=						+	
64	C11064T	S634=						+	
65	A11115G	A651=		+					
66	C11138T	P659L						+	1
67	C11161T	Q667*		+					

+ indicates the presence of the SNP in the wild relative.

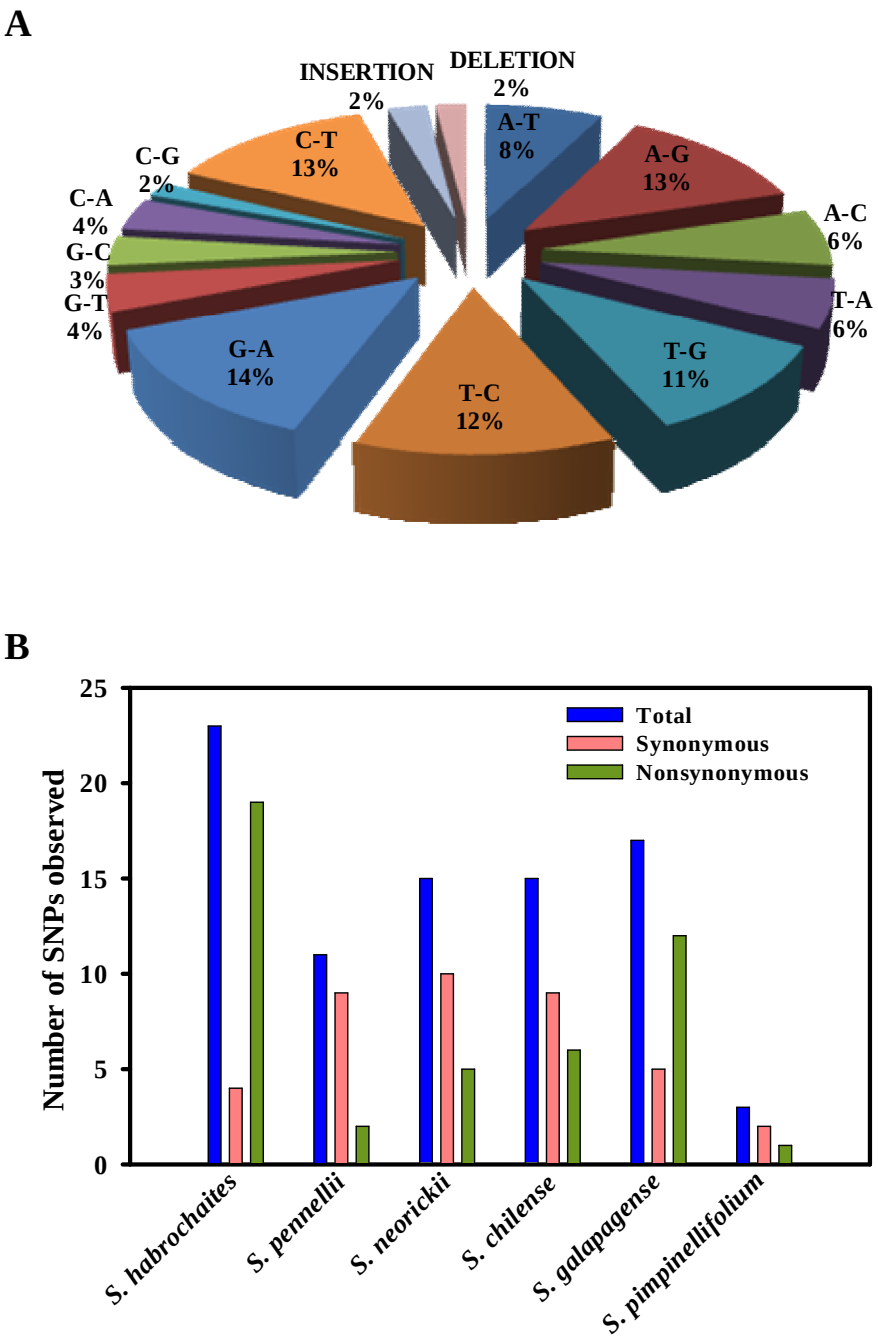


Figure 6.4 A, Nucleotide substitution patterns observed in tomato *COP1* gene. **B**, Frequency of SNPS observed in various wild relatives of tomato for *COP1* gene.

COP1 mutant identified by TILLING (chapter 4); we checked the effect of the synonymous changes observed in *COP1* gene on codon usage frequencies. As multiple synonymous changes were observed (in most of the species), some caused an increase, few resulted in decrease and few changes had no difference on codon usage efficiencies. Therefore, we could not conclude that the *COP1* gene function might have been affected due to these synonymous changes. Validation of these synonymous changes by further studies would reveal their importance in regulating gene function.

6.2.1.3 Neutrality test and phylogenetic analysis at the *COP1* locus

The value of Tajima's D based on the *COP1* locus was not significantly different from neutral expectations. The D value was negative at the *COP1* locus (**Table 6.4**), indicating a purifying selection i.e., polymorphisms would have occurred and accumulated at silent sites, but their frequency in the gene pool is very low which in turn reduces the chances of phenotypic variation in the next generations. Further, to assess the evolutionary divergence between the six wild relatives of tomato at *COP1* locus, we performed phylogenetic analysis using MEGA 4.0 software. A dendrogram was constructed based on the identified polymorphism in *COP1* gene (**Fig. 6.5**). The results showed conserved topology of the phylogenetic tree, in which the colored tomato species were close to the domesticated tomato and the green fruited species were represented as distant relatives at the bottom of the tree. Our results are consistent with the previous literature and the following patterns of relationships could be deduced from the *COP1* dendrogram: firstly, the red fruited wild relative *S. pimpinellifolium* is closely related to the domesticated *S. lyopersicum*. Secondly, all the green fruited species were found to be the ancestors of present day tomato. However, *S. chilense* was exceptionally placed in between the colored fruit species in *COP1* phylogram. This kind of discrepancy can be explained better by considering large number of accessions of wild relatives where the phylogenetic trees can be computed more stringently.

6.2.2 Sequence diversity of *SPA3LIKE* gene

Tomato *SPA3LIKE* gene has a single exon with no defined introns. The aligned length including the coding region and the 5′ and 3′ non coding regions

Table 6.4 Levels of nucleotide variation in the *COP1* and *SPA3LIKE* genes.

Gene	Fruit color	Indel	S	π_{tot}	D_{tot}
COP1	Orange /red	5	67	0.00389	-1.64465
	Green	12	1129	0.04965	
SPA3LIKE	Orange/ red	-	7	0.00151	-1.11158
	Green	1	27	0.00462	

S, number of segregating sites; π, average number of nucleotide differences per site between two sequences calculated on the total number of polymorphic sites (πtot); D, Tajima's D.

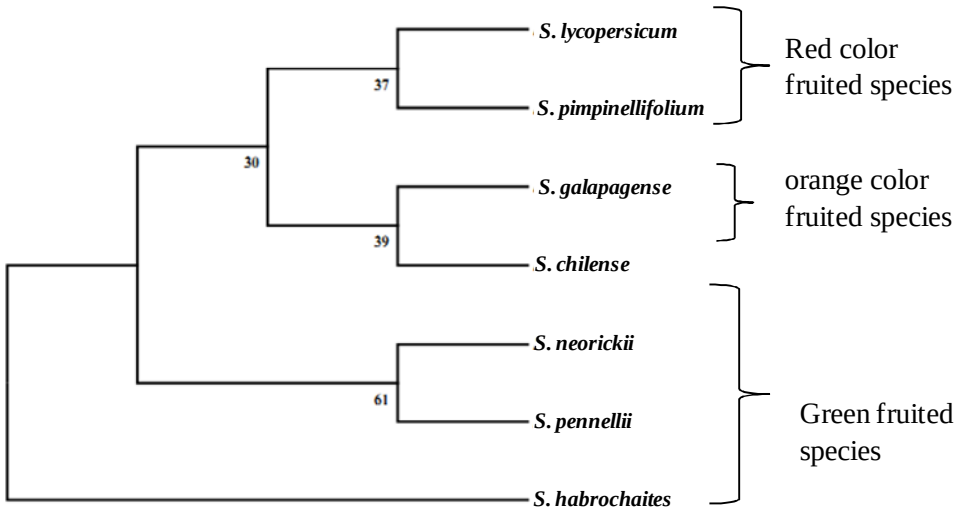


Figure 6.5 Dendrogram for six wild relatives and one domesticated species of tomato based on *COP1* gene sequence.

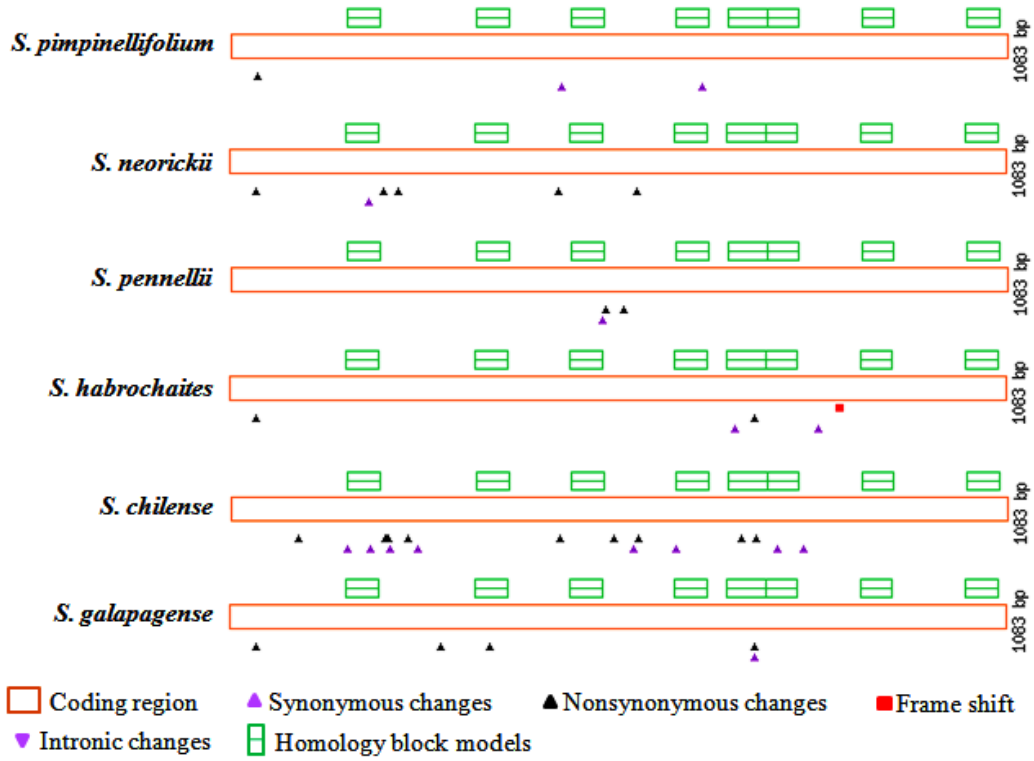


Figure 6.6 PARSESNP output file representing the distribution of SNPs for *SPA3LIKE* gene in wild relatives of tomato. Open orange boxes denote exons. Black upright triangle indicates changes in the DNA sequence that affect the amino acid. Purple upright triangle indicates changes in the DNA sequence that does not affect the amino acid. Purple downward triangles indicate changes in introns. Red boxes represent the frame shift changes.

considered for primer design was 3,100 nucleotides. A single INDEL of 1 nucleotide length was detected. The distribution of nucleotide variations among the sequenced regions of *SPA3LIKE* gene in different wild relatives considered are represented by PARSESNP output format (**Fig. 6.6**). The number of SNPs was higher than the number of INDELs, with a total of 32 SNPs at *SPA3LIKE* locus resulting in average of 1 SNP for 236 bp (**Table 6.2**). Like in *COP1*, the polymorphic diversity at *SPA3LIKE* locus was high in green fruited species than the color fruited species. Relatively fewer SNPs (7) and INDELs (0) were found in red fruited species than in the green fruited species (23 SNPs and 1 INDEL). On analyzing the polymorphism data per species, it was observed that the number of nucleotide polymorphisms for *SPA3LIKE* was high in *S. chilense* (green fruited) and least in the red fruited species - *S. pimpinellifolium* (**Fig. 6.7**).

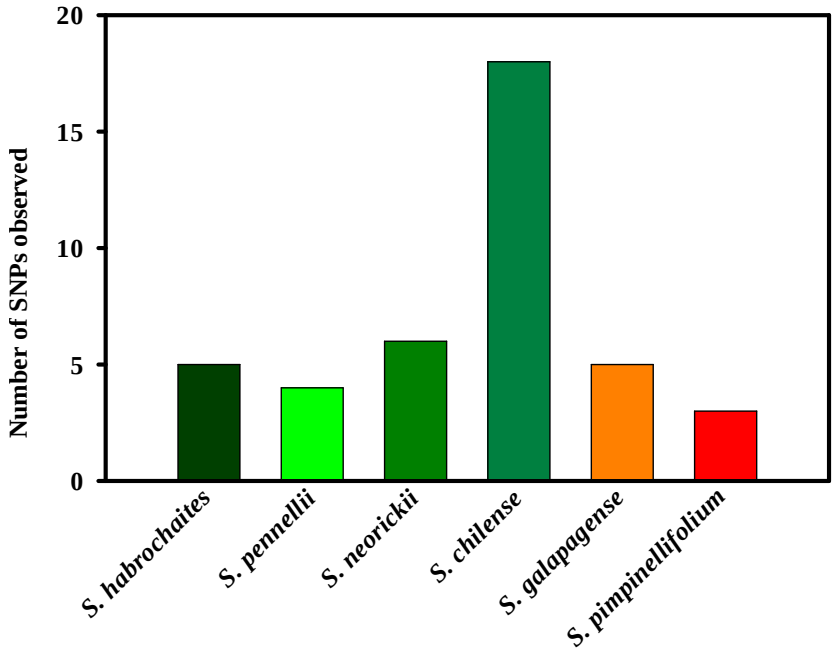


Figure 6.7 Single Nucleotide Polymorphisms observed in various wild relatives of tomato for *SPA3LIKE* gene.

The nucleotide diversity analyses showed that nucleotide diversity (**Fig. 6.8**) of green fruited varieties was higher (π (Pi 2 = 0.00462)) than the color fruited species (π (Pi 2 = 0.00151)). The sliding window analysis of the *SPA3LIKE* gene revealed high nucleotide diversity at the region located at 1,000 to 2,000 bp of the sequenced region in green fruited species. According to this plot, low nucleotide diversity was also observed in colored fruited species.

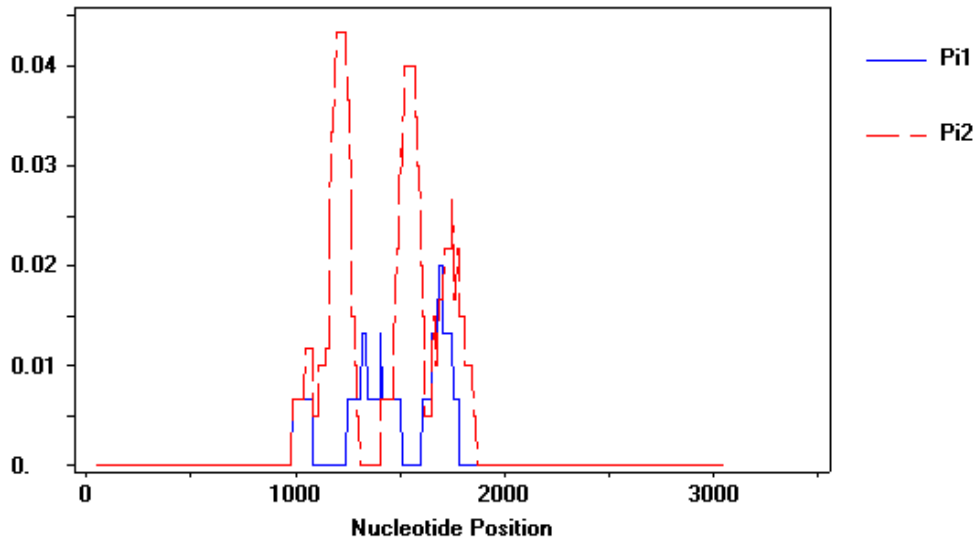


Figure 6.8 Nucleotide diversity (π) patterns along *SPA3LIKE* gene in sliding window among color fruited and green fruited wild species. Analysis was performed using a window length of 100 bp in DNASP v 5.0. Pi1 (blue color graph) and Pi2 (red colored graph) are the nucleotide diversities representing the colored fruited and green fruited species respectively. Note: in the overlaid graph, Pi2 is high compared to Pi1 on the entire length of the *SPA3LIKE* gene indicating the high nucleotide diversity in green fruited species.

6.2.2.1 Analysis of Nucleotide Base Changes

The point substitution analysis showed that the C/T transition was the most frequent type of substitution in *SPA3LIKE* gene with 22% frequency. We observed that G/A substitutions were the next prevalent base change with a frequency of 25% (**Fig. 6.9**).

Among 32 polymorphisms observed in the coding region, 14 were silent or synonymous and the rest were nonsynonymous (**Fig. 6.10**). The details of the SNPs, position on the gene and the corresponding amino acid change are listed in **Table 6.5**. Most of the polymorphisms were unique to each species excepting few which were commonly present in several species. Sequencing results revealed that the nucleotide change from A to T at 1036 position is common for *S. galapagense*, *S. pimpinellifolium*, *S. neorickii* and *S. habrochaites*. A single nucleotide insertion of G at 1851 caused frame shift in *S. habrochaites*.

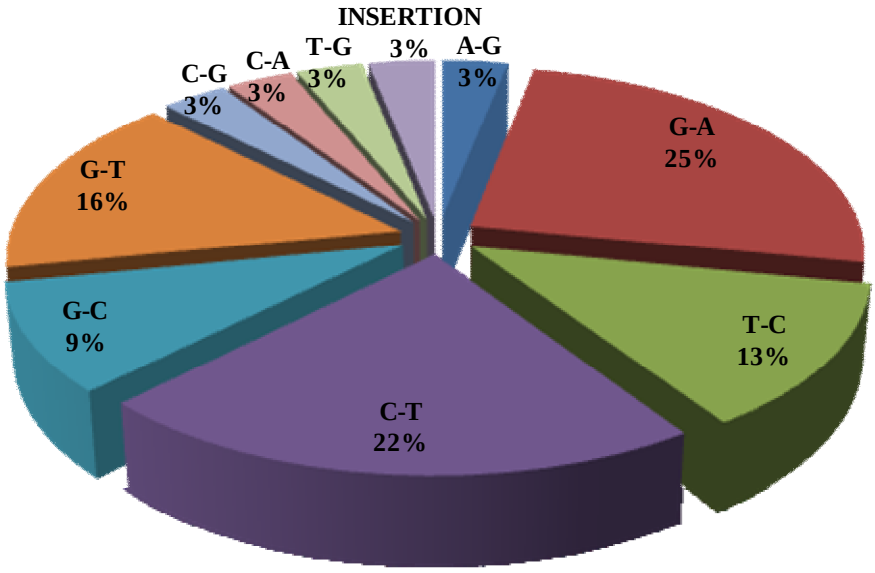


Figure 6.9 Nucleotide substitution patterns observed in tomato *SPA3LIKE* gene.

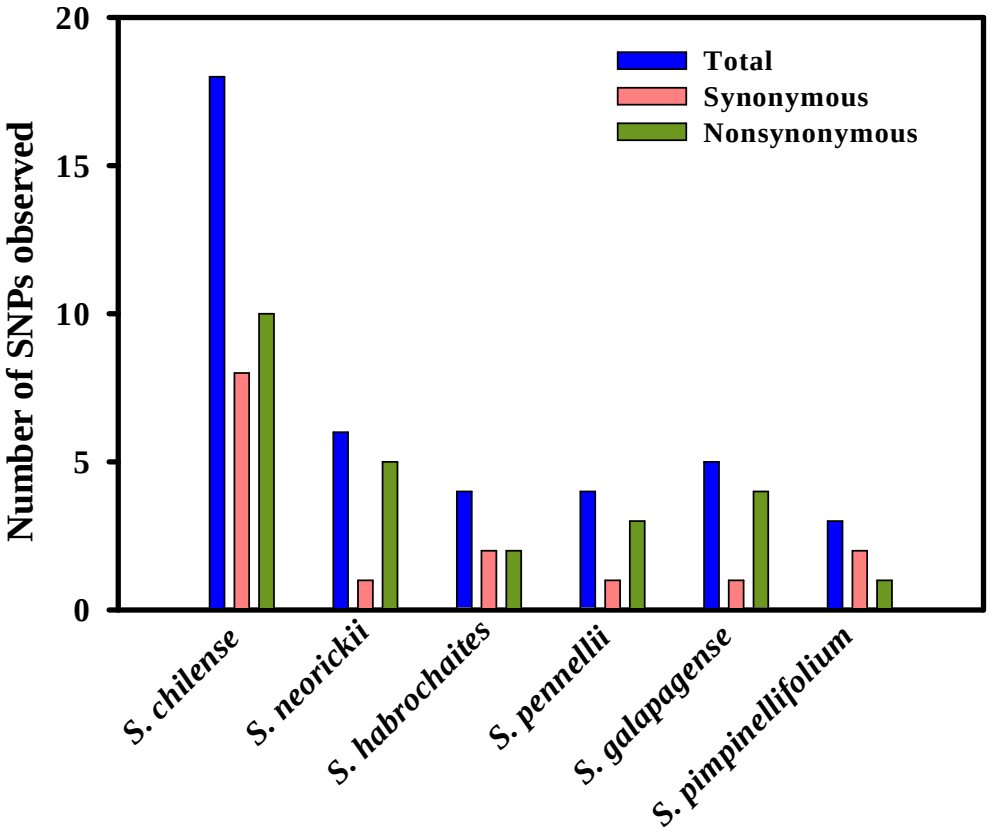


Figure 6.10 Frequency of SNPs observed in various wild relatives of tomato for *SPA3LIKE* gene.

Table 6.5 List of nucleotide polymorphisms in *SPA3LIKE* gene and their effect on amino acids.

S. No	Nucleotide change	Amino acid change	*S. habrochaites*	*S. pennellii*	*S. neorickii*	*S. chilense*	*S. galapagense*	*S. pimpinellifolium*	SIFT Score
1	A37G	K13E	+		+		+	+	0.00
2	G462C	S154=						+	
3	G657T	A219=						+	
4	G519T	P173=		+					
5	T523C	C175R		+					
6	G547T	A183S		+					
7	C548T	A183V		+					
8	G195A	R65=			+	+			
9	G214A	V72I			+	+			
10	G236A	G79D			+				
11	G459T	Q153H			+	+			
12	G569A	G190E			+	+			
13	T705C	H235=	+						
14	G733C	V245L	+			+	+		
15	G822A	K274=	+						
16	G852GG	frameshift	+						
17	G95T	C32F				+			
18	C162T	F54=				+			
19	C217G	Q73E				+			
20	G220A	G74S				+			
21	C222T	G74=				+			
22	C247A	L83I				+			
23	C261T	H87=				+			
24	G533A	S178N				+			
25	T561C	P187=				+			
26	C621T	I207=				+			
27	G712A	A238T				+			
28	C762T	S254=				+			
29	C798T	D266=				+			
30	G295C	A99P					+		
31	T363G	D121E					+		
32	T732C	F244=					+		

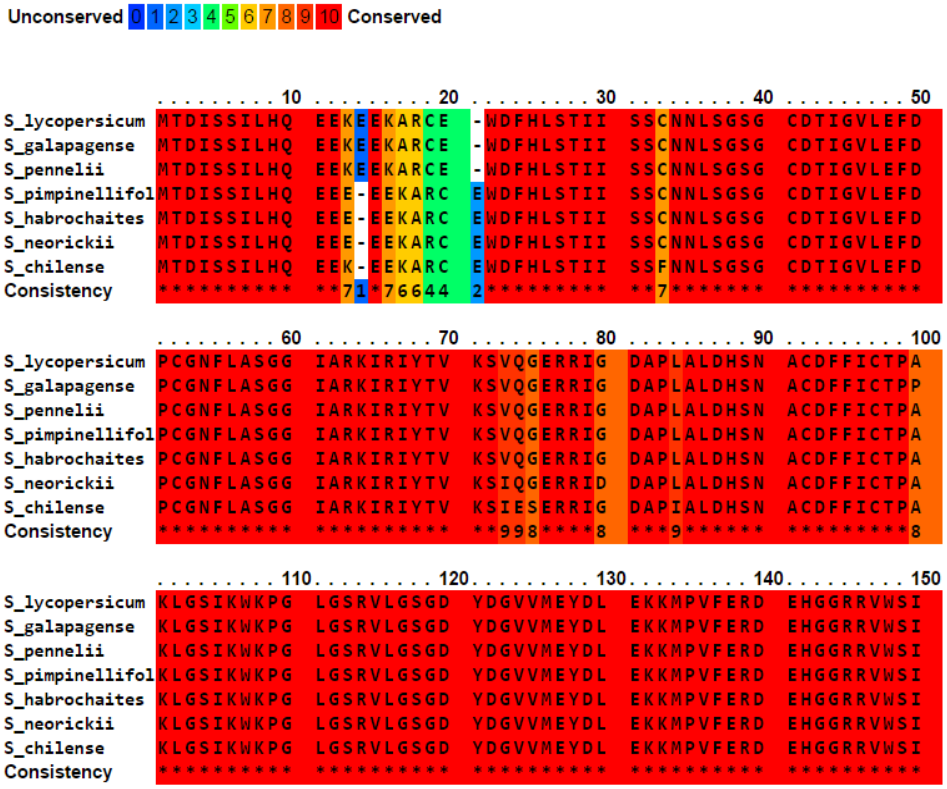
+ indicates the presence of the SNP in the wild relative.

6.2.2..2 Variations in SPA3LIKE protein sequence

SPA3LIKE is an intronless gene and any change in the coding region might affect the protein. Most of the nonsynonymous SNPs observed in *SPA3LIKE* did not show any drastic SIFT values and hence seem to be not detrimental to the protein function. However, an INDEL in *S. habrochaites* resulted in a premature stop codon at N301* in SPA3LIKE protein. The PRALINE alignment of the derived amino acids showed that the protein was highly conserved across the wild relatives (**Fig. 6.11**).

6.2.3 Neutrality test and phylogenetic analysis at the *SPA3LIKE* locus

The value of Tajima's D based on sequencing data for *SPA3LIKE* gene was negative and not significantly different from neutral expectations (**Table 6.4**). Akin to the data for *COP1* gene, the neutrality test for *SPA3LIKE* gene indicates a purifying selection at the locus. The dendrogram was constructed based on the identified polymorphism in *SPA3LIKE* gene among the six wild species (**Fig. 6.12**). With minor variations, the results were almost congruent with previous studies in tomato (Child 1990; Rick 1979; Peralta and Spooner, 2005).



Contd…

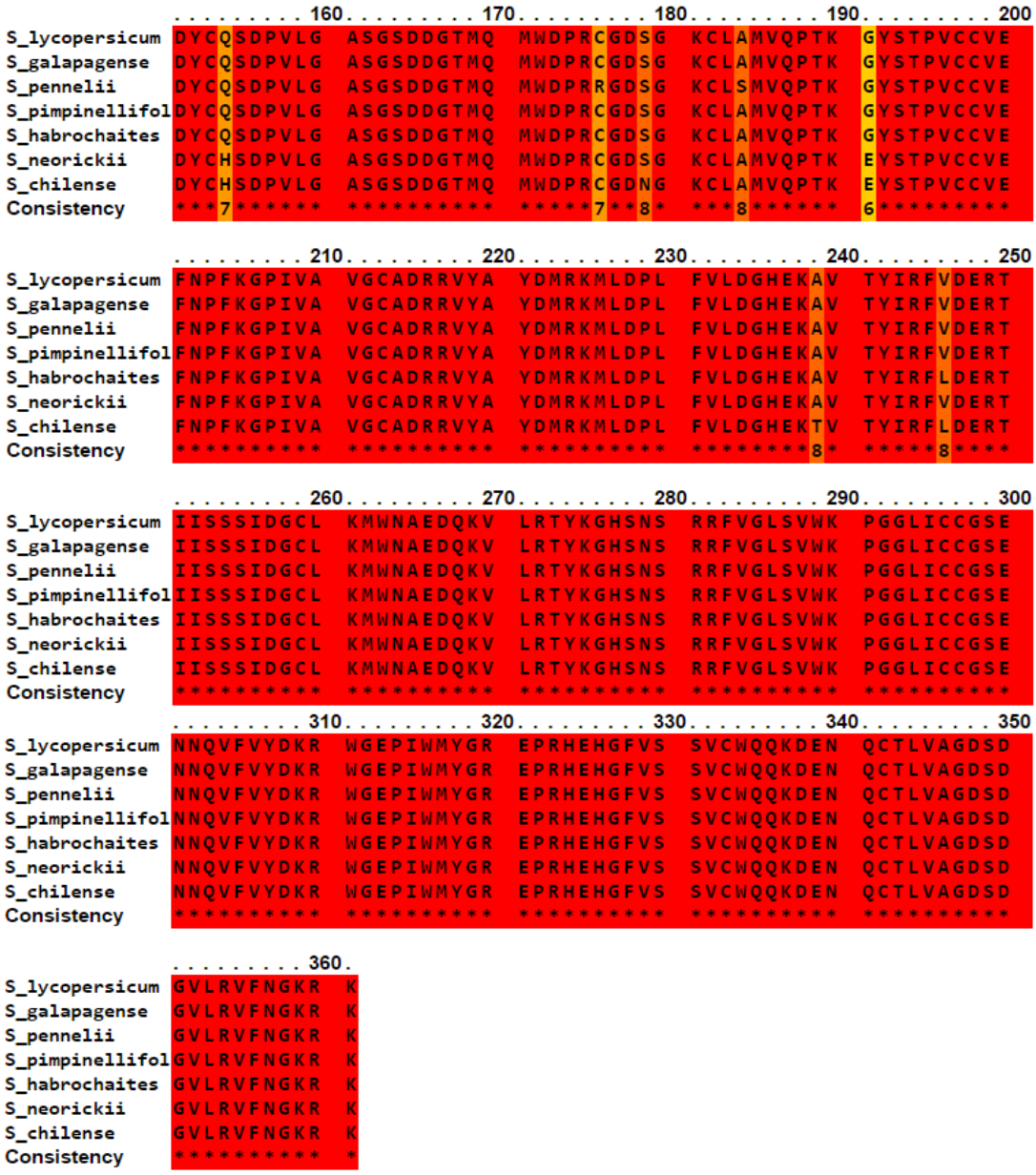


Figure 6.11 PRALINE multiple alignment output file of the derived SPA3LIKE protein sequences of the six wild relatives. The conservedness of the amino acids is denoted by a color code designed by the software. The highly unconserved moieties are denoted by dark blue color and on the contrary conserved moieties by deep red color.

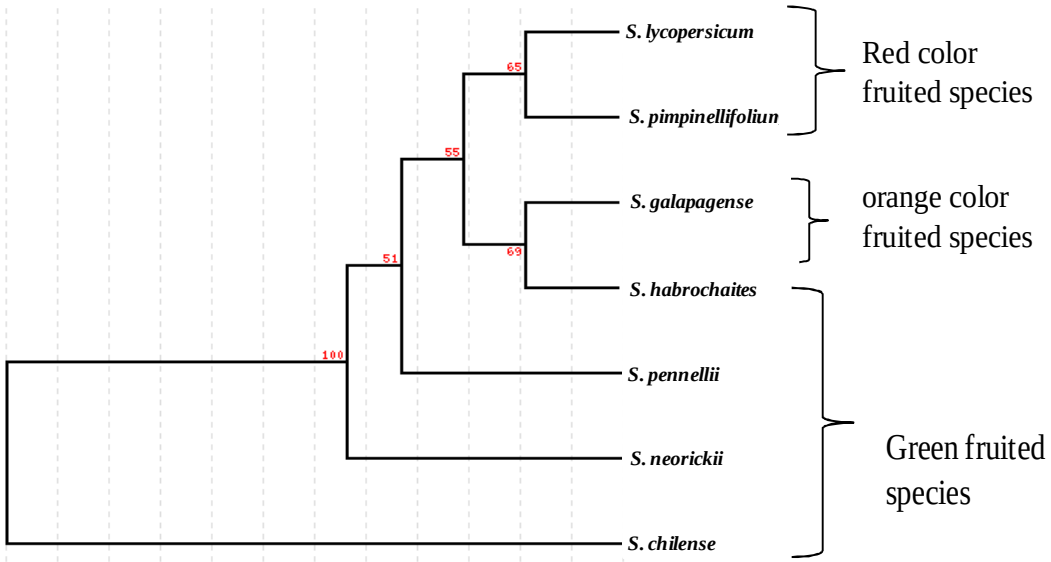


Figure 6.12 Dendrogram for six wild relatives and one domesticated species of tomato based on *SPA3LIKE* gene sequence.

6.3 Discussion

In general wild relatives of crop plants constitute an increasingly important resource for improving agronomic traits. With the recent improvements in tools available for molecular genome analysis, it is easier and quicker to identify useful traits in wild relatives and to develop new and improved varieties. Despite the enormous promise they hold for improving crops, wild relatives are not thoroughly studied. *S. lycopersicum* is the domesticated species of tomato that is widely cultivated all over the world. Apart from the widely distributed domesticated species *S. lycopersicum*, tomato has 9 wild relatives (Peralta and Spooner, 2005). These wild relatives are adapted to different geographical conditions, vary in morphological traits like leaf shape, structure and fruit color (ranging from green to red) etc. conferring their genetic potential to adapt to the ever changing environment. Understanding the genetic makeup of tomato wild relatives provides insights to evolutionary trends of different mechanisms underlying the agronomic traits (Bolger et al., 2014). In the present study we analyzed the nucleotide diversity of *COP1* and *SPA3LIKE* genes in tomato wild relatives by direct sequencing of the PCR products.

The SNPs identified were found to be distributed throughout the length of the gene with a high frequency of intronic mutations in *COP1*. As expected, the close wild relative of tomato, *S. pimpinellifolium* harbored fewer SNPs compared to the green fruited species in both *COP1* and *SPA3LIKE* genes. In a recent study on tomato by Aflitos et al., (2014), the estimated number of single nucleotide polymorphisms (SNPs) reported in wild relatives, especially in the green fruited species was 20-fold higher than found in most of the tomato accessions. This high number of nucleotide variations observed in the green fruited species like *S. habrochaites, S. pennellii, S. parviflorum* and *S. chilense* compared to the colored fruited (*S. pimpinellifolium* and *S. cheesmaniae*) species is an indicator of genetic potential of the wild species. Most probably, in the course of domestication, the variations were lost.

In terms of the overall density of polymorphisms across the regions considered for the study, an average of 1 SNP/45 Kb for *COP1* gene and 1 SNP/236 bp for *SPA3LIKE* gene were determined. As considered for any reverse genetic high-throughput technique for detecting mutations, the frequency of naturally occurring variations might also depend on the length of the gene, number of exons and introns, length of the exons, GC content, gene function etc. The low SNP frequency observed

in tomato *COP1* gene may be related to the functional importance of *COP1* gene in light mediated signalling pathway, as any spontaneous mutations could be lethal (as observed in our TILLING study). Secondly, it may be attributed to the above mentioned aspects of gene structure. In case of *SPA3LIKE* gene, the frequency seems to be relatively higher than for *COP1* gene, however, it has to be noted that *SPA3LIKE* in tomato is a very small gene (approx. 1kb) with no specified introns.

Nevertheless, the effect of genetic variations in the coding region can vary based on the position and nature of the SNPs or INDELs. Not necessarily all nucleotide changes affect the gene function. The change in nature of the amino acids determines the degree of loss of function. On the other hand, presence of INDELs in coding regions or regulatory domains leads to either truncation of the protein (due to stop codon), alter the binding sites (due to frameshifts) or affect the abundance and turnover (transcriptional regulation) of the protein, thus causing a loss of function. In spite of relatively high number of nonsynonymous changes (**Fig. 6.6** and **6.12**) observed for *COP1* and *SPA3LIKE* genes in wild relatives, the SIFT value for most of the changes was not significant to cause a drastic change in the function of both these gene products. However, the nucleotide changes scored for *COP1* at G8992A in *S. habrochaites* and C11161T in *S. pennellii* cause truncation of the protein due to premature stop codons. Similarly, a premature stop codon at N301 in SPA3LIKE protein of *S. habrochaites* is predicted to cause protein truncation. These changes may lead to loss of function of COP1 and SPA3LIKE in the respective wild relatives.

A negative Tajima's D value of -1.64465 and -1.11158 calculated from the sequences obtained reveal that *COP1* and *SPA3LIKE* genes are under purifying selection or stabilizing selection. These results were further supported by the protein conservedness as shown by PRALINE alignments. Also a Tajima's D value within the range -2 to +2 signifies that the genes have evolved neutrally. Potential explanations for such evolutionary pressure may include one or more recent genetic bottlenecks during domestication, a low mutation rate and a self-pollinating reproductive biology of *S. lycopersicum*. In corroboration with the studies of Peralta and Spooner, 2001 for *GBSSI* gene in tomato, the phylogenetic analyses of the two genes in wild relatives indicate the evolutionary trend of domesticated tomato i.e. from green fruited species to color fruited species. The position of *S. chilense* in red fruited clade in *SPA3LIKE* phylogram may be considered as an exception most

probably caused due to fewer accessions considered for the study. Nevertheless, analysis of more collections from diverse ecological niches will help to clarify this issue.

In conclusion, our investigations demonstrate that the two light signalling genes - *COP1* and *SPA3LIKE* are highly conserved in wild relatives of tomato in spite of the presence of nucleotide variations. In particular, though statistically significant signatures of selection were not detected in both target genes *COP1* and *SPA3LIKE*, low level of purifying selection that varied across the length of each gene was evident. The recalcitrance of the two genes for induced or natural polymorphisms was justified by these low diversity levels and negligible selection pressure as defined by the negative Tajima's D value. Notably, this study signifies the importance of *COP1* and *SPA3LIKE* genes in tomato fruit ripening and further investigations with the identification of multiple alleles or RNAi suppression lines would shed light on the underlying mechanisms and on identity of the interacting partners.

Chapter 7

Summary

Regulated proteolysis via ubiquitin mediated 26S proteasome is an essential process for plant growth and development. This process is carried out by three enzymes E1, E2 and E3 that sequentially activate, conjugate and transfer the ubiquitin molecule to target proteins. The plasticity of this mechanism is determined by the enormous number of E3 ligases specific to different target proteins. In plants, many physiological processes are regulated by this rapid mechanism including light mediated responses like photomorphogenesis, shade avoidance, flowering and recently fruit ripening too.

Fruit ripening is complex process involving several physiological, molecular, biochemical and metabolic changes that make the fruit nutritious and palatable. In climacteric fruits like tomato, the hierarchy of genetic and hormonal regulators (with ethylene as the major hormone) plays an important role in regulating ripening. Apart from these internal cues, light also contributes to fruit ripening (especially pigmentation) by regulating accumulation of carotenoids. The molecular and genetic characterization of *hp1* (Lieberman et al., 2004) and *hp2* (Mustilli, 1999) mutants revealed that these loci are encoded by the homologs of *DDB1* and *DET1*genes of Arabidopsis respectively. Both DDB1 and DET1 proteins are auxiliary to COP1 and together act as negative regulators of photomorphogenesis (Yi & Deng, 2005). Apart from these two auxiliary molecules, the RNAi repression of *Cullin4* (*CUL4*), a scaffolding protein subunit of E3 ligase and *COP1LIKE* (recently annotated as *SPA3LIKE*-whose protein is closely related to COP1) also exhibited altered photoresponses (Liu et al., 2004, Wang et al., 2008). Enhanced carotenoid content in these RNAi lines suggests that regulated proteolysis may be important for enhancing nutritional quality during tomato fruit ripening. The main aim of this work was to identify and isolate mutant alleles for *COP1* and *SPA3LIKE* genes using reverse genetic techniques- TILLING and Eco-TILLING.

The first objective was to isolate novel alleles of tomato *COP1* gene by TILLING. Our study highlights the need for optimizing PCR conditions for setting up cost effective TILLING. It also emphasizes the effect of synonymous changes on *COP1* gene function and difficulties in isolating *COP1* mutants.

Based on the CODDLE (Codons Optimized to Discover Deleterious Lesions) prediction analysis for the highly mutable regions of *COP1*, and TILLING of

approximately 10,000 M_2 EMS mutagenized tomato plants, it was observed that i) the exon 3 and 4 of *COP1* gene were hot spots of mutations, and ii) the frequency of mutations in *COP1* was very low, almost equal to the order of spontaneous mutations. Though nested strategy with labeled universal M13 primers in combination with M13 tagged gene specific primers was adapted to improve frequency of mutation detection, only three mutations were identified for *COP1* gene. Nonetheless, we obtained a single mutant line harboring a point mutation in exon 5, 6 and 7 regions. Sequencing of the mutated *COP1* gene revealed a substitution of C at 6964 to T on the genomic co-ordinates. PARSESNP (Project Aligned Related Sequences and Evaluate SNPs) analysis of the mutated sequence revealed a synonymous change i.e., no change in amino acid in the mutant line as both the codons, CUA (wild type) and UUG (altered/mutated) code for leucine.

Interestingly, the mutant plant exhibited an altered phenotype. The M_3 homozygous adult mutant plants showed delayed and stunted growth compared to the wild type Arka Vikas (AV). Importantly, the red ripe fruit of mutant showed two fold enrichment of lycopene over the WT. These results were consistent with the hypothesis that down regulation of *COP1* may have up-regulated the carotenoid levels in tomato (Liu et al., 2004). The examination of co-segregation of genotype and phenotype in subsequent generations led to interesting results. The homozygous mutant plant segregated into short (P1-S) and long (P1-L) plants. The short plant also showed an architectural difference in internode differentiation, notably after 5^{th} node. However, the carotenoid content of red ripe fruits of both the plants, P1-L and P1-S was similar and also twofold higher than AV.

In view of the synonymous mutation in *COP1* causing a distinct phenotype, the above intriguing results led us to examine the presence of additional mutations within the gene. Sequencing full length *COP1* gene from both the mutant lines (P1-S and P1-L) along with AV revealed another synonymous change (AC**T** to AC**C** coding for threonine) in exon 1 of the short plant (P1-S). Both the synonymous mutations led to non-preferred codons in the mutant lines as the frequency usage of both codons is 50% lower than the wild type codon. While there are no detailed studies on synonymous changes affecting phenotype in plant system, in animal system, studies on synonymous mutations (Wang et al., 2005) indicate that change of codon affects the gene function by regulating the mRNA folding stability and the transcript levels.

These in turn affect the protein abundance, structure and therefore function (Kimchi-Sarfaty et al., 2007). Considering that changes in codon may alter mRNA structure as well as efficiency of translation, we examined the *COP1* transcript levels in these mutant lines and AV. Quantification of *COP1* expression by RT-qPCR revealed a 60% and 80% reduction in *COP1* transcript levels in P1-L and P1-S respectively compared to AV (with respect to *β-actin* and *ubiquitin3* genes as internal controls).

Taking into account that Arabidopsis COP1 acts as a part of large complex and interacts with SPA1 (suppressor of phytochrome A1) which in turn mediates the degradation of HY5 (elongated hypocotyl 5), a photomorphogenesis promoting transcription factor, we examined *SPA1* and *HY5* transcript levels in the mutant lines. Strikingly, the levels of *SPA1* transcript were similarly reduced like *COP1* in both P1-S and P1-L lines. The observed increase of *HY5* transcript levels in the mutant lines indicates that the noticeable reduction in *COP1* transcript levels may have alleviated the repressive effect of COP1 on *HY5* expression. Our results suggest that the codon change in *COP1* may lower the COP1 protein abundance and thus stimulate the lycopene content. To specifically delineate the role of *COP1* on fruit pigmentation, the transcript levels of *PSY1* (*phytoene synthase 1*) and *CYCB* (*chromoplast- specific lycopene beta cyclase*) genes which encode the rate limiting enzymes of the carotenoid biosynthetic pathway were examined. The transcript levels of *PSY1* were up-regulated and positively correlated with increased lycopene accumulation in the mutant fruits.

Considering the above results, it is evident that the mutant line P1-S harboring two silent mutations is a promising line that lacks *COP1*. However, to rule out the possibilities of any other secondary mutations and to establish a stable mutant line, we made genetic cross between WT and P1-S. The F_2 segregation ratios (for the silent mutation L342=) was not in accordance with Mendelian ratio of monogenic trait. Moreover, we could not co-segregate the second silent mutation (T31=) due to difficulties faced during labeled PCR with SETVI primers and also sequencing. Hypothesizing that the second silent mutation may or may not co-segregate with the L342= in the present F_2 population, we are in the process of back crossing *COP1* mutant line with parent cultivar to examine a large number of F_2 population. Once a high lycopene tomato cultivar is confirmed in F_2, it can also be used to understand the significance of the underlying synonymous mutations.

The natural genetic variation existing in a population also serves as a valuable resource for examining fruit quality and yield related traits. Based on this we framed our second objective to detect the single nucleotide polymorphisms in *COP1* gene using AV as reference cultivar using a modified TILLING protocol- Eco-TILLING (Comai et al., 2004). SNPs in *COP1* gene were examined in nearly 600 tomato accessions obtained from NBPGR, New Delhi (300 accessions), TGRC, California (85 accessions), IIVR, Varanasi (140 accessions) and Bejo Sheetal Seeds, Jalna (60 accessions). In addition to the four sets of primers used for TILLING *COP1* gene, two additional sets of primers were designed to analyze exon 1 and exon 2 of the gene.

Based on sequencing of the amplicons, a total of 12 haplotypes were identified for *COP1*. Interestingly, most of the SNPs were located in the introns compared to exons, which indicated very low polymorphism at protein level. Among the exons, more SNPs were found in the exons 3 and 4 supporting the *in-silico* CODDLE prediction for *COP1* gene, wherein, exon 3 and 4 were predicted to be the highly mutable regions. Though the SNPs were also observed in the exons coding for WD40 domain, PARSESNP analysis revealed that these SNPs do not have any drastic effect on the protein function. Overall, 28 SNPs and 3 INDELS were detected with a mean of 1 SNP/160Kb and 1 INDEL/1497Kb. The accessions NB104, NB471, NB110 and NB515 belonging to haplotype 11 also showed enhanced levels of lycopene.

Considering the functional similarity of *SPA3LIKE* with *COP1*, polymorphism was also examined in *SPA3LIKE* which is an intronless gene. Four different haplotypes were identified for *SPA3LIKE* gene with an SNP frequency of 1/24 Kb and INDEL frequency of 1/124 Kb. PARSESNP analysis revealed three insertions in line NB518 and a single insertion in line NB182 resulting in frameshifts. In both these accessions, frameshifts caused stop codons resulting in premature termination of the protein after 169th amino acid in line NB518 and after 164th amino acid in NB182. Protein modeling and *in-silico* analysis of binding sites of SPA3LIKE protein from both reference and variants clearly indicated the loss of binding sites and distortion of protein structure in NB518 and NB182, thereby rendering the protein nonfunctional.

Since an earlier study showed that *COP1LIKE* RNAi line exhibits defects in photoresponses and increased lycopene levels (Liu et al., 2004), we also examined the phenotype of these lines. In congruence with the phenotype of *COP1LIKE* RNAi line, the seedlings of NB182 showed altered phenotype. When grown under white light, these seedlings exhibited short hypocotyls and a two day delay in opening of the

hook. However, unlike the *cop1* mutants of Arabidopsis, the above accessions did not show any increased anthocyanin accumulation in the seedlings. The phenotype of adult plants of both the accessions was similar to the wild type but fruits displayed increased lycopene content. The transcript levels of *SPA3LIKE* in the fruit tissue of these accessions were similar to AV, supporting that the frameshift mutations may have altered the protein function due to premature truncation.

Tomato has a wide range of geographical distribution and its wild relatives exhibit different fruit colors, ranging from green to red. This serves as an excellent genetic resource to understand the diversity of various loci in relation to evolutionary pressures. Taking a subset of green and red fruited wild relatives, the variation was examined in *COP1* and *SPA3LIKE* genes. Consistent with its closeness to tomato, red fruited *S. pimpinellifolium* harbored fewer SNPs compared to the green fruited species. In terms of the overall density of polymorphisms across the regions that were analyzed, we found an average of 1 SNP/45 Kb for *COP1* and 1/236 bp for *SPA3LIKE*. The nucleotide diversity for the *COP1*gene was observed to be very low.

The ratio of nonsynonymous (dN) to synonymous (dS) SNPs observed is considered as an important measure to estimate the selection pressures on a gene locus (Nielsen, 2001). Low dN/dS ratio was observed for both *COP1* and *SPA3LIKE* suggesting that the genes might be evolving primarily under purifying selection. A negative Tajima's D value of -1.64465 for *COP1* and -1.11158 for *SPA3LIKE* calculated using DnasP software suggested that the genes were under neutral evolution. The above results suggest that the variations might have occurred due to random events and also they may not have drastic effects on the protein function.

In conclusion, characterization of the *COP1* mutant allele and the *SPA3LIKE* natural variants indicated a probable role for these negative regulators of photomorphogenesis in increasing lycopene level in tomato fruits. In addition, we observed that the variation-either induced or natural was very low in *COP1* gene and a slightly high level of polymorphism was observed in *SPA3LIKE* gene. Our findings emphasize that both *COP1* and *SPA3LIKE* genes are essential for the plant development and therefore, are highly conserved.

References

Adams-Phillips L, Barry C, Giovannoni J (2004) Signal transduction systems regulating fruit ripening. Trends Plant Sci**9:** 331-338

Aflitos S, Schijlen E, Jong H, Ridder D, Smit S, Finkers R, Wang J, Zhang G, Li N, Mao L (2014) Exploring genetic variation in the tomato (Solanum section *Lycopersicon*) clade by whole-genome sequencing. Plant J **80:** 136-148

Alba R, Cordonnier-Pratt M-M, Pratt LH (2000) Fruit-localized phytochromes regulate lycopene accumulation independently of ethylene production in tomato. Plant Physiol **123:** 363-370

Al-Sady B, Ni W, Kircher S, Schäfer E, Quail PH (2006) Photoactivated phytochrome induces rapid PIF3 phosphorylation prior to proteasome-mediated degradation. Molecular cell **23:** 439-446

Ang L-H, Chattopadhyay S, Wei N, Oyama T, Okada K, Batschauer A, Deng X-W (1998) Molecular Interaction between COP1 and HY5 Defines a Regulatory Switch for Light Control of< i> Arabidopsis</i> Development. Molecular cell **1:** 213-222

Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics **22:** 195-201

Arnon DI (1949) Copper enzymes in isolated chloroplasts. Polyphenoloxidase in Beta vulgaris. Plant Physiol **24:** 1

Bach I, Ostendorff HP (2003) Orchestrating nuclear functions: ubiquitin sets the rhythm. Trends Biochem Sci **28:** 189-195

Baert J, Monte D, Verreman K, Degerny C, Coutte L, De Launoit Y (2010) The E3 ubiquitin ligase complex component COP1 regulates PEA3 group member stability and transcriptional activity. Oncogene **29:** 1810-1820

Bai M-Y, Shang J-X, Oh E, Fan M, Bai Y, Zentella R, Sun T-p, Wang Z-Y (2012) Brassinosteroid, gibberellin and phytochrome impinge on a common transcription module in Arabidopsis. Nature Cell Biol **14:** 810-817

Baker N, Nogués S, Allen D (1997) Photosynthesis and photoinhibition. *In* SEMINAR SERIES-SOCIETY FOR EXPERIMENTAL BIOLOGY, Vol 64. Cambridge University Press, pp 95-112

Banerjee R, Batschauer A (2005) Plant blue-light receptors. Planta **220:** 498-502

Bapat VA, Trivedi PK, Ghosh A, Sane VA, Ganapathi TR, Nath P (2010) Ripening of fleshy fruit: molecular insight and the role of ethylene. Biotechnology advances **28:** 94-107

Barkley NA, Wang M, Gillaspie A, Dean R, Pederson G, Jenkins T (2008) Discovering and verifying DNA polymorphisms in a mung bean [*V. radiata* (L.) R. Wilczek] collection by Eco-TILLING and sequencing. BMC Res Notes **1:** 28

Barkley NA, Wang ML, Gillaspie AG, Dean RE, Pederson GA, Jenkins TM (2008) Discovering and verifying DNA polymorphisms in a mung bean [V. radiata (L.) R. Wilczek] collection by EcoTILLING and sequencing. BMC research notes **1:** 28

Bauchet G, Causse M (2012) Genetic diversity in tomato (*Solanum lycopersicum*) and its wild relatives. Genetic diversity in plants**:** 133-162

Bauer D, Viczián A, Kircher S, Nobis T, Nitschke R, Kunkel T, Panigrahi KC, Ádám É, Fejes E, Schäfer E (2004) Constitutive photomorphogenesis 1 and multiple photoreceptors control degradation of phytochrome interacting factor 3, a transcription factor required for light signaling in Arabidopsis. Plant Cell **16:** 1433-1445

Bawono P, Heringa J (2014) PRALINE: A Versatile Multiple Sequence Alignment Toolkit. *In* Multiple Sequence Alignment Methods. Springer, pp 245-262

Beggs CJ, Wellmann E (1994) Photocontrol of flavonoid biosynthesis. *In* Photomorphogenesis in plants. Springer, pp 733-751

Bianchi E, Denti S, Catena R, Rossetti G, Polo S, Gasparian S, Putignano S, Rogge L, Pardi R (2003) Characterization of human constitutive photomorphogenesis protein 1, a RING

finger ubiquitin ligase that interacts with Jun transcription factors and modulates their transcriptional activity. J Biol Chem **278:** 19682-19690

Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acid Res 340

Boccalandro HE, Mazza CA, Mazzella MA, Casal JJ, Ballaré CL (2001) Ultraviolet B radiation enhances a phytochrome-B-mediated photomorphogenic response in Arabidopsis. Plant Physiol **126:** 780-788

Botticella E, Sestili F, Hernandez-Lopez A, Phillips A, Lafiandra D (2011) High resolution melting analysis for the detection of EMS induced mutations in wheat SbeIIa genes. BMC Plant Biol **11:** 156

Bradbury LM, Fitzgerald TL, Henry RJ, Jin Q, Waters DL (2005) The gene for fragrance in rice. Plant Biotech J **3:** 363-370

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem **72:** 248-254

Briggs W, Beck C, Cashmore A, Christie J, Hughes J, Jarillo J, Kagawa T, Kanegae H, Liscum E, Nagatani A (2001) The phototropin family of photoreceptors. Plant Cell **13:** 993-997

Briggs WR, Christie JM (2002) Phototropins 1 and 2: versatile plant blue-light receptors. Trends Plant Sci**7:** 204-210

Bush SM, Krysan PJ (2010) iTILLING: a personalized approach to the identification of induced mutations in Arabidopsis. Plant Physiol **154:** 25-35

Caldwell DG, McCallum N, Shaw P, Muehlbauer GJ, Marshall DF, Waugh R (2004) A structured mutant population for forward and reverse genetics in Barley (Hordeum vulgare L.). Plant J **40:** 143-150

Cashmore AR, Jarillo JA, Wu Y-J, Liu D (1999) Cryptochromes: blue light receptors for plants and animals. Science **284:** 760-765

Chamary J, Hurst LD (2009) The price of silent mutations. Scientific American **300:** 46-53

Chawade A, Sikora P, Bräutigam M, Larsson M, Vivekanand V, Nakash MA, Chen T, Olsson O (2010) Development and characterization of an oat TILLING-population and identification of mutations in lignin and β-glucan biosynthesis genes. BMC Plant Biol **10:** 86

Chen H, Huang X, Gusmaroli G, Terzaghi W, Lau OS, Yanagawa Y, Zhang Y, Li J, Lee J-H, Zhu D (2010) Arabidopsis CULLIN4-damaged DNA binding protein 1 interacts with CONSTITUTIVELY PHOTOMORPHOGENIC1-SUPPRESSOR OF PHYA complexes to regulate photomorphogenesis and flowering time. Plant Cell **22:** 108-123

Chen L, Wang S-Q, Hu Y-G (2011) Detection of Snps in the VRN-A1 Gene of Common Wheat ('Triticum aestivum'L.) by a Modified Ecotilling Method Using Agarose Gel Electrophoresis. Australian J Crop Sci **5:** 321

Chen M, Chory J (2011) Phytochrome signaling mechanisms and the control of plant development. Trends Cell Biol **21:** 664-671

Chen M, Chory J, Fankhauser C (2004) Light signal transduction in higher plants. Annu. Rev. Genet. **38:** 87-117

Cheung W, Friesen L, Rakow G, Seguin-Swartz G, Landry B (1997) A RFLP-based linkage map of mustard [*Brassica juncea* (L.) Czern. and Coss.]. Theort Appl Gen **94:** 841-851

Child A (1990) A synopsis of Solanum subgenus Potatoe (G. Don)(D'Arcy)(Tuberarium (Dun.) Bitter (s. 1.)). Feddes Repertorium **101:** 209-235

Christie JM (2007) Phototropin blue-light receptors. Annu. Rev. Plant Biol. **58:** 21-45

Christie JM, Arvai AS, Baxter KJ, Heilmann M, Pratt AJ, O'Hara A, Kelly SM, Hothorn M, Smith BO, Hitomi K (2012) Plant UVR8 photoreceptor senses UV-B by tryptophan-mediated disruption of cross-dimer salt bridges. Science **335:** 1492-1496

Christie JM, Murphy AS (2013) Shoot phototropism in higher plants: new light through old concepts. American journal of botany **100:** 35-46

Colbert T, Till BJ, Tompa R, Reynolds S, Steine MN, Yeung AT, McCallum CM, Comai L, Henikoff S (2001) High-throughput screening for induced point mutations. Plant Physiol **126:** 480-484

Coleman JR, Papamichail D, Skiena S, Futcher B, Wimmer E, Mueller S (2008) Virus attenuation by genome-scale changes in codon pair bias. Science **320:** 1784-1787

Comai L, Henikoff S (2006) TILLING: practical single-nucleotide mutation discovery. Plant J **45:** 684-694

Comai L, Young K, Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson JE, Burtner C, Odden AR (2004) Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. Plant J **37:** 778-786

Cordeiro G, Eliott FG, Henry RJ (2006) An optimized ecotilling protocol for polyploids or pooled samples using a capillary electrophoresis system. Anal Biochem **355:** 145-147

Corpet F (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acid Res**16:** 10881-10890

Dahmani-Mardas F, Troadec C, Boualem A, Leveˆque S, Alsadon AA, Aldoss AA, Dogimont C, Bendahmane A (2010) Engineering melon plants with improved fruit shelf life using the TILLING approach. PLoS One **5:** e15776

de Lucas M, Davière J-M, Rodríguez-Falcón M, Pontin M, Iglesias-Pedraz JM, Lorrain S, Fankhauser C, Blázquez MA, Titarenko E, Prat S (2008) A molecular framework for light and gibberellin control of cell elongation. Nature **451:** 480-484

Dehan K, Tal M (1978) Salt tolerance in the wild relatives of the cultivated tomato: responses of *Solanum pennellii* to high salinity. Irrigation Science **1:** 71-76

Deng X-W, Caspar T, Quail PH (1991) cop1: a regulatory locus involved in light-controlled development and gene expression in Arabidopsis. Genes Dev **5:** 1172-1182

Deng XW, Quail PH (1992) Genetic and phenotypic characterization of cop1 mutants of *Arabidopsis thaliana*. Plant J **2:** 83-95

Dentin R, Liu Y, Koo S-H, Hedrick S, Vargas T, Heredia J, Yates J, Montminy M (2007) Insulin modulates gluconeogenesis by inhibition of the co-activator TORC2. Nature **449:** 366-369

Deschamps S, Campbell MA (2010) Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery. Mol Breed **25:** 553-570

Devoto A, Muskett PR, Shirasu K (2003) Role of ubiquitination in the regulation of plant defence against pathogens. Curr. Opinion Plant Biol **6:** 307-311

Don R, Cox P, Wainwright B, Baker K, Mattick J (1991) 'Touchdown'PCR to circumvent spurious priming during gene amplification. Nucleic Acid Res**19:** 4008

Dong C, Vincent K, Sharp P (2009) Simultaneous mutation detection of three homoeologous genes in wheat by High Resolution Melting analysis and Mutation Surveyor®. BMC Plant Biol **9:** 143

Dornan D, Shimizu H, Mah A, Dudhela T, Eby M, O'Rourke K, Seshagiri S, Dixit VM (2006) ATM engages autodegradation of the E3 ubiquitin ligase COP1 after DNA damage. Science **313:** 1122-1126

Egashira H, Ishihara H, Takashina T, Imanishi S (2000) Genetic diversity of theperuvianum-complex'(Lycopersicon peruvianum (L.) Mill. and L. chilense Dun.) revealed by RAPD analysis. Euphytica **116:** 23-31

Eisinger WR, Bogomolni RA, Taiz L (2003) Interactions between a blue-green reversible photoreceptor and a separate UV-B receptor in stomatal guard cells. American J Bot **90:** 1560-1566

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One **6:** e19379

Feng S, Martinez C, Gusmaroli G, Wang Y, Zhou J, Wang F, Chen L, Yu L, Iglesias-Pedraz JM, Kircher S (2008) Coordinated regulation of *Arabidopsis thaliana* development by light and gibberellins. Nature **451:** 475-479

Franklin KA, Allen T, Whitelam GC (2007) Phytochrome A is an irradiance-dependent red light sensor. Plant J **50:** 108-117

Frary A, Nesbitt TC, Frary A, Grandillo S, Van Der Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB (2000) fw2. 2: a quantitative trait locus key to the evolution of tomato fruit size. Science **289:** 85-88

Fraser PD, Bramley PM (2004) The biosynthesis and nutritional uses of carotenoids. Progress in lipid research **43:** 228-265

Frerichmann SL, Kirchhoff M, Müller AE, Scheidig AJ, Jung C, Kopisch-Obuch FJ (2013) EcoTILLING in Beta vulgaris reveals polymorphisms in the FLC-like gene BvFL1 that are associated with annuality and winter hardiness. BMC Plant Biol **13:** 52

Fujisawa M, Nakano T, Shima Y, Ito Y (2013) A large-scale identification of direct targets of the tomato MADS box transcription factor RIPENING INHIBITOR reveals the regulation of fruit ripening. Plant Cell **25:** 371-386

Gady AL, Hermans FW, Van de Wal MH, van Loo EN, Visser RG, Bachem CW (2009) Implementation of two high through-put techniques in a novel application: detecting point mutations in large EMS mutated plant populations. Plant methods **5:** 13

Gady AL, Vriezen WH, Van de Wal MH, Huang P, Bovy AG, Visser RG, Bachem CW (2012) Induced point mutations in the phytoene synthase 1 gene cause differences in carotenoid content during tomato fruit ripening. Mol Breed **29:** 801-812

Galeano CH, Gomez M, Rodriguez LM, Blair MW (2009) CEL I nuclease digestion for SNP discovery and marker development in common bean (L.). Crop Science **49:** 381-394

Gilchrist EJ, Haughn GW, Ying CC, Otto SP, Zhuang J, Cheung D, Hamberger B, Aboutorabi F, Kalynyak T, Johnson L (2006) Use of Ecotilling as an efficient SNP discovery tool to survey genetic variation in wild populations of Populus trichocarpa. Mol Ecol **15:** 1367-1378

Giovannucci E, Rimm EB, Liu Y, Stampfer MJ, Willett WC (2002) A prospective study of tomato products, lycopene, and prostate cancer risk. J Nat Cancer Institute **94:** 391-398

Giuliano G (2014) Plant carotenoids: genomics meets multi-gene engineering. Curr. Opinion Plant Biol **19:** 111-117

Gruber H, Heijde M, Heller W, Albert A, Seidlitz HK, Ulm R (2010) Negative feedback regulation of UV-B–induced photomorphogenesis and stress acclimation in Arabidopsis. Proctl Nat Acad Sci **107:** 20132-20137

Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis **30:** S162-S173

Guiderdoni E (1991) Gametic selection in anther culture of rice (*Oryza sauva* L.). Theort Appl Gen **81:** 406-412

Gupta S, Charakana C, Sreelakshmi Y, Sharma R (2011) Fluorescent dye labeled DNA size standards for molecular mass detection in visible/infrared range. BMC Res Notes **4:** 12

Gupta SK, Sharma S, Santisree P, Kilambi HV, Appenroth K, Sreelakshmi Y, Sharma R (2014) Complex and shifting interactions of phytochromes regulate fruit development in tomato. Plant Cell Environ **37:** 1688-1702

Hajjar R, Hodgkin T (2007) The use of wild relatives in crop improvement: a survey of developments over the last 20 years. Euphytica **156:** 1-13

Hamilton A, Lycett G, Grierson D (1990) Antisense gene that inhibits synthesis of the hormone ethylene in transgenic plants.

Harborne JB (1967) Comparative biochemistry of the flavonoids.

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

Hartmann-Petersen R, Seeger M, Gordon C (2003) Transferring substrates to the 26S proteasome. Trends Biochem Sci **28:** 26-31

Hatfield PM, Gosink MM, Carpenter TB, Vierstra RD (1997) The ubiquitin-activating enzyme (E1) gene family in *Arabidopsis thaliana*. Plant J **11:** 213-226

Hecker KH, Roux KH (1996) High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR. Biotechniques **20:** 478-485

Heijde M, Ulm R (2012) UV-B photoreceptor-mediated signalling in plants. Trends Plant Sci**17:** 230-237

Hellmann H, Estelle M (2002) Plant development: regulation by protein degradation. Science **297:** 793-797

Henikoff JG, Greene EA, Pietrokovski S, Henikoff S (2000) Increased coverage of protein families with the blocks database servers. Nucleic Acid Res**28:** 228-230

Henikoff S, Comai L (2003) Single-nucleotide mutations for plant functional genomics. Annu. Rev Plant Biol **54:** 375-401

Henikoff S, Till BJ, Comai L (2004) TILLING. Traditional mutagenesis meets functional genomics. Plant Physiol **135:** 630-636

Henriques R, Jang I-C, Chua N-H (2009) Regulated proteolysis in light-related signaling pathways. Curr. Opinion Plant Biol **12:** 49-56

Hershko A (1998) The ubiquitin system. Springer

Hicke L (2001) Protein regulation by monoubiquitin. Nat Rev Mol Cell Biol **2:** 195-201

Hirschberg J (2001) Carotenoid biosynthesis in flowering plants. Curr. Opinion Plant Biol **4:** 210-218

Höblinger A, Grünhage F, Sauerbruch T, Lammert F (2009) Association of the c. 3972C> T variant of the multidrug resistance-associated protein 2 Gene (MRP2/ABCC2) with susceptibility to bile duct cancer. Digestion **80:** 36-39

Hoecker U (2005) Regulated proteolysis in light signaling. Curr Opinion Plant Biol **8:** 469-476

Hoecker U, Quail PH (2001) The Phytochrome A-specific Signaling Intermediate SPA1 Interacts Directly with COP1, a Constitutive Repressor of Light Signaling in Arabidopsis. J Biol Chem **276:** 38173-38178

Hohenlohe PA, Amish SJ, Catchen JM, Allendorf FW, Luikart G (2011) Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. Mol Ecol Res **11:** 117-122

Hohm T, Preuten T, Fankhauser C (2013) Phototropism: translating light into directional growth. American J Bot **100:** 47-59

Hohmann U, Jacobs G, Jung C (2005) An EMS mutagenesis protocol for sugar beet and isolation of non-bolting mutants. Plant Breed **124:** 317-321

Holland JJ, Roberts D, Liscum E (2009) Understanding phototropism: from Darwin to today. J Expt Bot 113

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

Holm M, Ma L-G, Qu L-J, Deng X-W (2002) Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in Arabidopsis. Genes Dev **16:** 1247-1259

Holsinger K (2012) Tajima's D, Fu's Fs, Fay and Wu's H, and Zeng et al.'s E. URL: http://darwin. eeb. uconn. edu/eeb348/lecture-notes/molevol-tajima. pdf

Huang X, Ouyang X, Deng XW (2014) Beyond repression of photomorphogenesis: role switching of COP/DET/FUS in light signaling. Curr Opinion Plant Biol **21:** 96-103

Hughes J (2010) Phytochrome three-dimensional structures and functions. Biochemical Society Transactions **38:** 710

Hunt RC, Simhadri VL, Iandoli M, Sauna ZE, Kimchi-Sarfaty C (2014) Exposing synonymous mutations. Trends Genet **30:** 308-321

Huq E, Al-Sady B, Hudson M, Kim C, Apel K, Quail PH (2004) Phytochrome-interacting factor 1 is a critical bHLH regulator of chlorophyll biosynthesis. Science **305:** 1937-1941

Ibiza VP, Cañizares J, Nuez F (2010) EcoTILLING in Capsicum species: searching for new virus resistances. BMC genomics **11:** 631

Inoue S-i, Matsushita T, Tomokiyo Y, Matsumoto M, Nakayama KI, Kinoshita T, Shimazaki K-i (2011) Functional analyses of the activation loop of phototropin2 in Arabidopsis. Plant Physiol **156:** 117-128

Jang I-C, Henriques R, Seo HS, Nagatani A, Chua N-H (2010) Arabidopsis PHYTOCHROME INTERACTING FACTOR proteins promote phytochrome B polyubiquitination by COP1 E3 ligase in the nucleus. Plant Cell **22:** 2370-2383

Jang I-C, Yang J-Y, Seo HS, Chua N-H (2005) HFR1 is targeted by COP1 E3 ligase for post-translational proteolysis during phytochrome A signaling. Genes Dev **19:** 593-602

Jarillo JA, Capel J, Tang R-H, Yang H-Q, Alonso JM, Ecker JR, Cashmore AR (2001) An Arabidopsis circadian clock component interacts with both CRY1 and phyB. Nature **410:** 487-490

Jenkins J (1948) The origin of the cultivated tomato. Economic Botany **2:** 379-392

Jones B, Frasse P, Olmos E, Zegzouti H, Li ZG, Latché A, Pech JC, Bouzayen M (2002) Down-regulation of DR12, an auxin-response-factor homolog, in the tomato results in a pleiotropic phenotype including dark green and blotchy ripening fruit. Plant J **32:** 603-613

Jones MO, Piron-Prunier F, Marcel F, Piednoir-Barbeau E, Alsadon AA, Wahb-Allah MA, Al-Doss AA, Bowler C, Bramley PM, Fraser PD (2012) Characterisation of alleles of tomato light signalling genes generated by TILLING. Phytochemistry **79:** 78-86

JOSSE E, Foreman J, Halliday KJ (2008) Paths through the phytochrome network. Plant Cell Environ **31:** 667-678

Kadaru SB, Yadav AS, Fjellstrom RG, Oard JH (2006) Alternative Ecotilling protocol for rapid, cost-effective single-nucleotide polymorphism discovery and genotyping in rice (*Oryza sativa* L.). Plant Mol Biol Reporter **24:** 3-22

Karlova R, Rosin FM, Busscher-Lange J, Parapunova V, Do PT, Fernie AR, Fraser PD, Baxter C, Angenent GC, de Maagd RA (2011) Transcriptome and metabolite profiling show that APETALA2a is a major regulator of tomato fruit ripening. Plant Cell **23:** 923-941

Karlova R, van Haarst JC, Maliepaard C, van de Geest H, Bovy AG, Lammers M, Angenent GC, de Maagd RA (2013) Identification of microRNA targets in tomato fruit development using high-throughput sequencing and degradome analysis. J Exp Bot **64:** 1863-1878

Kato S, Ding J, Pisck E, Jhala US, Du K (2008) COP1 functions as a FoxO1 ubiquitin E3 ligase to regulate FoxO1-mediated gene expression. J Biol Chem **283:** 35464-35473

Kevany BM, Taylor MG, Klee HJ (2008) Fruit-specific suppression of the ethylene receptor LeETR4 results in early-ripening tomato fruit. Plant Biotech J **6:** 295-300

Khudairi A, Arboleda O (1971) Phytochrome-mediated carotenoid biosynthesis and its influence by plant hormones. Physiol Plantarum **24:** 18-22

Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T (2009) The SWISS-MODEL Repository and associated resources. Nucleic Acid Res**37:** D387-D392

Kilambi HV, Kumar R, Sharma R, Sreelakshmi Y (2013) Chromoplast-specific carotenoid-associated protein appears to be important for enhanced accumulation of carotenoids in hp1 tomato fruits. Plant Physiol **161:** 2085-2101

Kim BC, Tennessen DJ, Last RL (1998) UV-B-induced photomorphogenesis in *Arabidopsis thaliana*. Plant J **15:** 667-674

Kim W-Y, Fujiwara S, Suh S-S, Kim J, Kim Y, Han L, David K, Putterill J, Nam HG, Somers DE (2007) ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. Nature **449:** 356-360

Kimchi-Sarfaty C, Oh JM, Kim I-W, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM (2007) A" silent" polymorphism in the MDR1 gene changes substrate specificity. Science **315:** 525-528

Kircher S, Gil P, Kozma-Bognár L, Fejes E, Speth V, Husselstein-Muller T, Bauer D, Ádám É, Schäfer E, Nagy F (2002) Nucleocytoplasmic partitioning of the plant photoreceptors phytochrome A, B, C, D, and E is regulated differentially by light and exhibits a diurnal rhythm. Plant Cell **14:** 1541-1555

Kiyosue T, Wada M (2000) LKP1 (LOV kelch protein 1): a factor involved in the regulation of flowering time in Arabidopsis. Plant J **23:** 807-815

Klee HJ, Giovannoni JJ (2011) Genetics and control of tomato fruit ripening and quality attributes. Annu. Rev Genet **45:** 41-59

Klose C, Viczián A, Kircher S, Schäfer E, Nagy F (2014) Molecular mechanisms for mediating light-dependent nucleo/cytoplasmic partitioning of phytochrome photoreceptors. New Phytologist

Knoll JE, Ramos ML, Zeng Y, Holbrook CC, Chow M, Chen S, Maleki S, Bhattacharya A, Ozias-Akins P (2011) TILLING for allergen reduction and improvement of quality traits in peanut (*Arachis hypogaea* L.). BMC Plant Biol **11:** 81

Kong S-G, Suetsugu N, Kikuchi S, Nakai M, Nagatani A, Wada M (2013) Both phototropin 1 and 2 localize on the chloroplast outer membrane with distinct localization activity. Plant Cell Physiol **54:** 80-92

Koornneef M, Bosma T, Hanhart C, Van der Veen J, Zeevaart J (1990) The isolation and characterization of gibberellin-deficient mutants in tomato. Theort Appl Gen **80:** 852-857

Korbie DJ, Mattick JS (2008) Touchdown PCR for increased specificity and sensitivity in PCR amplification. Nat Protocol **3:** 1452-1456

Kornienko A, Butorina A (2013) Induced mutagenesis in Sugar Beet (Beta vulgaris L.): Obtained results and prospects for use in development of TILLING project. Biol Bulletin Rev **3:** 152-160

Kraft E, Bostick M, Jacobsen SE, Callis J (2008) ORTH/VIM proteins that regulate DNA methylation are functional ubiquitin E3 ligases. Plant J **56:** 704-715

Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protocol **4:** 1073-1081

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:** 680-685

Lagarias J, Mercurio F (1985) Structure function studies on phytochrome. Identification of light-induced conformational changes in 124-kDa Avena phytochrome in vitro. J Biol Chem **260:** 2415-2423

Lau OS, Deng XW (2012) The photomorphogenic repressors COP1 and DET1: 20 years later. Trends Plant Sci**17:** 584-593

Laubinger S, Fittinghoff K, Hoecker U (2004) The SPA quartet: a family of WD-repeat proteins with a central role in suppression of photomorphogenesis in Arabidopsis. Plant Cell **16:** 2293-2306

Laubinger S, Hoecker U (2003) The SPA1-like proteins SPA3 and SPA4 repress photomorphogenesis in the light. Plant J **35:** 373-385

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

Lee JM, Joung JG, McQuinn R, Chung MY, Fei Z, Tieman D, Klee H, Giovannoni J (2012) Combined transcriptome, genetic diversity and metabolite profiling in tomato fruit reveals that the ethylene response factor SlERF6 plays an important role in ripening and carotenoid accumulation. Plant J **70:** 191-204

Leivar P, Monte E, Al-Sady B, Carle C, Storer A, Alonso JM, Ecker JR, Quail PH (2008) The Arabidopsis phytochrome-interacting factor PIF7, together with PIF3 and PIF4, regulates responses to prolonged red light by modulating phyB levels. Plant Cell **20:** 337-352

Li D-Q, Ohshiro K, Reddy SDN, Pakala SB, Lee M-H, Zhang Y, Rayala SK, Kumar R (2009) E3 ubiquitin ligase COP1 regulates the stability and functions of MTA1. Proctl Nat Acad Sci **106:** 17493-17498

Li QH, Yang HQ (2007) Cryptochrome Signaling in Plants†. Photochem Photobiol **83:** 94-101

Li Y-Y, Mao K, Zhao C, Zhao X-Y, Zhang H-L, Shu H-R, Hao Y-J (2012) MdCOP1 ubiquitin E3 ligases interact with MdMYB1 to regulate light-induced anthocyanin biosynthesis and red fruit coloration in apple. Plant Physiol **160:** 1011-1022

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25:** 1451-1452

Lieberman M, Segev O, Gilboa N, Lalazar A, Levin I (2004) The tomato homolog of the gene encoding UV-damaged DNA binding protein 1 (DDB1) underlined as the gene that causes the high pigment-1 mutant phenotype. Theort Appl Gen **108:** 1574-1581

Lin C, Shalitin D (2003) Cryptochrome structure and signal transduction. Annu. Rev Plant Biol **54:** 469-496

Lin C, Yang H, Guo H, Mockler T, Chen J, Cashmore AR (1998) Enhancement of blue-light sensitivity of Arabidopsis seedlings by a blue light receptor cryptochrome 2. Proctl Nat Acad Sci **95:** 2686-2690

Lin Z, Hong Y, Yin M, Li C, Zhang K, Grierson D (2008) A tomato HD-Zip homeobox protein, LeHB-1, plays an important role in floral organogenesis and ripening. Plant J **55:** 301-310

Liscum E, Hodgson DW, Campbell TJ (2003) Blue light signaling through the cryptochromes and phototropins. So that's what the blues is all about. Plant Physiol **133:** 1429-1436

Liu L, Jia C, Zhang M, Chen D, Chen S, Guo R, Guo D, Wang Q (2014) Ectopic expression of a BZR1-1D transcription factor in brassinosteroid signalling enhances carotenoid accumulation and fruit quality attributes in tomato. Plant Biotech J **12:** 105-115

Liu L, Wei J, Zhang M, Zhang L, Li C, Wang Q (2012) Ethylene independent induction of lycopene biosynthesis in tomato fruits by jasmonates. J Exp Bot **63:** 5751-5761

Liu Y, Koornneef M, Soppe WJ (2007) The absence of histone H2B monoubiquitination in the Arabidopsis hub1 (rdo4) mutant reveals a role for chromatin remodeling in seed dormancy. Plant Cell **19:** 433-444

Liu Y, Roof S, Ye Z, Barry C, van Tuinen A, Vrebalov J, Bowler C, Giovannoni J (2004) Manipulation of light signal transduction as a means of modifying fruit nutritional quality in tomato. Proctl Nat Acad Sci U S A **101:** 9897-9902

Lobell DB, Burke MB, Tebaldi C, Mastrandrea MD, Falcon WP, Naylor RL (2008) Prioritizing climate change adaptation needs for food security in 2030. Science **319:** 607-610

Lorrain S, Allen T, Duek PD, Whitelam GC, Fankhauser C (2008) Phytochrome-mediated inhibition of shade avoidance involves degradation of growth-promoting bHLH transcription factors. Plant J **53:** 312-323

Lu S, Li L (2008) Carotenoid metabolism: biosynthesis, regulation, and beyond. Journal of Integrative Plant Biology **50:** 778-785

Lucas M, Prat S (2014) PIFs get BRright: PHYTOCHROME INTERACTING FACTORs as integrators of light and hormonal signals. New Phytologist **202:** 1126-1141

Mancinelli AL (1994) The physiology of phytochrome action. *In* Photomorphogenesis in plants. Springer, pp 211-269

Mardis E (1999) Capillary electrophoresis platforms for DNA sequence analysis. J Biomol Tech JBT **10:** 137

Martel C, Vrebalov J, Tafelmeyer P, Giovannoni JJ (2011) The tomato MADS-box transcription factor RIPENING INHIBITOR interacts with promoters involved in numerous ripening processes in a COLORLESS NONRIPENING-dependent manner. Plant Physiol **157:** 1568-1579

Martínez-García JF, Huq E, Quail PH (2000) Direct targeting of light signals to a promoter element-bound transcription factor. Science **288:** 859-863

Maxted N, Kell S (2009) Establishment of a global network for the in situ conservation of crop wild relatives: status and needs. FAO commission on genetic resources for food and agriculture, Rome, Italy**:** 266

Mazzella MA, Casal JJ, Muschietti JP, Fox AR (2014) Hormonal networks involved in apical hook development in darkness and their response to light. Frontiers in plant science **5**

Mazzucotelli E, Belloni S, Marone D, De Leonardis A, Guerra D, Di Fonzo N, Cattivelli L, Mastrangelo A (2006) The E3 ubiquitin ligase gene family in plants: regulation by degradation. Curr Genomics **7:** 509

McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeted screening for induced mutations. Nature biotechnology **18:** 455-457

McClean PE, Hanson MR (1986) Mitochondrial DNA sequence divergence among Lycopersicon and related Solanum species. Genetics **112:** 649-667

Mejlhede N, Kyjovska Z, Backes G, Burhenne K, Rasmussen SK, Jahoor A (2006) EcoTILLING for the identification of allelic variation in the powdery mildew resistance genes mlo and Mla of barley. Plant Breeding **125:** 461-467

Menda N, Semel Y, Peled D, Eshed Y, Zamir D (2004) In silico screening of a saturated mutation library of tomato. Plant J **38:** 861-872

Miller J, Tanksley S (1990) RFLP analysis of phylogenetic relationships and genetic variation in the genus Lycopersicon. Theort Appl Gen **80:** 437-448

Minoia S, Petrozza A, D'Onofrio O, Piron F, Mosca G, Sozio G, Cellini F, Bendahmane A, Carriero F (2010) A new mutant genetic resource for tomato crop improvement by TILLING technology. BMC Res Notes **3:** 1-69

Moens CB, Donn TM, Wolf-Saxon ER, Ma TP (2008) Reverse genetics in zebrafish by TILLING. Briefings in functional Genom Prot **7:** 454-459

Möglich A, Yang X, Ayers RA, Moffat K (2010) Structure and function of plant photoreceptors. Annu. Rev Plant Biol **61:** 21-47

Moon J, Parry G, Estelle M (2004) The ubiquitin-proteasome pathway and plant development. Plant Cell **16:** 3181-3195

Murray M, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acid Res**8:** 4321-4326

Mustilli AC, Fenzi F, Ciliento R, Alfano F, Bowler C (1999) Phenotype of the tomato high pigment-2 mutant is caused by a mutation in the tomato homolog of DEETIOLATED1. Plant Cell **11:** 145-157

Nagatani A (2010) Phytochrome: structural basis for its functions. Curr Opinion Plant Biol **13:** 565-570

Nagy F, Schäfer E (2002) Phytochromes control photomorphogenesis by differentially regulated, interacting signaling pathways in higher plants. Annu. Rev Plant Biol **53:** 329-355

Neff MM, Fankhauser C, Chory J (2000) Light: an indicator of time and place. Genes Dev **14:** 257-271

Negrao S, Almadanim C, Pires I, McNally K, Oliveira M (2011) Use of EcoTILLING to identify natural allelic variants of rice candidate genes involved in salinity tolerance. Plant Genet Res **9:** 300-304

Nevo E, Chen G (2010) Drought and salt tolerances in wild relatives for wheat and barley improvement. Plant Cell Environ **33:** 670-685

Ng PC, Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. Nucleic Acid Res**31:** 3812-3814

Ni M, Tepperman JM, Quail PH (1998) PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop-helix protein. Cell **95:** 657-667

Nieto C, Piron F, Dalmais M, Marco CF, Moriones E, Gómez-Guillamón ML, Truniger V, Gómez P, Garcia-Mas J, Aranda MA (2007) EcoTILLING for the identification of allelic variants of melon eIF4E, a factor that controls virus susceptibility. BMC Plant Biol **7:** 34

Ninu L, Ahmad M, Miarelli C, Cashmore AR, Giuliano G (1999) Cryptochrome 1 controls tomato development in response to blue light. Plant J **18:** 551-556

Nozue K, Covington MF, Duek PD, Lorrain S, Fankhauser C, Harmer SL, Maloof JN (2007) Rhythmic growth explained by coincidence between internal and external cues. Nature **448:** 358-361

Oh E, Kim J, Park E, Kim J-I, Kang C, Choi G (2004) PIL5, a phytochrome-interacting basic helix-loop-helix protein, is a key negative regulator of seed germination in *Arabidopsis thaliana*. Plant Cell **16:** 3045-3058

Okabe Y, Ariizumi T, Ezura H (2013) Updating the Micro-Tom TILLING platform. Breed Sci **63:** 42-48

Omoni AO, Aluko RE (2005) The anti-carcinogenic and anti-atherogenic effects of lycopene: a review. Trends Food Sci Tech **16:** 344-350

Osorio S, Alba R, Damasceno CM, Lopez-Casado G, Lohse M, Zanor MI, Tohge T, Usadel B, Rose JK, Fei Z (2011) Systems biology of tomato fruit development: combined transcript, protein, and metabolite analysis of tomato transcription factor (nor, rin) and ethylene receptor (Nr) mutants reveals novel regulatory interactions. Plant Physiol **157:** 405-425

Osterlund MT, Ang L-H, Deng XW (1999) The role of COP1 in repression of< i> Arabidopsis</i> photomorphogenic development. Trends Cell Biol **9:** 113-118

Osterlund MT, Wei N, Deng XW (2000) The roles of photoreceptor systems and the COP1-targeted destabilization of HY5 in light control of Arabidopsis seedling development. Plant Physiol **124:** 1520-1524

Paiva SA, Russell RM (1999) β-Carotene and other carotenoids as antioxidants. Journal of the American College Nutrition **18:** 426-433

Park C-M, Bhoo S-H, Song P-S (2000) Inter-domain crosstalk in the phytochrome molecules. *In* Seminars in Cell Dev Biol, Vol 11. Elsevier, pp 449-456

Park E, Kim J, Lee Y, Shin J, Oh E, Chung W-I, Liu JR, Choi G (2004) Degradation of phytochrome interacting factor 3 in phytochrome-mediated light signaling. Plant Cell Physiol **45:** 968-975

Parks BM, Quail PH (1991) Phytochrome-deficient hy1 and hy2 long hypocotyl mutants of Arabidopsis are defective in phytochrome chromophore biosynthesis. Plant Cell **3:** 1177-1186

Parks BM, Quail PH (1993) hy8, a new class of Arabidopsis long hypocotyl mutants deficient in functional phytochrome A. Plant Cell **5:** 39-48

Parry MA, Madgwick PJ, Bayon C, Tearall K, Hernandez-Lopez A, Baudo M, Rakszegi M, Hamada W, Al-Yassin A, Ouabbou H (2009) Mutation discovery for crop improvement. J Exp Bot**60:** 2817-2825

Peralta I, Spooner D, Knapp S (2008) The taxonomy of tomatoes: a revision of wild tomatoes (Solanum section *Lycopersicon*) and their outgroup relatives in sections Juglandifolium and *Lycopersicoides*. Syst Bot Monogr **84:** 1-186

Peralta IE, Knapp S, Spooner DM (2006) Nomenclature for wild and cultivated tomatoes. TGC Report **56:** 6-12

Peralta IE, Spooner DM (2001) Granule-bound starch synthase (GBSSI) gene phylogeny of wild tomatoes (Solanum L. section *Lycopersicon* [Mill.] Wettst. subsection *Lycopersicon*). American J Bot **88:** 1888-1902

Peralta IE, Spooner DM, Knapp S (2008) Taxonomy of wild tomatoes and their relatives (Solanum sect. *Lycopersicoides*, sect. Juglandifolia, sect. *Lycopersicon*; Solanaceae).

Perrotta G, Ninu L, Flamma F, Weller JL, Kendrick RE, Nebuloso E, Giuliano G (2000) Tomato contains homologues of Arabidopsis cryptochromes 1 and 2. Plant Mol Biol **42:** 765-773

Perry JA, Wang TL, Welham TJ, Gardner S, Pike JM, Yoshida S, Parniske M (2003) A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume Lotus japonicus. Plant Physiol **131:** 866-871

Pickart CM (2001) Mechanisms underlying ubiquitination. Annu. Rev Biochem **70:** 503-533

Pillen K, Steinrücken G, Wricke G, Herrmann R, Jung C (1992) A linkage map of sugar beet (Beta vulgaris L.). Theort Appl Gen **84:** 129-135

Plotkin JB, Kudla G (2010) Synonymous but not the same: the causes and consequences of codon bias. Nat Rev Genetics **12:** 32-42

Pratt L, CORDONNIER-PRATT MM, Kelmenson P, Lazarova G, Kubota T, Alba R (1997) The phytochrome gene family in tomato (Solarium lycopersicum L.). Plant Cell Environ **20:** 672-677

Prischmann D, Dashiell K, Schneider D, Eubanks M (2009) Evaluating Tripsacum-introgressed maize germplasm after infestation with western corn rootworms (Coleoptera: Chrysomelidae). J Appl Entomol **133:** 10-20

Qi L, Heredia JE, Altarejos JY, Screaton R, Goebel N, Niessen S, MacLeod IX, Liew CW, Kulkarni RN, Bain J (2006) TRB3 links the E3 ubiquitin ligase COP1 to lipid metabolism. Science **312:** 1763-1766

Qin G, Wang Y, Cao B, Wang W, Tian S (2012) Unraveling the regulatory network of the MADS box transcription factor RIN in fruit ripening. Plant J **70:** 243-255

Quail PH (2002) Photosensory perception and signalling in plant cells: new paradigms? Curr Opinion Cell Biol **14:** 180-188

Ranjan A, Dickopf S, Ullrich KK, Rensing SA, Hoecker U (2014) Functional analysis of COP1 and SPA orthologs from Physcomitrella and rice during photomorphogenesis of transgenic Arabidopsis reveals distinct evolutionary conservation. BMC Plant Biol **14:** 178

Reed JW, Nagpal P, Poole DS, Furuya M, Chory J (1993) Mutations in the gene for the red/far-red light receptor phytochrome B alter cell elongation and physiological responses throughout Arabidopsis development. Plant Cell **5:** 147-157

Rick C (1979) Biosystematic studies in *Lycopersicon* and closely related species of Solanum. Hawkes, J, G, Lester, R, N,, Skelding, A, D ed (s). The biology and taxonomy of the Solanaceae. London, Academic Press for the Linnean Society**:** 667-678

Rick C, Chetelat R (1995) Utilization of related wild species for tomato improvement. *In* I International Symposium on Solanacea for Fresh Market 412, pp 21-38

Rick C, Holle M (1990) Andean Lycopersicon esculentum var. cerasiforme: genetic variation and its evolutionary significance. Economic Botany **44:** 69-78

Rick CM, Uhlig JW, Jones AD (1994) High alpha-tomatine content in ripe fruit of Andean Lycopersicon esculentum var. cerasiforme: developmental and genetic aspects. Proctl Nat Acad Sci **91:** 12877-12881

Rigola D, van Oeveren J, Janssen A, Bonné A, Schneiders H, van der Poel HJ, van Orsouw NJ, Hogers RC, de Both MT, van Eijk MJ (2009) High-throughput detection of induced mutations and natural variation using KeyPoint™ technology. PLoS One **4:** e4761

Rizzini L, Favory J-J, Cloix C, Faggionato D, O'Hara A, Kaiserli E, Baumeister R, Schäfer E, Nagy F, Jenkins GI (2011) Perception of UV-B by the Arabidopsis UVR8 protein. Science **332:** 103-106

Robert VJ, West MA, Inai S, Caines A, Arntzen L, Smith JK, Clair DAS (2001) Marker-assisted introgression of blackmold resistance QTL alleles from wild Lycopersicon cheesmanii to cultivated tomato (*L. esculentum*) and evaluation of QTL phenotypic effects. Mol Breed **8:** 217-233

Rockwell NC, Lagarias JC (2010) A brief history of phytochromes. ChemPhysChem **11:** 1172-1180

Rosati C, Aquilani R, Dharmapuri S, Pallara P, Marusic C, Tavazza R, Bouvier F, Camara B, Giuliano G (2000) Metabolic engineering of beta-carotene and lycopene content in tomato fruit. Plant J **24:** 413-420

Rosebrock TR, Zeng L, Brady JJ, Abramovitch RB, Xiao F, Martin GB (2007) A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. Nature **448:** 370-374

Sagar M, Chervin C, Mila I, Hao Y, Roustan JP, Benichou M, Gibon Y, Biais B, Maury P, Latché A (2013) SlARF4, an auxin response factor involved in the control of sugar metabolism during tomato fruit development. Plant Physiol **161:** 1362-1374

Saijo Y, Sullivan JA, Wang H, Yang J, Shen Y, Rubio V, Ma L, Hoecker U, Deng XW (2003) The COP1–SPA1 interaction defines a critical step in phytochrome A-mediated regulation of HY5 activity. Genes Dev **17:** 2642-2647

Saito T, Ariizumi T, Okabe Y, Asamizu E, Hiwasa-Tanase K, Fukuda N, Mizoguchi T, Yamazaki Y, Aoki K, Ezura H (2011) TOMATOMA: a novel tomato mutant database distributing Micro-Tom mutant collections. Plant Cell Physiol **52:** 283-296

Sakai T, Kagawa T, Kasahara M, Swartz TE, Christie JM, Briggs WR, Wada M, Okada K (2001) Arabidopsis nph1 and npl1: blue light receptors that mediate both phototropism and chloroplast relocation. Proctl Nat Acad Sci **98:** 6969-6974

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning, Vol 2. Cold spring harbor laboratory press New York

Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual (3-volume set), Vol 999. Cold spring harbor laboratory press Cold Spring Harbor, New York:

Santner A, Estelle M (2010) The ubiquitin-proteasome system regulates plant hormone signaling. Plant J **61:** 1029-1040

Sarikamis G, Marquez J, Maccormack R, Bennett RN, Roberts J, Mithen R (2006) High glucosinolate broccoli: a delivery system for sulforaphane. Mol Breed **18:** 219-228

Schäfer E, Bowler C (2002) Phytochrome-mediated photoperception and signal transduction in higher plants. EMBO reports **3:** 1042-1048

Schultz TF, Kiyosue T, Yanovsky M, Wada M, Kay SA (2001) A role for LKP2 in the circadian clock of Arabidopsis. Plant Cell **13:** 2659-2670

Schwechheimer C, Serino G, Callis J, Crosby WL, Lyapina S, Deshaies RJ, Gray WM, Estelle M, Deng X-W (2001) Interactions of the COP9 signalosome with the E3 ubiquitin ligase SCFTIR1 in mediating auxin response. Science **292:** 1379-1382

Schwechheimer C, Serino G, Deng X-W (2002) Multiple ubiquitin ligase–mediated processes require COP9 signalosome and AXR1 function. Plant Cell **14:** 2553-2563

Seo HS, Yang J-Y, Ishikawa M, Bolle C, Ballesteros ML, Chua N-H (2003) LAF1 ubiquitination by COP1 controls photomorphogenesis and is stimulated by SPA1. Nature **423:** 995-999

Serino G, Deng X-W (2003) The COP9 signalosome: regulating plant development through the control of proteolysis. Annu. Rev Plant Biol **54:** 165-182

Seymour GB, Granell A (2014) Fruit development and ripening. J Exp Bot **65:** 4489-4490

Shalata A, Tal M (1998) The effect of salt stress on lipid peroxidation and antioxidants in the leaf of the cultivated tomato and its wild salt-tolerant relative Lycopersicon pennellii. Physiologia Plantarum **104:** 169-174

Shalitin D, Yang H, Mockler TC, Maymon M, Guo H, Whitelam GC, Lin C (2002) Regulation of Arabidopsis cryptochrome 2 by blue-light-dependent phosphorylation. Nature **417:** 763-767

Sharma S, Kharshiing E, Srinivas A, Zikihara K, Tokutomi S, Nagatani A, Fukayama H, Bodanapu R, Behera RK, Sreelakshmi Y (2014) A Dominant Mutation in the Light-Oxygen and Voltage2 Domain Vicinity Impairs Phototropin1 Signaling in Tomato. Plant Physiol **164:** 2030-2044

Sharrock RA, Quail PH (1989) Novel phytochrome sequences in *Arabidopsis thaliana*: structure, evolution, and differential expression of a plant regulatory photoreceptor family. Genes Dev **3:** 1745-1757

Shen Y, Khanna R, Carle CM, Quail PH (2007) Phytochrome induces rapid PIF5 phosphorylation and degradation in response to red-light activation. Plant Physiol **145:** 1043-1051

Sherinmol Thomas (2012) Isolation of novel alleles for Phytochrome genes of tomato. Thesis submitted, University of Hyderabad

Shikazono N, Suzuki C, Kitamura S, Watanabe H, Tano S, Tanaka A (2005) Analysis of mutations induced by carbon ions in *Arabidopsis thaliana*. J Exp Bot **56:** 587-596

Shima Y, Fujisawa M, Kitagawa M, Nakano T, Kimbara J, Nakamura N, Shiina T, Sugiyama J, Nakamura T, Kasumi T (2014) Tomato FRUITFULL homologs regulate fruit ripening via ethylene biosynthesis. Biosci Biotech Biochem 1-7

Shinkle JR, Atkins AK, Humphrey EE, Rodgers CW, Wheeler SL, Barnes PW (2004) Growth and morphological responses to different UV wavebands in cucumber (Cucumis sativum) and other dicotyledonous seedlings. Physiologia Plantarum **120:** 240-248

Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC (2012) SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acid Res**40:** W452-W457

Singh VK, Kumar A (2001) PCR primer design. Mol Biol **2:** 27-32

Slade AJ, Knauf VC (2005) TILLING moves beyond functional genomics into crop improvement. Transgenic Res **14:** 109-115

Smalle J, Vierstra RD (2004) The ubiquitin 26S proteasome proteolytic pathway. Annu. Rev. Plant Biol. **55:** 555-590

Smith H (2000) Phytochromes and light signal perception by plants—an emerging synthesis. Nature **407:** 585-591

Smith H, Whitelam G (1997) The shade avoidance syndrome: multiple responses mediated by multiple phytochromes. Plant Cell Environ **20:** 840-844

Somers DE, Fujiwara S (2009) Thinking outside the F-box: novel ligands for novel receptors. Trends Plant Sci **14:** 206-213

Somers DE, Schultz TF, Milnamow M, Kay SA (2000) < i> ZEITLUPE</i> Encodes a Novel Clock-Associated PAS Protein from< i> Arabidopsis</i>. Cell **101:** 319-329

Sreelakshmi Y, Gupta S, Bodanapu R, Chauhan VS, Hanjabam M, Thomas S, Mohan V, Sharma S, Srinivasan R, Sharma R (2010) NEATTILL: A simplified procedure for nucleic acid extraction from arrayed tissue for TILLING and other high-throughput reverse genetic applications. Plant methods **6:** 3

Sridhar VV, Kapoor A, Zhang K, Zhu J, Zhou T, Hasegawa PM, Bressan RA, Zhu J-K (2007) Control of DNA methylation and heterochromatic silencing by histone H2B deubiquitination. Nature **447:** 735-738

Stacey MG, Kopp OR, Kim T-H, von Arnim AG (2000) Modular domain structure of Arabidopsis COP1. Reconstitution of activity by fragment complementation and mutational analysis of a nuclear localization signal in planta. Plant Physiol **124:** 979-990

Suesslin C, Frohnmeyer H (2003) An Arabidopsis mutant defective in UV-B light-mediated responses. Plant J **33:** 591-601

Sullivan J, Rozema J (1999) UV-B effects on terrestrial plant growth and photosynthesis. Stratospheric ozone depletion: the effects of enhanced UV-B radiation on terrestrial ecosystems**:** 39-57

Sullivan JA, Deng XW (2003) From seed to seed: the role of photoreceptors in< i> Arabidopsis</i> development. Dev Biol **260:** 289-297

Sullivan JA, Gray JC (2000) The pea light-independent photomorphogenesis1 mutant results from partial duplication of COP1 generating an internal promoter and producing two distinct transcripts. Plant Cell **12:** 1927-1937

Sun L, Sun Y, Zhang M, Wang L, Ren J, Cui M, Wang Y, Ji K, Li P, Li Q (2012) Suppression of 9-cis-epoxycarotenoid dioxygenase, which encodes a key enzyme in abscisic acid biosynthesis, alters fruit texture in transgenic tomato. Plant Physiol **158:** 283-298

Suzuki G, Yanagawa Y, Kwok SF, Matsui M, Deng X-W (2002) Arabidopsis COP10 is a ubiquitin-conjugating enzyme variant that acts together with COP1 and the COP9 signalosome in repressing photomorphogenesis. Genes Dev **16:** 554-559

Suzuki T, Eiguchi M, Kumamaru T, Satoh H, Matsusaka H, Moriguchi K, Nagato Y, Kurata N (2008) MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. Mol Gen Genome **279:** 213-223

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123:** 585-595

Talamè V, Bovina R, Sanguineti MC, Tuberosa R, Lundqvist U, Salvi S (2008) TILLMore, a resource for the discovery of chemically induced mutants in barley. Plant Biotech J **6:** 477-485

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular biology and evolution **24:** 1596-1599

Tanaka N, Itoh H, Sentoku N, Kojima M, Sakakibara H, Izawa T, Itoh J-I, Nagato Y (2011) The COP1 ortholog PPS regulates the juvenile–adult and vegetative–reproductive phase changes in rice. Plant Cell **23:** 2143-2154

Tanksley SD (2004) The genetic, developmental, and molecular bases of fruit size and shape variation in tomato. Plant Cell **16:** S181-S189

Taylor NE, Greene EA (2003) PARSESNP: a tool for the analysis of nucleotide polymorphisms. Nucleic Acid Res**31:** 3808-3811

Tepperman JM, Zhu T, Chang H-S, Wang X, Quail PH (2001) Multiple transcription-factor genes are early targets of phytochrome A signaling. Proctl Nat Acad Sci **98:** 9437-9442

Thomas RL, Jen JJ (1975) Phytochrome-mediated carotenoids biosynthesis in ripening tomatoes. Plant Physiol **56:** 452-453

Till BJ, Cooper J, Tai TH, Colowit P, Greene EA, Henikoff S, Comai L (2007) Discovery of chemically induced mutations in rice by TILLING. BMC Plant Biol **7:** 19

Till BJ, Jankowicz-Cieslak J, Sági L, Huynh OA, Utsushi H, Swennen R, Terauchi R, Mba C (2010) Discovery of nucleotide polymorphisms in the Musa gene pool by Ecotilling. Theort Appl Gen **121:** 1381-1389

Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson JE, Burtner C, Odden AR, Young K, Taylor NE (2003) Large-scale discovery of induced point mutations with high-throughput TILLING. Genome Res **13:** 524-530

Till BJ, Reynolds SH, Weil C, Springer N, Burtner C, Young K, Bowers E, Codomo CA, Enns LC, Odden AR (2004) Discovery of induced point mutations in maize genes by TILLING. BMC Plant Biol **4:** 12

Till BJ, Zerr T, Comai L, Henikoff S (2006) A protocol for TILLING and Ecotilling in plants and animals. Nat Protocol **1:** 2465-2477

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

Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proctl Nat Acad Sci **76:** 4350-4354

Trainotti L, Tadiello A, Casadoro G (2007) The involvement of auxin in the ripening of climacteric fruits comes of age: the hormone plays a role of its own and has an intense interplay with ethylene in ripening peaches. J Exp Bot **58:** 3299-3308

Triques K, Sturbois B, Gallais S, Dalmais M, Chauvin S, Clepet C, Aubourg S, Rameau C, Caboche M, Bendahmane A (2007) Characterization of *Arabidopsis thaliana* mismatch specific endonucleases: application to mutation discovery by TILLING in pea. Plant J **51:** 1116-1125

Tsai H, Missirian V, Ngo KJ, Tran RK, Chan SR, Sundaresan V, Comai L (2013) Production of a high-efficiency TILLING population through polyploidization. Plant Physiol **161:** 1604-1614

Ulijasz AT, Cornilescu G, Cornilescu CC, Zhang J, Rivera M, Markley JL, Vierstra RD (2010) Structural basis for the photoconversion of a phytochrome to the activated Pfr form. Nature **463:** 250-254

van Harten AM (1998) Mutation breeding: theory and practical applications. Cambridge University Press

Verwoerd TC, Dekker B, Hoekema A (1989) A small-scale procedure for the rapid isolation of plant RNAs. Nucleic Acid Res**17:** 2362

Vierstra RD (2003) The ubiquitin/26S proteasome pathway, the complex last chapter in the life of many plant proteins. Trends Plant Sci**8:** 135-142

Vierstra RD (2009) The ubiquitin–26S proteasome system at the nexus of plant biology. Nat Rev Mol Cell Biol **10:** 385-397

Vousden KH, Prives C (2009) Blinded by the light: the growing complexity of p53. Cell **137:** 413-431

Wang H, Deng XW (2003) Dissecting the phytochrome A-dependent signaling network in higher plants. Trends Plant Sci **8:** 172-178

Wang H, Ma L-G, Li J-M, Zhao H-Y, Deng XW (2001) Direct interaction of Arabidopsis cryptochromes with COP1 in light control development. Science **294:** 154-158

Wang J, Gu H, Yu H, Zhao Z, Sheng X, Zhang X (2012) Genotypic variation of glucosinolates in broccoli (< i> Brassica oleracea</i> var.< i> italica</i>) florets from China. Food chemistry **133:** 735-741

Wang S, Liu J, Feng Y, Niu X, Giovannoni J, Liu Y (2008) Altered plastid levels and potential for improved fruit nutrient content by downregulation of the tomato DDB1-interacting protein CUL4. Plant J **55:** 89-103

Wei N, Deng X-W (1996) The role of the COP/DET/FUS genes in light control of Arabidopsis seedling development. Plant Physiol **112:** 871

Wei N, Deng XW (1998) Characterization and purification of the mammalian COP9 complex, a conserved nuclear regulator initially identified as a repressor of photomorphogenesis in higher plants. Photochem Photobiol **68:** 237-241

Wei N, Deng XW (2003) The COP9 signalosome. Annu. Rev Cell Dev Biol **19:** 261-286

Weissman AM (2001) Themes and variations on ubiquitylation. Nat Rev Mol Cell Biol **2:** 169-178

Welsch R, Beyer P, Hugueney P, Kleinig H, von Lintig J (2000) Regulation and activation of phytoene synthase, a key enzyme in carotenoid biosynthesis, during photomorphogenesis. Planta **211:** 846-854

Whitelam GC, Johnson E, Peng J, Carol P, Anderson ML, Cowl JS, Harberd NP (1993) Phytochrome A null mutants of Arabidopsis display a wild-type phenotype in white light. Plant Cell **5:** 757-768

Wienholds E, van Eeden F, Kosters M, Mudde J, Plasterk RH, Cuppen E (2003) Efficient target-selected mutagenesis in zebrafish. Genome Res **13:** 2700-2707

Wilson ZA, Yang C (2004) Plant gametogenesis: conservation and contrasts in development. Reproduction **128:** 483-492

Winfield MO, Wilkinson PA, Allen AM, Barker GL, Coghill JA, Burridge A, Hall A, Brenchley RC, D'Amore R, Hall N (2012) Targeted re-sequencing of the allohexaploid wheat exome. Plant Biotech J **10:** 733-742

Winkler S, Schwabedissen A, Backasch D, Bökel C, Seidel C, Bönisch S, Fürthauer M, Kuhrs A, Cobreros L, Brand M (2005) Target-selected mutant screen by TILLING in Drosophila. Genome Res **15:** 718-723

Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ (2003) High-resolution genotyping by amplicon melting analysis using LCGreen. Clinical chemistry **49:** 853-860

Wu D, Hu Q, Yan Z, Chen W, Yan C, Huang X, Zhang J, Yang P, Deng H, Wang J (2012) Structural basis of ultraviolet-B perception by UVR8. Nature **484:** 214-219

Xin Z, Wang ML, Barkley NA, Burow G, Franks C, Pederson G, Burke J (2008) Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. BMC Plant Biol **8:** 103

Yamaguchi R, Nakamura M, Mochizuki N, Kay SA, Nagatani A (1999) Light-dependent translocation of a phytochrome B-GFP fusion protein to the nucleus in transgenic Arabidopsis. J Cell Biol **145:** 437-445

Yanagawa Y, Sullivan JA, Komatsu S, Gusmaroli G, Suzuki G, Yin J, Ishibashi T, Saijo Y, Rubio V, Kimura S (2004) Arabidopsis COP10 forms a complex with DDB1 and DET1

in vivo and enhances the activity of ubiquitin conjugating enzymes. Genes Dev **18:** 2172-2181

Yang J, Lin R, Sullivan J, Hoecker U, Liu B, Xu L, Deng XW, Wang H (2005) Light regulates COP1-mediated degradation of HFR1, a transcription factor essential for light signaling in Arabidopsis. Plant Cell **17:** 804-821

Yi C, Deng XW (2005) COP1–from plant photomorphogenesis to mammalian tumorigenesis. Trends Cell Biol **15:** 618-625

Yi C, Wang H, Wei N, Deng XW (2002) An initial biochemical and cell biological characterization of the mammalian homologue of a central plant developmental switch, COP1. BMC Cell Biol **3:** 30

Zhou D-X, Kim Y-J, Li Y-F, Carol P, Mache R (1998) COP1b, an isoform of COP1 generated by alternative splicing, has a negative effect on COP1 function in regulating light-dependent seedling development in Arabidopsis. Mol Gen Genet **257:** 387-391

Zhou L, Vandersteen J, Wang L, Fuller T, Taylor M, Palais B, Wittwer C (2004) High-resolution DNA melting curve analysis to establish HLA genotypic identity. Tissue antigens **64:** 156-164

Zhu D, Maier A, Lee J-H, Laubinger S, Saijo Y, Wang H, Qu L-J, Hoecker U, Deng XW (2008) Biochemical characterization of Arabidopsis complexes containing CONSTITUTIVELY PHOTOMORPHOGENIC1 and SUPPRESSOR OF PHYA proteins in light control of plant development. Plant Cell **20:** 2307-2323

Appendix: List of tomato accessions used in the study

S.No	Code No.	Accession	Details	Source
1	TGRC4	LA3538	*hp-1* (*high pihment-1*), *S. lycopersicum*,AC	TGRC
2	TGRC5	LA3247	*u* (*uniform ripening*), *S. lycopersicum* cv.Craigella,AC	TGRC
3	TGRC8	LA2089	*Epi* (*Epinastic*), *S. lycopersicum*, VFN8	TGRC
4	TGRC13	LA3530	*gs* (*green stripe*), *S. lycopersicum*, AC	TGRC
5	TGRC14	LA3534	*gf* (*green flesh*), *S. lycopersicum*, AC	TGRC
6	TGRC17	LA2529	*alc* (*alcobaca*), *S. lycopersicum* cv. alcobaca	TGRC
7	TGRC21	LA3539	*ug* (*uniform grey green*), *S. lycopersicum*, AC	TGRC
8	TGRC22	LA3770	*nor* (*Non ripening*), *S. lycopersicum* ,AC	TGRC
9	TGRC24	LA3255	*S. lycopersicum*	TGRC
10	TGRC26	LA3537	*Nr* (*Never ripe*), *S. lycopersicum*	TGRC
11	TGRC29	LA3532	*r* (*Yellow flesh*), *S. lycopersicum*, AC	TGRC
12	TGRC31	LA2056	*r* (*Yellow flesh*), *S. lycopersicum*	TGRC
13	TGRC32	LA2997	*r* (*Yellow flesh*), *S. lycopersicum*	TGRC
14	TGRC33	2-141	*r* (*Yellow flesh*), *S. lycopersicum*	TGRC
15	TGRC35	LA4025	*Bog* (*Beta old gold*), *S. lycopersicum*	TGRC
16	TGRC36	LA1795	*rin* (*ripening inhibitor*), *S. lycopersicum*	TGRC
17	TGRC39	LA0806	*Bc* (*Beta crimson*), *S. lycopersicum* cv. High Crimson	TGRC
18	TGRC40	LA3754	*rin* (*ripening inhibitor*), *S. lycopersicum*	TGRC
19	TGRC43	LA2374	*B* (*Beta*), *S. lycopersicum* cv. Caro Red	TGRC
20	TGRC47	LA0500	*rin* (*ripening inhibitor*), *S. lycopersicum*	TGRC
21	TGRC55	LA0292	*Bog* (*Beta old gold*), *S. lycopersicum*	TGRC
22	TGRC56	LA0505	*Od*, *S. lycopersicum* var. Cerasiforme	TGRC
23	TGRC57	LA3465	*Lx*, *S. lycopersicum* cv. Laketa	TGRC
24	TGRC58	LA0276	*S. lycopersicum* cv. Walter	TGRC
25	TGRC59	LA3203	*S. lycopersicum* Red Top VF	TGRC
26	TGRC60	LA2818	*S. lycopersicum* cv. Large Plum	TGRC
27	TGRC61	LA2715	*S. lycopersicum* cv. Porphyre	TGRC
28	TGRC62	LA3632	*S. lycopersicum* cv. Start 24	TGRC
29	TGRC65	LA0533	*S. lycopersicum* cv. Condine Red	TGRC
30	TGRC66	LA3129	*S. lycopersicum* cv. Ehovot 13	TGRC
31	TGRC70	LA3630	*S. lycopersicum* cv. Vrbikanske nizke	TGRC
32	TGRC71	LA0503	*S. lycopersicum* cv. Roumanian Sweet	TGRC
33	TGRC72	LA0012	*S. lycopersicum* cv. Pearson	TGRC
34	TGRC73	LA1022	*S. lycopersicum* cv. VFN8	TGRC
35	TGRC75	LA3030	*S. lycopersicum* cv. Gardener	TGRC
36	TGRC76	LA3246	*S. lycopersicum* cv. Vagabond	TGRC
37	TGRC77	LA3317	*S. lycopersicum* cv. Campvell 28, Canada	TGRC
38	TGRC78	LA3233	*S. lycopersicum* cv. Pritchard	TGRC
39	TGRC79	LA0502	*S. lycopersicum* cv. Marglobe	TGRC
40	TGRC82	LA4024	*S. lycopersicum* cv. E-6203	TGRC
41	TGRC83	LA0180	*S. lycopersicum* cv.San Marzano	TGRC
42	TGRC84	LA1089	*S. lycopersicum* cv. John Baer	TGRC
43	TGRC85	LA3121	*S. lycopersicum* cv. Chico Grande	TGRC
44	TGRC87	LA1504	*S. lycopersicum* cv. Marmande	TGRC
45	TGRC89	LA1506	*S. lycopersicum* cv. Stone	TGRC
46	TGRC92	LA0266	*S. lycopersicum* cv. Earli Pak	TGRC
47	TGRC93	LA3231	*S. lycopersicum* cv. Gulf State Market	TGRC
48	TGRC95	LA1091	*S. lycopersicum* cv. Stokedale	TGRC
49	TGRC97	LA3903	*S. lycopersicum* cv. Prima Bel	TGRC
50	TGRC99	LA3229	*S. lycopersicum* cv. Prospero	TGRC
51	TGRC100	LA3905	*S. lycopersicum* cv. Vantage	TGRC
52	TGRC101	LA1021	*S. lycopersicum* cv. Santa Cruz	TGRC

S.No	Code No.	Accession	Details	Source
53	TGRC102	LA3237	*S. lycopersicum* Homestead 24	TGRC
54	TGRC103	LA3343	*S. lycopersicum* Rio Grande	TGRC
55	TGRC104	LA0516	*S. lycopersicum* Ace	TGRC
56	TGRC106	LA3122	*S. lycopersicum* Vendor	TGRC
57	TGRC107	LA3234	*S. lycopersicum* Sioux	TGRC
58	TGRC108	LA3243	*S. lycopersicum* Platense	TGRC
59	TGRC109	LA1090	*S. lycopersicum* Rutgers	TGRC
60	TGRC117	LA2400	*sp u S. esc*. Cv. Castlemart	TGRC
61	TGRC118	LA0274	*d l w S. esc*.	TGRC
62	TGRC119	LA0842	*glu S. esc*.	TGRC
63	TGRC120	LA3554	*yv S.esc.*	TGRC
64	TGRC121	LA0854	*fa L esc.*	TGRC
65	TGRC122	LA3579	*Xa S. esc.*	TGRC
66	TGRC123	LA2065	*spl* W6 *S. esc*.	TGRC
67	TGRC124	LA0215	*at u y S. esc*.	TGRC
68	TGRC125	LA0458	*Cmr Lv* (*S. chilense*)	TGRC
69	TGRC126	LA2133	*S. parviflorum*	TGRC
70	TGRC127	LA1016	*dps S. lyc.*	TGRC
71	TGRC128	LA2999	*gf S. lyc.*	TGRC
72	TGRC129	LA3430	*Xa-3 S. lyc*	TGRC
73	TGRC133	LA3905	*S. lycopersicum* cv. Vantage	TGRC
74	TGRC134	LA0517	*S. lycopersicum* cv. Early Santa Clara	TGRC
75	TGRC135	LA3024	*S. lycopersicum* cv. Fireball	TGRC
76	TGRC136	LA4104	*sucr S. lycopersicum*	TGRC
77	TGRC138	LA3022	*d S. lycopersicum*	TGRC
78	TGRC140	LA2921	*Del S. lycopersicum*	TGRC
79	TGRC141	LA1996	*Aft S. lycopersicum*	TGRC
80	TGRC142	LA1088	*y S.esc*.cv. Ohio Globe A	TGRC
81	TGRC143	LA0744	*S. lycopersicum* cv. VF11	TGRC
82	TGRC145	LA0300	*bk d o p r s g S. lycopersicum*	TGRC
83	TGRC146	LA1500	*lp S. lycopersicum*	TGRC
84	IIVR1	EC520078		IIVR
85	IIVR2	WIR4361		IIVR
86	IIVR3	EC520065		IIVR
87	IIVR4	H-24		IIVR
88	IIVR5	WIR3957		IIVR
89	IIVR6	DVRT-2		IIVR
90	IIVR8	LA3957		IIVR
91	IIVR12	EC565216		IIVR
92	IIVR14	EC520076		IIVR
93	IIVR15	WIR3956		IIVR
94	IIVR17	WIR13717		IIVR
95	IIVR18	WIR3928		IIVR
96	IIVR19	EC8372		IIVR
97	IIVR21	Chiku Grande		IIVR
98	IIVR22	Pusa Rohini		IIVR
99	IIVR23	Type-1		IIVR
100	IIVR24	Pusa Sadabahar		IIVR
101	IIVR25	Shalimar-2		IIVR
102	IIVR26	M.local		IIVR
103	IIVR27	Arka Alok		IIVR
104	IIVR28	TLBR-3		IIVR
105	IIVR29	TLBR-5		IIVR
106	IIVR30	Mutant		IIVR
107	IIVR31	ALT-9797		IIVR
108	IIVR32	PDT-3-1		IIVR
109	IIVR33	TLBR-2		IIVR

S.No	Code No.	Accession	Details	Source
110	IIVR34	H-88-78-5		IIVR
111	IIVR35	A.Arka		IIVR
112	IIVR36	Riogrande		IIVR
113	IIVR37	Castle Rock		IIVR
114	IIVR38	KT-15		IIVR
115	IIVR39	Siberia		IIVR
116	IIVR40	H-88-87		IIVR
117	IIVR42	WIR-3928		IIVR
118	IIVR43	EC520077		IIVR
119	IIVR44	EC521080		IIVR
120	IIVR45	LA3996		IIVR
121	IIVR46	EC50-50		IIVR
122	IIVR47	EC3414425		IIVR
123	IIVR48	EC521078		IIVR
124	IIVR49	LA4003		IIVR
125	IIVR50	LA3956		IIVR
126	IIVR52	DT-10		IIVR
127	IIVR53	LA3995		IIVR
128	IIVR54	EC520079		IIVR
129	IIVR55	LA3980		IIVR
130	IIVR56	EC528365		IIVR
131	IIVR57	F-6102		IIVR
132	IIVR58	F-6030		IIVR
133	IIVR59	LA3928		IIVR
134	IIVR60	Gujrat Tomato		IIVR
135	IIVR61	WIR3969		IIVR
136	IIVR62	CHRT-4		IIVR
137	IIVR64	FLA7421		IIVR
138	IIVR66	H-86		IIVR
139	IIVR68	PDT-3-1		IIVR
140	IIVR69	VLT-34		IIVR
141	IIVR70	Mont favet		IIVR
142	IIVR71	IIVR-2200		IIVR
143	IIVR72	EC528367		IIVR
144	IIVR73	F-5070		IIVR
145	IIVR74	Cherry Red		IIVR
146	IIVR75	LA3967		IIVR
147	IIVR76	LA3971		IIVR
148	IIVR77	EC398405		IIVR
149	IIVR78	LA4040		IIVR
150	IIVR81	LA4024		IIVR
151	IIVR82	WIR5032		IIVR
152	IIVR83	LA3999		IIVR
153	IIVR84	BT-111-3-2-3		IIVR
154	IIVR85	LA4040-2		IIVR
155	IIVR86	LA3934		IIVR
156	IIVR87	Agata-30		IIVR
157	IIVR89	BL-1208		IIVR
158	IIVR90	DVKT-1		IIVR
159	IIVR91	Azad No.1		IIVR
160	IIVR92	H-88-78-2		IIVR
161	IIVR93	Arka Vikas		IIVR
162	IIVR94	P.Rohit		IIVR
163	IIVR95	PBC (Punjab Chuhara)		IIVR
164	IIVR96	BT-12		IIVR
165	IIVR97	Nandi		IIVR

S.No	Code No.	Accession	Details	Source
166	IIVR98	CLN-2998		IIVR
167	IIVR99	TLBR-4		IIVR
168	IIVR100	TLBR-12		IIVR
169	IIVR101	T.Local		IIVR
170	IIVR102	S-2-95-1-3-1		IIVR
171	IIVR103	Arka Saurabh		IIVR
172	IIVR104	M-88-78-3		IIVR
173	IIVR105	*Cerasiformae*		IIVR
174	IIVR106	PKM-1		IIVR
175	IIVR107	P.Pink		IIVR
176	IIVR108	IIHR-2201		IIVR
177	IIVR109	S. local		IIVR
178	IIVR111	P.Gaurav		IIVR
179	IIVR112	Vaibhav		IIVR
180	IIVR113	Sankranti		IIVR
181	IIVR114	Feb.4		IIVR
182	IIVR115	T-HL		IIVR
183	IIVR116	Feb.2		IIVR
184	IIVR117	Superbug		IIVR
185	IIVR118	FLA7171		IIVR
186	IIVR120	Sel-14		IIVR
187	IIVR121	?		IIVR
188	IIVR122	EC007785		IIVR
189	IIVR124	EC273966		IIVR
190	IIVR128	EC193538		IIVR
191	IIVR129	EC494372		IIVR
192	IIVR130	EC251643-3		IIVR
193	IIVR132	EC339058-A		IIVR
194	IIVR134	EC29933		IIVR
195	IIVR136	EC5627		IIVR
196	IIVR137	PUSA-SHEETAL		IIVR
197	IIVR138	PUSA-GAURAV		IIVR
198	IIVR143	EC2997		IIVR
199	IIVR152	EC362948		IIVR
200	IIVR153	EC241446-A		IIVR
201	IIVR156	EC009046		IIVR
202	IIVR158	EC5358139		IIVR
203	IIVR163	EC007345		IIVR
204	IIVR166	EC33878		IIVR
205	IIVR168	EC370867		IIVR
206	IIVR169	EC538153		IIVR
207	IIVR170	PUSA-UPKAR		IIVR
208	IIVR172	EC3176		IIVR
209	IIVR173	EC241446		IIVR
210	IIVR174	EC531800		IIVR
211	IIVR175	EC1914		IIVR
212	IIVR176	EC562073		IIVR
213	IIVR178	EC251068		IIVR
214	IIVR182	EC2791		IIVR
215	IIVR184	EC57442		IIVR
216	IIVR185	EC016786		IIVR
217	IIVR186	EC6053-1		IIVR
218	IIVR188	EC52106-B		IIVR
219	NBPGR2	EC252		NBPGR
220	NBPGR3	EC487		NBPGR

S.No	Code No.	Accession	Details	Source
221	NBPGR4	EC1087		NBPGR
222	NBPGR5	EC1154		NBPGR
223	NBPGR7	EC2673		NBPGR
224	NBPGR8	EC2790		NBPGR
225	NBPGR9	EC2798		NBPGR
226	NBPGR11	EC2977		NBPGR
227	NBPGR12	EC3176		NBPGR
228	NBPGR13	EC3216		NBPGR
229	NBPGR14	EC3668		NBPGR
230	NBPGR16	EC4506		NBPGR
231	NBPGR17	EC5863		NBPGR
232	NBPGR18	EC6192		NBPGR
233	NBPGR19	EC6486		NBPGR
234	NBPGR20	EC6845		NBPGR
235	NBPGR21	EC7912		NBPGR
236	NBPGR23	EC8591		NBPGR
237	NBPGR24	EC8822		NBPGR
238	NBPGR25	EC8936		NBPGR
239	NBPGR26	EC12689		NBPGR
240	NBPGR27	EC13736		NBPGR
241	NBPGR28	EC14073		NBPGR
242	NBPGR29	EC15127		NBPGR
243	NBPGR31	EC16786		NBPGR
244	NBPGR32	EC16790		NBPGR
245	NBPGR34	EC25265		NBPGR
246	NBPGR37	EC27910		NBPGR
247	NBPGR39	EC35244		NBPGR
248	NBPGR46	EC279088		NBPGR
249	NBPGR48	EC373378		NBPGR
250	NBPGR49	EC381554		NBPGR
251	NBPGR50	EC383117		NBPGR
252	NBPGR54	EC398701		NBPGR
253	NBPGR55	EC398707		NBPGR
254	NBPGR56	EC398711		NBPGR
255	NBPGR57	EC398715		NBPGR
256	NBPGR60	EC433607		NBPGR
257	NBPGR62	EC439542		NBPGR
258	NBPGR63	EC443369		NBPGR
259	NBPGR64	EC470413		NBPGR
260	NBPGR67	EC490128		NBPGR
261	NBPGR71	EC490141		NBPGR
262	NBPGR75	EC520046		NBPGR
263	NBPGR76	EC520075		NBPGR
264	NBPGR77	EC521039		NBPGR
265	NBPGR78	EC521043		NBPGR
266	NBPGR79	EC521046		NBPGR
267	NBPGR80	EC521083		NBPGR
268	NBPGR81	EC521086		NBPGR
269	NBPGR84	EC528365		NBPGR
270	NBPGR85	EC528372		NBPGR
271	NBPGR86	EC528373		NBPGR
272	NBPGR87	EC528374		NBPGR
273	NBPGR88	EC528388		NBPGR
274	NBPGR89	EC529083		NBPGR
275	NBPGR91	EC538141		NBPGR
276	NBPGR92	EC538146		NBPGR
277	NBPGR93	EC538148		NBPGR

S.No	Code No.	Accession	Details	Source
278	NBPGR94	EC538149		NBPGR
279	NBPGR96	EC538156		NBPGR
280	NBPGR97	EC546727		NBPGR
281	NBPGR101	EC170047		NBPGR
282	NBPGR104	EC1177297		NBPGR
283	NBPGR109	EC241446		NBPGR
284	NBPGR110	EC241446 A		NBPGR
285	NBPGR115	EC251646		NBPGR
286	NBPGR117	EC251649		NBPGR
287	NBPGR119	EC251581		NBPGR
288	NBPGR128	EC320583		NBPGR
289	NBPGR145	EC339066		NBPGR
290	NBPGR146	EC338717		NBPGR
291	NBPGR153	EC362941		NBPGR
292	NBPGR154	EC362949		NBPGR
293	NBPGR155	EC363863		NBPGR
294	NBPGR156	EC362933		NBPGR
295	NBPGR167	EC369020		NBPGR
296	NBPGR171	EC385654		NBPGR
297	NBPGR172	EC398614		NBPGR
298	NBPGR173	EC398684		NBPGR
299	NBPGR174	EC398685		NBPGR
300	NBPGR176	EC398688		NBPGR
301	NBPGR178	EC398687		NBPGR
302	NBPGR181	EC398695		NBPGR
303	NBPGR182	EC398697		NBPGR
304	NBPGR183	EC398699		NBPGR
305	NBPGR184	EC398704		NBPGR
306	NBPGR185	EC398710		NBPGR
307	NBPGR186	EC398712		NBPGR
308	NBPGR188	EC398714		NBPGR
309	NBPGR189	EC398716		NBPGR
310	NBPGR190	EC398717		NBPGR
311	NBPGR194	EC433607		NBPGR
312	NBPGR196	EC458213		NBPGR
313	NBPGR219	EC520059		NBPGR
314	NBPGR221	EC521048		NBPGR
315	NBPGR223	EC521076		NBPGR
316	NBPGR224	EC521077		NBPGR
317	NBPGR225	EC521078		NBPGR
318	NBPGR226	EC521079		NBPGR
319	NBPGR227	EC521080		NBPGR
320	NBPGR229	EC521082		NBPGR
321	NBPGR231	WIR1378		NBPGR
322	NBPGR232	WIR3768		NBPGR
323	NBPGR241	EC6488		NBPGR
324	NBPGR261	EC155		NBPGR
325	NBPGR262	EC429		NBPGR
326	NBPGR273	EC742		NBPGR
327	NBPGR275	EC1191		NBPGR
328	NBPGR282	EC2347		NBPGR
329	NBPGR285	EC25D		NBPGR
330	NBPGR289	EC2053		NBPGR
331	NBPGR293	EC2765		NBPGR
332	NBPGR295	EC2802		NBPGR
333	NBPGR296	EC2990		NBPGR
334	NBPGR298	EC3176-1		NBPGR

S.No	Code No.	Accession	Details	Source
335	NBPGR311	EC5888		NBPGR
336	NBPGR312	EC6053-1		NBPGR
337	NBPGR314	EC8630		NBPGR
338	NBPGR315	EC6875		NBPGR
339	NBPGR319	EC7317		NBPGR
340	NBPGR320	EC7345		NBPGR
341	NBPGR321	EC7785		NBPGR
342	NBPGR324	EC9046		NBPGR
343	NBPGR327	EC10662		NBPGR
344	NBPGR330	EC11309		NBPGR
345	NBPGR331	EC12692		NBPGR
346	NBPGR333	EC13274		NBPGR
347	NBPGR334	EC13574		NBPGR
348	NBPGR336	EC13904		NBPGR
349	NBPGR337	EC14181		NBPGR
350	NBPGR338	EC15416		NBPGR
351	NBPGR339	EC16343		NBPGR
352	NBPGR340	EC16368		NBPGR
353	NBPGR341	EC16788		NBPGR
354	NBPGR342	EC20636		NBPGR
355	NBPGR343	EC20639		NBPGR
356	NBPGR346	EC27885		NBPGR
357	NBPGR347	EC25563		NBPGR
358	NBPGR349	EC26150		NBPGR
359	NBPGR350	EC26676		NBPGR
360	NBPGR351	EC26750-1		NBPGR
361	NBPGR352	EC27251		NBPGR
362	NBPGR353	EC27911		NBPGR
363	NBPGR354	EC27960		NBPGR
364	NBPGR356	EC27995		NBPGR
365	NBPGR357	EC28356		NBPGR
366	NBPGR359	EC29914		NBPGR
367	NBPGR361	EC29933		NBPGR
368	NBPGR362	EC29969		NBPGR
369	NBPGR363	EC30303		NBPGR
370	NBPGR364	EC31764		NBPGR
371	NBPGR366	EC32019		NBPGR
372	NBPGR370	EC32287		NBPGR
373	NBPGR371	EC32481		NBPGR
374	NBPGR372	EC32557		NBPGR
375	NBPGR374	EC32614		NBPGR
376	NBPGR378	EC33878		NBPGR
377	NBPGR379	EC34477		NBPGR
378	NBPGR380	EC34480		NBPGR
379	NBPGR384	EC35236		NBPGR
380	NBPGR386	EC35240		NBPGR
381	NBPGR387	EC35242		NBPGR
382	NBPGR388	EC35252		NBPGR
383	NBPGR391	EC35272		NBPGR
384	NBPGR392	EC35293		NBPGR
385	NBPGR394	EC35322		NBPGR
386	NBPGR395	EC35360		NBPGR
387	NBPGR397	EC36238		NBPGR
388	NBPGR398	EC89248		NBPGR
389	NBPGR399	EC161645		NBPGR
390	NBPGR401	EC163598		NBPGR
391	NBPGR403	EC164660		NBPGR

S.No	Code No.	Accession	Details	Source
392	NBPGR404	EC164665		NBPGR
393	NBPGR406	EC27336		NBPGR
394	NBPGR410	EC129602		NBPGR
395	NBPGR411	EC129604		NBPGR
396	NBPGR412	EC141887		NBPGR
397	NBPGR413	EC144336 A		NBPGR
398	NBPGR415	EC490122		NBPGR
399	NBPGR416	IC447708		NBPGR
400	NBPGR417	EC496124		NBPGR
401	NBPGR419	EC357828		NBPGR
402	NBPGR421	EC381554 A		NBPGR
403	NBPGR424	EC490128		NBPGR
404	NBPGR427	EC490130		NBPGR
405	NBPGR431	IC447706		NBPGR
406	NBPGR432	IC469682		NBPGR
407	NBPGR433	IC469653		NBPGR
408	NBPGR434	IC469633		NBPGR
409	NBPGR435	IC469648		NBPGR
410	NBPGR436	IC469629		NBPGR
411	NBPGR437	IC469628		NBPGR
412	NBPGR438	IC469626		NBPGR
413	NBPGR439	IC469603		NBPGR
414	NBPGR441	EC57442		NBPGR
415	NBPGR442	IC469525		NBPGR
416	NBPGR443	IC469597		NBPGR
417	NBPGR469	EC520052		NBPGR
418	NBPGR471	IC469714		NBPGR
419	NBPGR477	EC398716		NBPGR
420	NBPGR479	EC531801		NBPGR
421	NBPGR480	EC529086		NBPGR
422	NBPGR482	EC398691		NBPGR
423	NBPGR483	EC531802		NBPGR
424	NBPGR484	EC398710		NBPGR
425	NBPGR485	EC6486		NBPGR
426	NBPGR486	EC6192		NBPGR
427	NBPGR490	EC521067 B		NBPGR
428	NBPGR491	EC521068		NBPGR
429	NBPGR495	EC368883		NBPGR
430	NBPGR502	EC362958		NBPGR
431	NBPGR504	EC363942		NBPGR
432	NBPGR514	EC538139		NBPGR
433	NBPGR515	EC528362		NBPGR
434	NBPGR516	EC538153		NBPGR
435	NBPGR517	EC538455		NBPGR
436	NBPGR518	EC529085		NBPGR
437	NBPGR520	EC526146		NBPGR
438	NBPGR521	EC531805		NBPGR
439	NBPGR523	EC529081		NBPGR
440	NBPGR526	EC368943		NBPGR
441	NBPGR533	EC168290		NBPGR
442	NBPGR537	EC164660		NBPGR
443	NBPGR539	EC16788		NBPGR
444	NBPGR542	EC177371		NBPGR
445	NBPGR543	EC25265		NBPGR
446	NBPGR547	EC16790		NBPGR
447	NBPGR548	EC16780		NBPGR
448	NBPGR549	EC168283		NBPGR

S.No	Code No.	Accession	Details	Source
449	NBPGR550	EC26684		NBPGR
450	NBPGR551	EC12692		NBPGR
451	NBPGR552	EC12689		NBPGR
452	NBPGR553	EC7912		NBPGR
453	NBPGR554	EC14078		NBPGR
454	NBPGR555	EC16654		NBPGR
455	NBPGR557	EC13736		NBPGR
456	NBPGR558	EC29933		NBPGR
457	NBPGR559	EC8936		NBPGR
458	NBPGR561	EC114375		NBPGR
459	NBPGR562	EC2673		NBPGR
460	NBPGR565	EC2798		NBPGR
461	NBPGR566	EC5888		NBPGR
462	NBPGR567	EC2977A		NBPGR
463	NBPGR568	EC1914		NBPGR
464	NBPGR569	EC3668		NBPGR
465	NBPGR570	EC3261		NBPGR
466	NBPGR571	EC31767		NBPGR
467	NBPGR572	EC2791		NBPGR
468	NBPGR573	EC3176		NBPGR
469	NBPGR574	EC3216		NBPGR
470	NBPGR579	EC35240		NBPGR
471	NBPGR581	EC1087		NBPGR
472	NBPGR582	EC129604		NBPGR
473	NBPGR584	EC14073		NBPGR
474	NBPGR585	EC27910		NBPGR
475	NBPGR586	EC251581		NBPGR
476	NBPGR589	EC35293		NBPGR
477	NBPGR590	EC27960		NBPGR
478	NBPGR592	EC141887		NBPGR
479	NBPGR596	EC2990		NBPGR
480	NBPGR597	EC34480		NBPGR
481	NBPGR598	EC2977-A		NBPGR
482	NBPGR600	EC8591		NBPGR
483	NBPGR601	EC8822		NBPGR
484	NBPGR602	EC135580		NBPGR
485	NBPGR603	EC2790		NBPGR
486	NBPGR605	EC52077		NBPGR
487	NBPGR607	EC276		NBPGR
488	NBPGR609	EC5863		NBPGR
489	NBPGR610	EC252		NBPGR
490	NBPGR611	EC372		NBPGR
491	NBPGR614	EC170047		NBPGR
492	NBPGR615	EC50055		NBPGR
493	BSS1	N2219	NA	BSS
494	BSS2	N2231	NA	BSS
495	BSS3	N2241	NA	BSS
496	BSS4	N2244	NA	BSS
497	BSS5	N2263	NA	BSS
498	BSS6	N2266	NA	BSS
499	BSS7	N2267	NA	BSS
500	BSS8	N2268	NA	BSS
501	BSS9	N2269	NA	BSS
502	BSS10	N2279	NA	BSS
503	BSS11	N2280	NA	BSS
504	BSS12	N2292	NA	BSS
505	BSS13	N2298	NA	BSS

S.No	Code No.	Accession	Details	Source
506	BSS14	N2369	NA	BSS
507	BSS15	N2370	NA	BSS
508	BSS16	N2387	NA	BSS
509	BSS17	N2389	NA	BSS
510	BSS18	N2403	NA	BSS
511	BSS19	N2411	NA	BSS
512	BSS20	N2414	NA	BSS
513	BSS21	N2448	NA	BSS
514	BSS22	N2457	NA	BSS
515	BSS23	N2465	NA	BSS
516	BSS24	N2466	NA	BSS
517	BSS25	N2469	NA	BSS
518	BSS26	N2480	NA	BSS
519	BSS27	N2481	NA	BSS
520	BSS28	N2485	NA	BSS
521	BSS29	N2523	NA	BSS
522	BSS30	N2535	NA	BSS
523	BSS31	N2536	NA	BSS
524	BSS32	N2537	NA	BSS
525	BSS33	N2673	NA	BSS
526	BSS34	N2717	NA	BSS
527	BSS35	N2730	NA	BSS
528	BSS36	N2731	NA	BSS
529	BSS37	N2733	NA	BSS
530	BSS38	N2735	NA	BSS
531	BSS39	N2764	NA	BSS
532	BSS40	N2812	NA	BSS
533	BSS41	N2829	NA	BSS
534	BSS42	N2832	NA	BSS
535	BSS43	4218	NA	BSS
536	BSS44	4219	NA	BSS
537	BSS45	4225	NA	BSS
538	BSS46	4228	NA	BSS
539	BSS47	4229	NA	BSS
540	BSS48	4240	NA	BSS
541	BSS49	4241	NA	BSS
542	BSS50	4243	NA	BSS
543	BSS51	4257	NA	BSS

TGRC: Tomato Genetics Resource Center at University of California, Davis (tgrc.ucdavis.edu/).
IIVR: Indian Institute of Vegetable Research, Varanasi, U. P., India (www.iivr.org.in/).
NBPGR: National Bureau of Plant Genetic Resources, New Delhi, India (www.**nbpgr**.ernet.in/).
BSS: Bejo Sheetal Seeds Pvt. Ltd. Jalna, India (http://www.bejosheetalseeds.com/).
NA: Not available

NB: The details of NBPGR accessions are available at the PGR (Plant Genetic Resources) portal of NBPGR site (http://www.nbpgr.ernet.in:8080/PGRPortal/(S(o3cy12bkoz5s5e55w0dacf2a))/AdvancePassportSearch.aspx).