ANALYSIS OF *CAPSANTHIN CAPSORUBIN SYNTHASE* GENE IN BYADAGI CHILLI (*CAPSICUM ANNUUM* L.) AND ELUCIDATION OF CAROTENOID METABOLIC PATHWAY

By

NARESH S.

(2014-11-107)

CENTRE FOR PLANT BIOTECHNOLOGY AND MOLECULAR BIOLOGY

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR – 680 656

KERALA, INDIA

2017

ANALYSIS OF *CAPSANTHIN CAPSORUBIN SYNTHASE* GENE IN BYADAGI CHILLI (*CAPSICUM ANNUUM* L.) AND ELUCIDATION OF CAROTENOID METABOLIC PATHWAY

By

NARESH S.

(2014-11-107)

THESIS

Submitted in partial fulfillment of the

requirements for the degree of

Master of Science in Agriculture

(Plant Biotechnology)

Faculty of Agriculture
Kerala Agricultural University, Thrissur

Centre for Plant Biotechnology and Molecular Biology

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR – 680 656

KERALA, INDIA

2017

DECLARATION

I hereby declare that the thesis entitled **"Analysis of *Capsanthin capsorubin synthase* gene in Byadagi chilli (*Capsicum annuum* L.) and elucidation of carotenoid metabolic pathway"** is a bonafide record of research work done by me during the course of research and the thesis has not been previously formed the basis for the award to me any degree, diploma, fellowship or other similar title, of any other University or Society.

Vellanikkara,

Date: 31/3/2017

Naresh S.

(2014-11-107)

CERTIFICATE

Certified that this thesis entitled "**Analysis of *Capsanthin capsorubin synthase* gene in Byadagi chilli (*Capsicum annuum* L.) and elucidation of carotenoid metabolic pathway**" is a bonafide record of research work done independently by **Naresh S. (2014-11-107)** under my guidance and supervision and that it has not previously formed the basis for the award of any degree, diploma, associateship or fellowship to him.

31/3/2017

Vellanikkara,

Date: 31/3/2017

Dr. M. R. Shylaja
(Major Advisor, Advisory Committee)
Professor and Head
Centre for Plant Biotechnology and
Molecular Biology
College of Horticulture

CERTIFICATE

We, the undersigned members of the advisory committee of **Mr. Naresh S. (2014-11-107),** a candidate for the degree of **Master of Science in Agriculture**, with major field in **Plant Biotechnology**, agree that the thesis entitled **"Analysis of *Capsanthin capsorubin synthase* gene in Byadagi chilli (*Capsicum annuum* L.) and elucidation of carotenoid metabolic pathway"** may be submitted by him in partial fulfillment of the requirement for the degree.

Dr. M. R. Shylaja
(Chairperson, Advisory Committee)
Professor and Head
Centre for Plant Biotechnology and
Molecular Biology
College of Horticulture,
Vellanikkara.

Dr. D. Girija
Professor and Head
Dept. of Agrl. Microbiology
College of Horticulture,
Vellanikkara.

Dr. Minimol J. S.
Assistant professor
Plant Breeding and Genetics
Cocoa Research Centre, KAU,
Vellanikara

Dr. P. Indira
Professor
Department of Olericulture.
College of Horticulture,
Vellanikkara.

EXTERNAL EXAMINER

Dr. P. A. Nazeem
Professor and Head (Retd.)
Centre for Plant Biotechnology and
Molecular Biology
College of Horticulture,
Vellanikkara

ACKNOWLEDGEMENT

I appraise my deep sense of gratitude and indebtness to my benevolent and perspicacious Professor, Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara, Chairperson of my advisory committee, ***Dr. M. R. Shylaja*** *Professor and Head, for her for the valuable guidance and help rendered in the conduct of the experiment. My sense of gratitude and indebtedness to her will remain till eternity.*

I consider it is my privilege to express my heartfelt gratitude to ***Dr. Girija D.*** *Professor and Head Dept. of Agrl. Microbiology, College of Horticulture, Vellanikkara, for her constant support and concern, suggestions throughout research programme and critical scrutiny and manuscript.*

It is a gratifying moment to express my indebtness with devotion to ***Dr. Minimol J. S.*** *Assistant professor, Plant Breeding and Genetics, Cocoa Research Centre, KAU, Vellanikkara, member of my advisory committee, for his inspiring guidance, support and encouragement.*

I express my heartiest gratitude to ***Dr. P. Indira,*** *Professor, Department of Olericulture, College of Horticulture and member of my Advisory Committee for her ever willing help, valuable guidance and above all good understanding during the period of my work.*

I think it is my privilege to express my heartfelt thanks to ***DR. Valsala P. A*** *Professor and Head Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara,* ***Dr. P. S. Abida,*** *Professor and Head, genetics and plant breeding RARS pattambi.* ***Dr. Deepu Mathew,*** *Assistant Professor, Centre for Plant Biotechnology and Molecular Biology,and Professor,* ***Dr. Nazeem P. A.*** *Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara, for their untiring support and help during my course of study.*

My heartfelt sincere extreme gratitude to ***Dr.sherin, kavya chechi, beautline chechi, Priyanka, Anjali chichi, Ramya chechi, Dhanusri chechi, sharathettan*** *for their care, love, good wishes, borderless support.*

My sincere thanks to ***Shiny chechi, Shylaja chechi, Lyla chechi, Shobha chechi, Manju chechi, Likitha chechi, Renu chechi, Pushpa chechi, Sahra chechi, Aiswarya, Mushna chechi*** *CPBMB, College of Horticulture,*

I express my sincere thanks to my classmates, ***Mangalam, Darshan, Kala, Vinu, Rose, Ashwini, Anju, Sandesh, Sujith, Saha, Saakre, Brinda, Sridhar,*** *whose prayers, love and affection rendered me a successful path which propped up my career all along. My duty is incomplete if I forget my Senior friends and also roommate* ***HarikumarV*** *for their affection and kind help offered during research work.* ***Naresh babu prathi, Naveen, Ajith, Amarnathreddy, Manju, Ramesh,*** *whose helping hands, love and affection fetched a remarkable place in my days in Kerala.*

I wish to express my thanks to my loving juniors ***Basil, Ashwin, Debashish, Giri*** *and others of PG from college of Horticulture, KAU.*

I am happy to thank special persons of my life ***Venky, Naveen K, V. Somu, Sudhakar reddy, Anand*** *for their encouragement and moral support which helped me got over all odds and tedious circumstances.*

Iam in derth of words to express my soulfull gratitude to my loving Parents ***Sri. Sreerama K, Ravanamma*** *and my soul sister* ***Anitha*** *and* ***Kerthi*** *and grandmother* ***kotamma*** *for their wishes, blessings, support, boundless patience, prayers, inspiration, sacrifices, unflagging interest and eternal love and blessings and support I would not have complete this work.*

I submit all eulogize grandeur and reverence to my loving relatives and my well wishers especially members of Nadapannagari family for their endurance, love, moral support and inspiration during the period of my study.

I appreciate and acknowledge the facilities provided at CPBMB and Bioinformatics centre of ***Kerala Agricultural University*** *for the successful completion of my work. I also acknowledge research grant and Junior Research Fellowship provided by the* ***Department of Biotechnology, Govt. of India, New Delhi,***

It would be impossible to list out all those who helped me in one or other way in the successful completion of this work. I once again express my heartfelt thanks to all those helped me in completing this venture of time.

Naresh S.

DEDICATED TO MY FAMILY AND MAJOR ADVISOR

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF ANNEXURES

ABBREVIATIONS

ASTA	American Spice Trade Association
ABA	Abscisic Acid
bp	Base pair
β	Beta
β-LCY	Lycopene β-cyclase
Ccs	Capsanthin Capsorubin synthase
CCD	Carotenoid cleavage dioxygenase
CPBMB	Centre for Plant Biotechnology and Molecular Biology
CDD	Conserved Domain Database
CRTISO	Carotenoid isomerase
CTAB	Cetyl Trimethyl Ammonium Bromide
^{0}C	Degree Celsius
Cm	Centimeter
DMAPP	Dimethylallyl diphosphate
DNA	Deoxyribo Nucleic Acid
DNTPs	Deoxyribo Nucleoside Triphosphate
EDTA	Ethylene Diamine Tetraacetic Acid
ESTs	Expressed Sequence Tags
ε-LCY	lycopene ε-cyclase
GA	Gibberellins
GGPP	Geranyl Geranyl Diphosphate
GRN	Gene regulatory network
g	Gram
IPP	Isopentenyl diphosphate
Kb	Kilo basepairs
KAU	Kerala Agricultural University
KEGG	Kyoto Encyclopedia of Genes and Genomes
L	Litre
M	molar

MAS	Marker Assisted Selection
MEP	Methylerythritol 4-phosphate
Mg	Milligram
µl	Microlitre
µM	Micro molar
ml	Millilitre
Mg	Megnesium
$MgCl_2$	Megnesium Chloride
MSA	Multiple sequence alignment
NaCl	Sodium Chloride
ng/ µl	Nanogram per micro litre
OD	Optical Density
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
P^H	Hydrogen Ion Concentration
Psy	Phytoene synthase
PDS	Phytoene desaturase
%	Percentage
PVP	Poly Vinyl Pyrrolidone
RNA	Ribo Nucleic Acid
RNase	Ribonuclease
Rpm	Revolutions per minute
Sec	Second
SSR	Simple Sequence Repeat
TAE	Tris Acetate EDTA
TE	Tris EDTA
UV	Ultra violet
V	Volts
v/v	Volume by Volume
w/v	Weight by Volume
ZDS	Zeta carotene desaturase

Introduction

1. Introduction

Chilli (*Capsicum* spp.) belonging to the family solanaceae is one of the major vegetable cum spice crop. It is originated in tropical America and is also known as red pepper or chili pepper. Chillies were first introduced to the European Union in fifteenth century by Columbus and then spread along the spice trading routes to Africa, India, China and Japan to rest of the globe. The Portuguese introduced this wonderful crop to India in seventeenth century, (Bosland and Votava, 2000).

India is the world's largest producer, consumer and exporter of chillies. About 42.4 per cent of total Indian spices exports were contributed by chillies and earned a foreign exchange Rs. 2482.83 Million during 2015-16. Globally, India shares 40 per cent of area and 39 per cent of production in chilli (http://www.indianspices.com 2016).

Carotenoids are important colour compounds found in fruits, which are normally fat soluble pigments derived from five carbon isoprene units and polymerized enzymatically to form 40 carbon structures. The characteristic red colour found in *Capsicum* spp. is derived from the capsanthin and capsorubin, both of which are involved in carotenoid biosynthetic pathway (Govindarajan, 1985). Capsanthin and capsorubin are the keto-carotenoids, present particularly in the thylakoid membranes of the chromoplasts. These are the major red pigment present in ripe fruits contributes up to sixty per cent of total carotenoids. Capsanthin and capsorubin increase consistently during ripening process (Moehs *et al.*, 2001). Capsanthin capsorubin synthase gene specifically expressed and synthesized carotenoid pigmens in mature fruits at chromoplast maturation stage by incorporating ketocarotenoids. (Bouvier *et al.*, 1994). Capsanthin is highly stable and used as natural red colour in food industry, confectionaries, cosmetics, beverages, pharmaceuticals and even as a dye in textile industries.

Byadagi chlli is famous for its deep red colour and negligible or zero pungency. Demand for Byadagi chilli has increased enormously as a source of natural red colour in food industry, confectionaries, cosmetics, beverages, pharmaceuticals and even as a dye in textile industries. Byadagi chilli is mainly exported as oleoresin which serves as a substitute for paprika oleoresin.

Byadagi chilli exhibited its specific characters like high colour value and low pungency under Kerala conditions also (Renuka 2014). The analysis done at Spices Board quality evaluation laboratory at Kochi showed that Byadagi cultivars recorded higher colour (108.92 ASTA) and lower pungency (0.0045% capsaicin) while KAU varieties Ujwala and Anugraha possessed lower colour value (59.1 ASTA) and higher pungency (0.32% capsaicin).

The inheritance of ripe fruit colour in chilli is regulated by three loci – *c,1c2* and *y* (Hurtado-hemandez and smith, 1985). The existence of dominant allels at all three loci leads to formation of red ripe colour fruit. The yellow fruit colour is ressesive to red and is regulated by y locus (Popovsky and Paran, 2000).Three kinds of mutation *i.e* deletion, frame shift and a premature stop codon in the coding sequence of Ccs gene leads to yellow fruit type (Ha *et al.,* 2007). A new Ccs variant was identified in the yellow fruit cultivar CK7 by Li *et al.* (2013) when they analysed the genetic and regulatory association between yellow ripe fruit colour and the Ccs gene in pepper. They used gene specific primers for amplifying the coding sequence *Cds* and promoter fragments of the Ccs gene from selected genotypes.

The variable increase of capsanthin and capsorubin are responsible for different fruit colour in chilli. The expression levels of some genes in the carotenoid metabolic pathway are comparatively high in red peppers while some genes are not expressed in peppers with lower levels of carotenoid (Ha *et al.,* 2007). Cloning and functional charactarization of *Ccs* gene from tiger lilly (*Lilium lancifolium* Thunb.) was reported by Jeknic *et al.* (2012) cloned the gene from flower tepals by rapid

amplification of cDNA ends (RACE). Shah *et al.* (2014) studied chemistry of capsanthin, classical genetics, metabolism of capsanthin and its regulatory genes in carotenoid biosynthetic pathway under different conditions to improve the carotenoid production.

Gene regulatory network analysis has provided a clear understanding of complex expression pattern between many genes. It has allowed to access components of strongly co-expressed genes, related with biological processes. Combined analysis of gene expression and metabolite aggregation has permitted us to assume the functions of genes linked with specific metabolic processes by Aoki *et.al* (2007).

Elucidation of carotenoid biosynthetic pathway for production of capsanthin and capsorubin gives information on the enzymes and genes involved in colour formation which could be further utilized for secondary metabolite production, metabolic pathway engineering and marker assisted selection. Integration of genomic and transcriptomic data with metabolic pathway using bioinformatic tools will further add information on DNA sequence variation and gene expression for colour formation in chilli.

Keeping the above in view, the present study on analysis of *Capsanthin–capsorubin synthase*gene in Byadagi chilli (*Capsicum annuum* L.) and elucidation of carotenoid metabolic pathway was taken up at the Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Kerala Agricultural University.

The objectives of the study were to analyse *Capsanthin-capsorubin synthase gene* (*Ccs*) in Byadagi chilli and to elucidate the carotenoid metabolic pathway for production of capsanthin and capsorubin.

Review of literature

2. REVIEW OF LITERATURE

The investigations on "Analysis of *Capsanthin-capsorubin synthase* gene in Byadagi chilli (*Capsicum annuum* L.) and elucidation of carotenoid metabolic pathway" were conducted at the Centre for Plant Biotechnology and Molecular Biology (CPBMB), College of Horticulture, Kerala Agricultural University, Thrissur. The objectives of the study were to analyse *Capsanthin-capsorubin synthase* gene (*Ccs*) in Byadagi chilli and to elucidate the carotenoid metabolic pathway for the production of capsanthin and capsorubin.

2.1 Origin and history

Chilli (*Capsicum annum*) is an important vegetable crop coming under the family Solanaceae. It originated from South and Central America. Chilli or pepper (*Capsicum annuum* L.) is cultivated extensively in tropical and temperate zones of the world. India is known to be the world's largest producer, consumer and exporter of chillies. About 42.4 per cent of total Indian spices exports was contributed by chillies and earned a foreign exchange of Rs. 2482.83 million during 2015-16. Globally, India shares 40 per cent of total cultivated area in chilli and 39 per cent of total chilli production (http://www.indianspices.com 2016). The major producing nations are Hungary, India, Mexico, China and Korea.

Chilli belonging to the genus *Capsicum* contains more than 25 species, among these five species namely *C. annuum* L., *C. chinense* Jacq., *C. frutescens* L, *C. baccatum* L. and *C. pubescens* L. are domesticated and cultivated (Bosland and Botava, 2000; Costa *et al.*, 2009). Among the five cultivated species of the genus *Capsicum*, *C. annuum* is extensively grown as pungent (hot pepper) and non-pungent (Sweet pepper or bell pepper) fruits. In India, Karnataka, Andhra Pradesh, Maharashtra, Tamil Nadu, and Orissa account for more than 75 per cent of the total area and total production in chilli.

Chilli fruits contain vitamin C, A and E and is a good source of oleoresin which is used in processed foods, beverage industries and in pharmaceuticals. Chilli has diverse uses as spice, condiment, culinary supplement, medicine, vegetable and as an ornamental plant (Berke and Shieh, 2001).

2.2 *Capsicum* spp. varieties and cultivars of chilli

2.2.1 *Capsicum annuum*

C. annuum is a species native to southern North America and northern South America. This is the most common and widely cultivated among the five domesticated *Capsicum* spp. The *C. annuum* encompasses a wide variety of size and shape of peppers, both mild and hot ranging from bell peppers to chilli peppers.

Byadagi Dabbi, Byadagi Kaddi, Ujwala and Anugraha are some of the promising varieties of *C. annuum.*

2.2.2 *Capsicum frutescens*

C. frutescens can be either annual or short lived perennial plant. Flowers are usually white with greenish white or greenish yellow petals. The plant is either self-pollinated or cross pollinated by insects. The berries grow erect, ellipsoid-conical to lanceoloid in shape. The fruits are usually very small and highly pungent (Bosland, 1996).

Vellayani Samrudhi is a variety coming under *C. frutescens.*

2.2.3 *Capsicum chinense*

C. chinense is commonly known as "Yellow Lantern Chilli" and is native to America. *C. chinense* is well known for its exceptional pungency. The appearance and characteristics of the plant can vary greatly within the species (Manju and Sreelathakumary, 2002). *C. chinense* usually does not survive in winter but in warm climate, it behaves as a perennial plant (Manju and Sreelathakumary, 2002).

Vellayani Thejus and CC8-1 are the varieties coming under *C. chinense.*

2.2.4 *Capsicum baccatum*

C. baccatum is native to ancient Peru and it is also known as Amarillo chilli which literary means "yellow chilli". *C. baccatum* has white or cream colored flower which is usually self-pollinated or insect pollinated. The pods typically hang downwards and can have citrus or fruity flavour (Finger *et al.*, 2009).

2.2.5 *Capsicum pubescens*

C. pubescens originated in Mexico which is also known as "Manzano" pepper that means "apple" due to its apple-shaped fruit. It is native to Central and South America. The species name is *pubescens,* means hairy because of the hairy leaves. This particular hairy character of the leaves along with the black seeds is distinguishable characters of this species. A much noticeable feature of this species is its capacity to withstand cooler climate compared to other *Capsicum* spp. It is the least widespread and systematically farthest from all the domesticated*Capsicum* spp.

2.3 Genotypes used in the study

2.3.1 Byadagi Kaddi

Byadagi chilli is a well-known chilli variety cultivated in Karnataka particularly in Haveri, Gadag and Dharwad districts. It is named after Byadagi town located in Haveri district. Byadagi chilli is known for its deep red colour with negligible pungency. The fruit has a length of 10 to 15 cm characterstic wrinkles on the pods, little pungency and sweet flavour. The fruit is slender, linear, light green in colour and upon ripen turns to deep red colour, developing characteristic wrinkles at the ripening stage. This variety possesing the highest ASTA colour value of 97.45 and it is promising under rain fed conditions. It is reasonably resistant to pests and diseases (Renuka, 2014)

2.3.2 Byadagi Dabbi

Byadagi Dabbi is a variant of Byadagi is being cultivated for green chilli and dry chilli purposes. The fruits are of moderate length of about 8 to 10 cm, little bent at the apex, and slightly expanded at the base of the calyx. The demand for Byadagi chilli in food industry has increased enormously as a source of natural red colour. It is having highest ASTA colour value of 120.4 (Renuka, 2014). Spices Board has taken steps to protect Byadagi chilli under GI registration.

2.3.3 Ujwala

Ujwala is a variety improved by single plant selection. Fruits are clustered, linear, erect, dark green at immature stage and turns red upon maturity and are highly pungent. It is resistant to bacterial wilt and tolerant to mosaic and leaf curl virus. It is cultivated all over India. It can be grown both as a rain fed crop and as an irrigated crop. This variety can be used both as dry and green chilli. The average green chilli yield is about 8-10 tonnes/ha. The dry chilli yield is 1800 -2200 kg/ha.

2.3.4 Anugraha

Anugraha is an improved chilli variety derived by back crossing between Ujwala and Pusa Jwala. It is having medium-statured habit and flowers early. The bright green fruits grow up to 12 cm in length and reach a girth of 3 cm. Fruits are elongated, overhanging that turns red upon ripening. Anugraha is resistant to bacterial wilt and insect pests. The fruits retain medium pungency with good chilli characteristics.

2.3.5 Vellayani Samrudhi

It is also known as bird chilli which produces attractive creamy white fruits which are medium pungent. It is tolerant to shade and is suitable for homesteads. It is perennial with long fruit period with low incidence of pests and diseases. It is high yielding with an average yield of 30t/ha.

2.3.6 Vellayani Thejus

Vellayani Thejus coming under*Capsicum chinense* is a shade tolerant chilly variety with highly pungent and round shaped fruits. The fruits are green at mature stage and turn red upon ripening.

2.3.7 CC8 -1

CC8-1 coming under *Capsicum chinense* bear fruits that are creamy yellow in colour when immature and turns fully yellow when mature. Fruits are medium pungent and round in shape.

2.4 Carotenoids

Carotenoids are a group of coloured compounds which are synthesized biologically by plants, algae and certain types of bacteria and fungi. These fat soluble pigments are derived from five carbon isoprene units that are polymerized enzymatically to form 40 carbon structures (Arathi *et al.,* 2015). In nature, at least 750 different carotenoids exercising important biological functions have been identified (Maoka, 2009).

Carotenoids are generally present intracellularly in the chloroplasts and chromoplasts of plants. Generally carotenoids are physically categorized as carotenes (*eg.* β-carotene and α-carotene) and xanthophylls (*eg.* neoxanthin, lutein, zeaxanthin, β-Cryptoxanthin, fucoxanthin, and violaxanthin) (Delgado-Vargas and Peredes-López, 2003).

Carotenoids function by protecting cells and organelles from photo oxidative damage by suppressing nacent reactive oxygen species and free radicals (Sandmann *et al.,* 1999). It is also involved in photosynthesis (light harvesting complex) (Delgado-Vargas and Peredes-López, 2003). It acts as precursor molecule for abscisic acid synthesis and also helps in pollination by attracting pollinators (Zhu *et al.,* 2009). In the nutraceutical industry, carotenoids are valued for their health enhancing effects

such as reduction of vitamin A deficiency, cancer, cardiovascular diseases, age-related macular degeneration *etc*. (Zeb and Mehmood, 2004).

2.4.1 Carotenoid metabolic pathway in Chilli

The carotenoid biosynthesis pathway and associated gene was extensively investigated in carotenoid enriched plants like arabidopsis, chilli, tomato, daffodil and citrus (Cunningham and Gantt, 1998; Hischberg, 2001; Fraser and Branley, 2004).

Chilli pepper is an important plant model to analyze the biosynthesis of carotenoids in plants. Chilli fruits secrete different colours like red, yellow or orange carotenoids during the period of ripening process. Chloroplast imparts green colour to chilli fruits in early stages of fruit development, undergo intense modifications during ripening stages and develop into chromoplasts which are carotenoid synthesizing organelles (Camara and Brangeon, 1981).

Carotenoids are mainlysynthesized from two isoprenoid molecules: isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). It also serves as precursor molecule for diverse range of compounds including tocopherols, chlorophylls, phylloquinone, gibberellins (GA), monoterpenes and plastoquinone (Nisar *et al*., 2015).

IPP molecule is mainly synthesized by two pathways namely mevalonic acid pathway (MVA) and methylerythritol 4-phosphate (MEP) pathway (Rodriguez and Boronat, 2002).The isopentenyl diphosphate (IPP) anddimethylallyl diphosphate (DMAPP) involved in carotenoid biosynthesis pathway are derived from the MEP pathway (Eisenreich *et al*., 2001 ; Rodriguez and Boronat, 2002).

The 2-C-methyl-D-erythritol 4-Phosphate (MEP) pathway utilizes glyceraldehyde-3-phosphate and pyruvate as primary substrates to synthesize deoxy-D-xylulose 5-phosphate (DXP) which is activated by DXP synthase (DXS). MEP is consequently made *via* intramolecular reorganization and decline of DXP by the

enzyme DXP reductoisomerase (DXR). DXS and DXR are important in carotenoid flux regulation. IPP and DMAPP are formed after a number of subsequent steps and undergo a successive series of condensation reactions to yield the precursor of carotenoid biosynthesis *i.e* geranylgeranyl diphosphate (GGPP).

The carotenoid biosynthesis starts with the formation of C_{40} carotenoid phytoeneby condensation of two molecules of the C_{20} geranylgeranyl diphosphate (GGPP) formed from IPP and DMAP, obtained from the methylerythritol-4-phosphate (MEP) pathway. The condensation of two geranylgeranyl diphosphate (GGPP) molecules catalyzed by the phytoene synthase (PSY) enzyme produces 15-cis-phytoene, the primary colourless carotenoid. PSY is a vital enzyme governing the pathway and forms the backbone of carotenogenesis (Cazzonelli and Pogson, 2010). Convertion of 15-cis-phytoene to tetra-cis-lycopene by introducing four double bonds into phytoene *via* two enzymes: phytoene desaturase (PDS) and ζ -carotene desaturase (ZDS) is asymmetric dehydrogenation reaction.

Convertion of tetra-cis-lycopene to all-trans-lycopene requires a specific isomerase enzyme called carotenoid isomerase (CRTISO). This enzyme isomerizes cis bondsat seventh, ninth, seventieth and nintyth positions to convert tetra-cislycopene to all-trans-lycopene. (Park *et al.*, 2002 ; Isaacson *et al.*, 2004).

For the cyclases in plants the desired substrate is all-trans-lycopene. Lycopene cyclization is an essential step in carotenoid metabolism and makes carotenoid diversity by changing cyclic end groups either by the addition of beta (β-ring) and/or epsilon (ε- ring). Lycopene β-cyclase (β-LCY) and lycopene ε-cyclase (ε-LCY) produce (β-ring) and/or epsilon (ε- ring) respectively (Cunningham *et al.*, 1993 ; Cunningham *et al.*, 1996 ; Pecker *et al.*, 1996 ; Ronen *et al.*, 1999).

Linear carotenoid lycopene undergoes cyclization and makes two essential branching points *i.e* one leading to beta carotene and its derivatives (xanthophylls, zeaxanthin, antheraxanthin, violaxanthin and neoxanthin) whereas the second leading

to α-carotene and lutein in carotenoid pathway type of the cyclic end group *i.e* cyclohexane having ε- or β-ionone ring added differentiates between the carotenoids (Goodwin, 1980).

The nature of the cyclase enzyme governs the type of end group formed. Two major cyclases in higher plants *i.e* β-LCY, which introduces β-rings and and ε-rings introduced by ε-LCY. β-LCY introduces two β-rings to form β- carotene whereas a combined interaction of β-LCY and ε-LCY produces α-carotene. β-ionone ring is needed to produce provitamin A activity (Send & Sundholm, 2007).

Thus, lycopene lacks provitamin activity because of the absence of β-ionone ring. α-Carotene is replaced upon by β-ring hydroxylase to make zeinoxanthin, which is then hydroxylated by ε-ring hydroxylase to produce lutein. In another branch, β-carotene is hydroxylated to make zeaxanthin with the help of β-carotene hydroxylase. Zeaxanthin can be epoxidized to antheraxanthin and violaxanthin, through the action of zeaxanthin and antheraxanthin epoxidase, separately. Neoxanthin is formed from Violaxanthin by the enzymatic action of neoxanthin synthase

Ccs enzyme synthesizes capsanthin and capsorubin as final carotenoid product. This enzyme catalyzes 5,6-epoxycarotenoids antheraxanthin and violaxanthin to form capsanthin and capsorubin (Camara, 1980, 1981). Chromoplast maturation was chiefly attributed to precise expression of Ccs gene in chilli pepper fruits. The yellow fruited mutants with impaired Ccs gene did not express red pigmemts capsanthin and capsorubin during biosynthetic process (Bouvier, *et al.*, 1994).

Apart from conversion of Violaxanthin into capsorubin it can also transform into xanthoxin, (a precursor of ABA) by the action carotenoid cleavage dioxygenase (CCD).

2.4.2 Pigments responsible for colour in chilli

Capsanthin and capsorubin are the major red coloured pigments present in the ripened fruits that contributes upto 60 per cent of the total carotenoids (Moehs *et al.*, 2001). In paparika powder, the fraction of red carotenoids is 70-80 per cent, while the orange carotenoids constitute 20-30 per cent of total carotenoids. These two compounds present very low abundance in green fruits and gradually increase upon ripening (Salmeron and Garrido, 1976).

Vialoxanthin occupies as major carotenoid 37-68 per cent of total carotenoids fallowed by lutein, antheraxanthin and cis-violaxanthin in yellow and orange pepper fruits belonging to *C. baccatum*, *C. pubescens* and *C. annuum*,

2.4.3 Genetics of *Capsanthin capsorubin synthase* gene

Capsanthin and capsorubin determines typical red colour in *Capsicum* spp. as both are involved in carotenoid biosynthesis pathway (Govindarajan, 1985). During fruit ripening the synthesized pigments get accumulated in the thylakoid membrenes of chrooplasts present in the pericarp of the fruit which is regulated by Ccs gene.

Molecular analysis of different genes involved in the carotenoid metabolic pathway showed polymorphism in the *Capsanthin capsorubin synthase* (*Ccs*) gene by crossing between orange and red coloured fruits. Orange colour fruits have a deletion in the upstream region of the *Ccs* gene and 211 bp region is conserved in the downstream region. Further TLC analysis of carotenoids showed that capsanthin is not detected in the orange fruits, suggesting that *Ccs* gene controls colour of the fruits (Lang *et al.*, 2004).

Capsanthin and capsorubin are variably accumulated that is responsible for different fruit colour in chilli. The expression levels of some genes in the carotenoid biosynthesis pathway in red peppers are relatively high while some genes remain unexpressed in peppers giving lower carotenoid yeild (Ha *et al.*, 2007).

In chilli, three loci *c1*, *c2* and *y* control the inheritance of mature fruit colour. (Hurtado-hemandez and smith, 1985). Production of red ripe colour fruits is governed by the presence of dominant allels at three loci. Yellow fruit colour controlled by *y* locus is recessive to red (Popovsky and Paran, 2000). Ccs gene mutations found in yellow friuted types include deletion, frame shift mutation and presence of premature stop codon (Ha *et al.*, 2007).

Control of colour development in orange coloured *Capsicum annuum* fruit occupied a complex process rathe than the presence of a deletion in the structural gene, transcription of *Psy* and/ or *Ccs* gene regulates orange ripe colour Rodriguez-Uribe *et al.* (2012).

Cloning and functional characterization of *Ccs* gene from tiger lily (*Lilium lancifolium* Thunb.) was reported by Jeknic *et al.* (2012). They cloned the gene from flower tepals by rapid amplification of cDNA ends (RACE).

CK7 (a new Ccs variant) was yellow fruited cultivar identified by Li *et al.* (2013) When they analysed the genetic and the regulatory association between yellow riped fruit colour and *Ccs* gene in pepper. Gene specific primers were used to amplify coding sequence and promoter fragments of Ccs gene from selected genotypes.

Capsanthin metabolism and regulation in *Capsicum* spp. under different conditions was reviewed by Shah *et al.* (2014).

Reddy *et al.* (2012) reported that genetic variability of carotenoids was detected along with the allelic variants of candidate gene *Ccs*. They shown that yellow pepper line LCA 1068 (Aparna) contained Ccs gene coding region thus concluding deletion of ccs gene is not a prerequisite for colour change from red to yellow. Introns were absent in Ccs coding region but exhibited polymorphism due to amino acid changes in chilli lines. Candidate or allele specific markers can be

developed from Novel allelic variants found in Byadagi Dabbi (dark red) and LCA1068 (yellow).

Analysis of carotenogenic genes in yellow and orange colour peppers suggested that, *Lcyb, Crtz, Psy* and *Ccs* genes were essential for capsanthin synthesis and their wider expression patterns resulted in varied fruit colours of pepper (Tian *et al.*, 2015).

Capsicum spp. produces fruit that synthesizes and acccumulate carotenoid pigments responsible for the yellow, orange and red fruit colour types. The carotenoid level differs between species which is influenced by environmental conditions. Capsanthin, capsorubin and capsanthin-5, 6 epoxide confer red colour. Red *Capsicum* spp. Synthesize red coloured capsanthin that is lipophilic. 50 per cent of total carotenoid content may be represented by capsanthin during the mturation stage. Eleven conjugated double bonds, a conjugated keto group and a cyclopentane ring constitute capsanthin structure. Due to these characterstics, capsanthin show antioxidant activity (Matsufuji, 1998).

The content and carotenoid profile during fruit ripening is determined by two metabolic processes (1) Transformation of pigments involved in photosynthesis (2) denovo carotenoid biosynthesis rate of carotenoid accumulation has been found to be in declining trend as follows: ketoxanthophylls including capsanthin and capsorubin; xanthophylls including zeaxanthin, neoxanthin and vialoxanthin; epoxyxanthophylls including capsanthin 5-6 epoxide and capsanthin-3 6 epoxide and hydrocarbons including lycopenes (Deli, 1996). Since denovo carotenoid biosynthesis can determine the carotenoid content and composition chilli pepper fruits. The elucidation of carotenoid pathway during this stage can be a useful tool for monitoring the production of capsanthin and capsorubin in *Capsicum annum*.

Recently the investigations on red orange and yellow chilli pepper fruits from 28 genotypes (*C. annum, C. frutescens* and *C. chinense* types) were conducted to study the association between fruit colour and β-carotene content. Also it was attempted to characterize six carotenoids present in orange fruited varieties and to analyze the nucleotide sequences of four structural biosynthetic genes. All cultivars shows wider variability in terms of β-carotene, capsanthin and total carotenoids. Orange colour was predominantly determined by the presence of β-carotene but two cases showed oange colour due to red and yellow carotenoids (Guzman, 2010). Analysis of *Psy, Lcyb, CrtZ-2* and *Ccs* gene sequences deduced that specific carotenoid biosynthetic enzymes coded by different allels were related to specific carotenoid accumulation in orange fruits. This could serve as valuable molecular markers for selection of genotypes with high beta carotene.

Mialoundama *et al.* (2010) found that, alanine, lysine and arginine substituted by glutamate-295 in the conserved 293-FLEET 295 motif of Ccs stopped the formation of β-carotene and k-cyclic carotenoids.

Capsanthin capsorubin synthase and ZEP zeaxanthin epoxidase are involved in Capsanthin and capsorubin biosynthesis. The expression of their corresponding genes *Ccs* and *ZEP* were monitored during fruit ripening to assess the capsanthin and capsorubin producing genes: *Psy*, *Lcyb*, *Crt-Z* and Ccs involved in capsanthin biosynthesis pathway. Formation of red colour pigments is initiated by the expression of Ccs gene which catalyzes zeaxanthin into capsanthin responsible for red colour pepper fruits. (Lefebvre *et al.,* 1998). Ccs gene deletion or mutation result in yellow colour pepper fruit as capsanthin synthesis blocked (Thorup *et al.,* 2000).

Borovsky *et al.* (2012) identified an amino acid transisition at 709th position to G in the cDNA sequence of β-carotene hydroxylase-2 in orange pepper and complete co-segregation of this SNP was observed with the mutated 'maor' phenotype. They propose that β-carotene hydroxylase-2 mutation governs the orange colour in pepper.

They found of β-carotene hydroxylase-2 expression was elevating additional carotenogenesis genes in the orange fruits indicating presence of feedback regulation of genes involved in the carotenoid biosynthesis pathway.

Recently virus induced gene silencing (VIGS) showed a association between fruit colour changes and target genes (*Psy, Lcyb, Crt-Z* and *Ccs*) silencing. The capsanthin content significantly decreased upon silencing a single gene or a group of genes. Capsanthin synthesis can be directly or indirectly influenced by silencing key genes (Tian *et al.*, 2014).

Wei *et al.* (2014) used HPLC and real time PCR to study the expression profile of 23 carotenoid biosynthesis gene family members to analyze relationship between carotenoids accumulation in flavedo, juice sacs, leaves of orange fruit at maturation. The result revealed that vioaloxanthin and leutin was mainly accumulated in young mature fruits and leaves. Vialoxanthin increase gradually in juice sacs, leutin content was three times more in ln leaves than in juice sacs. Ccs gene had similar expression in flavido and in juice sacs noticed. They also attempted comparision of expression patterns of 23 carotenoid pathway gene members with other plants and concluded these data is useful to describe tissue specific roles of these genes.

Ghosh *et al.* (2015) Detected SNPs in chalcone synthase gene which was is associated with to 6-gingerol content. They found alteration in the two amino acid sequence (aspargine to serine at 336 position) and (serine to leucine at 142 position). Asparagine at 336 is a critical amino acid catalytic triad of chalcone synthase gene responsible for substrate binding. They suggest that landraces with change in aspargine to serine causes low 6-gingerol content due to weak enzyme substrate association.

2.4.4 Conserved domain database (CDD)

Protein domains are distinct units of evolution associated with particular aspects of molecular function such as catalysis or binding. Identification of functionally characterized domains in protein sequences may give clues to their molecular and cellular function. The Conserved domain database (CDD) is the protein arrangement module of NCBI's Entrez query and retrieval system and is a faster communicating tool to identify conserved domains in new protein sequences. Bauer and Bryant (2004), described detection of structural and functional domains of sequences by using conserved domain search results that provided domain architecture, alignment and three dimensional graphic displays within the domain family.

2.4.5 BioEdit

Hall (2011) reviewed that BioEdit is biological sequence alignment tool that contains many features like Split window view, user defined color, information based shading and auto integration with other programs such as ClustalW and Blast.

Multiple alignment of nucleic acid and protein sequences for preparing phylogenetic trees using clustal series of programs generate output with high degree of robustness, portability and user-friendliness (Sugawara *et al.*, 2003).

Multiple sequence alignment is the most widely used method to align nucleotide or amino acid sequences using ClustalW and ClustalX which is a graphical alignment tool that displays sequence alignment on the screen which users can scroll easily on the aligned sequences. It contains a customizable colouring system used to highlight the conserved regions in the sequence. ClustalX is easy to install and is user-friendly (Thompson *et al.*, 1994).

2.4.6 Analysis of EST's from *Capsicum annuum* and transcriptome integration

Expressed Sequence Tags (ESTs) are most important tools for the discovery of new genes in plants and other organisms. Castillo *et al.* (2012) generated a genome source to help *Capsicum annuum* study by assembling of Expressed Sequence Tags (ESTs) from different parts of tissue like (root, stem, leaf, flower and fruit) which were sequenced by Sangers and GS-pyrosequencing methods. Validation by identity analysis with existing pepper sequences generated 32,314 contigs out of 1,324516 raw reads.

A MySQL database was constructed containing ESTs sequence, functional annotation, gene ontology classification, metabolic pathways and assembly information.

Expressed sequence tags (ESTs) provides the opportunity to evaluate sequence conservation and divergence on genomic scale (Rensink *et al.,* 2005). They analyzed expressed sequence tags (ESTs) from six *Solanaceae* species by clustering and assembling into gene indices and confirmed that at nucleotide level highly conserved sequences were found among *Solanaceae*.

For understanding more information in biosynthesis of capsaicinoids, they isolated total RNA from placenta and pericarp of pungent pepper (*Capsicum frutescens* L.). The RNA was sequenced and assessed for the effects of numerous assembly constraints using various assembly softwares. Found a suitable approach in order to de novo assemble transcriptome data. Around 54,045 transcripts (high quality uni genes) were obtained through Trinity software. Around 92.65 per cent genes showed sequence similarity to protein sequences genome of potato, tomato and pepper (*C. annuum*) ESTs databases.

Lu *et al.* (2011) reported comprehensive characterization of the transcriptomes of the mature fruits of two red pepper (*Capsicum annuum*) accessions,

YCM334 and Taean, using the 454 Genome Sequencer (GS)-FLX Titanium System. For YCM334, sequencing generated 121,152 reads, which were assembled into 11,770 contigs and 13,827 singletons. And Taean generated 154,088 reads, assembled into 14,200 contigs and 15,135 singletons. Using blast against the NCBI non-redundant (NR) and UniProt databases, unigenes were functionally annotated and expression pattern analysis revealed presence single nucleotide polymorphisms (SNPs) which provides valuable resources for depiction of some important agronomic traits and documentation of molecular mechanism such as phytopathogen resistance.

Kim *et al.* (2008) designed and constructed chilli pepper EST database in order to perform comprehensive analysis of gene expression with response to biotic and abiotic stresses. After clustering and assembling of ESTs into consensus cDNAs, the cDNAs were assigned to metabolic pathway by Gene Ontology (GO) and MIPS Functional Catalogue (FunCat). Pepper EST database contains data of numerous unigenes that help in analyzing expression patterns at different developmental stages under stress and comparison of ESTs with other *Solanaceae* family members.

2.4.7 Databases for metabolic pathway analysis

Metabolic pathway databases such as "Plant Metabolic Network/MetaCyc" and "KEGG" pathway are publicly accessible resources providing organism-specific information on reactions and metabolites. Kyoto Encyclopedia of Genes and Genomes (KEGG), is an online database that contains genomic, metabolomic and functional information. In particular, gene records from completely sequenced genomes are associated to systemic functions of the cell, the organism and the ecosystem.

Protein interaction network and chemical reactions involved in various cellular processes is mechanized by KEGG. KEGG reconstructs protein interaction networks from all those organisms with completely sequenced genome. It forms base for utilization in fuctional genomics, expression database and proteomic studies.

KEGG maps are graphical illustrations representing knowledge about genetic information processing, cellular processes, organismal systems, molecular interaction and metabolic reactions. Each graph is drawn manually by using kegSketch with experimental evidence in literature. KEGG pathway displays metabolic networks in the form of wired electronic circuit like maps. (Kanehisa *et al.,* 2011).

MetaCyc database is publicly available comprehensive database that explains metabolic pathways and enzymes from all domains of organisms. It contains experimentally determined 2400 small molecular metabolic pathways. BioCyc is a collection of genomic and metabolic networks of an organism containing metabolites, enzymes, reactions, metabolic pathways and transport systems (Caspi *et al.,* 2015).

BRENDA is an enzyme functional database that contains manually curated literature-based data on enzyme function, their metabolic role, genomic and protein sequence data and enzyme structure data (Chang *et al.*, 2015). BRENDA contains manually annotated data for 82,568 enzymes and 7.2 million enzyme sequences from UniProt. New interactive and intuitive BRENDA pathway maps give an overview on biochemical processes and facilitate the visualization of enzyme, ligand and organism information in the biochemical context (Placzek *et al.,* 2016).

2.4.8 Cytoscape

Cytoscape is an open source community software project used for integrating high troghput expresion data and other molecular state information with biomolecular interaction networks for analysis and visualization. "Core" software provides functionality to layout, query the network and visually integrate state data with network. It is powerful when utilized in conjugation with large protein –protein, protein - DNA and genetic interaction databases avilable for humans and model organisms. It visually integrates the network with phenotypes, expression profiles, other molecular state information and links to databases of functional annotations. (Baryshnikova, 2011).

Gene regulatory network (GRN) as an undirected graph represents biological systems and behavior using biological compounds that are generated from omics data. GRN is composed of nodes and edges representing genes and mutual co-expression relationships (Aoki *et al.,* 2007 ; vijesh *et al.,* 2013 ; Higasi and Saito, 2013). Gene regulatory network contains cluster of genes, proteins and metabolic network contains whole data of metabolites and their governing enzymes and reactions.

The complex of biological networks has a hierarchy (Oltvai and Barabasi, 2002). Among the levels of genome wide organization and individual molecular components, there are substructures such as modules, motifs and pathways characterized by topological properties such as degree distribution, network density and clustering coefficient (Watts and Strogatz, 1998). A module represent a set of genes having a distinct function that arises from interactions among them. The analysis of module is more informative compared with respect to regulatory mechanisms of the specific biological processes.

Perrin *et al.* (2003) used statistical machine learning approach to describe and identify gene regulatory networks from experimental data. It is acknowledged as a dynamic Bayesian network principally suited for tackling the random nature of gene regulation and gene expression measurement. When it was evaluated against S.O.S. DNA repair network data of *Escherichia coli*, it was able to extract exact regulations between genes involved in this network.

Analysis of gene regulatory network will provide a clear mode to describe gene expression pattern among several genes. It allows to access modules of strongly co-expressed genes related to biological processes. Functions of genes related to specific metabolic processes were hypothesized by integration of gene expression data and metabolomic data (Aoki *et al.,* 2007).

Hecker *et al.* (2008) used computational methods to reconstruct GRNs from experimental data. Methods such as standard GRN inference methods basically use gene expression data obtained from microarrays and focuses on capable modelling methods. It also provides a synopsis of modelling schemes, learning algorithms available in common and outlined current challenges.

Baerenfaller *et al.* (2012) described the investigation of *Arabidopsis* leaf growth and abiotic responses using systems-based omics analysis by mining and displaying transcriptome, proteome and phenome data.

Pan *et al.* (2013) discovered novel fruit related regulator gene involved in pigment accumulation in tomato by artificial neural network interface analysis and the function was validated in transgenic plants. Artifial neural network analysis and transcripon factor gene expression profiles at various developmental stages were used for GRN construction. It revealed that gene expression profiles with a sequence releated to *Arabidopsis thaliana i.e* ARBIDOPSIS PSEUDO RESPONSE REGULATOR2 LIKE GENE (APPR2) like. Overexpression of APRR2 gene in transgenic lines show upregulation of ripening related genes that link ripening process. APRR2-Like gene in *Capsicum annuum* sweet pepper is associated with pigment accumulation in fruits.

Vijesh *et al.* (2013) reviewed different methods of reconstructing gene regulatory network models and different classes of GRNs (1) Logical models, like Boolean network, Probabalastic Boolean network, Bayesian network (2) Continuous models (3) Single Molecule Level Model and (4) Hybrid Model *etc.*

Higashi, Y. and Saito, K. (2013) reviewed the network analysis in *Arabidopsis thaliana* and discovered genes involved in plant specialized metabolism in non-model plants by using gene coexpression network analysis and mass spectrometry-based metabolite profiles (gene-to-metabolite correlation).

Guerin *et al.* (2016) reported multilevel methodology in oil palm for combining gene co expression analysis, quantifying allel specific expression and joint multi variant analysis of transcriptomics and lipid data. It displayed fattyacid composition and contrasting oil contents. A tight transcriptional coordination of fatty acid synthesis in the plastid with sugar sensing, glycolysis, starch storage and carbo recapture pathways were revealed from the gene coexpression network.

2.5 Terminologies of gene regulatory network

Node – nodes represent genes usually equivalent to the corresponding transcripts and proteins and small molecules (metabolites)

Edges - edges between the nodes represents gene relationship between the genes and interacton of genes

Degree - number of links made by a node

Network density - a ratio of the observed number of links to all possible links

Clustering coefficient - clustering coefficient of node (n) is a ratio of observed number of links between neighbours to number of all posiible links between neighbours

Motif - statistically over-represented sub-graphs

Module - a group of nodes that linked more densely within the group

Materials and methods

3. Materials and Methods

The research work on "Analysis of *Capsanthin-capsorubin synthase* gene in Byadagi chilli (*Capsicum annuum* L.*)* and elucidation of carotenoid metabolic pathway." was under taken at CPBMB, CoH, Kerala Agricultural University, Thrissur during 2014 – 2016. The objective of the study was to analyse *Capsanthin-capsorubin synthase* gene (Ccs) in Byadagi chilli and to elucidate the carotenoid metabolic pathway for production of capsanthin and capsorubin. The experimental materials and methodology adopted for characterizing *Capsanthin-capsorubin synthase gene* (Ccs) at molecular level and elucidating the carotenoid metabolic pathway are presented in this chapter.

3.1 Materials

3.1.1 Plant materials

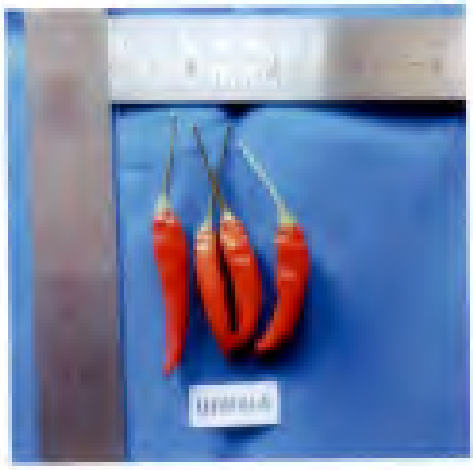
Seven genetically distinct chilli varieties /accessions were selected based on colour at fully ripe fruit stage. The accessions selected belong to three different *Capsicum* spp. Byadagi Kaddi, Byadagi Dabbi, Ujwala, Anugraha belong to *Capsicum annuum*, Vellayani Samrudhi to *Capsicum frutescens*, Vellayani Thejus, and CC8-1 to *Capsicum chinense* (Plate1a-1c).

Byadgi chilli is a well-known genotype of chilli particularly cultivated in Karnataka. It is named after Byadgi town located in haveri district. Fruits are dark red colour with negligible pungency. Byadgi Kaddi fruit has a length of 10 to 15 cm, categorized by wrinkles on the pods, and sweet flavour. Fruits are linear and light green at mature stage, at ripening stage changed to deep red colour by developing the characteristic wrinkles.

Byadgi Dabbi is another cultivar of Byadagi is being cultivated for green and dry chilli. The fruits are of medium length 8-10cm, a slight curved at the apex and bulged at the base of the calyx. Byadagi dabbi fruits are green at matre stage and

BYADAGI KADDI

BYADAGI DABBI



UJWALA

Plate 1a. Genotypes used for the study

ANUGRAHA

VELLAYANI SAMRUDHI

Plate1b. Genotypes used for the study

VELLAYANI THEJUS

CC8-1

Plate 1c. Genotypes used for the study

turns deep red colour bat ripening stage by developing characterstic wrinkles. Ujwala is a KAU released variety with clustered linear erect dark green fruit, resistant to bacterial wilt, tolerant to mosaic and leaf curl virus. Fruits are, dark green and turn red on ripening. Anugraha is a chilli variety released from KAU which is derived from the cross between Ujwala and Pusa Jwala. It produces elongated, hanging fruits having light green colour and turns red upon ripening. Vellayani samrudhi is a promising bird chilli. It is high yielding with an average yield 30t/ha. It produces attractive creamy white fruits at immature stage and turns orange colour upon ripening. Vellayani Thejus, is a shade tolerant chilly variety with highly pungent round shaped fruits. The fruit are green at mature stage and turns red upon ripening. CC8-1 is a Capsicum accession comes under *Capsicum chinense*. Fruits are creamy yellow at maturity and turns to fully yellow at ripening. Fruits are medium pungent and round in shape.

Seeds originally collected from Haveri district of Karanataka and maintained at CPBMB, College of Horticulture were used to raise seedlings in pots. Leaves were collected from one month old seedlings for DNA extraction.

3.1.2 Laboratory chemicals, glassware and equipment

The chemicals used for the study were pure and of good quality obtained from Merck India Ltd., HIMEDIA and SISCO Research Laboratories. The Taqpolymerase, dNTPs, Taq buffer and DNA ruler (Eco R1, 100bp ladder) were supplied by Invitrogen. The plastic wares were procured from Axygen and Tarson India Ltd. The Ccs gene specific primers were obtained from Sigma Aldrich Pvt. Ltd.

The equipment used for the investigations were high speed refrigerated centrifuge (KUBOTA 6500), NanoDrop ND-1000 spectrophotometer for estimation of quality and quantity of DNA, horizontal gel electrophoresis system (BIO-RAD, USA) and thermal cycler (Proflex) for DNA amplification. Details of equipment are provided in Annexure 1.

3.2 Methods

3.2.1 Raising chilli varieties /cultivars

The chilli genotypes used in the study were raised in the net house of CPBMB, KAU, Vellanikkara. The plants were raised in pots with one plant/pot, and five pots were maintained for each genotype. The seeds were sown in September 2014 and were germinated within ten days of sowing. Fourteen days after germination, tender, immature leaves were collected for DNA isolation.

3.2.2 Genomic DNA isolation

Young tender leaves were collected in ice box from individual plants of all seven genotypes. The surface was sterilized by wiping with 70 per cent ethanol. The fresh green leaves were ground to fine powder in liquid nitrogen along with β-mercaptoethanol and PVP in an ice cold mortar and pestle.

CTAB method developed by Rogers and Bendich (1994) was used for genomic DNA isolation

Reagents:

I. 2X CTAB buffer

II 10X CTAB buffer

III TE buffer

Procedure

- Ground one gram of leaf sample in a pre-chilled mortar and pestle in the presence of liquid nitrogen and β mercaptoethanol. (50µl) in an okridge tube
- Added 4ml of 2X CTAB extraction buffer

- Transferred the homogenized sample into an autoclaved 50ml centrifuge tube which contains 3ml of pre-warmed extraction buffer
- Mixed well and incubated the mixture at 65^0 c for 30 minutes with occasional mixing by gentle inversion.
- Added equal volume of chloroform: isoamyl alcohol (24:1) and mixed well by slight shaking to emulsify.
- Centrifuged at 10,000 rotation per minute for fifteen minutes at 4^0c
- After spinning, contents get separated into three distinct layers.

  Aqueous topmost layer - DNA

  Interphase - Fine particles, cell debris, and emulsified protein

  Lower layer - chloroform and pigments
- Transferred the top most aqueous layer to a clean centrifuge tube and added $1/10^{th}$ volume of 10 per cent CTAB solution and equal volume of chloroform: isoamyl alcohol (24:1) and mixed by gentle shaking
- Centrifuged at 10,000 rotation per minute for fifteen minutes at 4^0c
- Transferred the transparent content phase into a clean centrifuge tube and added 0.6 volume of pre chilled isopropanol and mixed by quick gentle inversion till the DNA precipitated. Kept at -20^0 for half an hour for complete precipitation
- Centrifuged at 10,000 rpm for 15 minutes at 4^0c and poured off supernatent
- Added 70 per cent ethanol to DNA pellet
- Spun for 5 minutes at 8000 rpm and discarded the ethanol
- Air dried the pellet, dissolved in 50µl of autoclaved distilled water and stored at -20^0c

3.2.2.1 Purification of DNA

The DNA isolated had RNA contamination which was purified by RNase treatment and further precipitated

50

Reagents

1.Chilled isopropanol

2. 70% ethanol

3. TE buffer

4. Chloroform: isoamyl alcohol (24:1, v/v)

One percent RNase was prepared by dissolving RNase in autoclaved distilled water. The solution was suspended into aliquots and stored at -20°c

Procedure

- To 100 µl DNA sample, added RNase solution (2µl) and incubated at 37°c in hot water bath for 45 minutes
- Added equal volume of chloroform: isoamyl alcohol (24:1) mixture and mixed gently
- Centrifuged at 10,000 rpm for 15 minutes at 4^0c
- Transferred the transparent layer into a fresh micro centrifuge tube and added same volume of chloroform : isoamylalcohol (24:1)
- Centrifuged at 10,000 rpm for 15 minutes at 4^0c
- Transferred the aqueous phase into a fresh microfuge tube and added 0.6 volume of chilled iso propanol and mixed by gentle inversion till the DNA was precipitated. Kept at -20^0c for half an hour for complete precipitation
- Centrifuged at 10,000 rpm for 15 minutes at 4^0c
- Gently poured off the supernatant and washed the pellet with 70 per cent ethanol
- Air dried the pellet, dissolved in autoclaved distilled water and stoted at -20^0c

3.2.2.2 Quantification of DNA

3.2.2.2.1 Assessment of quality and quantity of DNA by electrophoresis

The quantity and quality of isolated DNA were evaluated through agarose gel electrophoresis (sambrook *et al* 1989)

Materials used for agarose gel electrophoresis

I Agarose

II 50X TAE buffer (P^H 8.0)

III Tracking/ loading dye (6X)

IV Ethidium bromide

V Electrophoresis unit, power pack, gel casting tray and comb

VI UV transilluminator

VII Gel documentation and analysis system

Procedure

- The gel was made by sealing the ends with cello strip. The comb was positioned in gel tray about one centimeter at one end of the tray and positioned the comb vertically
- Prepared 0.8 per cent agarose (0.8g in 100ml) in a conical flask with 100ml 1X TAE buffer. Microwaved for 45 to 50 seconds, complete melting of agarose looks clear.
- Cool the solution to about 42 to 45^0 prior to pouring. After cooling Ethidium bromide (0.5µg/ml) was added
- Poured this warm gel into the plate to the depth of about 5mm and allowed to gel to solidify for about 30 to 45 minutes in room temperature
- Gently removed the comb and cello tape used for sealing, placed the gel containing tray in electrophoresis compartment and covered and 1X TAE buffer added to the chamber upto top of the gel

- Samples were prepared for loading by adding 3µl of 6X loading dye for 2µl of sample. Mixed gently and loaded 5µl of DNA per well. Loaded λDNA/ *Hind* III +*Eco*RI double digest (1000bp) in one lane.
- Electrophoresed at 100 volts until dye has migrated two thrd the length of the gel
- Intact DNA appeared as orange fluroscent bands when viewed under UV transilluminator. If degraded, it will appeared as because of the presence of a large number of bands, which differ in one or two bases.
- The image was documented and saved in gel documentation system.

3.2.2.2.2 Assessing the quality and quantity of DNA by nanodrop spectrophotometer

The purity of DNA was further cheked by using NanoDrop ND – 1000 spectrophotometer. Nucleic acid shows absoption maxima at 260nm where as proteins show peak absorbance was recorded both wave length and purity of DNA was indicated by the ratio OD_{260}/OD_{280}. The values between 1.8 to 2.0 indicate that the DNA was pure and free from proteins. The quantity of DNA was calculated using the relation 1 OD_{260} is equivalent to 50µg double standard DNA/ml sample.

1 OD at 260nm = 50µg DNA/ml

Therefore OD_{260} 50 gives the quantity of DNA µg/ml.

Procedure for nanodrop spectrophotometer

- Connected nanodrop spectrophotometer to the system and opened operating software ND-100
- Select the option Nucleic acid.
- Opened the sampling arm and pippeted 2µl of distilled water onto the lower measurement platform.

- Sampling arm was closed and spectral measurement initiated by using operating software on the PC. The sample was mechanically drawn between top and bottom pedestals and spectra of sample were made.
- Set the reading to zero with sample blank.
- 1µl sample was pipetted onto measurement pedestal and selected measure.
- After completion of measurement, opened the pedestal arm and clean the sample from both the top and bottom pedestals using tissue paper. Cleaning of sampling arm prevented sample carryover for further successive measurements.

3.2.3 Analysis of *Capsanthin capsorubin synthase* gene

3.2.3.1 Analysis of *Capsanthin capsorubin synthase* gene by gene specific primers

Analysis of *Capsanthin capsorubin synthase* gene in seven chilli genotypes (Byadagi Kaddi, Byadagi Dabbi, Ujwala, Anugraha, Vellayani Samrudhi, Vellayani Thejus, and CC8-1) was carried out with two SSR gene specific primers viz. *Ccs Cds* and *Ccs promoter* (Table1). (Li *et al.*, 2013)

3.2.3.2 Details of gene specific primers used for*Ccs* gene analysis

Sl. No	primer	Forward Sequence (5' to 3')	Reverse Sequence (5' to 3')
1	*Ccs cds*	5'-CCTTTTCCATCTCCTTTACTTTCCATT-3'	3'-AAGGCTCTCTATTGCTAGATTGCCCAG-5'
2	*Ccs promoter*	5'-TTGAACCTCCTTGATAAAA-3'	3'- GGAAAGTAAAGGAGATGGA-5'

Table1. Details of SSR *Ccs* gene specific primers used for DNA amplification

3.2.3.3 Analysis of Ccs gene using Ccs Cds primer

Good quality DNA (30-40 ng/µl) isolated from seven chilli genotypes using CTAB method and amplification was amplified using *Ccs cds* primer as per the procedure reported by Li *et al.*, (2013). The amplification was done in a Proflex Thermal Cycler. PCR reactions were performed using 20 µl reaction mixture.

Composition of the reaction mixture for PCR (20 µl)

Materials	Quantity (µl)
a) Genomic DNA (30 ng/µl)	-2.0
b) 10X taq assay buffer B	-2.0
c) $MgCl_2$	-2.0
d) $dNTP_S$	-1.5
e) Taq DNA polymerase (1U/ µl)	-0.4
f) Primer	-1.5
g) Autoclaved distilled water	-10.6
Total volume	**20.0 µl**

A master mix with all reagents for the required number of reactions was prepared and aliquots were dispensed into PCR tubes followed by addition of template DNA in each tube. One tube without template DNA was kept as blank.The PCR tubes were kept in the thermal cycler and the program as shown below was run

Step1: 94^0 C for 4 min - Initial denaturation

Step2: 94^0 C for 30 sec - Denaturation

Step3: 56^0 C for 30 sec - Primer annealing

Step4: 72^0 C for 30 sec - Primer extension

Step5: 72^0 C for 10 min - Final elongation

Step6: 4^0 C for infinity to hold the sample

The amplification was done in a (proflex) thermal cycler. PCR reaction was performed using 20µl reaction mixture. The amplified PCR samples were run on 1.8 per cent agarose gel using 50X TAE buffer. The gel was stained with ethidium bromide along with marker (100bp - 1kb ladder). Electrophoresis was performed at 70 volts for two hours. The gel picture was visualized and documented using gel documentation system.

3.2.3.4 Analysis of Ccs gene using Ccs Promoter primer

Good qualityDNA (30-40 ng/µl) isolated from seven chilli genotypes using CTAB method were amplified using *Ccs promoter* primer. Amplification was carried out in a Proflex Thermal Cycler. The PCR reactions were performed using 20 µl reaction mixture, composition of master mix are provided in page no 34.

Composition of the reaction mixture for PCR (20 µl)

Materials	Quantity (µl)
a) Genomic DNA (30 ng/µl)	-2.0
b) 10X taq assay buffer B	-2.0
c) $MgCl_2$	-2.0
d) $dNTP_S$	-1.5
e) Taq DNA polymerase (1U/ µl)	-0.4
f) Primer	-1.5
g) Autoclaved distilled water	-10.6
Total volume	**20.0 µl**

A master mix with all reagents for the required number of reactions was prepared first and aliquots were dispensed into PCR tubes followed by addition of template DNA in each tube. One tube without template DNA was kept as blank.The PCR tubes were kept in the thermal cycler and the programme as shown below was run:

Step1: 94^0 C for 7 min - Initial denaturation

Step2: 94^0 C for 30 sec - Denaturation

Step3: 47^0 C 30 sec - Primer annealing

Step4: 72^0 C for 30 sec - Primer extension

Step5: 72^0 C for 10 min - Final elongation

Step6: 4^0 C for infinity to hold the sample

Electrophoresis was performed at 70 volts for two hours. The gel picture was visualized and documented using gel documentation system.

3.2.3.5 Sequencing of PCR products

Two sets of PCR products, (1) amplified using Ccs *Cds* primer of seven chilli genotypes, (2) amplified using *Ccs promoter* primer of seven genotypes, confirmed to yield only single band on electrophoresis, were further sent for sequencing. Sequencing was carried out by outsourcing in Sci Genome Lab. Pvt. Ltd., Cochin.

3.2.4 DNA Sequence analysis

The DNA sequences obtained from Scigenom Biotech lab, Cochin were analyzed using BLASTN, BLASTX, CDD, ORF etc. Multiple sequence alignment of full length coding sequence cds was done by BioEdit software.

BLASTN program compares query nucleotide sequences to reference nucleotide sequence from databases and calculates the statistical significance. BLASTX is generally used to find protein coding genes in genomic DNA or to identify proteins encoded by transcripts

ORF finder was used to find out the number of open reading frames present in the coding sequence and to find out the proteins coded by ORFs by doing blast.

Conserved Domain Database (CDD) is a protein annotation resource that contains well-annotated multiple sequence alignment information for ancient domains and full-length proteins. CDD was used for fast identification of conserved domains in protein sequences and provide insights into sequence/structure/function relationships.

3.2.4.1 BioEdit

BioEdit is a biological sequence editor that runs in Windows 95/98/NT/2000/XP and is proposed to provide basic functions for protein and nucleic

acid sequence editing, alignment, manipulation and analysis. It provides many quick and easy functions for sequence editing, annotation and manipulation, as well as a few links of external sequence analysis programs. Sequence alignments of more than 100Mb have been edited on the desktop with good efficiency. BioEdit is spontaneous, menu-driven, and extremely graphical and offers a graphical interface for users to run external analysis programs.

Procedure for BioEdit

Opened the BioEdit software from server and imported the sequences in FASTA format by clicking on file and opened option and highlighted the sequences using the Edit Menu. All selected sequences aligned by using clustalw, a new window was appeared and, selected the first sequence present in new window, and opened feature window to edit query features. Coloring of the particular region was done and visualized by view- graphical features. By using first sequence as template, annotated the other sequences. All annotated sequences selected to remove gaps in the sequence by Sequence gaps and degap options and saved the file in FASTA format and saved the sequence in MS word which can be used for Blast analysis.

3.2.5 Elucidation of carotenoid metabolic pathway for production of capsanthin and capsorubin

Carotenoid metabolic pathway of chilli was downloaded from STRING pathway data base was used for elucidation of pathway. Analysis of pathway was done for the number of genes and enzymes involved.

3.2.5.1 Transcriptome data mining and analysis

In the study, a general search was performed for the retrieval of fruit ESTs from *Capsicum annum* and fifty ESTs were downloaded from NCBI dbEST database. The sequences were subjected to cleaning for the removal of vector contaminants and splice sites using SeqClean and Phrap programme and valid sequences were selected

for analysis. Assembly of the selected EST sequences was done using genome assembler and significant contigs were selected.

3.2.5.2 Metabolic network analysis

3.2.5.2.1 Cytoscape

Cytoscape is an open source community software project used for integrating high troghput expresion data and other molecular state information with biomolecular interaction networks for analysis and visualization. Core software provides central functionality to design and query the network and to visually integrate the network with state data.

It is powerful when utilized in conjugation with large protein-protein, protein -DNA and genetic interaction databases avilable for humans and model organisms. Based on these interactive partners and their confidence scores, protein-protein interaction network for the genes related to carotenoid metabolic pathway were constructed using Cytoscape, an open source platform for the analysis and visualization of complex biomolecular interaction networks with high throughput gene expression data and other molecular states.

3.2.5.3 Procedure for Cytoscape

3.2.5.3.1 Loading of Interaction Data:

Network data prepared in tab-delimited file containing genetic interaction and functional annotations formed the basic data for analysis. Imported the data file into cytoscape screen provided a preview that showed configuration spontaneously and generated network on network view window.

3.2.5.3.2 Network Editing

Network editing was done to build and modify the network interactions within the network. The network was edited by right-clicking on the network window; a new node was added to the network. Edges were added, by right-clicking on a node, then target node was selected and new edge was added between the two nodes. In same way annotation objects like shapes, pictures were added.

3.2.5.3.3 Organize the Network

Organization of network was performed by edge-weighted spring-embedded layout, reorganizing the network such that densely connected nodes are positioned close to each other, whereas disconnected nodes were spread farther apart.

3.2.5.3.4 Adjust the Visual Style of the Network

Adjusted network visualization style and colour by setting background color to black and changed the image edges into white and semitransparent, and made nodes into small, round shape by using control panel VizMapper.

3.2.5.3.5 Generation of Network clusters

The highly interconnected sub graphs for all the networks were analyzed using Molecular Complex Detection (MCODE) module. It is an automated method to find highly interconnected subgraphs as molecular complexes or sub networks from a whole network.

Final visualization of the network was performed manually by using a vector graphic editor such as Adobe Illustrator by exporting the network from Cytoscape into an editable PDF file view as graphics PDF File (*.pdf) and saved the file by specifying name and location.

Results

4. RESULTS

The study on "Analysis of *Capsanthin–capsorubin synthase* gene in Byadagi chilli (*Capsicum annuum* L.) and elucidation of carotenoid metabolic pathway." was carried out at the CPBMB. College of Horticulture, Vellanikkara during the period 2014 – 2016. The results of various aspects of the study are described in this chapter.

The research works included mainly analysis of *Capsanthin - capsorubin synthase* gene in Byadagi chilli (*Capsicum annuum* L.) and elucidation of carotenoid metabolic pathway. Seven genetically distinct chilli varieties /accessions selected based on colour at fully ripe fruit stage formed the study material. The genotypes studied were Byadagi Kaddi, Byadagi Dabbi, Ujwala, Anugraha, Vellayani Samrudhi, Vellayani Thejus, and CC8-1. Two gene specific primers of Ccs gene reported by (Li *et.al*, 2013) was used for PCR amplification. The PCR amplified products were sequenced and analysis of Ccs gene was done. Elucidation of metabolic pathway and metabolic network analysis were also attempted.

4.1 Genomic DNA isolation from different chilli genotypes

4.1.1 Source of DNA

Genomic DNA was isolated from the young tender leaf collected from potted plants of different genotypes maintained at centre for plant Biotechnology and Molecular Biology (CPBMB), College of Horticulture, Vellanikara. As reported young green coloured leaves gave good quality of DNA.

4.1.2 Isolation and purification of DNA

Genomic DNA isolated through the CTAB method developed by Rogers and Bendich (1994) was not pure and had RNA contamination. However RNase treatment after the DNA isolation resulted in good quality of DNA (Plate 2).

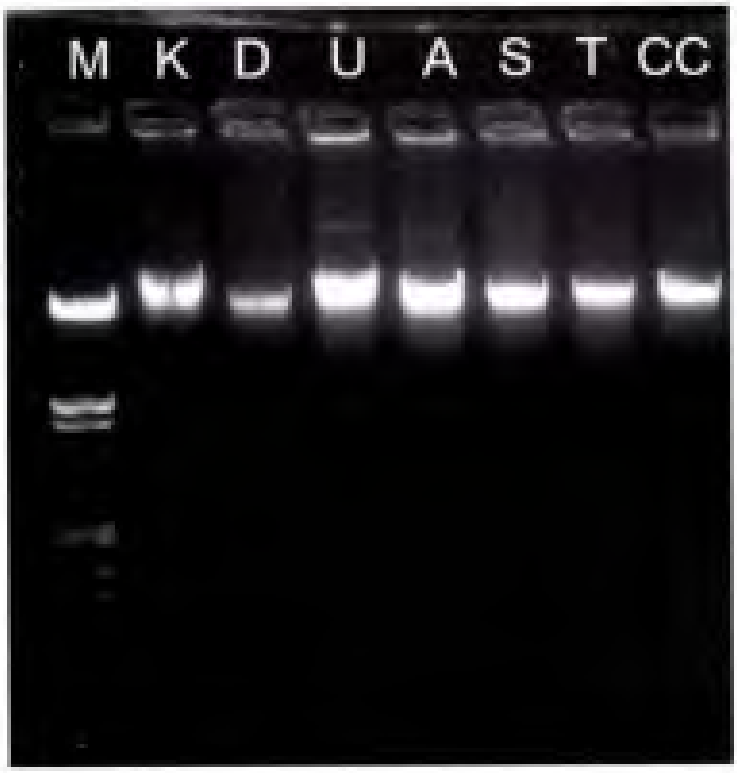


Plate 2. DNA samples from seven genotypes

M- Marker, K - BYADAGI KADDI, D - BYADAGI DABBI, U - UJWALA,
A -ANUGRAHA, S -VELLAYANI SAMRUDHI, T- VELLAYANI THEJUS, CC8-1

4.2 Quantification of DNA

The quality and quantity of isolated DNA were analyzed using electrophoresis and Nanodrop ND – 1000 spectrophotometer. Intact clear bands indicated that DNA was not degraded and was of good quality. The ratio of absorbance of the DNA isolated ranged from 1.5 to 2.1, which indicated that the quality of DNA was good(Table2). Appropriate dilutions of DNA were made and used as template for*Capsanthin capsorubin synthase* gene analysis in different genotypes.

Table 2. Quality and quantity of DNA estimated using NanoDrop spectrometer

Genotypes	Absorbance at 260 nm (A_{260})	Absorbance at 280 nm (A_{280})	Conc (ng/µl)	$OD_{260/280}$
Byadagi Kaddi	86.33	48.77	2213.71	1.82
Byadagi Dabbi	67.12	32.47	2098.11	1.89
Ujwala	73.49	38.87	3126.27	1.85
Anugraha	82.76	47.65	2134.76	1.83
Vellayani samrudhi	51.78	24.54	2986.27	1.91
Vellayani thejus	63.12	30.32	2454.76	1.84
CC 8-1	76.87	41.61	3198.23	1.88

4.3 Analysis of *Capsanthin capsorubin synthase* gene

Two gene specific primers viz., *Ccs cds* and *Ccs Promoter* reported by Li *et al.* (2013) were used for *Capsanthin Capsorubin synthase* (*Ccs)* gene analysis. Amplification was checked using bulked DNA from seven chilli genotypes.

4.3.1 Amplification of Ccs gene in different chilli genotypes using Ccs cds primer

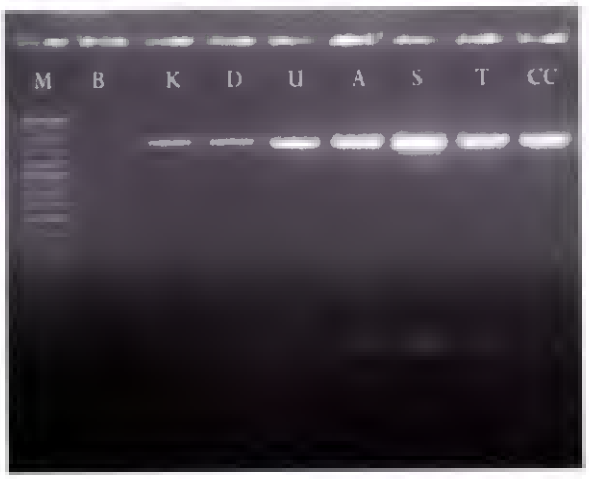
Amplification using *Ccs cds* primer was attempted in seven chilli genotypes to analyze the amplification pattern of *Capsanthin capsorubin synthase* gene in the seven genotypes. Ccs gene was found amplified in all the seven genotypes. Clear distinct band of size 1.5kb was observed in all the genotypes. The Ccs gene also found amplified in the yellow fruited genotype CC8-1. The amplification profile is presented in (Plate 3).

4.3.2 Amplification of Ccs gene in different chilli genotypes using Ccs promoter primer

PCR was performed to analyze the amplification of *Ccs promoter* primer of *Capsanthin capsorubin synthase* gene in the seven chilli genotypes. In all genotypes *Ccs promoter* sequence got amplified. Byadagi kaddi, Byadagi Dabbi, Ujwala, Anugraha, Vellayani Samrudhi produced 920bp as repoted by (Li*et al.*,2013). However vellayani Thejus and CC8-1 produced 1200bp bands. Amplification profile was presented in (Plate4)

4.4 Sequencing of PCR products

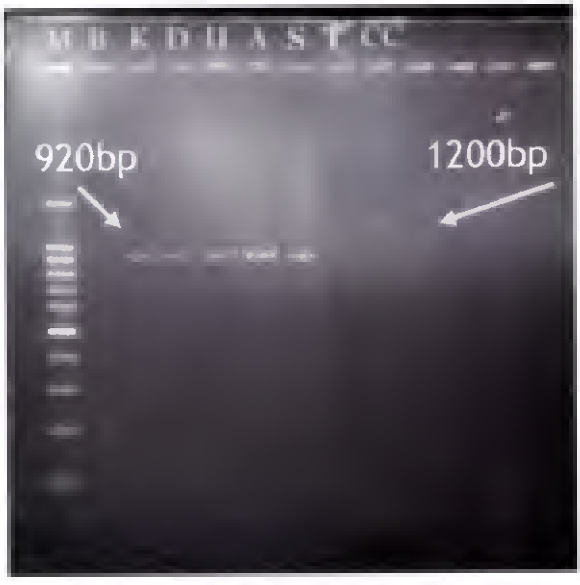
Two sets of PCR products, (1) amplified using *Ccs Cds* primer of seven chilli genotypes, (2) amplified using Ccs promoter primer of seven genotypes confirmed to yield only single band on electrophoresis, were further sent for sequencing. Sequencing was carried out by outsourcing in Sci Genome Lab. Pvt. Ltd., Cochin. The sequences obtained from sci genome shown in (Table3)



M: Marker (100bp plus), (100- 3000bp) B: Blank

K: ByadagiKaddi, D: ByadagiDabbi, U: Ujwala, A: Anugraha, S: VellayaniSamrudhi, T: VellayaniThejus, CC: CC8-1

Plate3. DNA amplification in Chilli genotypes with SSR gene specific primer *Ccs* cds



M: Marker (100bp),(100-1500bp) B: Blank

K: ByadagiKaddi, D: ByadagiDabbi, U: Ujwala, A: Anugraha, S: VellayaniSamrudhi, T: VellayaniThejus, CC: CC8-1

Plate 4. DNA amplification in Chilli genotypes with SSR gene specific primer *Ccs* promoter

Genotypes	Expected Size of Ccs cds amplicon (bp)	Size of the PCR product sequence (bp) with Ccs cds primer
Byadagi Kaddi	1497	1392
Byadagi Dabbi	1497	1410
Ujwala	1497	1383
Anugraha	1497	1404
Vellayani samrudhi	1497	1389
Vellayani thejus	1497	1425
CC 8-1	1497	1416

Table3: Size of the PCR products obtained in different chilli genotypes

The DNA sequences obtained from Scigenom Biotech lab, Cochin were analyzed using BLASTN, BLASTX, CDD, ORF *etc.* Multiple sequence alignment of full length coding sequence was done using clustalX software.

The sequences generated from each genotypes are furnished here under:

1. **Byadagi Kaddi**

ATGGAAACCCTTCTAAAGCCTTTTCCATCTCCTTTACTTTCCATTCCTACTCCTAACATGTATAGTTTCAA
ACACAACTCCACTTTTCCAAATCCAACCAAACAAAAAGATTCAAGAAAGTTCCATTTTAGAAACAAAA
GCAGTACACATTTTTGTAGCTTTCTTGATTTAGCACCCACATCAAAGCCAGAGTCTTTAGATGTTAACAT
CTCATGGGTTGATACTGATCTGGACGGGGCTGAATTCGACGTGATCATCATTGGAACTGGCCCTGCCGG
GCTTCGGCTAGCTGAACAAGTTTCTAAATATGGTATTAAGGTATGTTGCGTTGACCCTTCACCACTTTCC
ATGTGGCCAAATAATTATGGTGTTTGGGTTGATGAGTTTGAAAAGTTGGGATTAGAAGATTGTCTAGAT
CATAAGTGGCCTGTGAGTTGTGTTCATATAAGTGATCACAAGACTAAGTATTTGGACAGACCATATGGT
AGAGTAAGTAGAAAGAAGTTGAAGTTGAAATTGTTGAATAGTTGTGTTGAAAATAGAGTGAAGTTTTA
TAAAGCCAAGGTTTTGAAAGTGAAGCATGAAGAATTTGAGTCTTCGATTGTTTGTGATGATGGTAGGAA
GATAAGCGGTAGCTTGATTGTTGATGCAAGTGGCTATGCTAGTGATTTTATAGAGTATGACAAGCCAAG
AAACCATGGTTATCAAGTTGCTCATGGGATTTTAGCAGAAGTTGATAATCATCCATTTGATTTGGATAA
AATGATGCTTATGGATTGGAGGGATTCTCATTTAGGTAATGAGCCATATCTGAGGGTGAAGAATACTAA
AGAACCAACATTCTTGTATGCAATGCCATTTGATAGGAATTTGGTATTCTTGGAAGAGACTTCTTTAGT
GAGTCGGCCTATGTTATCGTATATGGAAGTGAAAAGAAGGATGGTAGCAAGATTAAGACATTTGGGGA
TCAAAGTGAGAAGTGTCCTTGAGGAAGAGAAGTGTGTGATCACTATGGGAGGACCACTTCCGCGGATT
CCTCAAAATGTTATGGCTATTGGTGGGACTTCAGGGATAGTTCATCCATCGTCTGGGTACATGGTGGCT
CGTAGCATGGCATTGGCACCAGTACTGGCTGAGGCCATCGTCGAAAGCCTTGGCTCAACAAGAATGAT

AAGAGGGTCTCAACTTTACCATAGAGTTTGGAATGGTTTGTGGCCTTCGGATAGAAGACGTGTTAGAGA
ATGTTATTGTTTCGGAATGGAGACTTTGTTAGAGCTTGATTTGGAAGGTACTAGGAGATTGTTTGATGCT
TTCTTTGATGTTGATCCCAAGTACTGGCACGGGTTCCTTTCTTCAAGATTGTCTGTCAAAGAACTTGCTG
TACTCAGTTTG

2. Byadagi Dabbi

ATGGAAACCCTTCTAAAGCCTTTTCCATCTCCTTTACTTTCCATTCCTACTCCTAACATGTATAGTTTCAA
ACACAACTCCACTTTTCCAAATCCAACCAAACAAAAAGATTCAAGAAAGTTCCATTATAGAAACAAAA
GCAGTACACATTTTTGTAGCTTTCTTGATTTAGCACCCACATCAAAGCCAGAGTCTTTAGATGTTAACAT
CTCATGGGTTGATACTGATCTGGACGGGGCTGAATTCGACGTGATCATCATTGGAACTGGCCCTGCCGG
GCTTCGGCTAGCTGAACAAGTTTCTAAATATGGTATTAAGGTATGTTGCGTTGACCCTTCACCACTTTCC
ATGTGGCCAAATAATTATGGTGTTTGGGTTGATGAGTTTGAAAAGTTGGGATTAGAAGATTGTCTAGAT
CATAAGTGGCCTGTGAGTTGTGTTCATATAAGTGATCACAAGACTAAGTATTTGGACAGACCATATGGT
AGAGTAAGTAGAAAGAAGTTGAAGTTGAAATTGTTGAATAGTTGTGTTGAAAATAGAGTGAAGTTTTA
TAAAGCCAAGGTTTTGAAAGTGAAGCATGAAGAATTTGAGTCTTCGATTGTTTGTGATGATGGTAGGAA
GATAAGCGGTAGCTTGATTGTTGATGCAAGTGGCTATGCTAGTGATTTTATAGAGTATGACAAGCCAAG
AAACCATGGTTATCAAGTTGCTCATGGGATTTTAGCAGAAGTTGATAATCATCCATTTGATTTGGATAA
AATGATGCTTATGGATTGGAGGGATTCTCATTTAGGTAATGAGCCATATCTGAGGGTGAAGAATACTAA
AGAACCAACATTCTTGTATGCAATGCCATTTGATAGGAATTTGGTATTCTTGGAAGAGACTTCTTTAGT
GAGTCGGCCTATGTTATCGTATATGGAAGTGAAAAGAAGGATGGTAGCAAGATTAAGACATTTGGGGA
TCAAAGTGAGAAGTGTCCTTGAGGAAGAGAAGTGTGTGATCACTATGGGAGGACCACTTCCGCGGATT
CCTCAAAATGTTATGGCTATTGGTGGGACTTCAGGGATAGTTCATCCATCGTCTGGGTACATGGTGGCT
CGTAGCATGGCATTGGCACCAGTACTGGCTGAGGCCATCGTCGAAAGCCTTGGCTCAACAAGAATGAT
AAGAGGGTCTCAACTTTACCATAGAGTTTGGAATGGTTTGTGGCCTTCGGATAGAAGACGTGTTAGAGA
ATGTTATTGTTTCGGAATGGAGACTTTGTTAGAGCTTGATTTGGAAGGTACTAGGAGATTGTTTGATGCT
TTCTTTGATGTTGATCCCAAGTACTGGCACGGGTTCCTTTCTTCAAGATTGTCTGTCAAAGAACTTGCTG
TACTCAGTTTGTACCTTTTTGGACATGCC

3. Ujwala

ATGGAAACCCTTCTAAAGCCTTTTCCATCTCCTTTACTTTCCATTCCTACTCCTAACATGTATAGTTTCAA
ACACAACTCCACTTTTCCAAATCCAACCAAACAAAAAGATTCAAGAAAGTTCCATTATAGAAACAAAA
GCAGTACACATTTTTGTAGCTTTCTTGATTTAGCACCCACATCAAAGCCAGAGTCTTTAGATGTTAACAT
CTCATGGGTTGATACTGATCTGGACGGGGCTGAATTCGACGTGATCATCATTGGAACTGGCCCTGCCGG
GCTTCGGCTAGCTGAACAAGTTTCTAAATATGGTATTAAGGTATGTTGCGTTGACCCTTCACCACTTTCC
ATGTGGCCAAATAATTATGGTGTTTGGGTTGATGAGTTTGAAAAGTTGGGATTAGAAGATTGTCTAGAT
CATAAGTGGCCTGTGAGTTGTGTTCATATAAGTGATCACAAGACTAAGTATTTGGACAGACCATATGGT
AGAGTAAGTAGAAAGAAGTTGAAGTTGAAATTGTTGAATAGTTGTGTTGAAAATAGAGTGAAGTTTTA
TAAAGCCAAGGTTTTGAAAGTGAAGCATGAAGAATTTGAGTCTTCGATTGTTTGTGATGATGGTAGGAA
GATAAGCGGTAGCTTGATTGTTGATGCAAGTGGCTATGCTAGTGATTTTATAGAGTATGACAAGCCAAG
AAACCATGGTTATCAAGTTGCTCATGGGATTTTAGCAGAAGTTGATAATCATCCATTTGATTTGGATAA
AATGATGCTTATGGATTGGAGGGATTCTCATTTAGGTAATGAGCCATATCTGAGGGTGAAGAATACTAA
AGAACCAACATTCTTGTATGCAATGCCATTTGATAGGAATTTGGTATTCTTGGAAGAGACTTCTTTAGT
GAGTCGGCCTATGTTATCGTATATGGAAGTGAAAAGAAGGATGGTAGCAAGATTAAGACATTTGGGGA
TCAAAGTGAGAAGTGTCCTTGAGGAAGAGAAGTGTGTGATCACTATGGGAGGACCACTTCCGCGGATT

CCTCAAAATGTTATGGCTATTGGTGGGACTTCAGGGATAGTTCATCCATCGTCTGGGTACATGGTGGCT
CGTAGCATGGCATTGGCACCAGTACTGGCTGAGGCCATCGTCGAAAGCCTTGGCTCAACAAGAATGAT
AAGAGGGTCTCAACTTTACCATAGAGTTTGGAATGGTTTGTGGCCTTCGGATAGAAGACGTGTTAGAGA
ATGTTATTGTTTCGGAATGGAGACTTTGTTGAAGCTTGATTTGGAAGGTACTAGGAGATTGTTTGATGCT
TTCTTTGATGTTGATCCCAAGTACTGGCACGGGTTCCTTTCTTCAAGATTGTCTGTCAAAGAACTTGCTG
TA

4. Anugraha

ATGGAAACCCTTCTAAAGCCTTTTCCATCTCCTTTACTTTCCATTCCTACTCCTAACATGTATAGTTTCAA
ACACAACTCCACTTTTCCAAATCCAACCAAACAAAAAGATTCAAGAAAGTTCCATTATAGAAACAAAA
GCAGTACACATTTTTGTAGCTTTCTTGATTTAGCACCCACATCAAAGCCAGAGTCTTTAGATGTTAACAT
CTCATGGGTTGATACTGATCTGGACGGGGCTGAATTCGACGTGATCATCATTGGAACTGGCCCTGCCGG
GCTTCGGCTAGCTGAACAAGTTTCTAAATATGGTATTAAGGTATGTTGCGTTGACCCTTCACCACTTTCC
ATGTGGCCAAATAATTATGGTGTTTGGGTTGATGAGTTTGAAAAGTTGGGATTAGAAGATTGTCTAGAT
CATAAGTGGCCTGTGAGTTGTGTTCATATAAGTGATCACAAGACTAAGTATTTGGACAGACCATATGGT
AGAGTAAGTAGAAAGAAGTTGAAGTTGAAATTGTTGAATAGTTGTGTTGAAAATAGAGTGAAGTTTTA
TAAAGCCAAGGTTTTGAAAGTGAAGCATGAAGAATTTGAGTCTTCGATTGTTTGTGATGATGGTAGGAA
GATAAGCGGTAGCTTGATTGTTGATGCAAGTGGCTATGCTAGTGATTTTATAGAGTATGACAAGCCAAG
AAACCATGGTTATCAAGTTGCTCATGGGATTTTAGCAGAAGTTGATAATCATCCATTTGATTTGGATAA
AATGATGCTTATGGATTGGAGGGATTCTCATTTAGGTAATGAGCCATATCTGAGGGTGAAGAATACTAA
AGAACCAACATTCTTGTATGCAATGCCATTTGATAGGAATTTGGTATTCTTGGAAGAGACTTCTTTAGT
GAGTCGGCCTATGTTATCGTATATGGAAGTGAAAAGAAGGATGGTAGCAAGATTAAGACATTTGGGGA
TCAAAGTGAGAAGTGTCCTTGAGGAAGAGAAGTGTGTGATCACTATGGGAGGACCACTTCCGCGGATT
CCTCAAAATGTTATGGCTATTGGTGGGACTTCAGGGATAGTTCATCCATCGTCTGGGTACATGGTGGCT
CGTAGCATGGCATTGGCACCAGTACTGGCTGAGGCCATCGTCGAAAGCCTTGGCTCAACAAGAATGAT
AAGAGGGTCTCAACTTTACCATAGAGTTTGGAATGGTTTGTGGCCTTCGGATAGAAGACGTGTTAGAGA
ATGTTATTGTTTCGGAATGGAGACTTTGTTGAAGCTTGATTTGGAAGGTACTAGGAGATTGTTTGATGCT
TTCTTTGATGTTGATCCCAAGTACTGGCACGGGTTCCTTTCTTCAAGATTGTCTGTCAAAGAACTTGCTG
TACTCAGTTTGTACCTTTTTGGA

5. Vellayani Samrudhi

ATGGAAACCCTTCTAAAGCCTTTTCCATCTCCTTTACTTTCCATTCCTACTCCTAACATGTATAGTTTCAA
ACACAACTCCACTTTTCCAAATCCAACCAAACAAAGAGATTCAAGAAAGTTCCATTCTAGAAACAAAA
GCAGTACACATTTTTGTAGCTTTCTTGATTTAGCACCCACATCAAAGCCAGAGTCTTTAGATGTTAACAT
CTCATGGGTTGATACTGATCTGGACCGGGCTGAATTCGACGTGATCATCATTGGAACTGGCCCTGCCGG
GCTTCGGCTAGCTGAACAAGTTTCTAAATATGGTATTAAGGTATGTTGCGTTGACCCTTCACCACTTTCC
ATGTGGCCAAATAATTATGGTGTTTGGGTTGATGAGTTTGAAAAGTTGGGATTAGAAGATTGTCTAGAT
CATAAGTGGCCTGTGAGTTGTGTTCATATAAGTGATCACAAGACTAAGTATTTGGACAGACCATATGGT
AGAGTAAGTAGAAAGAAGTTGAAGTTGAAATTGTTGAATAGTTGTGTTGAAAATAGAGTGAAGTTTTA
TAAAGCCAAGGTTTTGAAAGTGAAGCATGAAGAATTTGAGTCTTCGATTGTTTGTGATGATGGTAGGAA
GATAAGTGGTAGCTTGATTGTTGATGCAAGTGGCTATGCTAGTGATTTTATAGAGTATGACAAGCCAAG
AAACCATGGTTATCAAGTTGCTCATGGGATTTTAGCAGAAGTTGATAATCATCCATTTGATTTGGATAA
AATGATGCTTATGGATTGGAGGGATTCTCATTTAGGTAATGAGCCATATCTGAGGGTGAAGAATACTAA
GAA~CCAACATTCTTGTATGCAATGCCATTTGATAGGAATTTGGTATTCTTGGAAGAGACTTCTTTAGTG

AGTCGGCCTATGTTATCGTATATGGAAGTGAAAAGAAGGATGGTAGCAAGATTAAGACATTTGGGGAT
CAAAGTGAGAAGTGTCCTTGAGGAAGAGAAGTGTGTGATCACTATGGGAGGACCACTTCCGCGGATTC
CTCAAATGTTATGGCTAT~GGTGGGACTTCAGGATAGTTCATCCATCGTCTGGGTACATGGTGGCTCGT
AGCATGGCATTGGCACCAGTACTGGCTGAGGCCATCGTCGAAAGCCTTGGCTCAACAAGAATGATAAG
AGGGTCTCAACTTTACCATAGAGTTTGGAATGGTTTGTGGCCTTCGGATAGAAGACGTGTTAGAGAATG
TTATTGTTTCGGAATGGAGACTTTGTTGAAGCTTGATTTGGAAGGTACTAGGAGATTGTTTGATGCTTTC
TTTGATGTTGATCCCAAGTACTGGCACGGGTTCCTTTCTTCAAGATTGTCTGTCAAAGAACTTGCTGTAC
TCAGT

6. Vellayani Thejus

ATGGAAACCCTTCTAAAGCCTTTTCCATCTCCTTTACTTTCCATTCCTACTCCTAACATGTATAGTTTCAA
ACACAACTCCACTTTTCCAAATCCAACCAAACAAAGAGATTCAAGAAAGTTCCATTCTAGAAACAAAA
GCAGTACACATTTTTGTAGCTTTCTTGATTTAGCACCCACATCAAAGCCAGAGTCTTTAGATGTTAACAT
CTCATGGGTTGATACTGATCTGGACCGGGCTGAATTCGACGTGATCATCATTGGAACTGGCCCTGCCGG
GCTTCGGCTAGCTGAACAAGTTTCTAAATATGGAATTAAGGTATGTTGCGTTGACCCTTCACCACTTTCC
ATGTGGCCAAATAATTATGGTGTTTGGGTTGATGAGTTTGAAAAGTTGGGATTAGAAGATTGTCTAGAT
CATAAGTGGCCTGTGAGTTGTGTTCATATAAGTGATCACAAGACTAAGTATTTGGACAGACCATATGGT
AGAGTAAGTAGAAAGAAGTTGAAGTTGAAATTGTTGAATAGTTGTGTTGAAAATAGAGTGAAGTTTTA
TAAAGCCAAGGTTTTGAAAGTGAAGCATGAAGAATTTGAGTCTTCGATTGTTTGTGATGATGGTAGGAA
GATAAGTGGTAGCTTGATTGTTGATGCAAGTGGCTATGCTAGTGATTTTATAGAGTATGACAAGCCAAG
AAACCATGGTTATCAAGTTGCTCATGGGATTTTAGCAGAAGTTGATAATCATCCATTTGATTTGGATAA
AATGATGCTTATGGATTGGAGGGATTCTCATTTAGGTAATGAGCCATATCTGAGGGTGAAGAATACTAA
AGAACCAACATTCTTGTATGCAATGCCATTTGATAGGAATTTGGTATTCTTGGAAGAGACTTCTTTAGT
GAGTCGGCCTATGTTATCGTATATGGAAGTGAAAAGAAGGATGGTAGCAAGATTAAGACATTTGGGGA
TCAAAGTGAGAAGTGTCCTTGAGGAAGAGAAGTGTGTGATCACTATGGGAGGACCACTTCCGCGGATT
CCTCAAAATGTTATGGCTATTGGTGGGACTTCAGGGATAGTTCATCCATCGTCTGGGTACATGGTGGCT
CGTAGCATGGCATTGGCACCAGTACTGGCTGAGGCCATCGTCGAAAGCCTTGGCTCAACAAGAATGAT
AAGAGGGTCTCAACTTTACCATAGAGTTTGGAATGGTTTGTGGCCTTCGGATAGAAGACGTGTTAGAGA
ATGTTATTGTTTCGGAATGGAGACTTTGTTGAAGCTTGATTTGGAAGGTACTAGGAGATTGTTTGATGCT
TTCTTTGATGTTGATCCCAAGTACTGGCACGGGTTCCTTTCTTCAAGATTGTCTGTCAAAGAACTTGCTG
TACTCAGTTTGTACTTTTTGGACATGCCTCTAATTTGGCTAGG

7. CC8-1

ATGGAAACCCTTCTAAAGCCTTTTTCATCTCCTTTACTTTCCATTCCTACTCCTAACATGTATAGTTTCAA
ACACAACTCCACTTTTCCAAATCCAACCAAACAAAGAGATTCAAGAAAGTTCCATTCTAGAAACAAAA
GCAGTACACATTTTTGTAGCTTTCTTGATTTAGCACCCACATCAAAGCCAGAGTCTTTAGATGTTAACAT
CTCATGGGTTGATACTGATCTGGACCGGGCTGAATTCGACGTGATCATCATTGGAACTGGCCCTGCCGG
GCTTCGGCTAGCTGAACAAGTTTCTAAATATGGAATTAAGGTATGTTGCGTTGACCCTTCACCACTTTCC
ATGTGGCCAAATAATTATGGTGTTTGGGTTGATGAGTTTGAAAAGTTGGGATTAGAAGATTGTCTAGAT
CATAAGTGGCCTGTGAGTTGTGTTCATATAAGTGATCACAAGACTAAGTATTTGGACAGACCATATGGT
AGAGTAAGTAGAAAGAAGTTGAAGTTGAAATTGTTGAATAGTTGTGTTGAAAATAGAGTGAAGTTTTA
TAAAGCCAAGGTTTTGAAAGTGAAGCATGAAGAATTTGAGTCTTAGATTGTTTGTGATGATGGTAGGAA
GATAAGTGGTAGCTTGATTGTTGATGCAAGTGGCTATGCTAGTGATTTTATAGAGTATGACAAGCCAAG

AAACCATGGTTATCAAGTTGCTCATGGGATTTTAGCAGAAGTTGATAATCATCCATTTGATTTGGATAA
AATGATGCTTATGGATTGGAGGGATTCTCATTTAGGTAATGAGCCATATCTGAGGGTGAAGAATACTAA
AGAACCAACATTCTTGTATGCAATGCCATTTGATAGGAATTTGGTATTCTTGGAAGAGACTTCTTTAGT
GAGTCGGCCTATGTTATCGTATATGGAAGTGAAAAGAAGGATGGTAGCAAGATTAAGACATTTGGGGA
TCAAAGTGAGAAGTGTCCTTGAGGAAGAGAAGTGTGTGATCACTATGGGAGGACCACTTCCGCGGATT
CCTCAAAATGTTATGGCTATTGGTGGGACTTCAGGGATAGTTCATCCATCGTCTGGGTACATGGTGGCT
CGTAGCATGGCATTGGCACCAGTACTGGCTGAGGCCATCGTCGAAAGCCTTGGCTCAACAAGAATGAT
AAGAGGGTCTCAACTTTACCATAGAGTTTGGAATGGTTTGTGGCCTTCGGATAGAAGACGTGTTAGAGA
ATGTTATTGTTTCGGAATGGAGACTTTGTTGAAGCTTGATTTGGAAGGTACTAGGAGATTGTTTGATGCT
TTCTTTGATGTTGATCCCAAGTACTGGCACGGGTTCCTTTCTTCAAGATTGTCTGTCAAAGAACTTGCTG
TACTCAGTTTGTACCTTTTTGGACATGCCTCTAAT

4.4.1 Analysis of Ccs gene sequence using BLASTn

The Blastn result of Ccs Cds gene sequences briefed in (Table4)

4.4.1.1 BLASTn analysis of sequence of Byadagi Kaddi

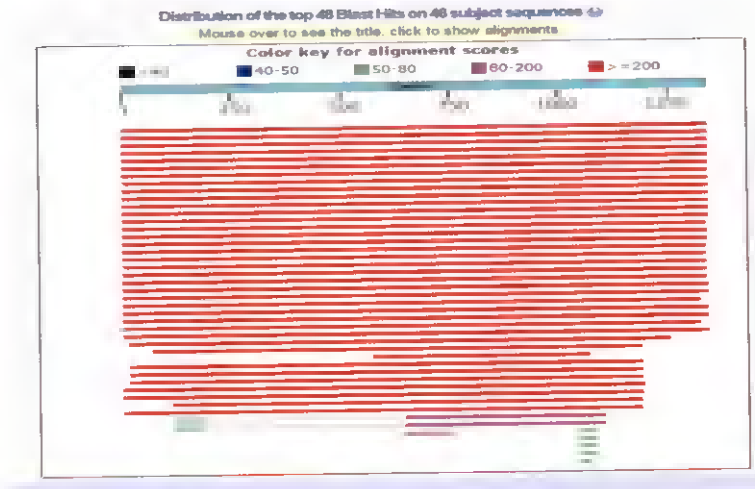
The sequences of Byadagi kaddi was analysed using BLASTn programme. The result showed that the sequence of Byadagi kaddi showed 99 per cent similarity with Ccs gene accession KM037690.1 (Fig1a) a *Capsicum annuum* cultivar landrace velarde, *Capsanthin capsorubin synthase* gene C-allele, complete cds.

4.4.1.2 BLASTn analysis of sequence of Byadagi Dabbi

The sequences of Byadagi dabbi was analysed using BLASTn programme. The result showed that the sequence of Byadagi dabbi showed 99 per cent similarity with Ccs gene accession KM037690.1 (Fig1b) a *Capsicum annuum* cultivar landrace velarde, Capsanthin capsorubin synthase gene C-allele, complete cds.

4.4.1.3 BLASTn analysis of sequence of Ujwala

The sequence of Ujwala was analysed using BLASTn programme. The result showed that the sequence of Ujwala showed 100 per cent similarity with Ccs gene accession KM037687.1 (Fig1c) a *Capsicum annuum* cultivar landrace Jemez, Capsanthin capsorubin synthase (Ccs) gene complete cds.



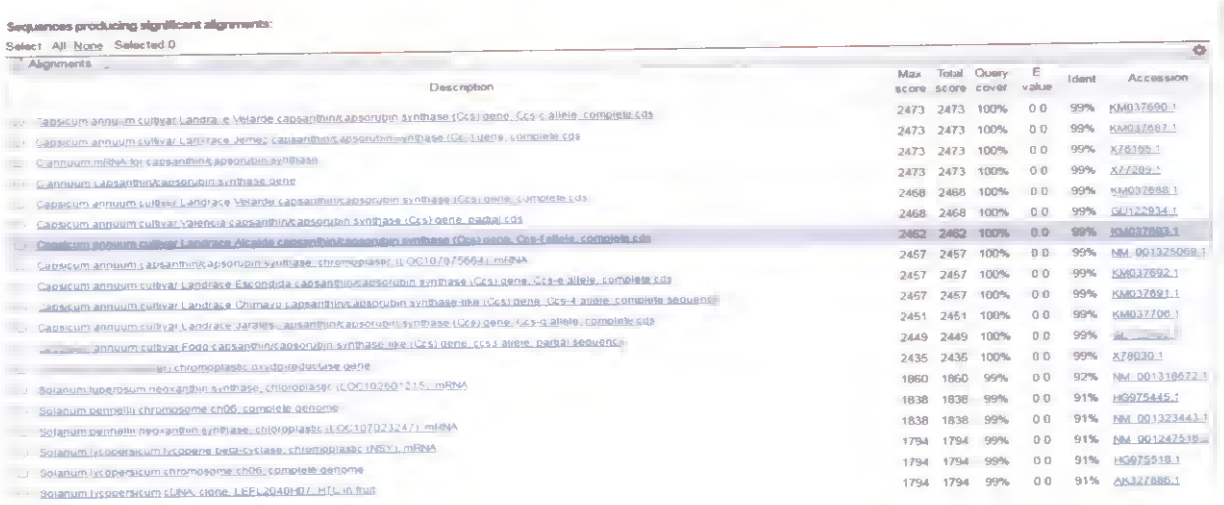


Fig 1a: BLASTn of Byadagi Kaddi *Ccs* gene sequence showing 99 per cent similarity with Ccs gene accession KM037690.1

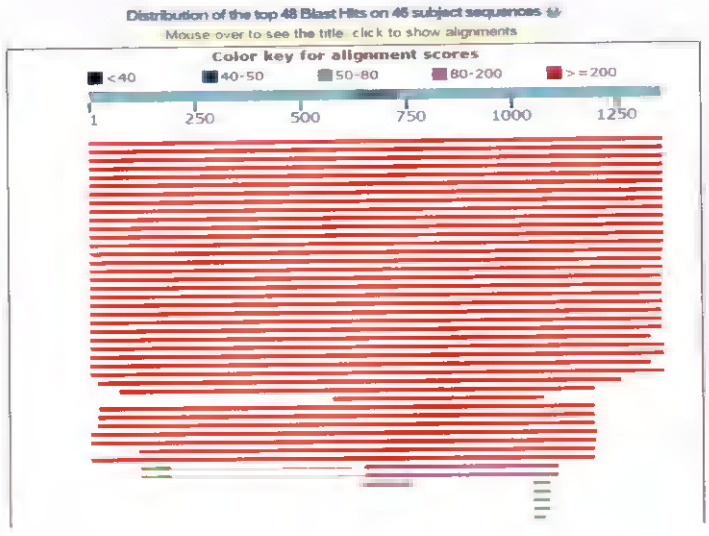
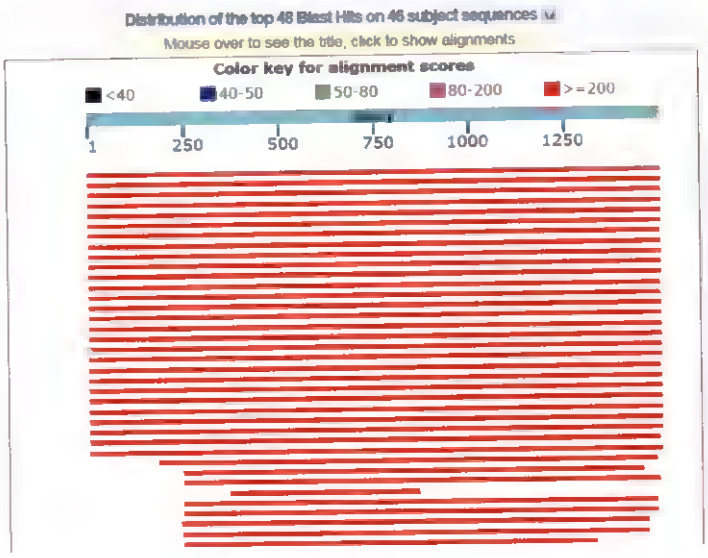


Fig 1b: BLASTn of Byadagi Dabbi *Ccs* gene sequence showing 99 per cent similarity with Ccs gene accession KM037690.1



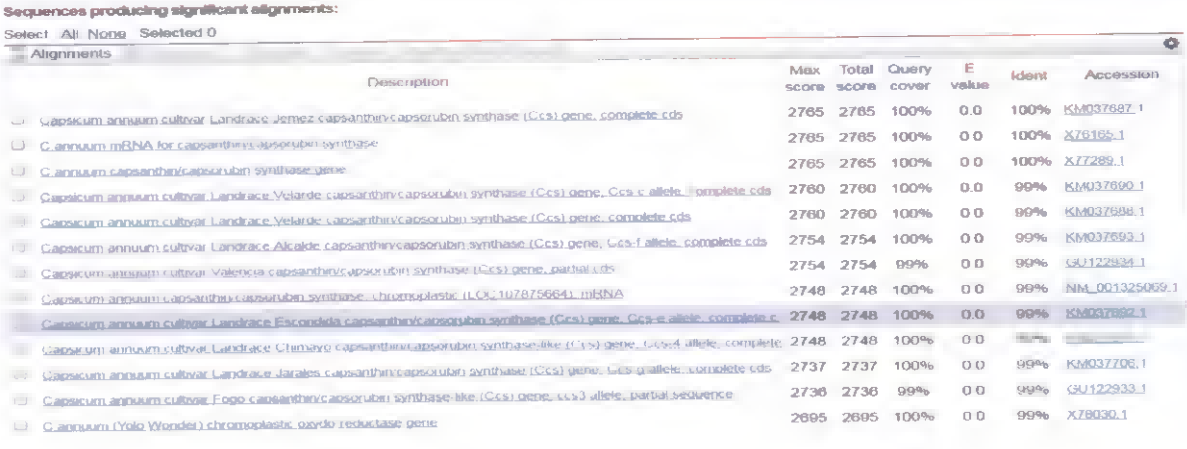


Fig 1c: BLASTn of Ujwala Ccs gene sequence showing 100 per cent similarity with Ccs gene accession KM037687.1

Genotypes	Expected Size of Ccs gene amplicon	Size of the amplified sequence obtained with Ccs cds primer	Similarity %	Query coverage	e- value	Accession showing highest similarity
Byadagi Kaddi	1497	1392	99	100	0.0	KM037690.1
Byadagi Dabbi	1497	1410	100	100	0.0	KM037690.1
Ujwala	1497	1383	100	100	0.0	KM037687.1
Anugraha	1497	1404	100	100	0.0	KM037687.1
Vellayani Samrudhi	1497	1389	99	100	0.0	KM037693.1
Vellayani Thejus	1497	1425	99	100	0.0	KM037693.1
CC 8-1	1497	1416	99	100	0.0	KM037693.1

Table4: BLASTn analysis of *Ccs cds* gene sequence in different chilli genotypes

4.4.1.4 BLASTn analysis of sequence of Anugraha

The sequences of Anugraha was analysed using BLASTn programme. The result showed that the sequence of Anugraha showed 100 per cent similarity with Ccs gene accession KM037687.1 (Fig1d) a *Capsicum annuum* cultivar landrace alcalde, *Capsanthin capsorubin synthase* gene f-allele, complete cds.

4.4.1.5 BLASTn analysis of sequence of Vellayani Samrudhi

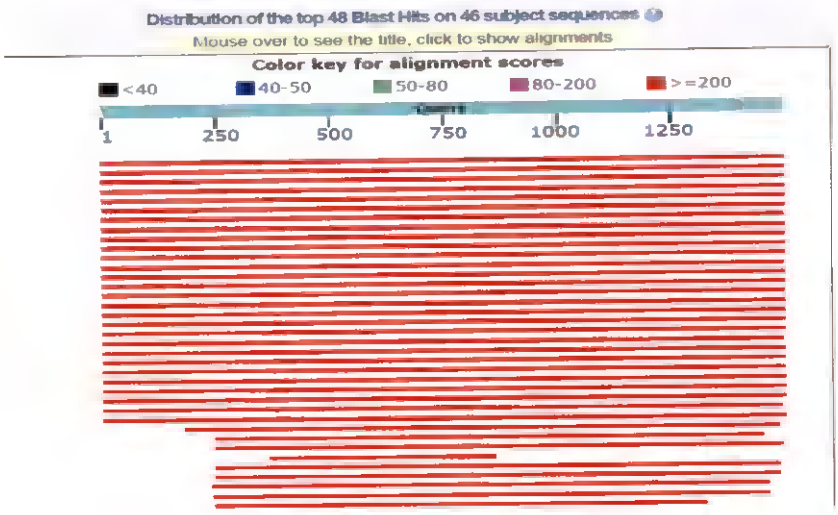
The sequence of Vellayani Samrudhi was analysed using BLASTn programme. The result showed that the sequence of Vellayani Samrudhi had shown 99 per cent similarity with Ccs gene accession KM037693.1 (Fig1e) a *Capsicum annuum* cultivar landrace alcalde, Capsanthin capsorubin synthase gene f-allele, complete cds.

4.4.1.6 BLASTn analysis of sequence of Vellayani Thejus

The sequences of Vellayani Thejus were analysed using BLASTn programme. The result showed that the sequence of Vellayani Thejus showed 99 per cent similarity with Ccs gene accession KM037693.1 (Fig1f) a *Capsicum annuum* cultivar landrace alcalde, Capsanthin capsorubin synthase gene f-allele, complete cds.

4.4.1.7 BLASTn analysis of sequence of CC8-1

The sequences of CC8-1 were analysed using BLASTn programme. The result showed that the sequence of CC8-1 showed 99 per cent similarity with Ccs gene accession KM037693.1 (Fig1g) a *Capsicum annuum* cultivar landrace alcalde, Capsanthin capsorubin synthase gene f-allele, complete cds.



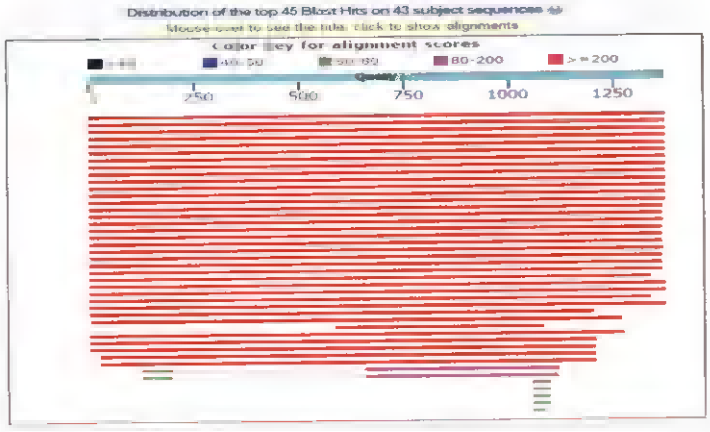
Sequences producing significant alignments:

Select: All None Selected:0

Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
Capsicum annuum cultivar Landrace Jerez capsanthin/capsorubin synthase (Ccs) gene, complete cds	2765	2765	100%	0.0	100%	KM037687.1
C.annuum mRNA for capsanthin/capsorubin synthase	2765	2765	100%	0.0	100%	X76165.1
C.annuum capsanthin/capsorubin synthase gene	2765	2765	100%	0.0	100%	X77289.1
Capsicum annuum cultivar Landrace Velarde capsanthin/capsorubin synthase (Ccs) gene, Ccs-c allele, complete cds	2760	2760	100%	0.0	99%	KM037690.1
Capsicum annuum cultivar Landrace Velarde capsanthin/capsorubin synthase (Ccs) gene, complete cds	2760	2760	100%	0.0	99%	KM037688.1
Capsicum annuum cultivar Landrace Alcalde capsanthin/capsorubin synthase (Ccs) gene, Ccs-f allele, complete cds	2754	2754	100%	0.0	99%	KM037693.1
Capsicum annuum cultivar Valencia capsanthin/capsorubin synthase (Ccs) gene, partial cds	2754	2754	99%	0.0	99%	GU122934.1
Capsicum annuum capsanthin/capsorubin synthase, chromoplastic (LOC107875664), mRNA	2748	2748	100%	0.0	99%	NM_001325069.1
Capsicum annuum cultivar Landrace Escondida capsanthin/capsorubin synthase (Ccs) gene, Ccs-e allele, complete cc	2748	2748	100%	0.0	99%	KM037692.1
Capsicum annuum cultivar Landrace Chimayo capsanthin/capsorubin synthase-like (Ccs) gene, Ccs-4 allele, complete	2748	2748	100%	0.0	99%	KM037691.1
Capsicum annuum cultivar Landrace Jarales capsanthin/capsorubin synthase (Ccs) gene, Ccs-g allele, complete cds	2737	2737	100%	0.0	99%	KM037708.1
Capsicum annuum cultivar Fogo capsanthin/capsorubin synthase-like (Ccs) gene, ccs3 allele, partial sequence	2736	2736	99%	0.0	99%	GU122933.1
C.annuum (Yolo Wonder) chromoplastic oxydo-reductase gene	2695	2695	100%	0.0	99%	X78030.1

Fig 1d: BLASTn of Anugraha Ccs gene sequence showing 100 per cent similarity with Ccs gene accession KM037687.1



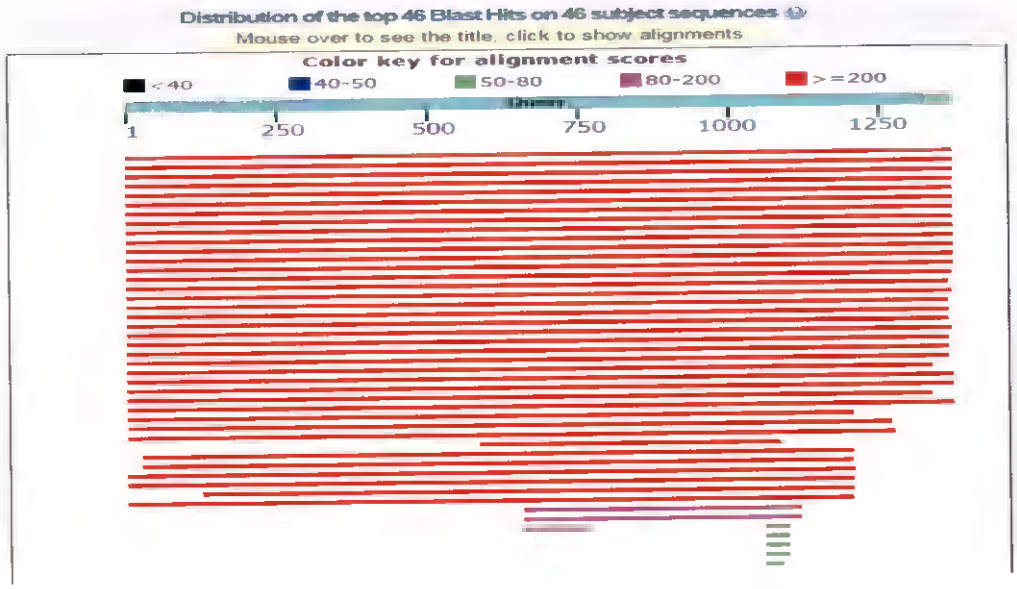
Sequences producing significant alignments

Select All None Selected 0

Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
Capsicum annuum cultivar Landrace Alcalde capsanthin/capsorubin synthase (Ccs) gene, Ccs-f allele, complete cds	2508	2508	100%	0.0	99%	KM037693.1
Capsicum annuum capsanthin/capsorubin synthase, chromoplastic (LOC107875664), mRNA	2503	2503	100%	0.0	99%	NM_001325069.1
Capsicum annuum cultivar Landrace Escondida capsanthin/capsorubin synthase (Ccs) gene, Ccs-e allele, complete cds	2503	2503	100%	0.0	99%	KM037692.1
Capsicum annuum cultivar Landrace Chimayo capsanthin/capsorubin synthase-like (Ccs) gene, Ccs-4 allele, complete sequence	2503	2503	100%	0.0	99%	KM037691.1
Capsicum annuum cultivar Landrace Velarde capsanthin/capsorubin synthase (Ccs) gene, complete cds	2503	2503	100%	0.0	99%	KM037688.1
Capsicum annuum cultivar Landrace Jarales capsanthin/capsorubin synthase (Ccs) gene, Ccs-g allele, complete cds	2497	2497	100%	0.0	99%	KM037706.1
Capsicum annuum cultivar Landrace Velarde capsanthin/capsorubin synthase (Ccs) gene, Ccs-c allele, complete cds	2497	2497	100%	0.0	99%	KM037690.1
Capsicum annuum cultivar Landrace Jemez capsanthin/capsorubin synthase (Ccs) gene, complete cds	2497	2497	100%	0.0	99%	KM037687.1
C.annuum mRNA for capsanthin/capsorubin synthase	2497	2497	100%	0.0	99%	X76365.1
C.annuum capsanthin/capsorubin synthase gene	2497	2497	100%	0.0	99%	X77289.1
Capsicum annuum cultivar Valencia capsanthin/capsorubin synthase (Ccs) gene, partial cds	2492	2492	100%	0.0	99%	GU122934.1
Capsicum annuum cultivar Fogo capsanthin/capsorubin synthase-like (Ccs) gene, ccs3 allele, partial sequence	2473	2473	100%	0.0	99%	GU122933.1
[redacted] oid Monger) chromoplastic oxydo-reductase gene	2446	2446	100%	0.0	99%	X78030.1
Solanum tuberosum neoxanthin synthase, chloroplastic (LOC102601215), mRNA	1862	1862	99%	0.0	91%	NM_001318672.1
Solanum pennellii chromosome ch06, complete genome	1851	1851	99%	0.0	91%	HG975445.1
Solanum pennellii neoxanthin synthase, chloroplastic (LOC107023247), mRNA	1851	1851	99%	0.0	91%	NM_001323443.1

Fig 1e: BLASTn of Vellayani Samrudhi Ccs gene sequence showing 99 per cent similarity with Ccs gene accession KM037693.1



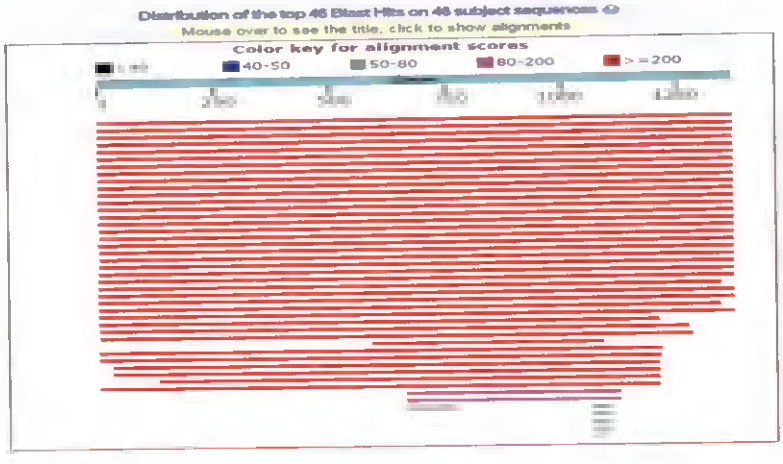
Sequences producing significant alignments:

Select All None Selected 0

Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
Capsicum annuum cultivar Landrace Alcalde [illegible] capsorubin synthase (Ccs) gene, Ccs-f allele, complete cds	2495	2495	100%	0.0	99%	KM037693.1
Capsicum annuum capsanthin/capsorubin synthase, chromoplastic (LOC107875664), mRNA	2490	2490	100%	0.0	99%	NM_001325069.1
Capsicum annuum cultivar Landrace Escondida capsanthin/capsorubin synthase (Ccs) gene, Ccs-e allele, complete cds.	2490	2490	100%	0.0	99%	KM037692.1
Capsicum annuum cultivar Landrace Chimayo capsanthin/capsorubin synthase-like (Ccs) gene, Ccs-4 allele, complete sequence	2490	2490	100%	0.0	99%	KM037691.1
Capsicum annuum cultivar Landrace Velarde capsanthin/capsorubin synthase (Ccs) gene, complete cds	2490	2490	100%	0.0	99%	KM037688.1
Capsicum annuum cultivar Landrace Jarales capsanthin/capsorubin synthase (Ccs) gene, Ccs-g allele, complete cds	2484	2484	100%	0.0	99%	KM037706.1
Capsicum annuum cultivar Landrace Velarde capsanthin/capsorubin synthase (Ccs) gene, Ccs-c allele, complete cds.	2484	2484	100%	0.0	99%	KM037690.1
Capsicum annuum cultivar Landrace Jemez capsanthin/capsorubin synthase (Ccs) gene, complete cds.	2484	2484	100%	0.0	99%	KM037687.1
C.annuum mRNA for capsanthin/capsorubin synthase	2484	2484	100%	0.0	99%	X76165.1
C.annuum capsanthin/capsorubin synthase gene	2484	2484	100%	0.0	99%	X77289.1
Capsicum annuum cultivar Valencia capsanthin/capsorubin synthase (Ccs) gene, partial cds	2479	2479	100%	0.0	99%	GU122934.1
Capsicum annuum cultivar Fogo capsanthin/capsorubin synthase-like (Ccs) gene, ccs3 allele, partial sequence	2460	2460	100%	0.0	99%	GU122933.1
C.annuum (Yolo Wonder) chromoplastic oxydo-reductase gene	2433	2433	100%	0.0	99%	X78030.1
Solanum tuberosum neoxanthin synthase, chromoplastic (LOC102601215), mRNA	1857	1857	99%	0.0	91%	NM_001318572.1
Solanum pennellii chromosome ch06, complete genome	1845	1845	99%	0.0	91%	HG975445.1

Fig 1f: BLASTn of Vellayani Thejus Ccs gene sequence showing 99 per cent similarity with Ccs gene accession KM037693.1



Sequences producing significant alignments

Select All Selected 0

Alignments

Description	score	score	Query cover	E value	Ident	Accession
Capsicum annuum cultivar Landrace Arcalde capsanthin/capsorubin synthase (Ccs) gene, Ccs-f allele, complete cds	2486	2486	100%	0.0	99%	KM037693.1
Capsicum annuum capsanthin/capsorubin synthase, chromoplastic (LOC107875664), mRNA	2481	2481	100%	0.0	99%	NM_001325069.1
Capsicum annuum cultivar Landrace Escondida capsanthin/capsorubin synthase (Ccs) gene, Ccs-e allele, complete cds	2481	2481	100%	0.0	99%	KM037692.1
Capsicum annuum cultivar Landrace Chimayo capsanthin/capsorubin synthase-like (Ccs) gene, Ccs-4 allele, complete sequence	2481	2481	100%	0.0	99%	KM037691.1
Capsicum annuum cultivar Landrace Velarde capsanthin/capsorubin synthase (Ccs) gene, complete cds	2481	2481	100%	0.0	99%	KM037688.1
Capsicum annuum cultivar Landrace Jaralas capsanthin/capsorubin synthase (Ccs) gene, Ccs-g allele, complete cds	2475	2475	100%	0.0	99%	KM037706.1
Capsicum annuum cultivar Landrace Velarde capsanthin/capsorubin synthase (Ccs) gene, Ccs-c allele, complete cds	2475	2475	100%	0.0	99%	KM037690.1
Capsicum annuum cultivar Landrace Jemez capsanthin/capsorubin synthase (Ccs) gene, complete cd	2475	2475	100%	0.0	99%	KM037687.1
C.annuum mRNA for capsanthin/capsorubin synthase	2475	2475	100%	0.0	99%	X76165.1
C.annuum capsanthin/capsorubin synthase gene	2475	2475	100%	0.0	99%	X77289.1
Capsicum annuum cultivar Valencia capsanthin/capsorubin synthase (Ccs) gene, partial cds	2470	2470	100%	0.0	99%	GU122934.1
Capsicum annuum cultivar Fogo capsanthin/capsorubin synthase-like (Ccs) gene, ccs3 allele, partial sequence	2451	2451	100%	0.0	99%	GU122933.1
C.annuum (Yolo Wonder) chromoplastic oxydo-reductase gene	2418	2418	100%	0.0	99%	X78030.1
[illegible]num tuberosum neoxanthin synthase, chloroplastic (LOC102601215), mRNA	1853	1853	99%	0.0	91%	NM_001318672.1
Solanum pennellii chromosome ch06, complete genome	1842	1842	99%	0.0	91%	HG975445.1
Solanum pennellii neoxanthin synthase, chloroplastic (LOC107023247), mRNA	1842	1842	99%	0.0	91%	NM_001323443.1

Fig 1g: BLASTn of CC8-1 Ccs gene sequence showing 99 per cent similarity with Ccs gene accession KM037693.1

4.4.2.1 Analysis of Ccs gene sequence using BLASTX

The Blastx result of Ccs gene sequences briefed in (Table5)

4.4.2.1.1 BLASTX analysis of sequence of Byadagi Kaddi

The sequences of Byadagi kaddi was analysed using BLASTX programme. The result showed that the sequence of Byadagi kaddi had shown 99 per cent similarity with amino acid sequence of Ccs gene in the accession Q42435.1 shown in (Fig2a) a Capsanthin/capsorubin synthase, chromoplastic enzyme catalyzes 5,6-epoxycarotenoids antheraxanthin and violaxanthin to form capsanthin and capsorubin.

4.4.2.1.2 BLASTX analysis of sequence of Byadagi Dabbi

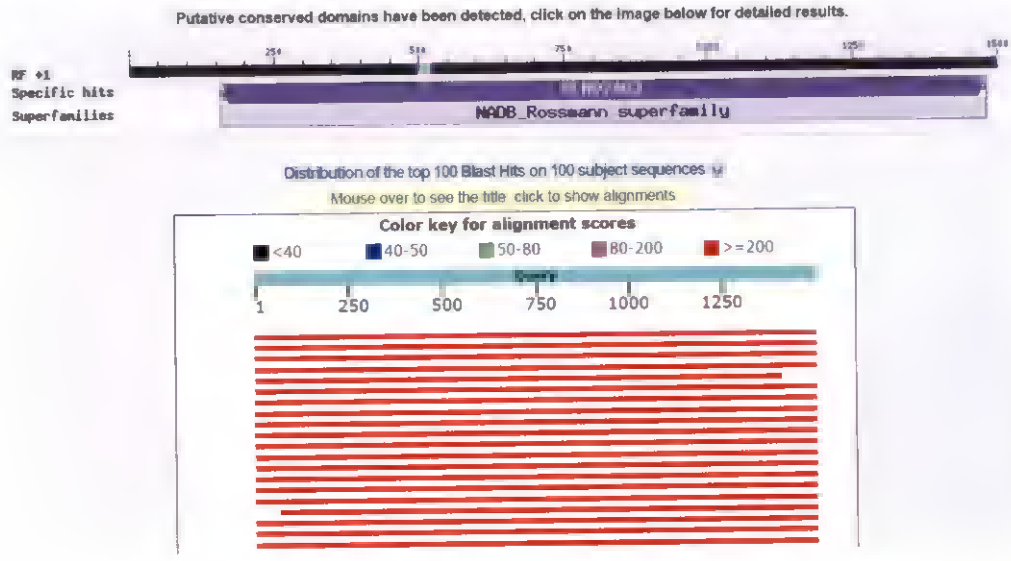
The sequences of Byadagi Dabbi were analysed using BLASTX programme. The result showed that the sequence of Byadagi Dabbi had shown 99 per cent similarity with amino acid sequence of Ccs gene in the accession Q42435.1 shown in (Fig2b) a Capsanthin capsorubin synthase, chromoplastic enzyme catalyzes 5,6-epoxycarotenoids antheraxanthin and violaxanthin to form capsanthin and capsorubin.

4.4.2.1.3 BLASTX analysis of sequence of Ujwala

The sequence of variety Ujwala was analysed using BLASTX programme. The result showed that the sequence of Ujwala had shown 99 per cent similarity with amino acid sequence of Ccs gene in the accession Q42435.1 shown in (Fig2c) a Capsanthin/capsorubin synthase, chromoplastic enzyme.

4.4.2.1.4 BLASTX analysis of sequence of variety Anugraha

The sequence of variety Anugraha was analysed using BLASTX programme. The result showed that the sequence of Anugraha had shown 100 per cent similarity with



Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
RecName: Full=Capsanthin-capsorubin synthase, chromoplastic; Flags: Precursor	1010	1010	99%	0.0	99%	Q42435.1
capsanthin/capsorubin synthase [Capsicum annuum]	1009	1009	99%	0.0	99%	ADH04285.1
capsanthin-capsorubin synthase, chromoplastic [Capsicum annuum]	1007	1007	99%	0.0	99%	NP_001311998.1
capsanthin-capsorubin synthase [Capsicum annuum]	1003	1003	99%	0.0	99%	AIQ82733.1
putative chromoplastic oxydo-reductase [Capsicum annuum]	910	910	93%	0.0	96%	CAA54961.1
neoxanthin synthase, chloroplastic [Solanum pennellii]	876	876	99%	0.0	87%	NP_001310372.1
chromoplast-specific lycopene beta-cyclase [Solanum lycopersicum]	868	868	99%	0.0	85%	AAG21133.1
neoxanthin synthase, chloroplastic [Solanum lycopersicum]	868	868	99%	0.0	86%	NP_001234445.2
PREDICTED: neoxanthin synthase, chloroplastic [Nicotiana tabacum]	867	867	99%	0.0	85%	XP_016480982.1

Fig 2a: BLASTX of Byadagi Kaddi *Ccs* cds gene sequence showing 99 per cent similarity with Ccs gene accession Q42435.1

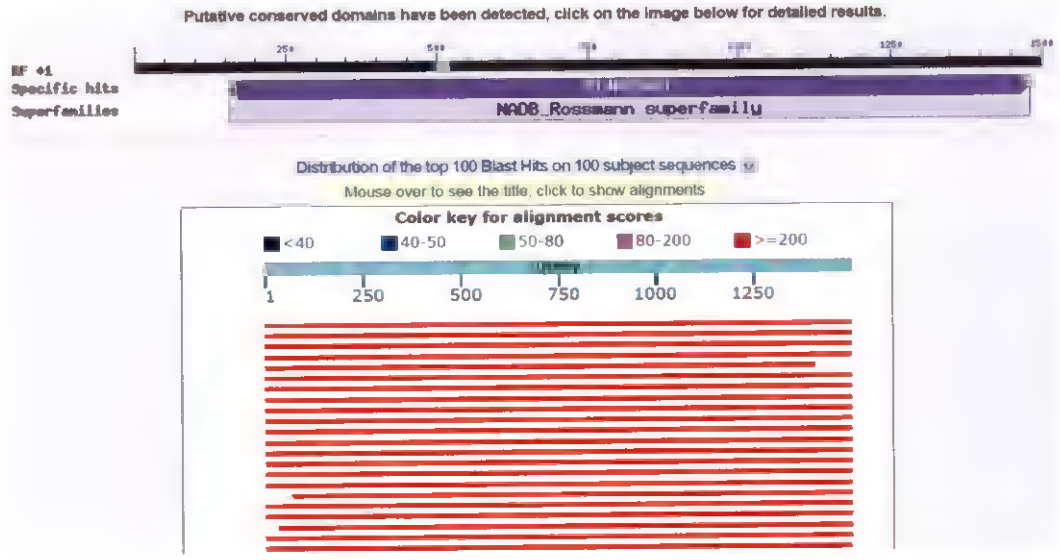
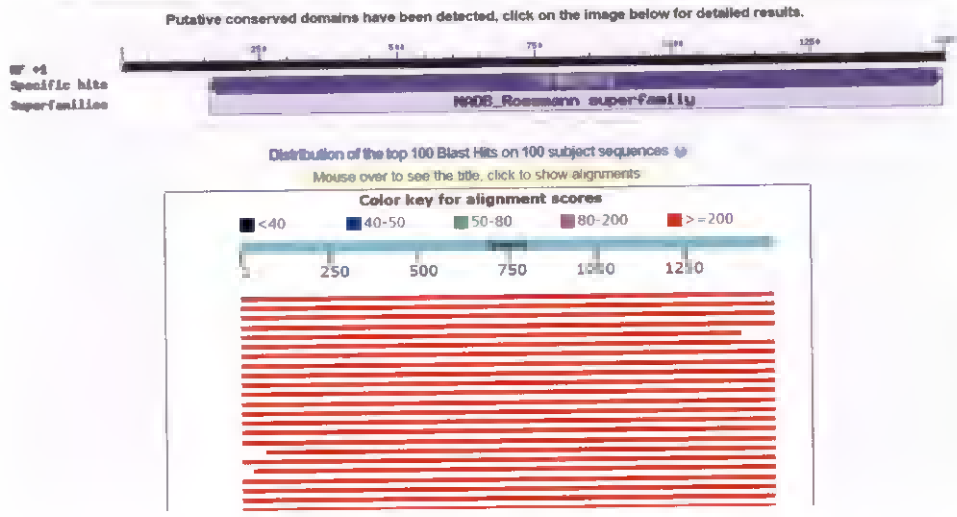


Fig 2b: BLASTX of Byadagi Dabbi *Ccs* cds gene sequence showing 99 per cent similarity with Ccs gene accession Q42435.1

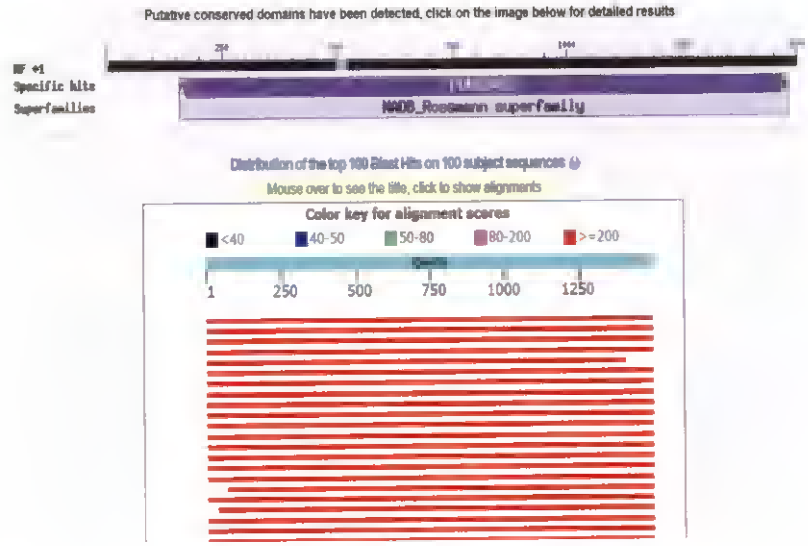


Select All None Selected 0

Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
RecName: Full=Capsanthin/capsorubin synthase, chromoplastic; Flags: Precursor	1013	1013	99%	0.0	100%	Q42435.1
capsanthin/capsorubin synthase [Capsicum annuum]	1011	1011	99%	0.0	99%	ADH04285.1
capsanthin-capsorubin synthase, chromoplastic [Capsicum annuum]	1009	1009	99%	0.0	99%	NP_001311998.1
capsanthin/capsorubin synthase [Capsicum annuum]	1006	1006	99%	0.0	99%	AIQ82733.1
putative chromoplastic oxydo-reductase [Capsicum annuum]	913	913	93%	0.0	97%	CAA54961.1
neoxanthin synthase, chloroplastic [Solanum pennellii]	877	877	99%	0.0	87%	NP_001310372.1
PREDICTED: neoxanthin synthase, chloroplastic [Nicotiana tabacum]	870	870	99%	0.0	86%	XP_016480982.1
chromoplast-specific lycopene beta-cyclase [Lycium barbarum]	869	869	99%	0.0	85%	AIX87499.1
chromoplast-specific lycopene beta-cyclase [Solanum lycopersicum]	868	868	99%	0.0	86%	AAG21133.1

Fig 2c: BLASTX of Ujwala *Ccs* cds gene sequence showing 99 per cent similarity with Ccs gene accession Q42435.1



Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
RecName: Full=Capsanthin/capsorubin synthase, chromoplastic; Flags: Precursor	1013	1013	99%	0.0	100%	Q42435.1
capsanthin capsorubin synthase [Capsicum annuum]	1011	1011	99%	0.0	99%	ADH04285.1
capsanthin capsorubin synthase, chromoplastic [Capsicum annuum]	1009	1009	99%	0.0	99%	NP_001311998.1
capsanthin capsorubin synthase [Capsicum annuum]	1006	1006	99%	0.0	99%	AIQ82733.1
putative chromoplastic oxydo-reductase [Capsicum annuum]	913	913	93%	0.0	97%	CAA54961.1
neoxanthin synthase, chloroplastic [Solanum pennellii]	877	877	99%	0.0	87%	NP_001310372.1
PREDICTED: neoxanthin synthase, chloroplastic [Nicotiana tabacum]	870	870	99%	0.0	86%	XP_016490982.1
chromoplast-specific lycopene beta-cyclase [Lycium barbarum]	869	869	99%	0.0	85%	AIX87499.1
chromoplast-specific lycopene beta-cyclase [Solanum lycopersicum]	868	868	99%	0.0	86%	AAG21133.1

Fig 2d: BLASTX of Anugraha *Ccs* cds gene sequence showing100 per cent similarity with Ccs gene accession Q42435.1

Genotypes	Expected size of Ccs cds amplicon	Size of the amplified sequence obtained with Ccs cds primer	Query coverage	e-value	Similarity %	Accession showing highest similarity
Byadagi Kaddi	1497	1392	99	0.0	99	Q42435.1
Byadagi Dabbi	1497	1410	99	0.0	99	Q42435.1
Ujwala	1497	1383	99	0.0	100	Q42435.1
Anugraha	1497	1404	99	0.0	100	Q42435.1
Vellayani Samrudhi	1497	1389	99	0.0	99	NP_001311998.1
Vellayani Thejus	1497	1425	99	0.0	99	NP_001311998.1
CC 8-1	1497	1416	99	0.0	99	NP_001311998.1

Table5: BLASTx analysis of Ccs amino acid sequence in different chilli genotypes

amino acid sequence of Ccs gene in the accession Q42435.1 shown in (Fig2d) a Capsanthin capsorubin synthase, chromoplastic enzyme.

4.4.2.1.5 BLASTX analysis of sequence of Vellayani Samrudhi

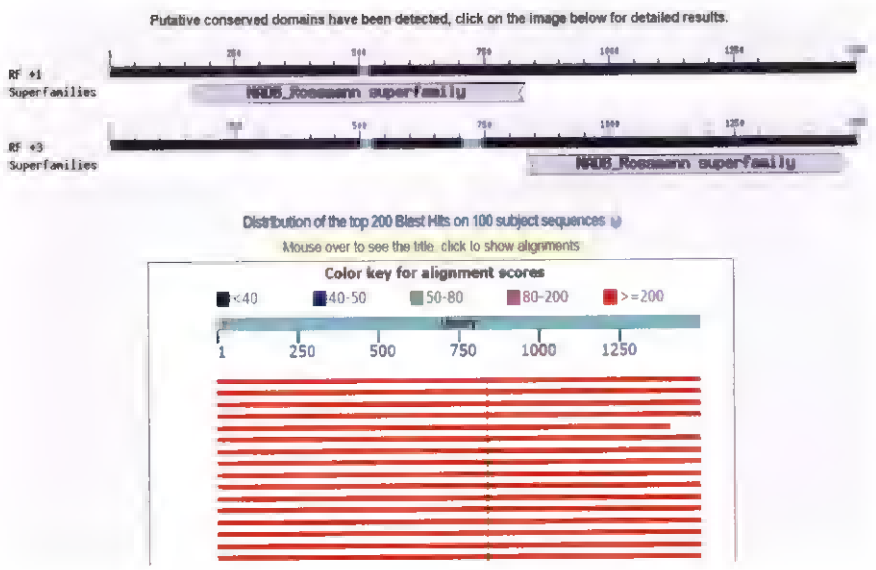
The sequence of variety Vellayani samrudhi was analysed using BLASTX programme. The result showed that the sequence of Vellayani samrudhi had shown 99 per cent similarity with amino acid sequence of Ccs gene in the accession NP_001311998.1 (Fig2e) *Capsicum annuum* cultivar specific capsanthin capsorubin synthase, chromoplastic protein contains 498 amino acids sequence involved incarotenoid biosynthetic process

4.4.2.1.6 BLASTX analysis of sequence of Vellayani Thejus

The sequence of variety Vellayani Thejus was analysed using BLASTX programme. The result showed that the sequence of Vellayani Thejus had shown 99 per cent similarity with amino acid sequence of Ccs gene in the accession NP001311998.1 (Fig2f) *Capsicum annuum* cultivar specific capsanthin capsorubin synthase, chromoplastic protein contains 498 amino acids sequence involved in carotenoid biosynthetic process

4.4.2.1.7 BLASTX analysis of sequence of CC8-1

The sequence of accession CC8-1was analysed using BLASTX programme. The result showed that the sequence of CC8-1had shown 99 per cent similarity with amino acid sequence of Ccs gene in the accession NP001311998.1 shown in (Fig2g) *Capsicum annuum* cultivar specific capsanthin capsorubin synthase, chromoplastic protein contains 498 amino acids sequence involved in carotenoid biosynthetic process

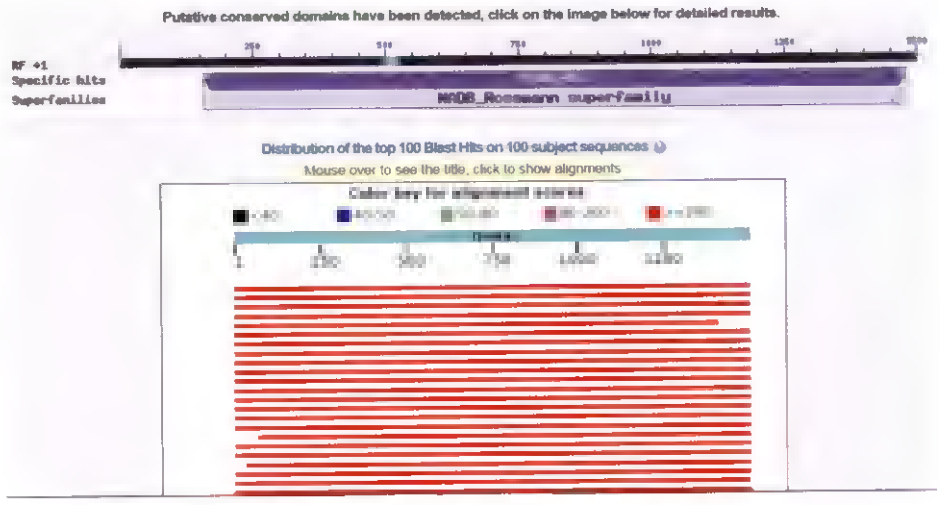


Select All None Selected 0

Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
capsanthin/capsorubin synthase, chromoplastic [Capsicum annuum]	559	981	99%	0.0	99%	NP_001311998.1
RecName: Full=Capsanthin/capsorubin synthase, chromoplastic; Flags: Precursor	556	977	99%	0.0	99%	Q42435.1
capsanthin/capsorubin synthase [Capsicum annuum]	556	975	99%	0.0	99%	ADH04285.1
capsanthin/capsorubin synthase [Capsicum annuum]	555	974	99%	0.0	99%	AIQ82733.1
putative chromoplastic oxydo-reductase [Capsicum annuum]	553	877	93%	0.0	99%	CAA54961.1
PREDICTED: neoxanthin synthase, chloroplastic [Nicotiana tabacum]	474	848	99%	0.0	85%	XP_016480982.1
PREDICTED: neoxanthin synthase, chloroplastic [Nicotiana tomentosiformis]	471	845	99%	0.0	85%	XP_009609006.1
PREDICTED: neoxanthin synthase, chloroplastic [Nicotiana sylvestris]	470	845	99%	0.0	85%	XP_009789371.1
neoxanthin synthase, chloroplastic [Solanum pennellii]	469	854	99%	0.0	84%	NP_001310372.1
PREDICTED: neoxanthin synthase, chloroplastic [Nicotiana attenuata]	469	840	99%	0.0	84%	XP_019260188.1
chromoplast-specific lycopene beta-cyclase [Lycium barbarum]	468	847	99%	0.0	83%	AIX87499.1

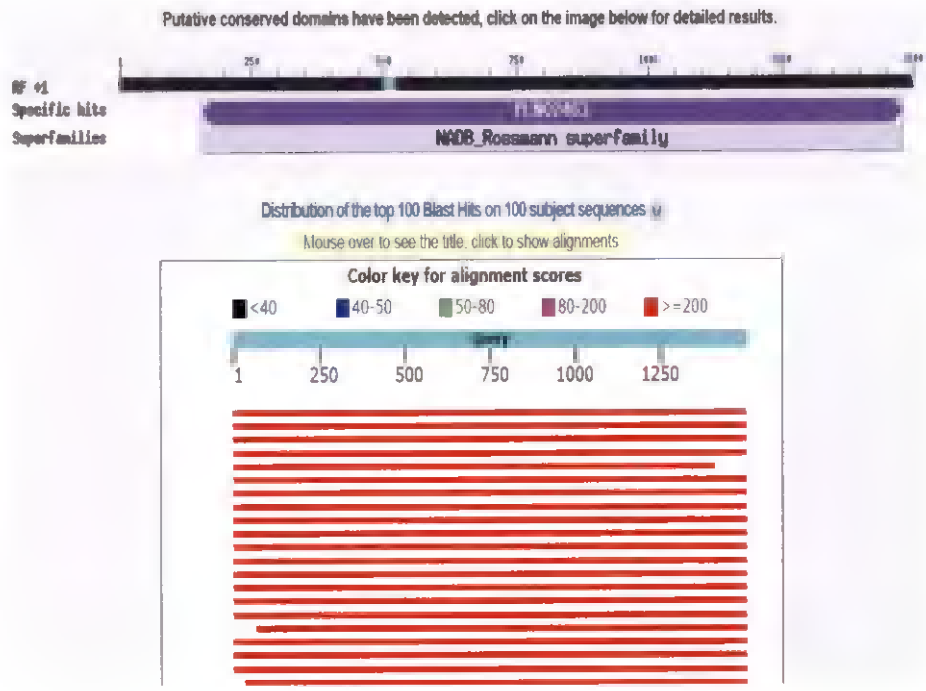
Fig 2e: BLASTX of Vellayani Samrudhi *Ccs* cds gene sequence showing 99 per cent similarity with Ccs gene accession NP001311998.1



Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
capsanthin-capsorubin synthase, chromoplastic [Capsicum annuum]	1009	1009	99%	0.0	99%	NP_001311998.1
RecName: Full=Capsanthin-capsorubin synthase, chromoplastic; Flags: Precursor	1005	1005	99%	0.0	99%	Q42435.1
capsanthin/capsorubin synthase [Capsicum annuum]	1005	1005	99%	0.0	99%	AIQ62733.1
capsanthin-capsorubin synthase [Capsicum annuum]	1004	1004	99%	0.0	99%	ADH04285.1
putative chromoplastic oxydo-reductase [Capsicum annuum]	906	906	93%	0.0	96%	CAA54961.1
neoxanthin synthase, chloroplastic [Solanum pennellii]	879	879	99%	0.0	87%	NP_001310372.1
chromoplast-specific lycopene beta-cyclase [Lycium barbarum]	872	872	99%	0.0	85%	AIX87499.1
PREDICTED: neoxanthin synthase, chloroplastic [Nicotiana tabacum]	871	871	99%	0.0	86%	XP_016480962.1
chromoplast-specific lycopene beta-cyclase [Solanum lycopersicum]	871	871	99%	0.0	86%	AAG21133.1
neoxanthin synthase, chloroplastic [Solanum lycopersicum]	871	871	99%	0.0	86%	NP_001234445.2
PREDICTED: neoxanthin synthase, chloroplastic [Nicotiana sylvestris]	870	870	99%	0.0	86%	XP_009789371.1

Fig 2f: BLASTX of Vellayani Thejus Ccs cds gene sequence showing 99 per cent similarity with Ccs gene accession NP001311998.1



Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
capsanthin-capsorubin synthase, chromoplastic [Capsicum annuum]	1006	1006	99%	0.0	99%	NP_001311998.1
RecName: Full=Capsanthin-capsorubin synthase, chromoplastic; Flags: Precursor	1002	1002	99%	0.0	99%	Q42435.1
capsanthin/capsorubin synthase [Capsicum annuum]	1001	1001	99%	0.0	99%	AIQ82733.1
capsanthin-capsorubin synthase [Capsicum annuum]	1001	1001	99%	0.0	99%	ADH04285.1
putative chromoplastic oxydo-reductase [Capsicum annuum]	902	902	93%	0.0	96%	CAA54961.1
neoxanthin synthase, chloroplastic [Solanum pennellii]	875	875	99%	0.0	87%	NP_001310372.1
chromoplast-specific lycopene beta-cyclase [Lycium barbarum]	869	869	99%	0.0	85%	AIX87499.1
PREDICTED: neoxanthin synthase, chloroplastic [Nicotiana tabacum]	868	868	99%	0.0	86%	XP_016480982.1
chromoplast-specific lycopene beta-cyclase [Solanum lycopersicum]	867	867	99%	0.0	85%	AAG21133.1
neoxanthin synthase, chloroplastic [Solanum lycopersicum]	867	867	99%	0.0	86%	NP_001234445.2

Fig 2g: BLASTX of CC8-1 Ccs cds gene sequence showing 99 per cent similarity with Ccs gene accession NP001311998.1

4.4.3 Analysis of Ccs gene sequences using conserved domain database (CDD)

The CDD result of Ccs gene sequences briefed in (Table6)

4.4.3.1 CDD analysis of sequence of Byadagi kaddi

Analysis of Byadagi kaddi sequence carriedout for the presence of conserved domain by using CDD (conserved domain database). The sequence contain conserved domain Lycopene beta cyclase or carotene cyclase, spanned from 157 to 1392bp. (Fig3a) Conserved domain comes under NADB Rossmann super family. It catalyzes the biosynthesis of β-carotene and its derivatives.

4.4.3.2 CDD analysis of sequence of Byadagi Dabbi

Analysis of Byadagi dabbi sequence carriedout for the presence of conserved domain by using CDD (conserved domain database). The sequence contain conserved domain Lycopene beta cyclase or carotene cyclase, spanned from 157 to 1410bp (Fig3b) Conserved domaincome under NADB Rossmann super family. It catalyzes the biosynthesis of β-carotene and its derivatives.

4.4.3.3 CDD analysis of sequence of Ujwala

Analysis of Ujwala sequence carriedout for the presence of conserved domain by using CDD (conserved domain database). The sequence contain conserved domain Lycopene beta cyclase or carotene cyclase, spanned from 157 to 1383bp. (Fig3c) Conserved domaincome under NADB Rossmann super family. It catalyzes the biosynthesis of β-carotene and its derivatives.

4.4.3.4 CDD analysis of sequence of Anugraha

Analysis of Anugraha sequence carriedout for the presence of conserved domain by using CDD (conserved domain database). The sequence contain conserved domain Lycopene beta cyclase, spanned from 157 to 1404bp. (Fig3d) Conserved domaincome under NADB Rossmann super family. It catalyzes the biosynthesis of β-carotene and its derivatives.

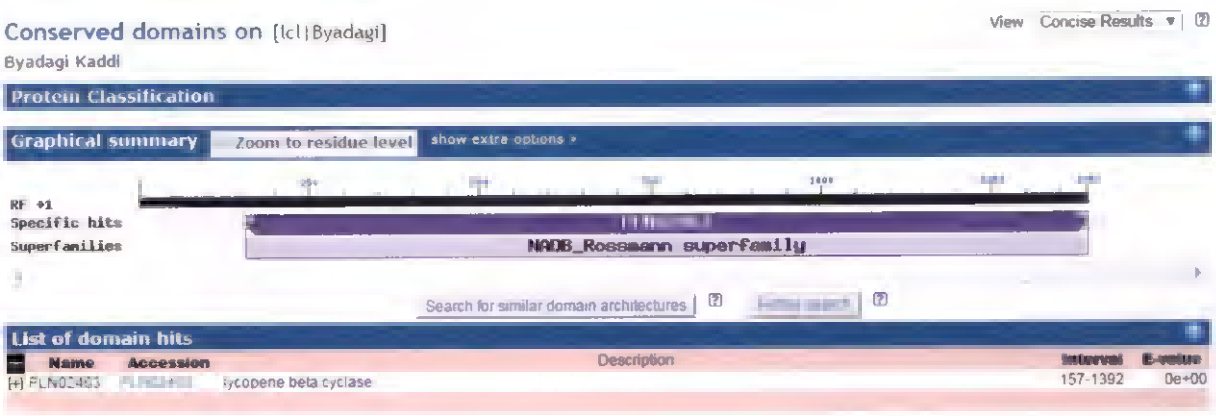


Fig 3a: Conserved domain identified from Ccs cds sequence of Byadagi Kaddi

Fig 3b: Conserved domain identified from Ccs cds sequence of Byadagi Dabbi

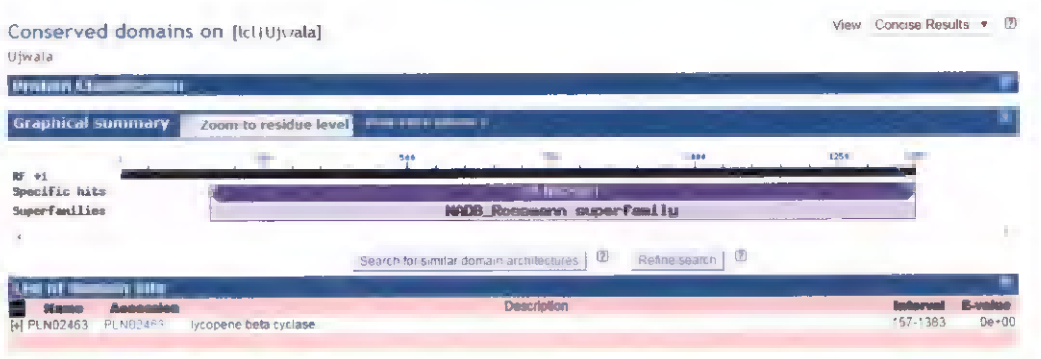


Fig 3c: Conserved domains identified from Ccs *Cds* sequence of Ujwala

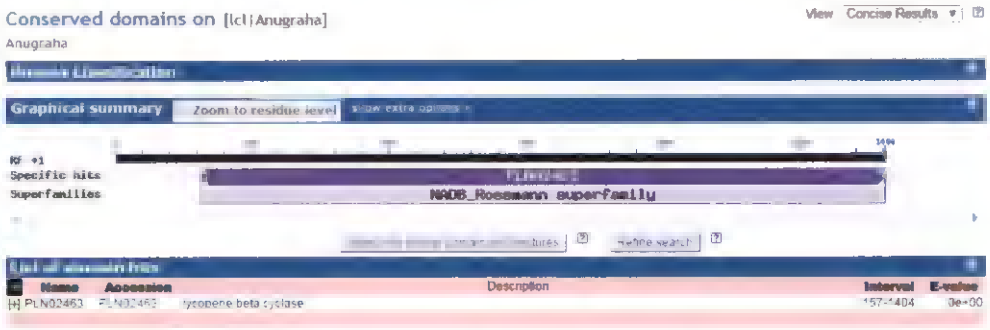


Fig 3d: Conserved domains identified from Ccs *Cds* sequence of Anugraha

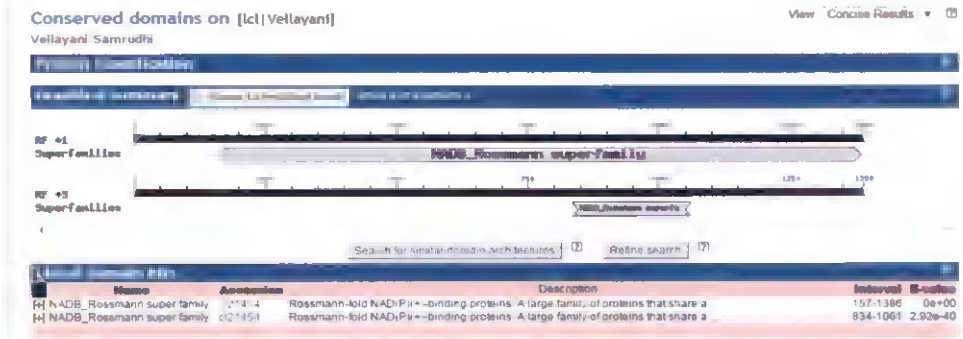


Fig 3e: Conserved domains identified in Ccs *Cds* sequence of Vellayani samrudhi

Genotypes	Size of the analyzed sequence with Ccs cds primer	Domain position	Domain name
Byadagi Kaddi	1392	157 to 1392bp	Lycopene beta cyclase
Byadagi Dabbi	1410	157 to 1410bp	Lycopene beta cyclase
Ujwala	1383	157 to 1387bp	Lycopene beta cyclase
Anugraha	1404	157 to 1404bp	Lycopene beta cyclase
Vellayani Samrudhi	1389	157 to 1389bp	NADP super family protein
		834 to 1061bp	NADP super family protein
Vellayani Thejus	1425	157 to 1425bp	Lycopene beta cyclase
CC 8-1	1416	157 to 1416bp	Lycopene beta cyclase

Table 6: Conserved domains present in Ccs gene sequence of different chilli genotypes

4.4.3.5 CDD analysis of sequence of Vellayani samrudhi

Analysis of Vellayani samrudhi sequence carriedout for the presence of conserved domain by using CDD (conserved domain database). The sequence contains Rossmann-family NAD(P) proteins; (157-1386) (Fig3e) this family contains large number of proteins with NADB binding domain. The NADB domain was present in various biological pathways such as glycolysis, and several other redox enzymes.

4.4.3.6 CDD analysis of sequence of Vellayani Thejus

Analysis of Vellayani Thejus sequence carriedout for the presence of conserved domain by using CDD (conserved domain database). The sequence contain conserved domain Lycopene beta cyclase or carotene cyclase, spanned from 157 to 1425bp (Fig3f). Conserved domaincome under NADB Rossmann super family. It catalyzes the biosynthesis of β-carotene and its derivatives.

4.4.3.7 CDD analysis of sequence CC8-1

Analysis of CC8-1 sequence carriedout for the presence of conserved domain by using CDD (conserved domain database). The sequence contain conserved domain Lycopene beta cyclase or carotene cyclase, spanned from 157 to 1416bp (Fig3g). Conserved domaincome under NADB Rossmann super family. It catalyzes the biosynthesis of β-carotene and its derivatives.

4.4.4 Analysis of sequences using ORF Finder

Sequences were analyzed using ORF finder to find out the open reading frames in the sequence obtained by the amplification of two gene specific primers from seven genotypes.

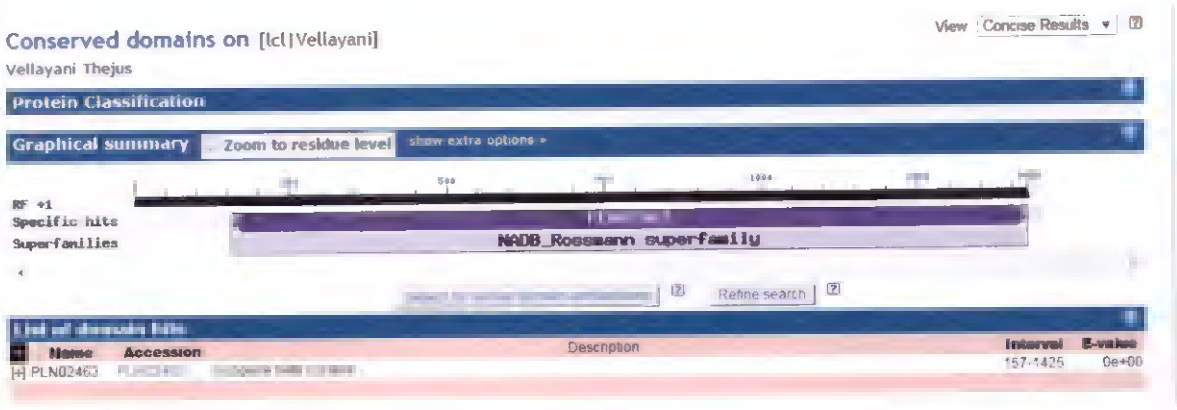


Fig 3f: Conserved domains identified from Ccs *Cds* sequence of Vellayani Thejus

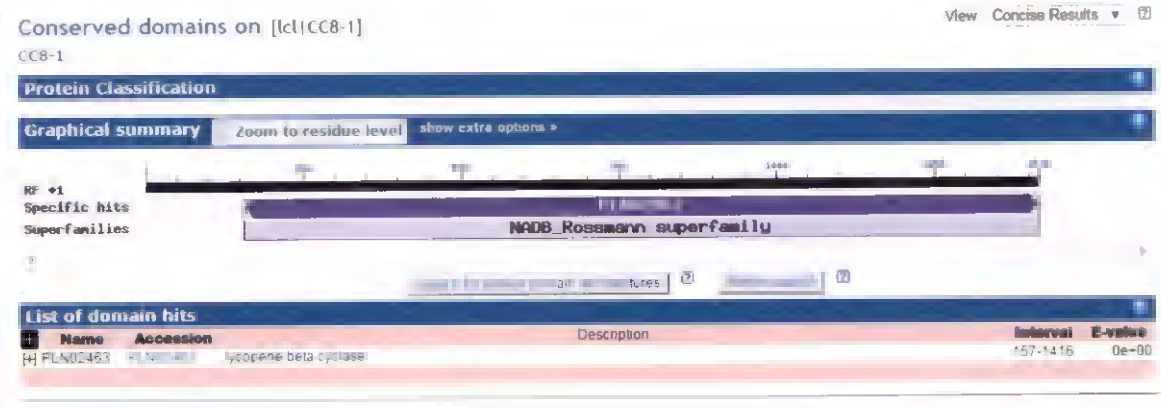


Fig 3g: Conserved domains identified from Ccs *Cds* sequence of CC8-1

4.4.4.1 Open Reading Frames in the PCR product amplified with *Ccs* cds primer

The ORF result of Ccs gene sequences briefed in (Table7)

4.4.4.2 Analysis of sequence of Byadagi Kaddi for ORFs

When the Byadagi Kaddi sequence was analysed for the ORFs, it was found that the region contains 6 ORFs. Among them the ORF1 was the longest, coding for 463 aminoacids. The ORF1 spanned from 1 to 1392bp (Fig4a). The amino acid sequences coded by this ORF are given below. BLASTp had shown that this sequence stands for the Capsanthin capsorubin synthase, chromoplastic; Flags: Precursor protein.

ORF1

METLLKPFPSPLLSIPTPNMYSFKHNSTFPNPTKQKDSRKFHFRNKSSTHFCSFLDLAPTSKPESLDVNISW
VDTDLDGAEFDVIIIGTGPAGLRLAEQVSKYGIKVCCVDPSPLSMWPNNYGVWVDEFEKLGLEDCLDHK
WPVSCVHISDHKTKYLDRPYGRVSRKKLKLKLLNSCVENRVKFYKAKVLKVKHEEFESSIVCDDGRKIS
GSLIVDASGYASDFIEYDKPRNHGYQVAHGILAEVDNHPFDLDKMMLMDWRDSHLGNEPYLRVKNTKE
PTFLYAMPFDRNLVFLEETSLVSRPMLSYMEVKRRMVARLRHLGIKVRSVLEEEKCVITMGGPLPRIPQN
VMAIGGTSGIVHPSSGYMVARSMALAPVLAEAIVESLGSTRMIRGSQLYHRVWNGLWPSDRRRVRECYC
FGMETLLELDLEGTRRLFDAFFD VDPKYWHGFLSSRLSVKELAVLSL

4.4.4.3Analysis of sequence of Byadagi Dabbi for ORFs

When the Byadagi Dabbi sequence was analysed for the ORFs, it was found that the region contains 7 ORFs. Among them the ORF1 was the longest, coding for 469 amino acids. The ORF1 spanned from 1 to 1410bp (Fig4b). The amino acid sequences coded by this ORF are given below. BLASTp had shown that this sequence stands for the Capsanthin capsorubin synthase, chromoplastic; Flags: Precursor protein.

ORF1

METLLKPFPSPLLSIPTPNMYSFKHNSTFPNPTKQKDSRKFHYRNKSSTHFCSFLDLAPTSKPESLDVNISW
VDTDLDGAEFDVIIIGTGPAGLRLAEQVSKYGIKVCCVDPSPLSMWPNNYGVWVDEFEKLGLEDCLDHK
WPVSCVHISDHKTKYLDRPYGRVSRKKLKLKLLNSCVENRVKFYKAKVLKVKHEEFESSIVCDDGRKIS
GSLIVDASGYASDFIEYDKPRNHGYQVAHGILAEVDNHPFDLDKMMLMDWRDSHLGNEPYLRVKNTKE
PTFLYAMPFDRNLVFLEETSLVSRPMLSYMEVKRRMVARLRHLGIKVRSVLEEEKCVITMGGPLPRIPQN
VMAIGGTSGIVHPSSGYMVARSMALAPVLAEAIVESLGSTRMIRGSQLYHRVWNGLWPSDRRRVRECYC
FGMETLLELDLEGTRRLFDAFFDVDPKYWHGFLSSRLSVKELAVLSLYLFGHA

Fig 4a: ORFs identified from the sequence of the *Ccs* cds in Byadagi Kaddi

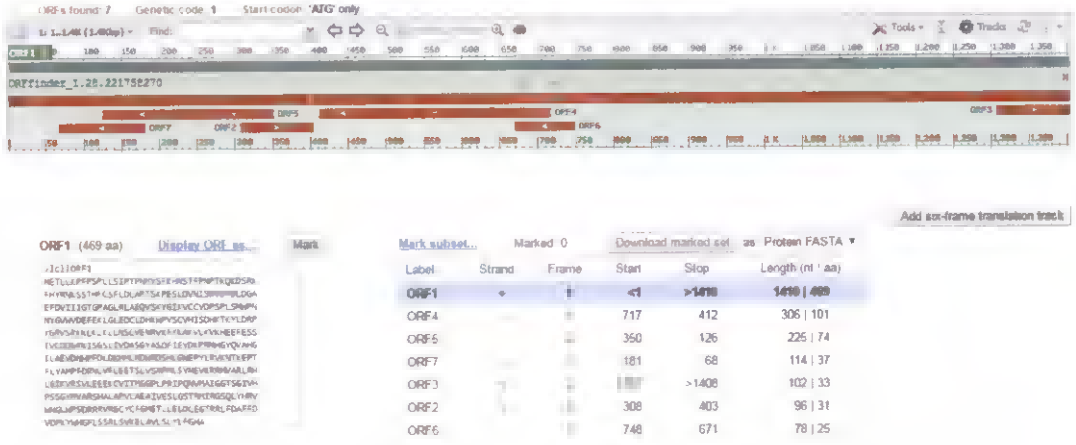


Fig 4b: ORFs identified from the sequence of the *Ccs* cds in Byadagi Dabbi

Genotypes	Size of the sequence obtained with Ccs cds primer	No of ORFs present in sequence	Length of longest ORF	Longest ORF	Number of amino acids coded	Protein coded by longest ORF
Byadagi Kaddi	1392	6	1392	ORF1	463	Ccs
Byadagi Dabbi	1410	7	1410	ORF1	469	Ccs
Ujwala	1383	7	1383	ORF1	460	Ccs
Anugraha	1404	7	1404	ORF1	467	Ccs
Vellayani Samrudhi	1389	7	897	ORF1	298	Ccs
Vellayani Thejus	1425	6	1425	ORF1	474	Ccs
CC 8-1	1416	7	654	ORF2	217	Ccs

Table 7: Number of ORFs identified in the *Ccs* gene sequence

4.4.4.4 Analysis of sequence of Ujwala for ORFs

When the Ujwala sequence was analysed for the ORFs, it was found that the region contains 15 ORFs. Among them the ORF1was the longest, coding for 460 aminoacids. The ORF1 spanned from 1 to 1383bp (Fig4c). The amino acid sequences coded by this ORF are given below. BLASTp had shown that this sequence stands for the Capsanthin capsorubin synthase, chromoplastic; Flags: Precursor protein.

ORF1

METLLKPFPSPLLSIPTPNMYSFKHNSTFPNPTKQKDSRKFHYRNKSSTHFCSFLDLAPTSKPES
LDVNISWVDTDLDGAEFDVIIIGTGPAGLRLAEQVSKYGIKVCCVDPSPLSMWPNNYGVWVD
EFEKLGLEDCLDHKWPVSCVHISDHKTKYLDRPYGRVSRKKLKLKLLNSCVENRVKFYKAK
VLKVKHEEFESSIVCDDGRKISGSLIVDASGYASDFIEYDKPRNHGYQVAHGILAEVDNHPFDL
DKMMLMDWRDSHLGNEPYLRVKNTKEPTFLYAMPFDRNLVFLEETSLVSRPMLSYMEVKR
RMVARLRHLGIKVRSVLEEEKCVITMGGPLPRIPQNVMAIGGTSGIVHPSSGYMVARSMALAP
VLAEAIVESLGSTRMIRGSQLYHRVWNGLWPSDRRRVRECYCFGMETLLKLDLEGTRRLFDA
FFDVDPKYWHGFLSSRLSVKELAV

4.4.4.5 Analysis of sequence of Anugraha for ORFs

When the Anugraha sequence was analysed for the ORFs, it was found that the region contains 7 ORFs. Among them the ORF1was the longest, coding for 467 aminoacids. The ORF1 spanned from 1 to 1404bp (Fig4d). The amino acid sequences coded by this ORF are given below. BLASTp had shown that this sequence stands for the Capsanthin capsorubin synthase, chromoplastic; Flags: Precursor protein.

ORF1

METLLKPFPSPLLSIPTPNMYSFKHNSTFPNPTKQKDSRKFHYRNKSSTHFCSFLDLAPTSKPES
LDVNISWVDTDLDGAEFDVIIIGTGPAGLRLAEQVSKYGIKVCCVDPSPLSMWPNNYGVWVD
EFEKLGLEDCLDHKWPVSCVHISDHKTKYLDRPYGRVSRKKLKLKLLNSCVENRVKFYKAK
VLKVKHEEFESSIVCDDGRKISGSLIVDASGYASDFIEYDKPRNHGYQVAHGILAEVDNHPFDL
DKMMLMDWRDSHLGNEPYLRVKNTKEPTFLYAMPFDRNLVFLEETSLVSRPMLSYMEVKR
RMVARLRHLGIKVRSVLEEEKCVITMGGPLPRIPQNVMAIGGTSGIVHPSSGYMVARSMALAP
VLAEAIVESLGSTRMIRGSQLYHRVWNGLWPSDRRRVRECYCFGMETLLKLDLEGTRRLFDA
FFDVDPKYWHGFLSSRLSVKELAVLSLYLFG

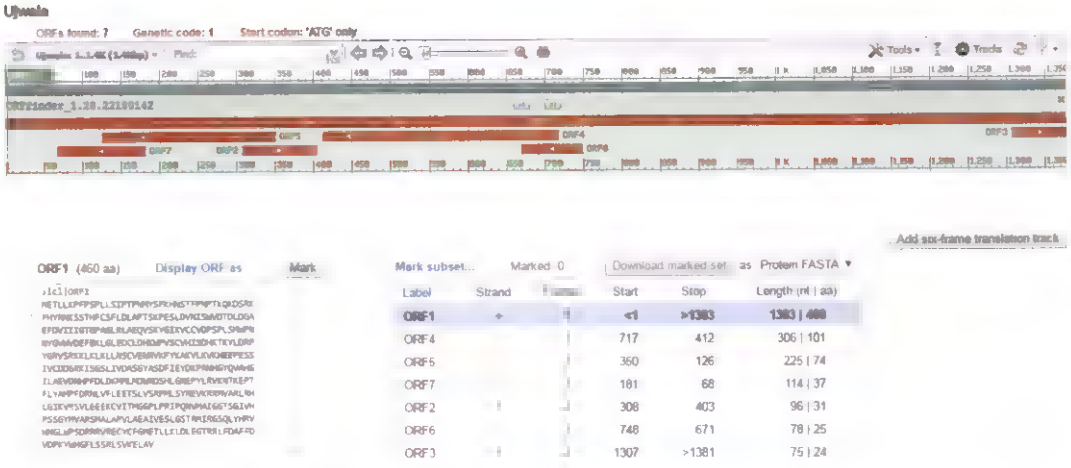


Fig 4c: ORFs identified from the sequence of the *Ccs* cds in Ujwala

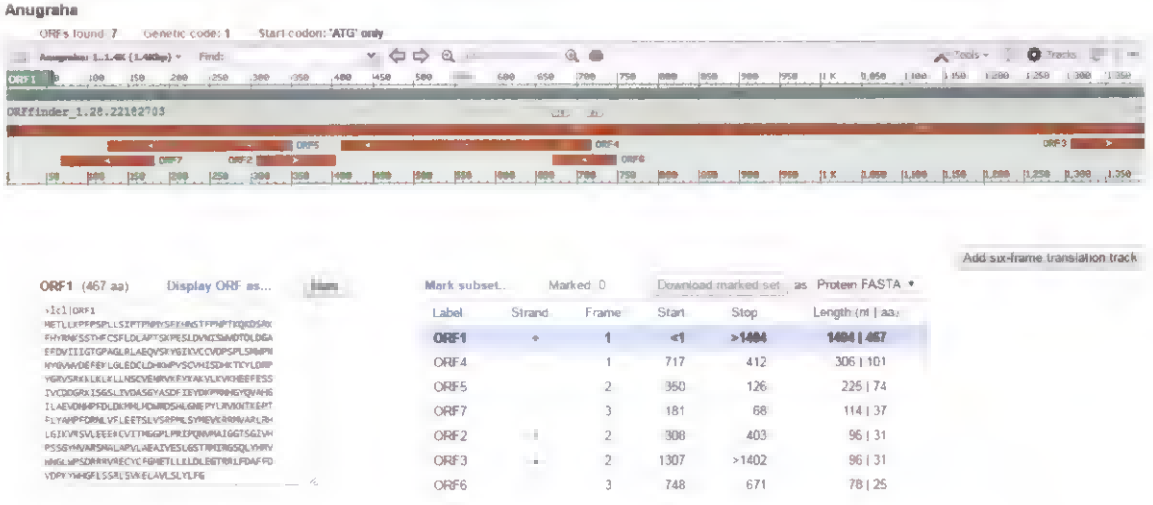


Fig 4d: ORFs identified from the sequence of the *Ccs* cds in Anugraha

4.4.4.6 Analysis of sequence of Vellayani samrudhi for ORFs

When the Vellayani samrudhi sequence was analysed for the ORFs, it was found that the region contains 7 ORFs. Among them the ORF1 was the longest, coding for 298 amino acids. The ORF1 spanned from 1 to 897bp (Fig4e). The amino acid sequences coded by this ORF are given below. BLASTp had shown that this sequence stands for the Capsanthin capsorubin synthase, chromoplastic; Flags: Precursor protein.

ORF1

METLLKPFPSPLLSIPTPNMYSFKHNSTFPNPTKQRDSRKFHSRNKSSTHFCSFLDLAPTSKPES
LDVNISWVDTDLDRAEFDVIIIGTGPAGLRLAEQVSKYGIKVCCVDPSPLSMWPNNYGVWVD
EFEKLGLEDCLDHKWPVSCVHISDHKTKYLDRPYGRVSRKKLKLKLLNSCVENRVKFYKAK
VLKVKHEEFESSIVCDDGRKISGSLIVDASGYASDFIEYDKPRNHGYQVAHGILAEVDNHPFDL
DKMMLMDWRDSHLGNEPYLRVKNTKNQHSCMQCHLIGIWYSWKRLL

4.4.4.7Analysis of sequence of Vellayani Thejus for ORFs

When the Vellayani Thejus sequence was analysed for the ORFs, it was found that the region contains 6 ORFs. Among them the ORF1 was the longest, coding for 474 aminoacids. The ORF1 spanned from 1 to 1425bp (Fig4f). The amino acid sequences coded by this ORF are given below. BLASTp had shown that this sequence stands for the Capsanthin capsorubin synthase, chromoplastic; Flags: Precursor protein.

ORF1

METLLKPFPSPLLSIPTPNMYSFKHNSTFPNPTKQRDSRKFHSRNKSSTHFCSFLDLAPTSKPES
LDVNISWVDTDLDRAEFDVIIIGTGPAGLRLAEQVSKYGIKVCCVDPSPLSMWPNNYGVWVD
EFEKLGLEDCLDHKWPVSCVHISDHKTKYLDRPYGRVSRKKLKLKLLNSCVENRVKFYKAK
VLKVKHEEFESSIVCDDGRKISGSLIVDASGYASDFIEYDKPRNHGYQVAHGILAEVDNHPFDL
DKMMLMDWRDSHLGNEPYLRVKNTKEPTFLYAMPFDRNLVFLEETSLVSRPMLSYMEVKR
RMVARLRHLGIKVRSVLEEEKCVITMGGPLPRIPQNVMAIGGTSGIVHPSSGYMVARSMALAP
VLAEAIVESLGSTRMIRGSQLYHRVWNGLWPSDRRRVRECYCFGMETLLKLDLEGTRRLFDA
FFDVDPKYWHGFLSSRLSVKELAVLSLYLFGHASNLAR

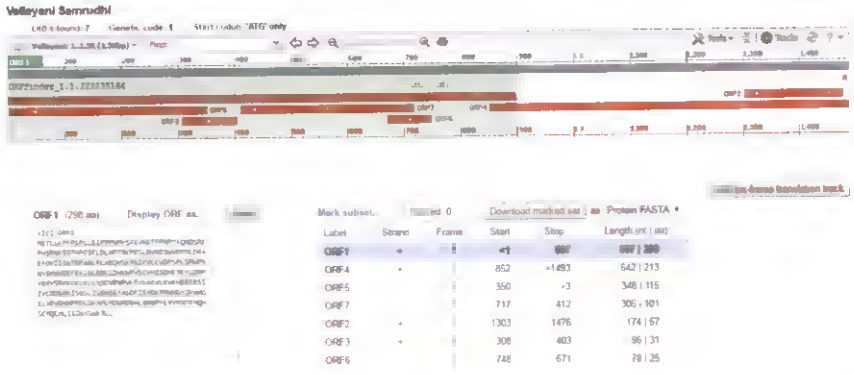


Fig 4e: ORFs identified from the sequence of the *Ccs* cds in Vellayani Samrudhi

Fig 4f: ORFs identified from the sequence of the *Ccs* cds in Vellayani Thejus

Fig 4g: ORFs identified from the sequence of the *Ccs* cds in CC8-1

4.4.4.8 Analysis of sequence of CC8-1 for ORFs

When the CC8-1 sequence was analysed for the ORFs, it was found that the region contains 7 ORFs. Among them the ORF2 was the longest, coding for 217 amino acids. The ORF2 spanned from 763 to 1416bp (Fig4g). The amino acid sequences coded by this ORF are given below. BLASTp had shown that this sequence stands for the Capsanthin capsorubin synthase, chromoplastic; Flags: Precursor protein.

ORF2

MMLMDWRDSHLGNEPYLRVKNTKEPTFLYAMPFDRNLVFLEETSLVSRPMLSYMEVKRRM
VARLRHLGIKVRSVLEEEKCVITMGGPLPRIPQNVMAIGGTSGIVHPSSGYMVARSMALAPVL
AEAIVESLGSTRMIRGSQLYHRVWNGLWPSDRRRVRECYCFGMETLLKLDLEGTRRLFDAFF
DVDPKYWHGFLSSRLSVKELAVLSLYLFGHASN

4.5 Multiple sequence alignment of Capsanthin *capsorubin synthase* gene

Multiple sequence alignment from the coding regions was performed to identify the presence of structural changes in the sequences (Fig5). From the study it was found that Byadagi kaddi had two SNP's T instead of A was found at 129bp and (G) instead of (A) replacing adenine was found at 1273bp positions. Byadagi Dabbi had one SNP *i.e.* guanine (G) replacing adenine (A) at 1273bp. compared with the reference GI: 719762908 cds sequence.

Vellayani samrudhi, an Orange fruit line of *Capsicum frutescens* had several SNP's G 107 A , C 129 A, C 235 G, T 630 C, A 312 T, G 831 A, A 834 G, A 1071 G, G 1272 A, A 1273 G and nucleotide deletions observed at 834,1043,1056,1072 positions in the coding region of the gene.

Vellayani Thejus, and CC8-1 had several SNP's G 107 A, C 129 A, C 235 G, T 630 C, G 831 A, A 834 G, A 1071 G, G 1272 A, A 1273 G positions in the coding region of the gene (Table8).

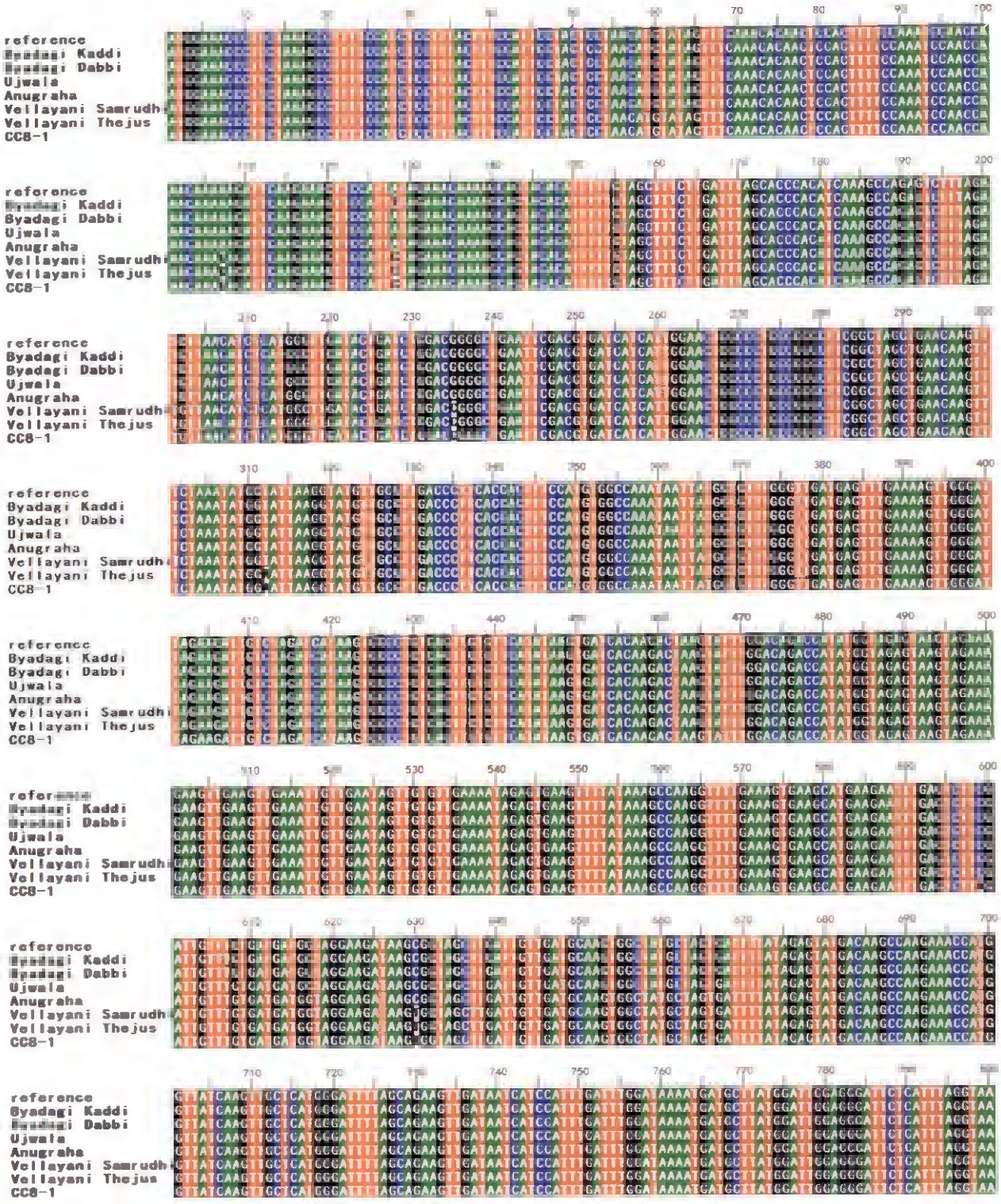


Fig.5. Multiple sequence alignment of nucleotide sequences of *Ccs* coding region with NCBI GI: 719762908 sequences

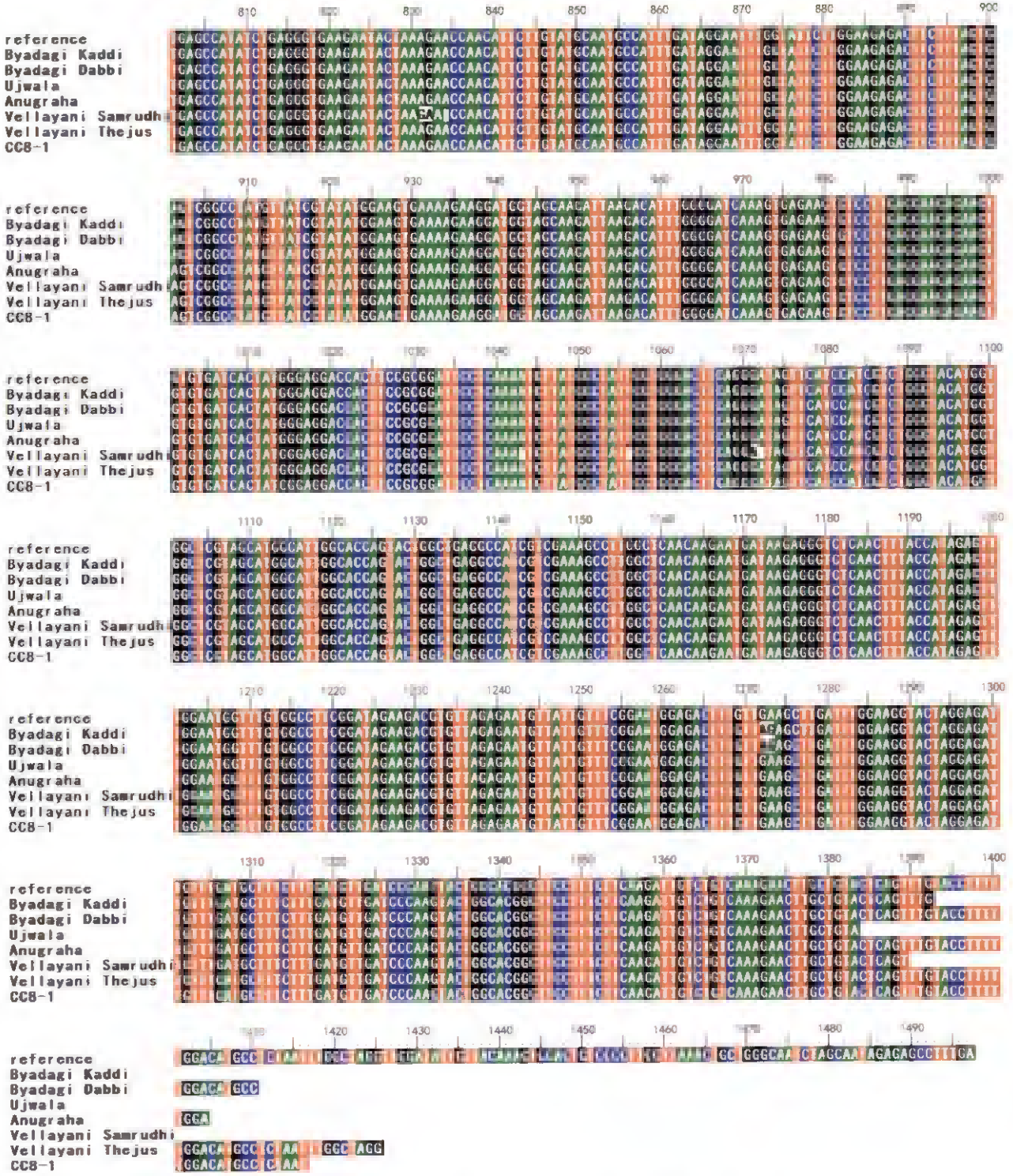


Fig.5. cont... Multiple sequence alignment of nucleotide sequences of *Ccs* coding region with NCBI GI: 719762908 sequences

Multiple sequence alignment of the inferred amino acid sequences revealed that occurrence of Single Nucleotide Polymorphisms in diverse colour genotypes showed in (Table9).

In deep red colour line Byadagi kaddi, it was observed that two SNPs was found with change in amino acids sequence. Tyrosine (Y) replaced by Phenyl alanine (F) at 43th position, and Lysine (K) replaced by Glutamic acid (E) at 425th position. In Byadagi Dabbi, one SNP correlated with change in amino acids sequence, Lysine (K) replaced by Glutamic acid (E) at 425th position.

In Vellayani samrudhi (Orange fruit line of *Capsicum frutescens*), two SNPs were found with change in amino acids sequence. Arginine (R) present instead of Lysine at 36th position, tyrosine (Y) replaced by serine (S) at 43th position, arginine (R) replaced glycine (G) at 79th position. Amino acid deletions were seen at 278, 348, 352, 358th positions and Glutamic acid (E) was replaced by Lysine (K) at 425th position.

Vellayani thejus had two SNPs with change in amino acids sequence.Arginine (R) present instead of Lysine at 36th position, Serine (S) replaced by tyrosine (Y) at 43th position, Arginine (R) replaced glycine (G), at 79th position and Glutamic acid (E) replaced by Lysine (K) at 425th position. In yellow fruited line CC8-1 *Capsicum chinense* one SNP was found with change in amino acids sequence. Arginine (R) present instead of Lysine at 36th position Serine (S) replaced by tyrosine (Y) at 43th position, Arginine (R) replaced glycine (G), at 79th position. Serine (S) to stop codon was present at 200 position, and Glutamic acid (E) replaced by Lysine (K) at 425th position.

4.5.1 Camparision of Ccs gene sequence between the genotypes

Byadgi kaddi had two SNPs which lead to change in amino acid sequence at positions 43 and 425 of Capsanthin capsorubin synthase peptide. Tyrosine (Y) was replaced by Phenyl alanine (F) and Lysine (K) was replaced by Glutamic acid (E).

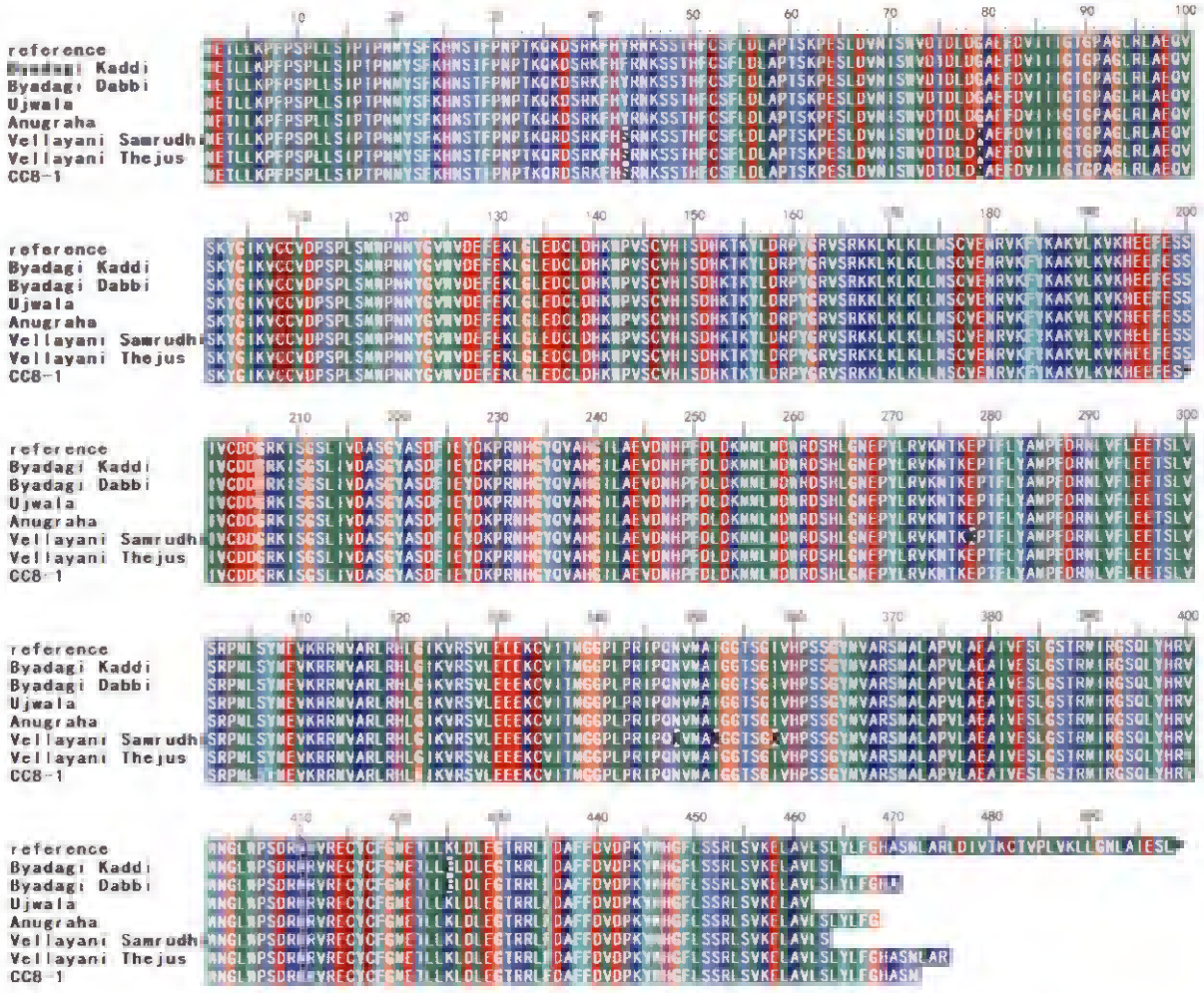


Fig.6. Multiple sequence alignment of deduced amino acid sequences of Ccs coding region with NCBI GI: 719762908

Byadagi dabbi had one SNP at position 425 Lysine (K) replaced by Glutamic acid (E).

In Vellayani samrudhi (Orange fruit line of *Capsicum frutescens*), two SNPs were found with change in the amino acid sequence tyrosine (Y) replaced by serine (S) at 43th position, arginine (R) replaced glycine (G), 79th position and amino acid deletions are present at 278, 348, 352, 358 positions and Glutamic acid (E) replaced by Lysine (K) at 425th position when Compared with Byadagi sequence.

Vellayani Thejus had two SNPs with change in amino acids sequence. Serine (S) replaced by tyrosine (Y) at 43th position, Arginine (R) replaced glycine (G), at 79th position and Glutamic acid (E) replaced by Lysine (K) at 425th position. In yellow fruited line CC8-1 *Capsicum chinense* one SNP was found with change in amino acids sequence Serine (S) 200 stop codon was present, which might have resulted in change in phenotypic expression from red to yellow color when compared with Byadagi sequence shown in (Table9).

Genotypes	Nucleotide position (bp)														
	107	129	235	312	598	630	831	832	834	1043	1056	1071	1072	1272	1273
Byadagi Kaddi	A	T	G	T	C	C	A	G	A	A	T	G	A	A	G
Byadagi Dabbi	A	A	G	T	C	C	A	G	A	A	T	G	A	A	G
Ujwala	A	A	G	T	C	C	A	G	A	A	T	G	A	G	A
Anugraha	A	A	G	T	C	C	A	G	A	A	T	G	A	G	A
Vellayani Samrudhi	G	C	C	T	C	T	G	A	-	-	-	A	-	G	A
Vellayani Thejus	G	C	C	A	C	T	A	G	A	A	T	G	A	G	A
CC8-1	G	C	C	A	G	T	A	G	A	A	T	G	A	G	A

Table 8: Single Nucleotide Polymorphisms identified in the coding region of *Capsanthin capsorubin synthase* gene (107 – 1273bp)

Genotypes	Amino acid position (bp)													
	36	43	79	104	200	210	277	278	348	352	357	358	424	425
Byadagi Kaddi	K	F	G	G	S	S	T	E	N	I	S	I	L	E
Byadagi Dabbi	K	Y	G	G	S	S	T	E	N	I	S	I	L	E
Ujwala	K	Y	G	G	S	S	T	E	N	I	S	I	L	K
Anugraha	K	Y	G	G	S	S	T	E	N	I	S	I	L	K
Vellayani Samrudhi	R	S	R	G	S	S	T	-	-	-	S	-	L	K
Vellayani Thejus	R	S	R	G	S	S	T	E	N	I	S	I	L	K
CC8-1	R	S	R	G	*	S	T	E	N	I	S	I	L	K

Table 9: Amino acid substitutions in the coding region of *Capsanthin capsorubin synthase* gene (36 – 425bp)

4.5.2 Camparision of Ccs gene sequence between the *Capsicum* spp.

Multiple sequence alignment of the inferred amino acid sequences shown in (Fig6) revealed that there was change in the *Cpasanthin capsorubin synthase* peptide, sequence in diffrent *Capsicum* spp. studied.

Capsicum annuum cultivars (Byadagi Kaddi, Byadagi Dabbi, Ujwala, Anugraha) had changes in Capsanthin capsorubin synthase peptide, at positions 43 and 425 positions tyrosine (Y), replaced by Phenyl alanine (F) and Lysine (K) replaced by Glutamic acid (E).

In *Capsicum frutescence* (Vellayani samrudhi) it was observed that two SNPs with change in amino acids sequence. Tyrosine (Y) replaced by serine (S) at 43th position, arginine (R) replaced glycine (G) at 79th position. Amino acid deletions are present at 278, 348, 352, 358 positions and Lysine (K) present instead of glutamic acid at 425th position when compared with Byadagi cultivars (*Capsicum annuum*).

Capsicum chinense (Vellayani Thejus) had two SNPs correlated with change in amino acids sequence. Serine (S) replaced by tyrosine (Y) at 43th position, Arginine (R) replaced glycine (G), at 79th position and Glutamic acid (E) replaced by Lysine (K) at 425th position. In yellow fruited line CC8-1 *Capsicum chinense* it was found that one SNPs correlated with change in amino acids sequence Serine (S) at the position 200 stop codon was present, which might have resulted in change in phenotypic expression from red to yellow color when Compared with Byadagi sequence (*Capsicum annuum*).

Sequence comparison between *Capsicum* spp. showed that amino acid variation present at positions 43 thyrosine (Y), 79 glycine (G) in *Capsicum annuum* (Byadagi kaddi, Byadagi dabbi, Ujwala, Anugraha) whereas serine (S), and Arginine (R) in *Capsicum frutescence,* (Vellayani samrudhi), and *Capsicum chinense*, Vellayani thejus and CC8-1.

The Ccs promoter sequences from each genotype are furnished here under:

Byadagi Kaddi

TAAATTGCAAAAATTTCAGTTCGTTTTTTAATTTCTGTTTTGGGAAGAGGAATACTAC
AAGGCCTACAACAATCACCAATACCTAAATTGCAAAAATTTTAGTTTGTTTTTTAGTT
TCTGTTTTGGGAAGAGGAATACTACAAGGTAAAGGCCTCCAACAATCACCAATACCT
AAATTGCAAAAATTTCAGTTCGTATTTTCGTTTCTATTTTGGGAAGTGGAATAGTATA
AGGTAGGACCTCCAACAATCACCAATACCTAAATTGCAAAAGTTCCGATTCATTTTT
TAGTTTCTGTTTTGGAAAGAGAAATACTACAAGGTAGGGCCTCCAACAATCACCAGT
ACCTAAATTGTAAAAATTTCAGTTCGTTTTTTAGTTTCTATTTTGGGAAGTGGAATAG
TATAAGGTAGGACCTCCAACAATCACCAATACCTAAATTGCAAAAGTTCCGATTCTT
TTTTTAGTTTCTGTTTTGGAAAGAGAAATACTACAAGATAGGACCTTCAACAATCAC
CAATACCTAAATTGCAAAAACTTCAGTTCATTTTTTAGTTTCTGTTTTGGGAAGAAGA
ATACTTCAAGGTAACAATCACCAATACCTAAATTAAAAATTTCAGTTAGTTTTTTAGT
TTCTGTTTTTGGGAAGAGGGAATACTTTCTTTTGCTATATAAAGCCAAAGTAGGTAC
CTATAAGCATCAATATTTTGTATTGCTTAGTGATTCCCCTAGTTCGGTATTTCATTTT
TTTTTCACTATACTATATCACCTCCTCTCATAAATAGCCATTATAAATCTTGCATTTT
CTCTA

Byadagi Dabbi

TAAATTGCAAAAATTTCAGTTCGTTTTTTAATTTCTGTTTTGGGAAGAGGAATACTAC
AAGGCCTCCAACAATCACCAATACCTAAATTGCAAAAATTTTAGTTTGTTTTTTAGTT
TCTGTTTTGGGAAGAGGAATACTACAAGGTAAGGCCTCCAACAATCACCAATACCTA
AATTGCAAAAATTTCAGTTCGTATTTTCGTTTCTATTTTGGGAAGTGGAATAGTATAA
GGTAGGACCTCCAACAATCACCAATACCTAAATTGCAAAAGTTCCGATTCATTTTTT
AGTTTCTGTTTTGGAAAGAGAAATACTACAAGGTAGGGCCTCCAACAATCACCAGTA
CCTAAATTGTAAAAATTTCAGTTCGTTTTTTAGTTTCTATTTTGGGAAGTGGAATAGT
ATAAGGTAGGACCTCCAACAATCACCAATACCTAAATTGCAAAAGTTCCGATTCTTT
TTTTAGTTTCTGTTTTGGAAAGAGAAATACTACAAGATAGGACCTTCAACAATCACC
AATACCTAAATTGCAAAAACTTCAGTTCATTTTTTAGTTTCTGTTTTGGGAAGAAGAA
TACTTCAAGGTAACAATCACCAATACCTAAATTAAAAATTTCAGTTAGTTTTTTAGTT
TCTGTTTTTGGGGAAGAGGAATACTTTCTTTTGCTATATAAAGCCAAAGTAGGGTAC
CTATAAGCATCAATATTTTGTATTGCTTAGTGATTCCCCTAGGTTCGGTATTTCATTT
TTTTTTCACTATACTATATCACCTCCTCTCATAAATAGCCATTATAAATCTTGCATTTT
CTCTA

Ujwala

TAAATTGCAAAAATTTCAGTTCGTTTTTTAATTTCTGTTTTGGGAAGAGGAATACTAC
AAGGCCTCCAACAATCACCAATACCTAAATTGCAAAAATTTTAGTTTGTTTTTTTAGTT
TCTGTTTTGGGAAGAGGAATACTACAAGGTAAGGCCTCCAACAATCACCAATACCTA
AATTGCAAAAATTTCAGTTCGTATTTTCGTTTCTATTTTGGGAAGTGGAATAGTATAA
GGTAGGACCTCCAACAATCACCAATACCTAAATTGCAAAAGTTCCGATTCATTTTTT
AGTTTCTGTTTTGGAAAGAGAAATACTACAAGGTAGGGCCTCCAACAATCACCAGTA

CCTAAATTGTAAAAATTTCAGTTCGTTTTTTAGTTTCTATTTTGGGAAGTGGAATAGT
ATAAGGTAGGACCTCCAACAATCACCAATACCTAAATTGCAAAAGTTCCGATTCTTT
TTTTAGTTTCTGTTTTGGAAAGAGAAATACTACAAGATAGGACCTTCACACAATCAC
CAATACCTAAATTGCAAAAACTTCAGTTCATTTTTTAGTTTCTGTTTTGGGAAGAAGA
ATACTTCAAGGTAACAATCACCAATACCTAAATTAAAAATTTCAGTTAGTTTTTTAGT
TTCTGTTTTTGGGGAAGAGGAATACTTTTCTTTTTGCTATATAAAGACAAAGTAGGT
ACCTATAAGCATCAATATTTTGTATTGCTTAGTGATTCCCCTAGGTTCGGTATTTCAT
TTTTTTTTCACTATACTATATCACCTTCCTCTCATAAATAGCCATTATAAATCTTGCAT
TTTCTCTA

Anugraha

TAAATTGCAAAAATTTCAGTTCGTTTTTTAATTTCTGTTTTGGGAAGAGGAATACTAC
AAGGCCTCCAACAATCACCAATACCTAAATTGCAAAAATTTTAGTTTGTTTTTTAGTT
TCTGTTTTGGGAAGAGGAATACTACAAGGTAAGGCCTCCAACAATCACCAATACCTA
AATTGCAAAAATTTCAGTTCGTATTTTCGTTTCTATTTTGGGAAGTGGAATAGTATAA
GGTAGGACCTCCAACAATCACCAATACCTAAATTGCAAAAGTTCCGATTCATTTTTT
AGTTTCTGTTTTGGAAAGAGAAATACTACAAGGTAGGGCCTCCAACAATCACCAGTA
CCTAAATTGTAAAAATTTCAGTTCGTTTTTTAGTTTCTATTTTGGGAAGTGGAATAGT
ATAAGGTAGGACCTCCAACAATCACCAATACCTAAATTGCAAAAGTTCCGATTCTTT
TTTTAGTTTCTGTTTTGGAAAGAGAAATACTACAAGATAGGACCTTCAACAATCACC
AATACCTAAATTGCAAAAACTTCAGTTCATTTTTTAGTTTCTGTTTTGGGAAGAAGAA
TACTTCAAGGTAACAATCACCAATACCTAAATTAAAAATTTCAGTTAGTTTTTTAGTT
TCTGTTTTTGGGAAGAGGAATACTTTCTTTTTGCTATATAAAGCCAAAGTAGGTACC
TATAAGCATCAATATTTTGTATTGCTTAGTGATTCCCCTAGGTTCGGTATTTCATTTT
TTTTTCACTATACTATATCACCTTCCTCTCATAAATAGCCATTATAAATCTTGCATTTT
CTCTA

Vellayani Samrudhi

TAAATTGCAAAAATTTCAGTTCGTTTTTTAATTTCTGTTTTGGGAAGAGGAATACTAC
AAGGCCTCCAACAATCACCAATACCTAAATTGCAAAAATTTTAGTTTGTTTTTTAGTT
TCTGTTTTGGGAAGAGGAATACTACAAGGTAAGGCCTCCAACAATCACCAATACCTA
AATTGCAAAAATTTCAGTTCGTATTTTCGTTTCTATTTTGGGAAGTGGAATAGTATAA
GGTAGGACCTCCAACAATCACCAATACCTAAATTGCAAAAGTTCCGATTCATTTTTT
AGTTTCTGTTTTGGAAAGAGAAATACTACAAGGTAGGGCCTCCAACAATCACCAGAA
CCTAGATTGTAAAAATTTCAGTTCGGTTTTTAGTTCTATTTTGGGAAGTGGAATAGTA
TAAGGTAGGACCTCCAACAATCACCACAATCTAAATTGCAAAAGTTCCGATTCTTTT
TTTAGTTTCTGTTTTGGAAAGAGAAATACTACAAGATAGGACCTTCAACAATCACCA
ATACCTAAATTGCAAAAATTTCAGTTCATTTTTTAGTTTCTGTTTTGGGAAGAAGAAT
ACTAGAAGATAACAATCACCAATACCTAAATTAAAAATTTCAGTTAGTTTTTTTAGTTT
CTGTTTTTGGGAAGAGGAATACTTTTCTTTTGCTATATAAAGCCAAGTAGGTACCTA
TAAGCATCAATATTTTGTATTGCTTAGTGATTCCCCTAGGTTCGGTATTTCATTTTTT
TTTCACTATACTATATCACCTTCCTCTCATAAATAGCCATTATAAATCTTGCATTTTC
TCTA

Vellayani Thejus

AATTGCAAAAATTTCAGTTCGTTTTTTAATTTCTATTTTGGGAAGAGGGTACTACAAG
GACTCCAACAATCACCAATACCTAAATTGCAAAAATTTCAGTTTGTTTTTTAGTTTCT
GTTTTGGGAAGAGGAATACTACAAGGTAGGGCCTCCAACAATCACCAATACCTAAAT
TGCAACGTTTTTTTAGGTTCTGTTTTGGGAAGAGGAACACTACAAGGAAGGGCCTCCA
ACGATCACCAATACCTAAATTGCAAAAATTTCAGGTCGTATTTTCGTTTCTATTTTGG
GAAGTGGAAGAGTATAAGGTAGGACCTCCAACAATCAACAATACCTAAATTGCAAAA
GTTCCGATTCATTTTTTAGTTTCTGTTTTGGAAAGAGAAATACTACAAGGTAGGGCC
TCCAACAATCACCAGAACCTAGATTGTAAAAATTTCAGTTCGGTTTTTAGGTTCTATT
TTGGGAAGTGGAATAGTATAAGGTAGGACCTCCAACAATCACCACAATCTAAATTGC
AAAAGTTCCGATTCTTTTTTTTTAGTTTCTGTTTTGGAAAGAGAAATACTACAAGGTAG
GGCTTTTCAACAATCACCAGTACCTAAATTGTAAAAATTTCAGTTCATTTTTTTTAGTT
TCTATTTTGGGAAGTGGAATAGTATAAGGTAGGACCTCCAACAATCACCAATACCTA
AATTGCAAAAGTTCCGATC

CC8-1

ATCTAAATTGCAAAAGTTCCGATTCTTTTTTTAGTTTCTGTTTTGGAAAGAGAAATAC
TACAAGGTAGGGCTTTCAACAATCACCAGTACCTAAATTGTAAAAATTTCAGTTCAT
TTTTTTTAGTTTCTATTTTTGGGAAGTGGAATAGTATAAGGGAGGACCTCCAACAATC
ACCAATACCTAAAATTGCAAAAGGTTCCGATTCTTTT-
TTTAGTTTCTGTTTCGGAAAGAGAAATACTACAAG

4.6 Analysis of Ccs promoter sequence using BLASTn

The Blastn result of Ccs promoter sequences briefed in (Table10)

4.6.1 BLASTn analysis of Ccs promoter sequence of Byadagi Kaddi

The sequences of Byadagi kaddi was analysed using BLASTn programme. The result showed that the sequence of Byadagi kaddi had 99 per cent similarity with Ccs gene accession DQ907615.1 shown in (Fig7a) a *Capsicum annuum* cultivar landrace Nockwang, *Capsanthin capsorubin synthase* promoter region partial sequence.

4.6.2 BLASTn analysis of Ccs promoter sequence of Byadagi Dabbi

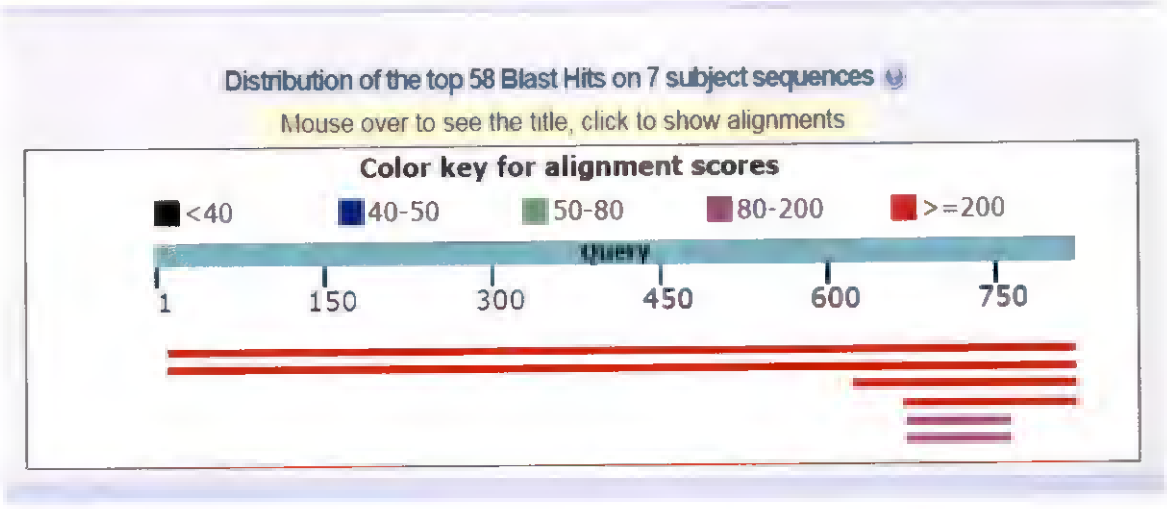
The sequences of Byadagi dabbi was analysed using BLASTn programme. The result showed that the sequence of Byadagi dabbi had 99 per cent similarity with Ccs gene accession DQ907615.1 shown in (Fig7b) a *Capsicum annuum* cultivar landrace Nockwang, *Capsanthin capsorubin synthase* promoter region partial sequence.

4.6.3 BLASTn analysis of Ccs promoter sequence of Ujwaja

The sequences of Ujwala were analysed using BLASTn programme. The result showed that the sequence of Ujwala had 99 per cent similarity with Ccs gene accession DQ907615.1 shown in (Fig7c) a *Capsicum annuum* cultivar landrace Nockwang, *Capsanthin capsorubin synthase* promoter region partial sequence.

4.6.4 BLASTn analysis of Ccs promoter sequence of Anugraha

The sequences of Anugraha were analysed using BLASTn programme. The result showed that the sequence of Anugraha had 91 per cent similarity with Ccs gene accession DQ907615.1 shown in (Fig7d) a *Capsicum annuum* cultivar landrace Nockwang, *Capsanthin capsorubin synthase* promoter region partial sequence.



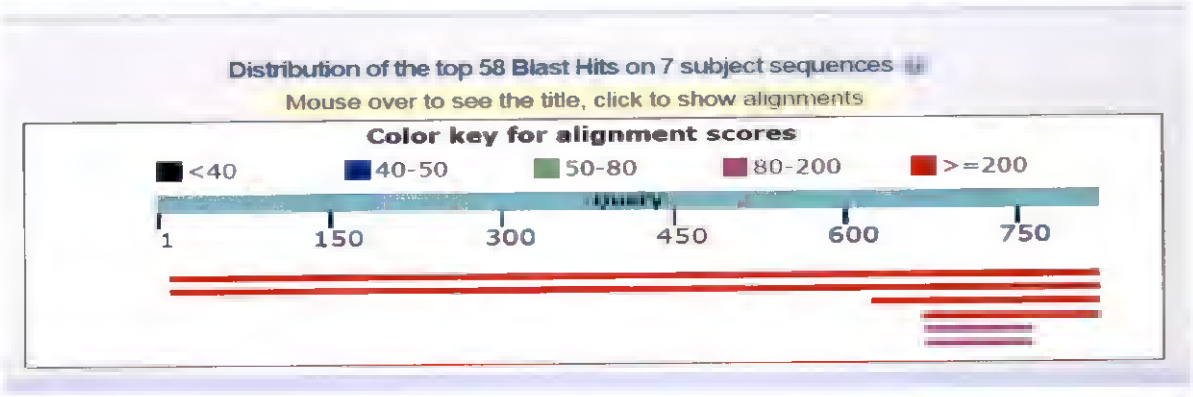
Sequences producing significant alignments:

Select All None Selected 0

Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
Capsicum annuum cultivar Nockwang capsanthin/capsorubin synthase, promoter region and partial sequence	1450	11041	98%	0.0	99%	DQ907615.1
Capsicum annuum gene encoding capsanthin/capsorubin synthase, promoter region	1448	10744	98%	0.0	99%	Y14165.1
C.annuum capsanthin-capsorubin synthase gene	344	344	24%	2e-90	98%	X77289.1
C.annuum (Yolo Wonder) chromoplastic oxydo-reductase gene	239	239	18%	1e-58	95%	X78030.1
C.annuum mRNA for capsanthin/capsorubin synthase	115	115	7%	2e-21	100%	X76165.1
Solanum lycopersicum cultivar Purdue_89281 chromoplast-specific lycopene beta-cyclase (Cyc-B) gene, promoter region al	104	104	11%	4e-18	87%	KP233163.1
Solanum arcanum cultivar G 32591 Cyc B protein gene, 5' UTR and partial cds	104	104	11%	4e-18	87%	EU937155.1

Fig 7a: BLASTn of Byadagi Kaddi *Ccs promoter* sequence showing 99 per cent similarity with Ccs gene accession DQ907615.1



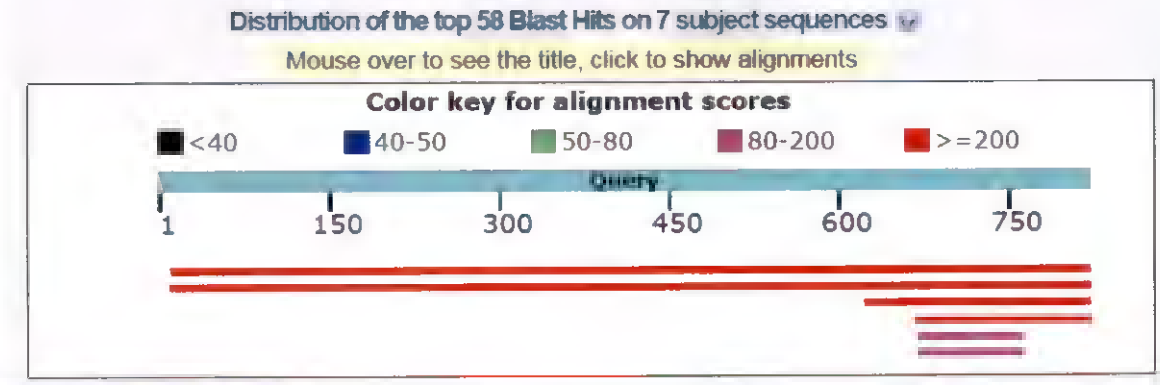
Sequences producing significant alignments:

Select: All None Selected 0

Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
Capsicum annuum cultivar Nockwang capsanthin-capsorubin synthase gene, promoter region and partial sequence	1450	11041	98%	0.0	99%	DQ907615.1
Capsicum annuum gene encoding capsanthin-capsorubin synthase, promoter region	1448	10744	98%	0.0	99%	Y14165.1
C.annuum capsanthin-capsorubin synthase gene	344	344	24%	2e-90	98%	X77289.1
C.annuum (Yolo Wonder) chromoplastic oxido-reductase gene	239	239	18%	1e-58	95%	X78030.1
C.annuum mRNA for capsanthin-capsorubin synthase	115	115	7%	2e-21	100%	X76165.1
Solanum lycopersicum cultivar 'Purdue_89281' chromoplast-specific lycopene beta cyclase (Cyc-B) gene, promoter region a	104	104	11%	4e-18	87%	KP233163.1
Solanum arcanum cultivar G 32591 Cyc B protein gene, 5' UTR and partial cds	104	104	11%	4e-18	87%	EU937155.1

Fig 7b: BLASTn of Byadagi Dabbi *Ccs promoter* sequence showing 99 per cent similarity with Ccs gene accession DQ907615.1



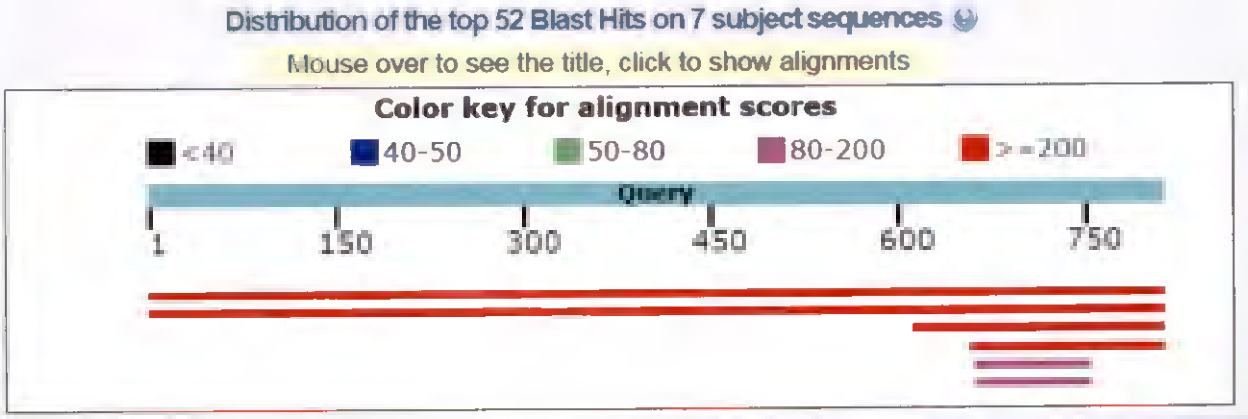
Sequences producing significant alignments:

Select All None Selected 0

Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
[redacted] m annuum cultivar Nockwang capsanthin-capsorubin synthase, promoter region and partial sequence	1450	11041	98%	0.0	99%	DQ907615.1
C.apsicum annuum gene encoding capsanthin:capsorubin synthase, promoter region	1448	10744	98%	0.0	99%	Y14165.1
C.annuum capsanthin:capsorubin synthase gene	344	344	24%	2e-90	98%	X77289.1
C.annuum (Yolo Wonder) chromoplastic oxydo-reductase gene	239	239	18%	1e-58	95%	X78030.1
C.annuum mRNA for capsanthin:capsorubin synthase	115	115	7%	2e-21	100%	X76105.1
[redacted] lycopersicum cultivar Purdue, 89281 chromoplast-specific lycopene beta-cyclase (Cyc-B) gene, promoter region ar	104	104	11%	4e-18	87%	KP233163.1
Solanum arcanum cultivar G.32591 Cyc B protein gene, 5' UTR and partial cds	104	104	11%	4e-18	87%	EU937155.1

Fig 7c: BLASTn of Ujwala *Ccs promoter* sequence showing 99 per cent similarity with Ccs gene promoter accession DQ907615.1



Sequences producing significant alignments:

Select: All None Selected:0

Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
Capsicum annuum cultivar Nockwang capsanthin/capsorubin synthase, promoter region and partial sequence	1079	10560	100%	0.0	91%	DQ907615.1
Capsicum annuum gene encoding capsanthin/capsorubin synthase, promoter region	1066	10437	100%	0.0	91%	Y14165.1
C.annuum capsanthin/capsorubin synthase gene	339	339	24%	9e-89	98%	X77289.1
C.annuum (Yolo Wonder) chromoplastic oxydo-reductase gene	233	233	19%	4e-57	94%	X78030.1
C.annuum mRNA for capsanthin/capsorubin synthase	115	115	7%	2e-21	100%	X76185.1
Solanum lycopersicum cultivar Purdue_89281 chromoplast-specific lycopene beta-cyclase (Cyc-B) gene, promoter region ar	99.0	99.0	11%	2e-16	86%	KP233163.1
Solanum arcanum cultivar G 32591 Cyc B protein gene, 5' UTR and partial cds	99.0	99.0	11%	2e-16	86%	EU937155.1

Fig 7d: BLASTn of Anugraha *Ccs promoter* sequence showing 91 per cent similarity with Ccs promoter gene accession DQ907615.1

Genotypes	Expected size of the amplicon	Size of the amplified sequence obtained	Similarity per cent	Query coverage	e-value	Accession showing highest similarity
Byadagi Kaddi	920	811	99	99	0.0	DQ907615.1
Byadagi Dabbi	920	812	99	99	0.0	DQ907615.1
Ujwala	920	815	99	99	0.0	DQ907615.1
Anugraha	920	812	91	100	0.0	DQ907615.1
Vellayani samrudhi	920	810	90	100	0.0	DQ907615.1
Vellayani thejus	1200	709	91	100	0.0	DQ907615.1
CC 8-1	1200	244	99	100	2e-109	Y14165.1

Table 10: BLASTn analysis of Ccs promoter sequence in different genotypes

4.6.5 BLASTn analysis of Ccs promoter sequence of Vellayani Samrudhi

The sequences of Vellayani samrudhi was analysed using BLASTn programme. The result showed that the sequence of Vellayani samrudhi had 91 per cent similarity with Ccs gene accession DQ907615.1 shown in (Fig7e) a *Capsicum annuum* cultivar landrace Nockwang, *Capsanthin capsorubin synthase* promoter region partial sequence.

4.6.6 BLASTn analysis of Ccs promoter sequence of Vellayani Thejus

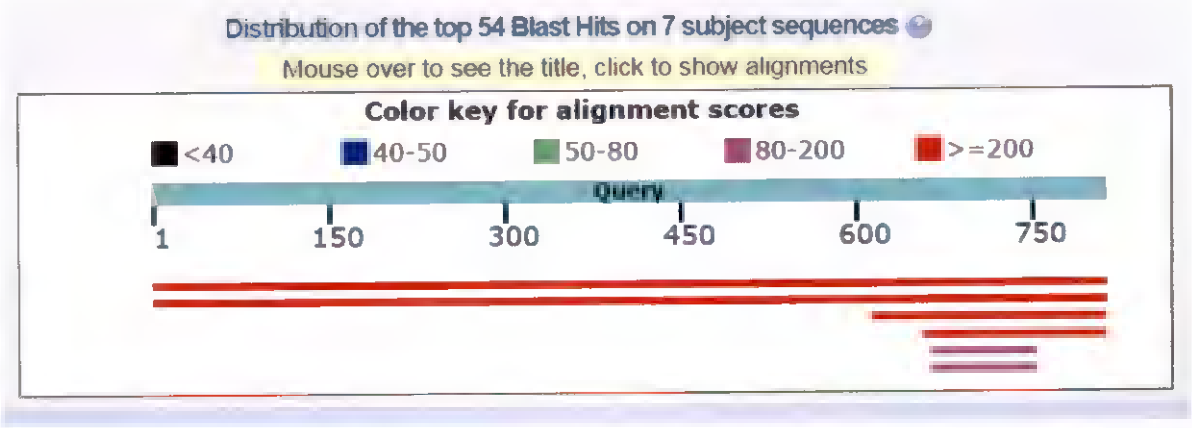
The sequences of Vellayani Thejus were analysed using BLASTn programme. The result showed that the sequence of Vellayani Thejus had 90 per cent similarity with Ccs gene accession DQ907615.1 shown in (Fig7f) a *Capsicum annuum* cultivar landrace Nockwang, *Capsanthin capsorubin synthase* promoter region partial sequence.

4.6.7 BLASTn analysis of Ccs promoter sequence of CC8-1

The sequences of CC8-1 were analysed using BLASTn programme. The result showed that the sequence of CC8-1 had 97 per cent similarity with Ccs gene accession Y14165.1 shown in (Fig7g) a *Capsicum annuum* gene encoding *Capsanthin capsorubin synthase* promoter region.

4.7 Multiple sequence alignment of *Ccs* promoter sequence

Multiple sequence alignment of promoter regions were performed (Fig8) to identify the presence of structural changes in the sequences (Table 11a- 11d). From the study it was found Byadagi kaddi had deletion in the sequence between 147-151bp region and also having SNP's adenine (A) at 239th position and Adenine instead of thymine was found at 764th position compared with reference sequence Y14165.1



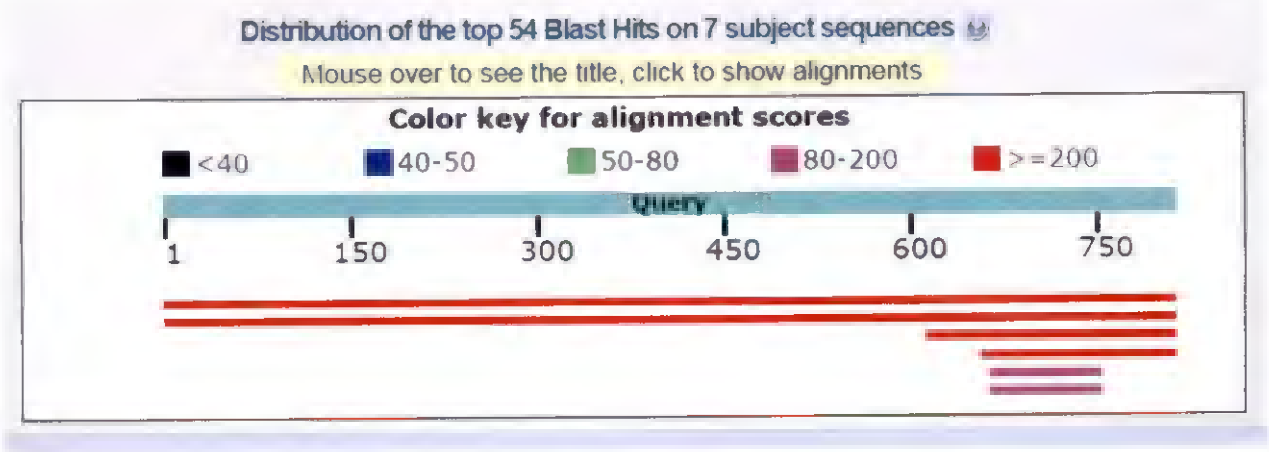
Sequences producing significant alignments:

Select All None Selected 0

Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
Capsicum annuum cultivar Nockwang capsanthin/capsorubin synthase, promoter region and partial sequence	1062	10795	100%	0.0	91%	DQ907615.1
Capsicum annuum gene encoding capsanthin/capsorubin synthase, promoter region	1050	10641	100%	0.0	90%	Y14165.1
C.annuum capsanthin/capsorubin synthase gene	344	344	24%	2e-90	98%	X77289.1
C.annuum (Yolo Wonder) chromoplastic oxydo-reductase gene	233	233	19%	4e-57	94%	X76160.1
C.annuum mRNA for capsanthin/capsorubin synthase	115	115	7%	2e-21	100%	X76165.1
Solanum lycopersicum cultivar Purdue_89281 chromoplast-specific lycopene beta-cyclase (Cyc-B) gene, promoter region ar	102	102	11%	1e-17	87%	KP233163.1
Solanum arcanum cultivar G 32591 Cyc B protein gene, 5' UTR and partial cds	102	102	11%	1e-17	87%	EU937155.1

Fig 7e: BLASTn of Vellayani Samrudhi *Ccs promoter* sequence showing 91 per cent similarity with Ccs gene accession DQ907615.1



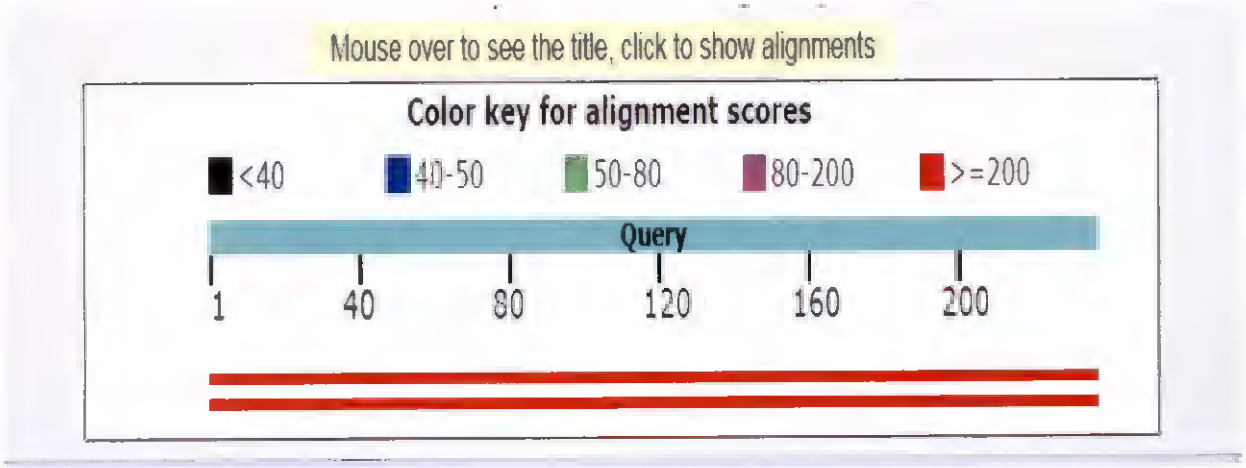
Sequences producing significant alignments:

Select All None Selected 0

Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
Capsicum annuum cultivar Nockwang capsanthin/capsorubin synthase, promoter region and partial sequence	1035	9079	100%	0.0	90%	DQ907615.1
Capsicum annuum gene encoding capsanthin/capsorubin synthase, promoter region	1022	8892	100%	0.0	90%	Y14165.1
C.annuum capsanthin/capsorubin synthase gene	344	344	24%	2e-90	98%	X77289.1
C.annuum (Yolo Wonder) chromoplastic oxydo-reductase gene	239	239	19%	9e-59	95%	X78030.1
C.annuum mRNA for capsanthin/capsorubin synthase	115	115	7%	2e-21	100%	X76165.1
Solanum lycopersicum cultivar Purdue_89281 chromoplast-specific lycopene beta-cyclase (Cyc-B) gene, promoter region and	104	104	11%	4e-18	87%	KP233163.1
Solanum arcanum cultivar G 32591 Cyc B protein gene, 5' UTR and partial cds	104	104	11%	4e-18	87%	EU937155.1

Fig 7f: BLASTn of Vellayani Thejus *Ccs promoter* sequence showing 90 per cent similarity with Ccs promoter gene accession DQ907615.1



Sequences producing significant alignments:

Select All None Selected 0

Alignments

Description	Max score	Total score	Query cover	E value	Ident	Accession
Capsicum annuum gene encoding capsanthin/capsorubin synthase, promoter region	405	3055	100%	2e-109	97%	Y14165.1
Capsicum annuum cultivar Nockwang capsanthin/capsorubin synthase, promoter region and partial sequence	399	3068	100%	1e-107	97%	DQ907615.1

Fig 7g: BLASTn of CC8-1 *Ccs promoter* sequence showing 97 per cent similarity with Ccs promoter gene accession Y14165.1

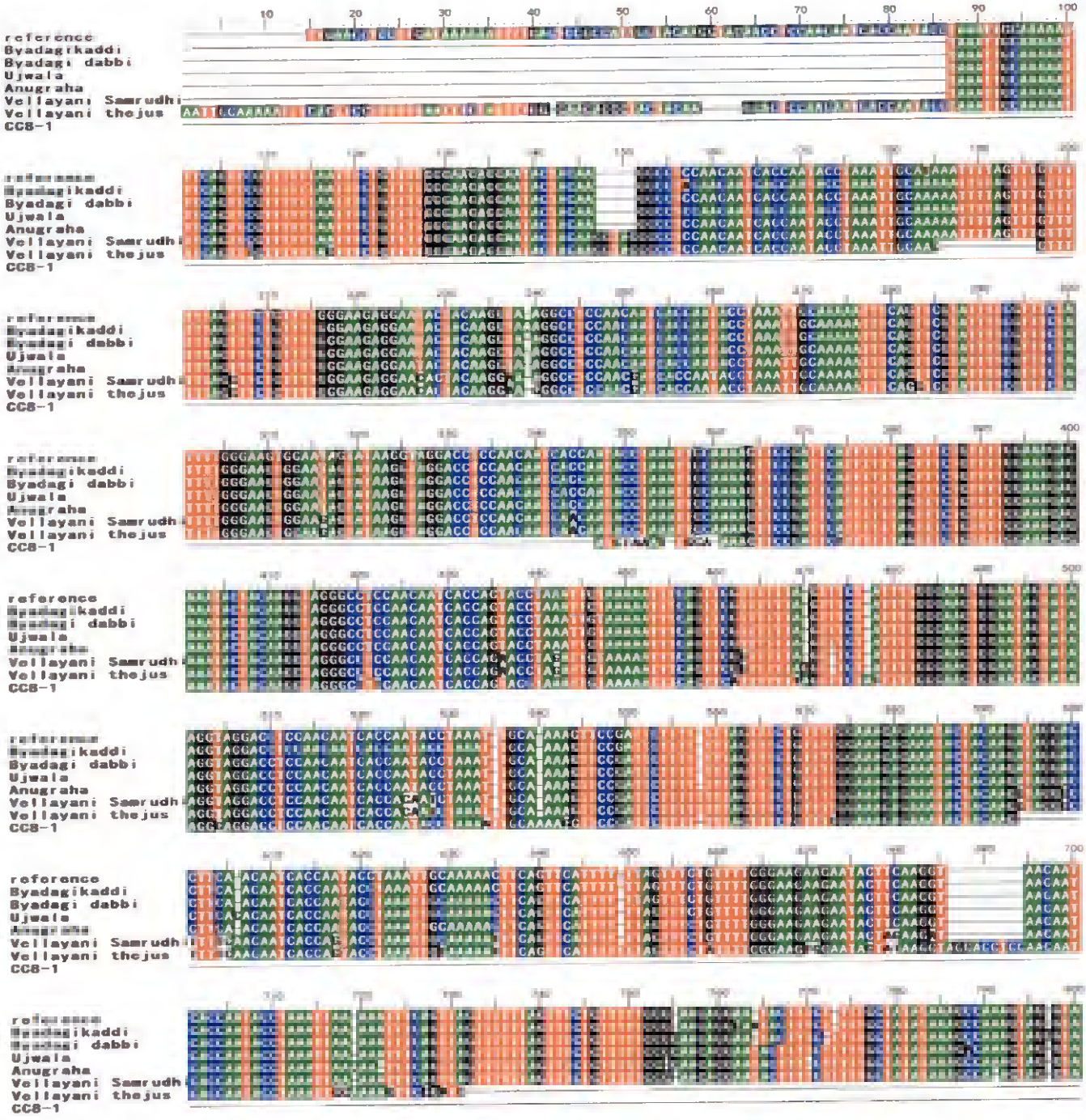


Fig. 8: Multiple sequence alignment of nucleotide sequences of Ccs promoter region with reference sequence (NCBI GI: Y14165.1)

Byadagi Dabbi had one SNP *i.e* guanine at 797th position and several deletions was noticed in the sequence between 147-151bp, 239, 470, 719, 764, 772th positions.

Ujwala had 5bp deletion at 147-151th position and several single nucleotide deletions were found at 239, 470, 719, 764, 772, 797bp positions. It was found that Ujwala had one SNP adenine instead of cytosine (C) at 788th position.

Anugraha had one SNP adenine (A) at 719th positon and several single nucleotide deletions was found at 239, 470, 764, 722, 754, 772, 797bp positions.

Vellayani Samrudhi, an Orange fruit line of *Capsicum frutescens* had several SNP's T436A, A442G, T463G, C604T, A605C, A617G, T679A, C680G and T found at 773rd position, several deletions was found at 239,470,473,719,755,791,797bp positions in the promoter region compared with reference sequence.

Vellayani Thejus, had several SNP's T206G, T227C, T237A, A240G, A251G, T283G, T316G, C344A, T436A, A442G, T463G, 470G, A599G, C601T, C604T, A605C, A617G, C629T, C659T, A671G, C677G, T679A, C680T, A717G, 719A, T723G, A727C, T729A, A731C and several deletions was found at 186-196bp, 239, 473, 754, 755, 764, 772, 773, 788, 791, 797th positions in the promoter region of the gene.

In yellow fruited line CC8-1 *Capsicum chinense* SNPs was found at C350T, A374T, C420T, C422T, G462A, A469T and (A) at 470th positions were noticed and several deletions was present throughout the promoter sequence.

4.7.1 Analysis of Ccs *promoter* region in chilli genotypes

Sequences contain a tandem repeat structure in the promoter region. Tandem sequence repeated three times in the sequence with insertion or deletion of nucleotides. Some *cis* regulatory elements like heat stress related *cis*-elements (HSE), Myb binding site (MYBPZM), Light responsive elements, TATA box, CAAT box in the promoter regions from different species (Fig.9a – 9b).

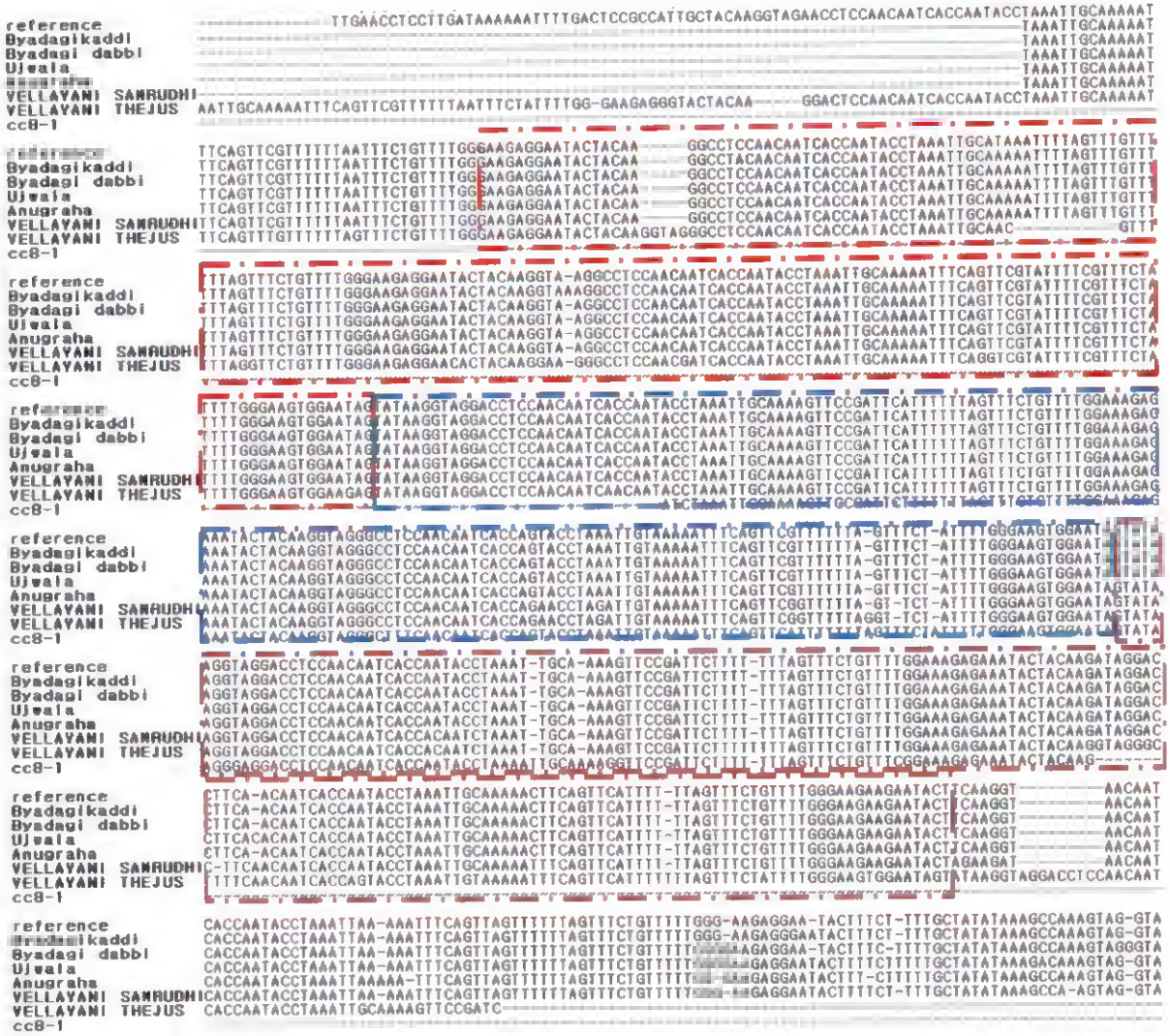


Fig 9a: Sequence comparisons in the *Ccs* promoter region among the *Capsicum* genotypes, boxes (Red, purple, orange) shown the three repeat units found in this region.

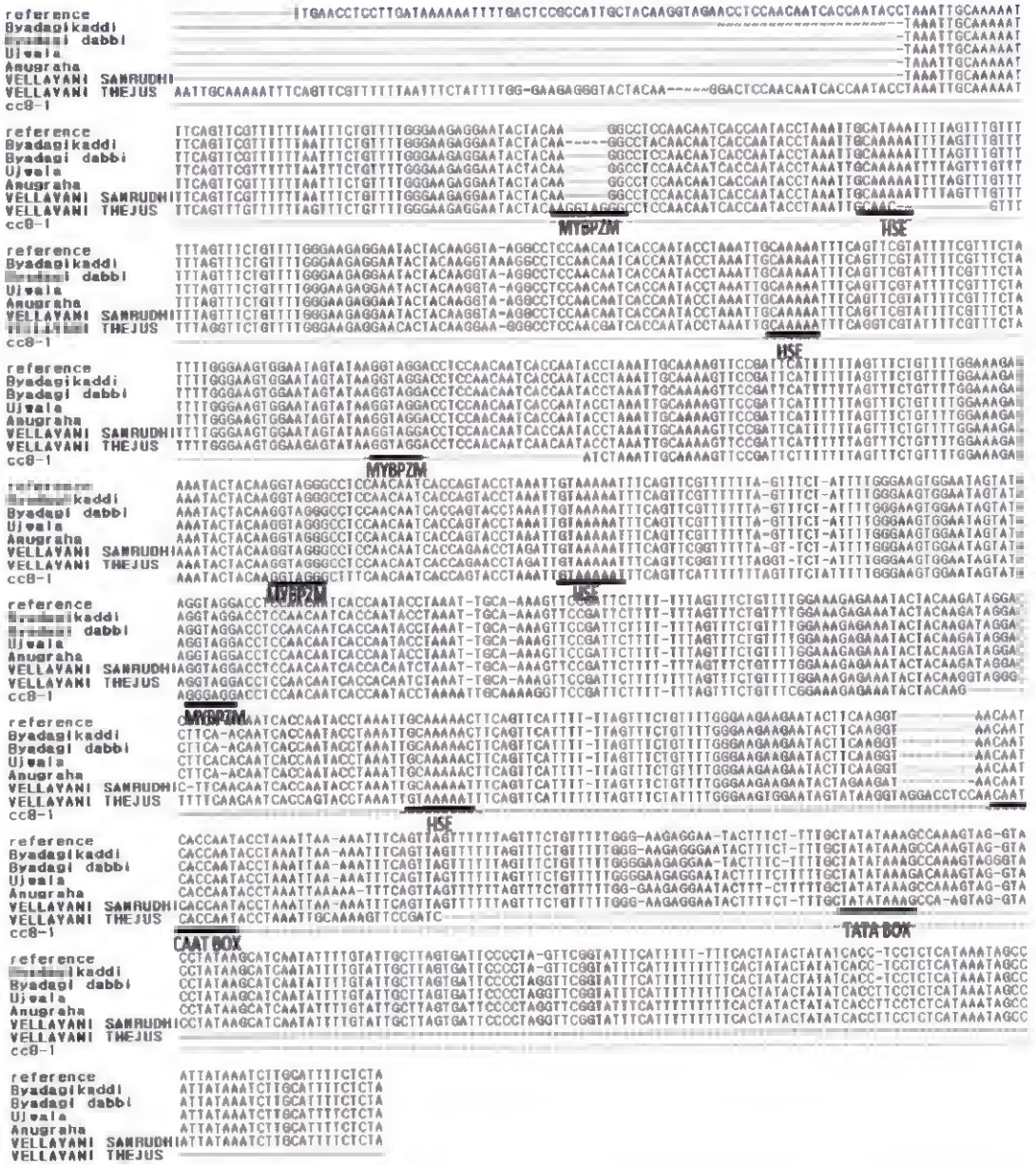


Fig. 9b: Multiple sequence alignment of *Ccs* promoter sequences showing Cis regulatory elements (MYBPZM, HSE, TATA box and CAAT box).

GENOTYPES	Nucleotide position bp									
	147-151	186-196	206	227	237	239	240	251	283	316
Byadagi Kaddi	-	AAATTTCAGTT	T	T	T	A	A	A	T	T
Byadagi Dabbi	-	AAATTTCAGTT	T	T	T	-	A	A	T	T
Ujwala	-	AAATTTCAGTT	T	T	T	-	A	A	T	T
Anugraha	-	AAATTTCAGTT	T	T	T	-	A	A	T	T
Vellayani Samrudhi	AGTAG	AAATTTCAGTT	T	T	T	-	A	A	T	T
Vellayani Thejus	AGTAG	Deletion	G	C	A	-	G	G	G	G
CC8-1	-	-	-	-	-	-	-	-	-	-

Table 11a: Single Nucleotide Polymorphisms identified in promoter region sequence of Capsanthin capsorubin synthase gene (147-316bp)

GENOTYPES	Nucleotide position bp												
	344	350	374	420	422	436	442	462	463	469	470	473	599
Byadagi Kaddi	C	C	A	C	C	T	A	G	T	A	-	T	A
Byadagi Dabbi	C	C	A	C	C	T	A	G	T	A	-	T	A
Ujwala	C	C	A	C	C	T	A	G	T	A	-	T	A
Anugraha	C	C	A	C	C	T	A	G	T	A	-	T	A
Vellayani Samrudhi	C	C	A	C	C	A	G	G	G	A	-	-	A
Vellayani Thejus	A	C	A	C	C	A	G	G	G	A	**G**	-	**G**
CC8-1	-	T	T	T	T	T	A	A	T	T	**A**	T	-

Table 11b: Single Nucleotide Polymorphisms identified in promoter region sequence of Capsanthin capsorubin synthase gene (344-599bp)

GENOTYPES	Nucleotide position bp																			
	601	604	605	617	629	635	659	670	671	677	599	601	604	605	617	629	635	659	670	671
Byadagi Kaddi	C	C	A	A	C	C	G	A	A	C	A	C	C	A	A	C	C	G	A	A
Byadagi Dabbi	C	C	A	A	C	C	G	A	A	C	A	C	C	A	A	C	C	G	A	A
Ujwala	C	C	A	A	C	C	G	A	A	C	A	C	C	A	A	C	C	G	A	A
Anugraha	C	C	A	A	C	C	G	A	A	C	A	C	C	A	A	C	C	G	A	A
Vellayani Samrudhi	C	T	C	G	C	C	G	A	A	C	A	C	T	C	G	C	C	G	A	A
Vellayani Thejus	**T**	T	C	G	**T**	**T**	**A**	T	G	G	**G**	**T**	T	C	G	**T**	**T**	**A**	T	G
CC8-1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 11c: Single Nucleotide Polymorphisms identified in promoter region sequence of Capsanthin capsorubin synthase gene (601-671bp)

GENOTYPES	Nucleotide position bp																		
	677	679	680	717	718	719	722	723	727	729	731	754	755	764	772	773	788	791	797
Byadagi Kaddi	C	T	C	A	A	-	A	T	A	T	A	G	-	**A**	T	-	C	A	-
Byadagi Dabbi	C	T	C	A	A	-	A	T	A	T	A	G	G	-	-	T	C	A	**G**
Ujwala	C	T	C	A	A	-	A	T	A	T	A	G	G	T	T	T	A	A	-
Anugraha	C	T	C	A	A	**A**	-	T	A	T	A	-	G	T	T	T	C	A	-
Vellayani Samrudhi	C	A	**G**	A	A	-	A	T	A	T	A	G	-	T	T	T	C	-	-
Vellayani Thejus	G	A	**T**	G	C	**A**	A	G	C	**A**	C	-	-	-	-	-	-	-	-
CC8-1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 11d: Single Nucleotide Polymorphisms identified in promoter region sequence of Capsanthin capsorubin synthase gene (677- 797bp)

4.8 Primer binding and amplification of *Ccs* gene

4.8.1 Mapping of Capsanthin capsorubin synthase gene position

Capsanthin capsorubin synthase, (*Ccs*) chromoplastic *Capsicum annuum* gene located in Chromosome six between 9497216 – 9500911bp. The total size of *Ccs* gene is 3695bp. *Ccs* cds gene specific primer binding site was 18bp downstream region of the sequence. The *Ccs promoter* lies upstream 5'of the protein coding region. Genome sequence of *Ccs* is obtained from *Capsicum annuum* cultivar Zunla-1 chromosome 6, Pepper Zunla 1 Ref NC_029982.1 whole genome sequence.

CATACAATCCCAGCAAATAAAATTATATCTACGAACTACGAAAAAGAAAAAAGATGAAG
AAGAAAATAATAATAAAAACTAAAATAAAAAAAGGTGTGTTTACCTTTTTGGGGGCAGC
AAAACGAGACTACGAGCTTCGTTGGAATCAACCACCACCGAAAAACGTCGCCGGTAAAT
TATATTTAAGTAAATCCGATATTTTGTAAAATAAATATTTAAGTAACAAATTATATAATCT
CGTATAACTGAACATGATAGTATATAATCGCTACTTAAAATGAAAAAATCACGTACGAAA
AATATTTTAAGTTTTTTTTTTTAAATTTGGATGAAATAATTATTTTGGTTTATTAAAAACTG
AAGAAACTCAGGATTAAATTTAGTCGTGGAGGGCAAAAATTAGGTGTCAACAGTAATTGT
TCAGGAATAGGAAATGTTGAAATTTAGTAATCCATTAGGAAGTTAGTTAATGTTACGTGA
GCTTAATAAGGCTCCGGCTTAGCATGTCAGTAGCAGACTCATTGTATAGGCCACAGCTTG
GCCTCTGTTTCAGTCATCTTAGGCTTGCATCAGTTGGTCTTGCACTTGTGCTGTCTGTGTCA
TAATTCAACCACCCAAGGGTGTGGCCTAGTGGTTCAGTGAAGTGGGTTGAGTACCATGAG
GTCTCAGGTTCAATCTCCAACAGAGACAAACATTAGGTGAATTATTTTCATCTGTTCTAAG
CTTGGTAGACAGGTACCTGTTGTTGGTGGGAGGTGGCAGGTATCCGTGGAATTAGTCGAG
TGCACGCAATCTGGCCCGAACACCATAGTTATCCAAAAAAACAAAAAACAAAGAACTAA
TTCAACTAATCTAACTACTGTAATTACATTGAGATATGCAACACGTGGAATTTCTTTTCCA
AGAAAGTTCTCAGTTTGAAGCATTTTGGTCTTATATATTCGTCAGTCTGAGGTTCGTAATT
TTTGCTTGTACAGGGGTTTGTTAACATGCTTGATTGTGCTCTTCCTTTACTTGATAATTGCT
GCTTGTTGTGGGGGCATCACTCTACCTTCCTGCAGATCATGAATTTTCTGACAAGGAACAC
AGTGAAAGGTTATTGTATAACTAATCCAGTGGATTCTCATTCTGGCACCTTTAAAAGTACA
TGTTCAGAAAAGAATGATAAGGTTTGTATTGTTGTTGACGAGGCTCTGTGCCTTTTCAAAT
TTGTAAATGTTCTGGTGGACTCCTAAACTGCCCTAAAGCTGGTATTGGTCTTCCATATTAT
GTATGTTATAGAGAAGTTCAGAACTTACATAAGTGCCTGTTTGTAAATATAAGGCCGTGTT
GTGACTTACCTTTTGCAAATATTTACTATGGCAATTCACATTTTTCAAATACGTGGTGGTT
GAGAGTTTGTGTATAGATGTTGCTTAGCAATAATTTGAGTTTGTGGGAGGTATTTCAAGAG
TTGTTAGGAAGTGGATTACCTTTGACTTTAAGGGGTAGTTTGGTAGAGTGTATTGAAAAA
GTTAATGCATGCATTAGCTTAATGTACTATTAGTACCTTGTTTGGTACATTGTTTTGCCCTA
TGTATAAGTAATGTTTCCATTACTTATACACACTATAGTGTATTGGGTGATGTGTTACTAA
TACCATCATTTTCCTATGTATTAGTAATGCAAAAGATTTCAACACATGCATTAACTTGATA
ATGACAATCTTGCCCCTCAAATTCAACATTTCTCAAGTTCTTTTGTTTGTGGATTTTCTAAA
AAAAAAAAATAGATTAATATGTAAAATAATTGGAAAAAAATGAATATTAAATTTTTTATA
TATTGAAAATGAAAACACTGGAGTGTTGTTTTATTGAAAAACTTCATCTTTGCATAGATGC
ATTTTTGTGAGGAAGTAGTAGAAGATATCAACCATTTCTTCATTAATTGAAAATCTTTAGT
CAACTGTAGAATTTAAAATTCAACATTTTGGTGTATGCCTGGATCTTTTTCGGAATTTCTG

TTTACACGATAGAGAAATGGAATAACTAAACCACTGCTTCCCCTGGCACACCACTCTTGG
AGCCATATGTTGGGTTAGTTAGAAAGAAAGGAACACTAAATTTTTCGAAGACGAATGATA
TTGTACGGCAACTCAACACAAGTGGAGAGGAAATTGCTTATCTTAGAATAAAACATAAAA
AAGAATTTACCTTGTGTTTTAGAATAAAACACAGAAAAGGAGTTATCTTAAAAAAAAAAA
AAAGGAAGAAAGGATCAACTTTGAAAAAATTATGCGTTTTACCATTTAAAATTGAAAAGA
TAAATATTCTTTTAAAAGAAGTTGTTGAATGGAAAATATTGGAAGAATTTCATTTCATTTT
ACAAAAATAAAGAGTGTAGAGGGTATTTTTGTAAATCAATATTTTTTCTATAAAAAATAT
ATAAGAAATATTATTTTAATACATCAAATCAAATACTGTATAAGAAATAATGTTAACATA
ATTAATGCAAGTATAGCTAATACCAACATTACTAATGCAAGTATTACTAATACACCATAT
TCTATATTAATCTTATATACTCTACCAAACGACCCTAAGTGTGTATCTATATCCTCCGAGA
ATTTGGAATTTGCAAATTCCAAGTTTTGTATCTCCCTTTCCCAGAAATTAAGATAATTCTG
GTGCTTTTAGCATTAGAAAAGTATTTATTGGGTAGGGAAATGTCATGACTTCACAGCATTA
AGCATCAAGGGTATAACTTAATGAAATAGTGGTCAATGAATTATATTGAGAATGACGAGG
TCTCTGTTCCAACTTTGGTAGACTTTGGAAATGCTCGTCTGGACGCCGCCCATTCTTTCTA
GTCTTGGTGCCATTCTATTTGGTCTGAGAATGGCATGATGCCAAATTCTACCTTTTCACAA
TGAGCATTCGACCTACTCTTCTTTTTTCGACTCATTTGACCTACTAGGCATTGGCCAACTTG
GCTAACCACTTGAGGAACTAGAGTTCGGATTCAATAGAATCTAATAATTTTAATCAAAAG
ACTTCATGTATATTGAAAAATCTATTTATAACTAACTTTAAATCGGCCTTTACGTATCGAC
GTAATCAAAATTGTGTCAGCTTGCCACGTGGGGTCTAGTATGAGTTTGAAATTGGTCATA
GGGGCCCCAATTCCACTAATACAGCTGCCGTCCATGCACTACAAGACAAATACACCACTA
TGTTTGTTAGTGCTTGGTAAATGTAAAACAAACTTTTGATGAGAATCTATTCGTGGCATCG
AAGTGCTGCAAATTGGCTTTTACCTCTGCTACTTCAAGCCTCACTGATTTTCACCCCAACT
TTCTCATTTCCCTTTCAAGGATTTGATTTTCCAGTTGGGCATGTTAAAAACAACAATTTTCC
TCAAAACTGTAGAAATGATTTCTCATATTTTAATCAGTCAAATTATTTAAACAAGAAGTTG
ATTTTTTTTTAATTTTTTTTTTTACAAAAAAATTTCAAATGTCAAGTAAGATTTTTCAAATT
GAAACTGAATAAGCTGCGACTTTAGAAACAAAAAACTAAGATAAGTAAAAATACCAAAA
AGAGTGAATCACATCAATTGAATTCTTCCAACAGTTCGTTTTTTAGTTTCTGTTTTGGGAA
GAGGAGTACTACAAGGTAGGACCTCCAACAATCAACAATATCTAAGTTGCAAAAGTTTTT
GTGCGTTTTTTAGTTTCTGTTTCGAGAAGAGGAATACTACAAGTTCGTTTTTTAGTTTCTGT
TTTGGGAAGAGGAGTACTGCAAGGTAGGACCTCCAACAATTATCAATATCTAAATTGCAA
AAATTTCAGTTCGTTTTTAGTTTCTGTTTCGGGAAAAGGAATACTACAAGTTCGTTTTTTA
GTTTCTATTTTGGGAAGAGGAGTACTACAAGGTAGGACCTCCAATACCTAAATTGCAAAA
ATTTCAGTTCGTTTTTTAGTTTCAGTTTAGGGAAGAGGAATACTACAAGGTAGGACCTCCA
ACAATCATCAATACCTAAATTGCAAAAATTTCAGTTCGTTTTTTAGTTTATGTTTTGGGAA
GAAGAATACTACAAGGCAGTGGTGGAGCTACCTTATGATTAGGGGGTTCATCCGAACCTC
CTTCGACGGAAAATTATACTATTTTTATAAGTGAAAATTATTTTTTATGTATATATAATTG
ATGTTGAACCCCCTTCGGTTAGTTCATGTATCTATATTTTTTTATTTTGAACCCCGATGAAA
ATTTGGGCTCCGCCACTGCTACAAGGTAGGACCTCCAACAATCACCAATACCTAAATTGC
AAAAATTTCAGTTTGCTTTTTAGTTTCTGTTTTGGGAAGAGGAATACTACAAGGTAGGACC
TCCAACAATCACCAATACCTAAATTGCAACGGTTTTTTTAGTTTCTGTTTTGGGAAGAGGAA
TACTACATGGTAGGGCCTCCAACAATCACCAATACCTAAATTGCAAAAATTTCAGTTCGT
ATTTTCGTTTCTATTTTGGGAAGTGGAATAGTATAAGGTAGGACCTCCAACAATCACCAAT
ACCTAAATTGCAAAAGTTCCGATTCATTTTTTAGTTTCTGTTTTGGAAAGAGAAATACTAC
AAGGTAGGGCCTACAACAATCACCAGTACCTAAATTGTAAAAATTTCAGTTCGTTTTTTA
GTTTCTATTTTGAGAAGAGGAATGCTACAAGGTAGGGCCTACAACAATCACCAGTACCTA
AATTGTAAAAATTTCAGTTCGTTTTTTAGTTTCTGTTTTGGGAAGAGGAATACTACAAGGT
AGGGCCTCCAACAATCAGCAATACCTAAATTACAAAAATTTCAATTCGTTTTTTAGTTTCT

GTTTTGGGAAGAGGAATACTACAAGGCAGTGGCGGAGCTACCTTATGATTAGGGGTTCAT
CCGAACCTCCTTCGACGGAAAATTATACTATTTTTGTATAGTAAAAATTATTTTTTATGTA
TATATAATTGATGTTGAACCCTCTTCGGTTAGTTTGTGTATCTATATTTTTTTATT
TTTTGACTCCGCCATTGCTACAAGGTAGAACCTCCAACAATCACCA
ATACCTAAATTGCAAAAATTTCAGTTCGTTTTTTAATTTCTGTTTTGGGAAGAGGAATACTAC
AAGGCCTCCAACAATCACCAATACCTAAATTGCATAAATTTTAGTTTGTTTTTTAGTTTCTGT
TTTGGGAAGAGGAATACTACAAGGTAAGGCCTCCAACAATCACCAATACCTAAATTGCAAAA
ATTTCAGTTCGTATTTTCGTTTCTATTTTGGGAAGTGGAATAGTATAAGGTAGGACCTCCAA
CAATCACCAATACCTAAATTGCAAAAGTTCCGATTCATTTTTTAGTTTCTGTTTTGGAAAGAG
AAATACTACAAGGTAGGGCCTCCAACAATCACCAGTACCTAAATTGTAAAAATTTCAGTTCG
TTTTTTAGTTTCTATTTTGGGAAGTGGAATAGTATAAGGTAGGACCTCCAACAATCACCAAT
ACCTAAATTGCAAAAGTTCCGATTCTTTTTTTAGTTTCTGTTTTGGAAAGAGAAATACTACAA
GATAGGACCTTCAACAATCACCAATACCTAAATTGCAAAAACTTCAGTTCATTTTTTAGTTTC
TGTTTTGGGAAGAAGAATACTTCAAGGTAACAATCACCAATACCTAAATTAAAAATTTCAGT
TAGTTTTTTAGTTTCTGTTTTTGGGAAGAGGAATACTTTCTTTTGCTATATAAAGCCAAAGTA
GGTACCTATAAGCATCAATATTTTGTATTGCTTAGTGATTCCCCTAGTTCGGTATTTCATTTT
TTTTCACTATACTATATCACCTCCTCTCATAAATAGCCATTATAAATCTTGCATTTTCTCTAAT
*GGAAACCCTTCTAAAG*CCTTTTCCATCTCCTTTACTTTCCATTCCTACTCCTAACATGT
ATAGTTTCAAACACAACTCCACTTTTCCAAATCCAACCAAACAAAAAGATTCAAGAA
AGTTCCATTATAGAAACAAAAGCAGTACACATTTTTGTAGCTTTCTTGATTTAGCACC
CACATCAAAGCCAGAGTCTTTAGATGTTAACATCTCATGGGTTGATACTGATCTGGA
CGGGGCTGAATTCGACGTGATCATCATTGGAACTGGCCCTGCCGGGCTTCGGCTAG
CTGAACAAGTTTCTAAATATGGTATTAAGGTATGTTGCGTTGACCCTTCACCACTTTC
CATGTGGCCAAATAATTATGGTGTTTGGGTTGATGAGTTTGAAAAGTTGGGATTAGA
AGATTGTCTAGATCATAAGTGGCCTGTGAGTTGTGTTCATATAAGTGATCACAAGAC
TAAGTATTTGGACAGACCATATGGTAGAGTAAGTAGAAAGAAGTTGAAGTTGAAATT
GTTGAATAGTTGTGTTGAAAATAGAGTGAAGTTTTATAAAGCCAAGGTTTTGAAAGT
GAAGCATGAAGAATTTGAGTCTTCGATTGTTTGTGATGATGGTAGGAAGATAAGCGG
TAGCTTGATTGTTGATGCAAGTGGCTATGCTAGTGATTTTATAGAGTATGACAAGCC
AAGAAACCATGGTTATCAAGTTGCTCATGGGATTTTAGCAGAAGTTGATAATCATCC
ATTTGATTTGGATAAAATGATGCTTATGGATTGGAGGGATTCTCATTTAGGTAATGA
GCCATATCTGAGGGTGAAGAATACTAAAGAACCAACATTCTTGTATGCAATGCCATT
TGATAGGAATTTGGTATTCTTGGAAGAGACTTCTTTAGTGAGTCGGCCTATGTTATC
GTATATGGAAGTGAAAAGAAGGATGGTAGCAAGATTAAGACATTTGGGGATCAAAG
TGAGAAGTGTCCTTGAGGAAGAGAAGTGTGTGATCACTATGGGAGGACCACTTCCG
CGGATTCCTCAAAATGTTATGGCTATTGGTGGGACTTCAGGGATAGTTCATCCATCG
TCTGGGTACATGGTGGCTCGTAGCATGGCATTGGCACCAGTACTGGCTGAGGCCAT
CGTCGAAAGCCTTGGCTCAACAAGAATGATAAGAGGGTCTCAACTTTACCATAGAGT
TTGGAATGGTTTGTGGCCTTCGGATAGAAGACGTGTTAGAGAATGTTATTGTTTCGG
AATGGAGACTTTGTTGAAGCTTGATTTGGAAGGTACTAGGAGATTGTTTGATGCTTT
CTTTGATGTTGATCCCAAGTACTGGCACGGGTTCCTTTCTTCAAGATTGTCTGTCAA
AGAACTTGCTGTACTCAGTTTGTACCTTTTTGGACATGCCTCTAATTTGGCTAGGTT
GGATATTGTTACAAAGTGCACTGTCCCCTTGGTTAAACTGCTGGGCAATCTAGCAAT
AGAGAGCCTTTGAATTAATATGATAGTTTTGAAGCACTGTTTTCATTTTAATTTCTTAGGT
TATTTTCATCTTTTCTCAATGCAAAAGTGAAACAAAAGCTATACACATTGTCATCGTTGTT
CAAACTCAGACAAGTTTGCCTAGCTCTATGTATTTATCCTTAACATATGTATTCATCAAAT
TCGAAATATACAATGCATTGGACAAAAGTATAGAGCCACAATCCGATACCAAGTCTGTAT

137

TTGGAAGCACTGTCTAATTGTTATGGTTACCAAACACTTTGAATTGGCTGGATAATAACAA
ACAGGAAATTTATGTTTTTAATCATTAACAGCAAATTGGGAAAGCAAGAATTATTAGGAA
AGTTAATATAGTGTCTTGGTTATTCTAATGGAGTGGGTTATGCAAATTAAGTTCCCTTATC
AAAGTTTGGTTTATGAACTGCTCCACTCTTGTCCCTCTTAAAAGCCTTAATCCCAACATGT
ACCACCAAAGAATTGAGCTGCTCCATCAGATCCTTTGAGAATGTTAATATGTTATTTAAAT
GAAGGACTGAATGATTATGAGGATGCAATGCATAGGTTTAATTACCAGTTATCTGTAAAT
TGTCTTCTTTGCCATTATTTTAAAAGTTTAATAACAAGTGTAACATCTACAAAGAGTTGAT
AATTACAAAGCAGCTACTAGTTTTAGGTTCCTCCACCAACTAGACTAAATAAGCCATGGC
CCCAAGTAACAAGACAAAGGTTTGATCAAAAGAAAAAAAAAAAGTAAAAAGATTAAATA
TCACATGAGATTAATAGAAGCGAATGAGAACTTAAGAGTAATTTGATAAAGAAATTCGA
AAACCAACCATTTTGGAAAGTTCTTGTATAGTTTTACGTCTTAACTCCTCCTGTAACTTTTA
AGTGTCAAGTGGGCCGGGCCGGTTCCGGTTCAGGTTGGGGTTAGGTGGGGCGGGCCGGTT
CCGGGTTCAACAGTAAAAATATTGAAAACAACCCTAAACCGGCCCACTTAACCCAACCCG
GTTAAACTCACCTAAATCCGGTGAAACCCGATTAAAAATCGGTCAAACCCGTTTAAAAAA
AATGAAAAATCCAATATACACTACATACACTACAATTATATACACTTCAATATACATTAT
ATACACTACAATTATGCAATATATATTTTTTAAAATATATATAAACAATTACACATATATA
CGTATCATTCAACACATAATTATATATGACTATACGTACTACTTTAAATAAAACATATACT
ACAAATCTACAATAGTACAATATATATATATATATACACTACAATATATATATATATATAG
TTATATACTAATTTATAAAACATATACACATTTATACGTTAAATATATACGCTTATAGAAG
ATATATACACATATATACGCACCATTCAATA

Green bold - *Ccs* cds

Blue bold italicized - *Ccs* promoter

Red bold - Primer binding region of *Ccs* cds

[redacted] - Primer binding region of *Ccs* promoter

136

4.9 Elucidation of carotenoid metabolic pathway for production of capsanthin and capsorubin

Carotenoid metabolic pathway was downloaded from STRING pathway data base and analysis of pathway for the number of genes and enzymes involved was done. Details are provided in (Table12).

Carotenoid pathway genes	Carotenoid pathway enzymes
IDI	Iso pentenyl di phosphate isomerase
Psy	phytoene synthase
Pds	phytoene desaturase
Zds	ζ-carotene desaturase
Lcyβ	Lycopene β cyclase
Lcyε	Lycopene epsilon cyclase
Zep	Zeaxanthin epoxidase
LCY	lycopene cyclase
CCD	Carotenoid cleavage dioxygenase
Ccs	Capsanthin capsorubin synthase
NCED	9-cis-epoxycarotenoid dioxygenase
CrtISO	ζ-carotene isomerase
CrtL-e1	Lycopene epsilon cyclase
CYP97A29	β-carotene hydroxylase (P450) 97A29
VDE	violaxanthin de-epoxidase

Table 12: Genes and enzymes present in the carotenoid metabolic pathway

4.9.1 Transcriptome data mining and analysis

In this study, a general search was performed for the retrieval of fruit ESTs from *Capsicum annum* and fifty ESTs were downloaded from NCBI's dbEST database. The sequences were subjected to cleaning for the removal of vector contaminants and splice sites using SeqClean and Phrap programs and valid sequences were selected for analysis. Assembly of the selected EST sequences was done using genome assembler and significant contigs were selected.

>Contig1
GTCGACCTTTGGTACGCTTGATACAGATTATGTTTTCAATTTTTTAGAGTAAAGCTAGGGAA
CTTCTTATGTAAGCCTTGAAATTTTGAACCCATGATCGGGGAATAGAACAAGTGACCATCC
CCTTGTTATCCAGCTGCAGAGTATATTCCTGAATTTTGTGACAACAATAAATTTCTTTTCCA
TCTAATAAT

>Contig2
GTAGACGATTAGTACGCTTGATTCGAGCGGCCGCCCGGGCAGGTAATTCGAATATTTCCTT
AGATATACGTAGTCTACGTATTGAGATTTGAGTCAGAGGGGAATTTTCGGAAGGAGAACCC
CTTGTAGCAATTCCCTTGATGTTGTTGACCAAATTTCTCTGTAGTATAGTAAGTAAATGTGA
TTTCATTTCATTTGTAATAAAAAAATAATATTATCTC

>Contig3
CCCCGTCTGCAGGAATTCGATACTTTTAACTTGGTGCTGACATAATTAGTGATTCAATTTTC
GGGGAAGTCTAGTTTGTCTACTTTTTTTTTTTTTTTTTGGGGGGTATTATTTTCTGTTGTCTA
ATTTATCACACTAGCAAGTGGCAAAAATGTTTTAACTACG

>Contig4
CCTGCAATTGAATCACATATTACAGGCGGAGTTAAAGTGTAATTCATCCTAAACATTATAC
TATTAGAACTTTCTGATGGTGCTTTGAATTTTTCTTTTAGATAGTGCACTTTTATGTCGTATG
AAGTTTTAAATGAGAGAAGTTTTTCAAGTCAAGATGCTGACAATTAGATTTATTTCATCTGA
AACTTTGTTTCATGTAGTGCACTGATGTTGGAAGTAACCATGCAGTGCTGGTAACCATTTAG
ACATGTTATTATGTCAATACTAGCTCTGATTTTCGAAGTATAATACAAACTTGTGTTCCTTC
GTCATTTTCC

>Contig5
GGAATGAAATGAAATGAAATTACTTACTAACTGCAGAGAAATTTGGTCAACAACATCAAG
GGAATTGCTACAAAAAATTCTCCTTCCGAAAATTTGTCTCTGACTCAAATCTCAATACGTAG
ACTACGTATATCTAAGGAAATATTCGAATTACCTGCCCCGGGAGACTTTTTTCCATGTCTTG
CTGGCCGGCTTTTAACAACCTGCGTATTAGTTCCTTACGCCCTTTTCAAAGGCCCTTCCAAA
AAATAACGGTTGGGCGTGTATTCCCCCATGTTACCCGCTCCAGTTGCCCGGGCCCTTTACAC
CCGTCTTTGA

>Contig6
TTAACCCCTTCTGCAGGCAATTCGATACGAGGGGGTCTGAGAAAGAAGCATAAGGCATCTT
TTGCAACTTGAGGGGTAGTAATGTTGGTATTGAGTATGGAATGTTGTGGCTTCCATCTCCCA
GTGCTTGTTTTGTGCTTGTGATGGACATTCTTTTTCGTTTTTCATGTAGTAATTCCAATTTTA
GTAACTTAAAAGTTGGTGTTCCTCAAGTATAAAAGCAATGTTTCTTGCTCCCGTGAACCACT
AAGGGGACCCCTAACCCCAACGGCCCCCCCGTGCCCAAATTACCCTCTCTCCCCCCCCCTT
ACCGGGGGCCCCCTGTCCCCATTTAAACCCCTTGGGAAACCCAAAAACCCCCCCTTTCACC
ACATTTCCTTTTCCCCCAGGAGCTGCAAAAACCCCCAGCCCCCCAAACTAATACCCCTCCC

140

ACAACCCAACCCCCCCTCACAACCGGATGGGAACCCATAAATGGCCCCACCCAATTTCCCC
CTCACCAAAATTTTTACACCCCCCCAAGGGTAACCCCGCCCCCCCCCCC

>Contig7
TTGACGCCTTACAGATTATGTTTTCAATTTTTTAGAGTAAAGCTAGGGAATTTCTTATGTAA
GCCTTGAAATTTTGAACCCATGATCCAAAATAGAACAAGTGATCATTGACTTGTTATCCAG
CTGCAGAGTATATTCTGCATTTTGTGACAACAATAAATTCCTTTTCCAGC

These contigs were further searched for identification of their respective genes using Blastn against *Capsicum annum* genome. The gene names of the seven contigs and their corresponding chromosomal location are given in Table13.

Contigs	Name	Chromosomal Location
Contig 1	Bromo domain containing protein 9 (LOC107866786)	4
Contig 2	Uncharacterized (LOC107853383) mRNA	Un
Contig 3	E3 Ubiquitin Ligase SUD1 (LOC107840255)	8
Contig 4	YIF-1B like (LOC107875322)	6
Contig 5	Uncharacterized (LOC107853383) mRNA	Un
Contig 6	Zinc finger A20 and AN1 domain containing stress-associated protein 8-like (LOC107839016) transcript variant X3	1
Contig 7	Bromo domain containing protein 9 (LOC107866786)	4

Table13: Details of contigs generated with EST's using genome assembler

Further mapping of these contigs with carotinoid biosynthesis pathway using KEGG mapping showed none of the seven genes in the pathway.

4.10 Gene regulatory network analysis

Gene regulatory network is an undirected graph representing biological systems and behavior using biological compounds that are generated from omics data. Gene regulatory network composed of nodes and edges representing genes and mutual co-expression relationships.

In this study genes associated with the carotenoid biosynthesis pathway and transcriptomics data for carotenoid biosynthetic genes from STRING database were used. The transcriptomic data were analyzed by gene regulatory network provided a network of transcription factors with several major highly interacting nodes by using cytoscape a web based software.

Carotenoid Gene regulatory network contained 94 nodes and 100 edges and included many genes associated with the general Carotenoid biosynthesis pathway, Steroid pathway, kinases, transcription factors and several corresponding to proteins of unknown function. Gene annotation analysis showed that many of the genes were associated with carotenoid biosynthesis processes. The main Carotenoid biosynthetic pathway genes (Phytoene Synthase, Phytoene desaturase, Zeta carotene isomerase, Zetacarotene desaturase, Carotene isomerase, lycopene cyclase, zeaxnthin epoxidase, vialoxathin epoxidase, Capsanthin capsorubin synthase, neoxanthion epoxidase), as well as some transcription factors and transferase/transport proteins were densely connected, while several other genes involved in further upstream and downstream pathways were also within the network (Fig10).

In pepper (*Capsicum annuum*) carotenoids are synthesized from five carbon isoprenoid units like isopentenyl diphosphate (IPP) and its double-bond isomer dimethylallyl diphosphate (DMAPP). Produced by the plastidal MEP (2-C-methyl-d-erythritol-4-phosphate) pathway. The first step in carotenoid biosynthesis is condensation of two GGPP molecules catalyzed by PSY (phytoene synthase) to form phytoene, and interacting protein partners like DXS (1-deoxy-d-xylulose-5-phosphate

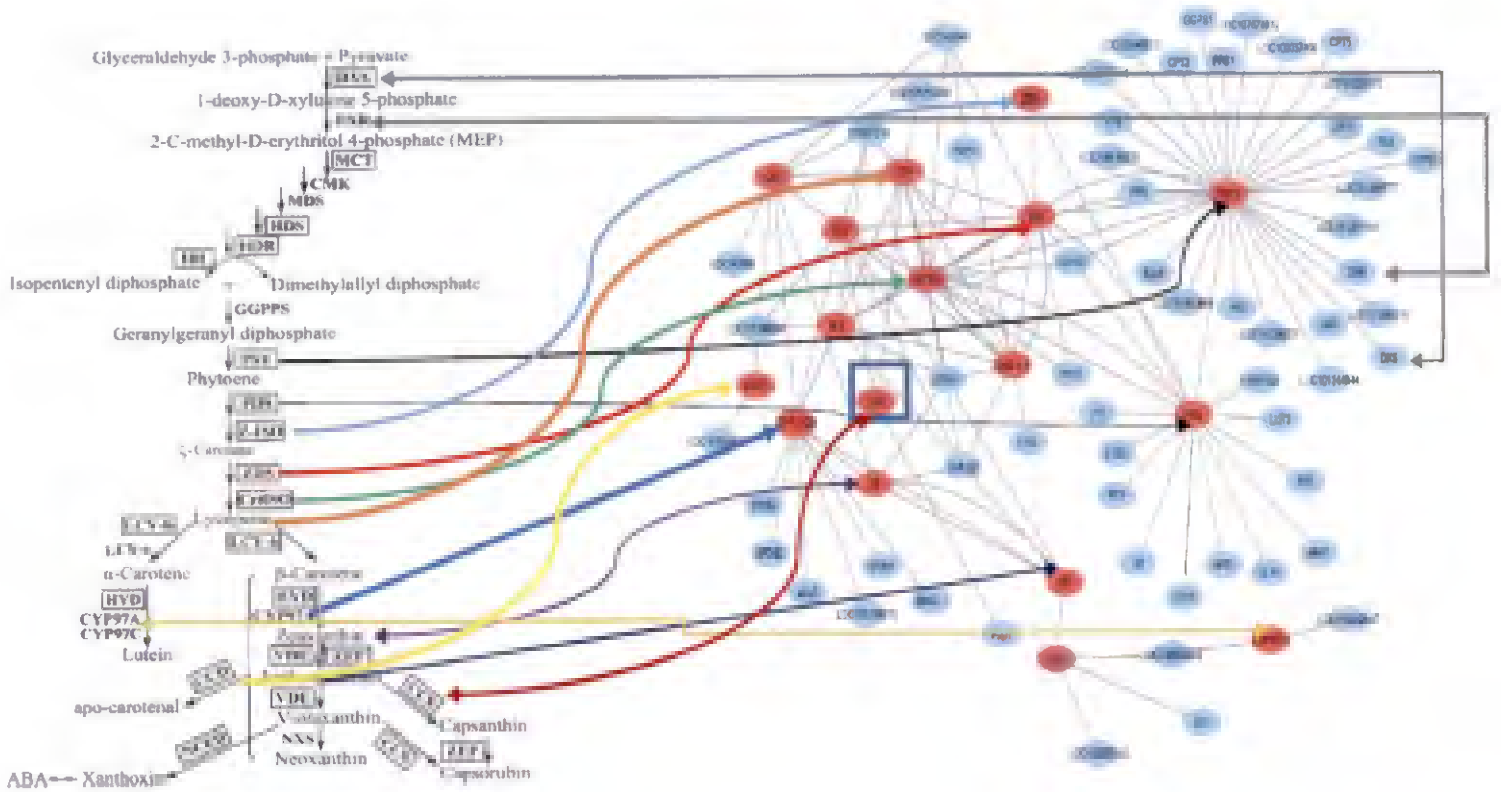


Fig 10: Comparision of carotenoid biosynthesis pathway, with gene regulatory Network and their interaction partners

synthase), GGPS1 (Geranylgeranyl pyrophosphate synthase1), GGPS2 (Geranylgeranyl pyrophosphate synthase 2), DXR(1-deoxy-d-xylulose-5-phosphate reductase), GGR (Geranyl geranyl reductase), and other proteins shown in (Table14).

Phytoene is colour less compound undergoes series of desaturation and isomerization reaction to synthesize conjugated double bonds and transform 15 cis-phytoene into all-trans-lycopene by PDS (Phytoene desaurase) and correlated enzymes Z-ISO (Zeta carotene isomerase), ZDS (Zeta-carotene desaturase), CrtISO (Carotenoid isomerase), LYC (Lycopene cyclase), LUT2 (Lutein deficient 2), ABA1 (ABA deficient1), IM (Immutans), CYP97A3, (Cytochrome P450, family 97, subfamily A, polypeptide 3), CLA1 (Chloroplastos alterados 1), and CCD1,(Carotenoid cleavage dioxygenase 1).

All-trans-lycopene requires CRTISO (Carotenoid isomerase), capable of isomerizing cis bonds at 7, 9 and 70, 90 positions to convert tetra-cislycopene to all-trans-lycopene. CRTISO (Carotenoid isomerase), enzyme interacted with LYC (Lycopene cyclase) undergoescyclization of lycopene distinguished by different end groups by addition of (β ring) or (ε-type ring). Rings synthesized by enzymeslycopene β-cyclase andcrtL-e-1 (lycopene epsilon cyclase), β -chain interacted with CrtRb2 (β - carotene hydroxylase2) protein and convert β-carotene into zeaxanthin. And CrtRb2 interacting with zeaxnthin epoxidase, vialoxathin epoxidase enzyme to convert zeaxanthin into vialoxanthin it is a reversible reaction. And vialoxanthin converted into capsorubin with the help of capsanthin capsorubin synthase enzyme. CCS interrelated to pap (fibrillin) chromoplatic protein involved in colour formation. Vialoxantin again converted to neoxanthin via neoxanthin synthase. ϵ –strand correlated with crtL-e-1 (lycopene epsilon cyclase), and catalyzes single cyclization reaction which converts lycopene into delta carotene and alpha carotene and converted into leutin by CYP97A29, CRTRB1, CRTRB2 and CYP97C11 enzymes.

144

4.11 Enzymes involved in carotenoid metabolic pathway in *Capsicum annuum* Chilli

(*Capsicum annuum*) comprised more than 20 core enzymes involved in carotenoid metabolism through 2-C-methyl-d-erythrito1-4-phosphate (MEP) pathway (Table14). Analysis of chromoplast proteins from chilli identified 17 enzymes present in the carotenoid metabolic pathway. The MEP pathway uses glyceraldehyde 3-phosphate and pyruvate as primary substrate to produce deoxy-D-xylulose 5-phosphate (*DXP*); this is catalyzed by *DXP* synthase, and *DXS*, 1-deoxy-d-xylulose-5-phosphate synthase. MEP is consequently formed via reduction of DXP by the enzyme DXP reductoisomerase (*DXR*). IPP and DMAPP are formed after a number of subsequent steps and successively undergo a sequential series of condensation with the help of MCT, 4-hydroxy-3- methylbut-2-en-1-y1 diphosphate synthase (*HDS*), HDR, and ID enzymes to yield, geranylgeranyl diphosphate (*GGPP*) the precursor of carotenoid biosynthesis.

Carotenoid biosynthesis started with the condensation of two *GGPP* molecules by phytoene synthase (*PSY*) to form phytoene. (Tomato contains three tissue-specific isoforms with PSY1) Phytoene is converted into tetra-cis-lycopene by the introduction of four double bonds via enzymes: phytoene desaturase (*PDS*) and z-carotene desaturase (ZDS). Tetra-cislycopene is converted into all-trans-lycopene via isomerization of cis bonds at 7, 9 and 70, 90 positions by carotenoid isomerase (*CRtISO*) enzyme.

For the cyclases in plants the desired substrate is all-trans-lycopene. The cyclization of lycopene is a critical step in carotenoid metabolism and generates carotenoid diversity distinguished by different cyclic end groups: either by adding beta 03-ring) or epsilon (a-ring). These rings are generated by the enzymes lycopene 13-cyclase (β-*LCY*) and lycopene c-cyclase (ε-LCY), respectively.β-LCY catalyzes cyclization of both ends of lycopene, c-LCY typically cyclizes only one end, forming

the monocyclic (a-carotene).and other enzymes include production of keto cyclic carotenoids *i.e.* capsanthin and capsorubin, which are the signature pigments of the pepper family catalyzed by capsanthin-capsorubin synthase (Ccs) by the convertion of epoxy-carotenoids antheraxanthin into capsanthin and violaxanthin into capsorubin.

Pathway genes	Interacting proteins
(Psy) phytoene synthase	PDS, CrtISO,GGPS2 ZDS, CYP51 LOC101250973, FPS1, SppS, DXS, GGPS1 LOC101246371, LOC101249319, LOC100037490, LOC101264777, LOC104649012, LOC100736533, LOC107873488, LOC101254144, CrtISO, GGR, LOC101244544, DXR, HY5, FTB, FPS, FTA, LOC107874612, LOC107862834, CPT6, CPT3, PSY1
(Pds) phytoene desaturase	PSY, Z-ISO, CrtISO, LYC, ZDS, LUT2, ABA1, IM, CYP97A3, CLA1, PEP, CCD1, PDS
(ZISO) ζ-carotene isomerase	ZDS
(Zds) ζ-carotene desaturase	AL2, CrtISO, LCY1, PSY1, LOC543649, CrtL-e-1, PSY2
(CrtISO) ζ-carotene isomerase carotenoid isomerase	ZDS, PSY11, CrtL-e-1, PSY2, LCY1, LOC543649 GGPS2
(LCY)lycopene cyclase	CYP97A29, ZDS, CrtISO, CrtR-b2, CrtL-e-1, CYP97C11 LOC101266560, CrtR-b1
(CrtL-e1)Lycopene epsilon cyclase	CYP97C1, CrtISO, ZDS, LCY1, LOC543649, CrtR-b2, CrtR-b1
(ZE) Zeaxanthin epoxidase	Gene B, VDE, CYP97A29, CrtR-b1
(Ccs) Capsanthin capsorubin synthase	NSY
(VDE) violaxanthin de-epoxidase	ZE, CYP97A29, CrtR-b1
(NSY) neoxanthin synthase	
(CYP97A29) β-carotene hydroxylase(P450) 97A29	AL2, NSY, CA1, SMO2-1, LOC101249215, SPDS2, SPDS3, ACL5
(CCD) Carotenoid cleavage dioxygenase	CCD7
(Crt-Z) Chloroplast-specific lycopene beta-cyclase	CYP97A29, CYP97C11, LCY1, ZE, LOC543696, LOC101266560 ,CrtL-e-1

Table14: Enzymes involved in carotenoid biosynthesis pathway and interacting proteins in *Capsicum annuum* based on network analysis

147

Discussion

5. DISCUSSION

Chilli (*Capsicum annuum* L.) is the most important solanaceous spice cum vegetable crop grown throughout the world. Carotenoids are important coloured compounds found in fruits which are normally fat soluble pigments derived from five carbon isoprene units that polymerized enzymatically to form 40 carbon structures. The characteristic red colour present in *Capsicum* spp. is resulting from capsanthin and capsorubin compounds, both these enzymes involved in carotenoid biosynthesis pathway (Govindarajan, 1985).

Capsanthin and capsorubin are the keto-carotenoids present in the thylakoid membrane of the chromoplasts. These are the major red pigments present in ripened fruits contributing upto 60 per cent of the total carotenoids. During ripening process capsanthin and capsorubin increase constantly (Moehs *et al.,* 2001). *Capsanthin capsorubin synthase* (*Ccs*) gene specifically express and synthesise carotenoid pigments in mature fruits at chromoplasts maturation stage by incorporating ketocarotenoids (Bouvier *et al.,* 1994). Capsanthin is highly stable and used as natural red colour in food industry, confectionaries, cosmetics, beverages, pharmaceuticals and even as a dye in textile industries.

Byadagi is an important *C. annuum* chilli type known for its deep red colour (108.92 ASTA) and negligible pungency (0.0045 per cent capsaicin).Demand for Byadagi chilli has increased enormously as a source of natural red colour. It is mainly exported as oleoresin which serves as a substitute to paprika oleoresin. In Byadagi chilli, the studies on characterization using molecular markers and evaluation were reported by Sandeep (2007) and Renuka (2014).

Investigations of carotenoid biosynthesis genes and their molecular mechanism for the control of carotenoid biosynthesis is an active research area for a number of vegetables and spices. Variation in specific alleles of carotenoid

biosynthetic enzymes have been altered for carotenoid production (Welsch *et al.*,2010 ; Yan *et al.*, 2010).

The present study was conducted to analyze *Capsanthin-capsorubin synthase* (*Ccs*) gene in Byadagi chilli and to elucidate the carotenoid metabolic pathway for production of capsanthin and capsorubin. The studies were focused on seven genetically distinct chilli varieties/accessions of three different *Capsicum* spp. based on colour at fully ripe fruit stage.

The results of the present investigations were discussed under two different headings

1. Analysis of *capsanthin capsorubin synthase* gene
2. Elucidation of carotenoid metabolic pathway

5.1 Analysis of capsanthin capsorubin synthase gene in *Capsicum*spp.

5.1.1 Isolation, Purification and Quantification of DNA

Young, immature leaves were used for isolating DNA from seven chilli genotypes. CTAB method reported by Roger and Bendich (1994) was used to extract the DNA from the samples. This method yielded sufficient quantity of good quality DNA. The electrophoresed DNA showed distinct band without shearing.

Grinding in liquid nitrogen was found to improve the quality of isolated DNA. For preventing the phenolic oxidation, the addition of antioxidants like β-mercaptoethanol during DNA isolation was found effective.

The disruption of the cell membrane of the sample was done with the help of detergents which were used in the extraction buffer of CTAB (Cetyl Trimethyl Ammonium Bromide). As a result of the action of detergent, the nucleic acid inside the cell was released into the extraction buffer and prevented the co-precipitation of the polysaccharides with nucleic acid by acting as a selective precipitant of nucleic

acid. CTAB being a cationic detergent solubilised the cell membrane and formed a complex with DNA (Sghaier and Mohammed, 2005).

CTAB along with PVP had an advantageous effect on the quality of DNA. (Sreenath *et al.*, 1992; Gallego and Martinez, 1996). It acts along with the NaCl which effectively disrupts the cell membrane and separate the polysaccharides. The EDTA in the extraction buffer chelates the Mg^{2+} ions, protecting the DNA from endonuclease activity. In the DNA extraction protocol, double treatment with chloroform: isoamylalcohol mixture and centrifugation removed the pigments and proteins from the nucleic acid. The DNA got precipitated with the addition of chilled isopropanol and the pellet obtained after the centrifugation was washed with 70 per cent alcohol followed by absolute alcohol to clear the traces of CTAB. The DNA was rehydrated in TE buffer (Wettasinghe and Peffley, 1998; Babu, 2000).

According to Wettasinghe and Peffley (1998), a DNA sample is reported to be of high quality if it has a band of high molecular weight with little smearing and a low amount of RNA. To remove the RNA contamination from the isolated DNA samples, RNase treatment was given (Gallego and Martinez, 1996 ; Wettasinghe and Peffley, 1998). In the present study, on electrophoresis the RNase treated DNA samples had shown a higher molecular weight, forming a single band just below the well. This indicated that the DNA was of good quality.

For checking the quality of DNA, the absorbance ratio was calculated as OD at 260/280 nm using spectrophotometer. Those samples which were having the ratio ranging 1.8-2.0 were considered to be of high quality DNA. In samples with RNA contamination, the absorbance ratio went beyond 2.0 and with protein contamination, the ratio went below 1.8. High quality DNA was used for the molecular analyses.

5.1.2 Analysis of *capsanthin capsorubin synthase* gene in *Capsicum* spp.

In the present study an investigation was made to analyze genetic variation of *Capsanthin capsorubin synthase* (*Ccs*) gene in different chilli varieties with varied

colour. In the study, two *Ccs* gene specific SSR markers, Ccs Cds and Ccs promoter (Table1). (Li *et al.,* 2013) were used for detecting specific variation between chilli genotypes.

PCR amplification with primer Ccs Cds generated single, clear, distinct band of about 1.5kb size in all red and yellow varieties (Plate5). The primer specifically amplified the *Capsanthin capsorubin synthase* (*Ccs*) region of the *y* locus in the genome of the *Capsicum* (Lefebvre *et al.,* 1998; Li *et al.*, 2013).

Sequence analysis of the amplicons of *Ccs Cds* primer set had showed 99 per cent identity with the *Capsanthin capsorubin synthase* gene (Fig1a-1g). Comparitive study of *Ccs* sequences using clustalX in different genotypes revealed several SNPs identified between the genotypes.

SNPs identified in Byadagi Kaddi and Byadagi Dabbi showed that the specific substitution of a novel allele at 425th amino acid position as lysine replaced by glutamic acid in the gene locus and this may lead to the increase of colour and increase in the catalytic activity of carotenoid cyclases in chillies as reported by (Mialoundama *et al.,* 2010).

Vellayani Samrudhi, an orange fruited genotype (*Capsicum frutescens*), had 10 SNPs G 107 A , C 129 A, C 235 G, T 630 C, A 312 T, G 831 A, A 834 G, A 1071 G, G 1272 A, A 1273 G and nucleotide deletions observed at 834,1043,1056,1072 positions in the coding region of the gene that led to change in amino acid coded *i.e* tyrosine (Y) replaced by serine (S) at 43th position, arginine (R) replaced glycine (G) 79th position and amino acid deletions are present at 278, 348, 352, 358 positions and glutamic acid (E) replaced by lysine (K) at 425th position of *Ccs* gene. Orange colour fruit is due to deletions in the *Capsanthin capsorubin synthase* gene which is in accordance with reports by Lang *et al.* (2004). Formation of red colour pepper fruits requires the regular expression of *Ccs* gene and adequate amount of substrate for *Ccs*, whereas deficiency of *Ccs* gene expression or substrate leads to synthesis of orange

or yellow fruit colours. Rodriguez-Uribe *et al.* (2012) reported that orange fruit colour formation was due to changes in the carotenoid biosynthesis genes and transcriptional and post transcriptional regulation of the genes.

Vellayani Thejus, a *Capsicum chinense* genotype had nine SNPs G 107 A, C 129 A, C 235 G, T 630 C, G 831 A, A 834 G, A 1071 G, G 1272 A, A 1273 G positions in the coding region of the gene that led to change in the amino acids coding sequence *i.e* Serine (S) replaced by tyrosine (Y) at 43th position, arginine (R) replaced glycine (G) at 79th position and glutamic acid (E) replaced by Lysine (K) at 425th position.

CC8-1 is a yellow colour fruit genotype (*Capsicum chinense*). Analysis of Ccs gene sequence showed presence of premature stop codon at 200th position *via* single base change at 599 bp that led to transcriptional silencing of *Capsanthin capsorubin synthase* gene. This mutation was preventing the Ccs expression to produce capsanthin and capsorubin in yellow cultivar in accordance with Ha *et al.* (2007) and Rodriguez-Uribe *et al.* (2012). Li *et al.* (2013) concluded that coding and promoter regions were present in yellow peppers and deletions or mutations in the *Ccs* gene is the cause for yellow colour.

Gene regulation and expression is a primary area for the study of chilli fruit colour, experiments were made to analyze the promoter sequence of *Ccs.* Amplification with *Ccs* promoter primer in all seven varieties were performed, Amplicon of 920bp in *Capsicum annuum*, (Byadagi kaddi and Byadagi dabbi, Ujwala, anugraha) and *Capsicum frutescens* (Vellayani samrudhi) genotypes was found whereas in *Capsicum chinense* genotypes (Vellayani Thejus and CC8-1), larger amplicon of 1200 bp size was found that was also in accordance with Ha *et al.* (2007) and Li *et al.* (2013).

The sequence of the amplicons of *Ccs* promoter primer set had shown 99 per cent identity with the *Capsanthin capsorubin synthase* gene promoter region. The

sequence analysis showed that Ccs promoter region having a single set repeat structure was located near transcriptional start point. It is also having heat responsive elements in each repeat which may explain increasing of capsanthin. Capsorubin increases actively to more sunlight (Kuntz *et al.,* 1998). *Ccs* promoter sequence containing ABA (abscisic acid-responsive), cis-acting elements like Myb binding sites and heat stress elements (HSE), suggest the possibility of Ccs expression regulation by all of these factors (Bouvier *et al.,* 1994).

5.1.3 Trnscriptome data mining and analysis

Fruit ESTs of *Capsicum annuum* retrieved from NCBI dbEST database. Sequences were cleaned and removed vector contaminants, splice sites. Valid sequences were selected for analysis. Contigs were searched for identification of genes using Blastn. Mapping of the contigs with carotenoid biosynthesis pathway using KEGG mapping showed none of the seven genes in the pathway. The available EST's in the data base may be from immature fruits and that may be the reason for the non availability of genes of carotenoid metabolic pathway when mapped with KEGG.

5.1.4 Elucidation of carotenoid metabolic pathway

Carotenoid gene regulatory network (GRN) contained 94 nodes and 100 edges and included many genes associated with the general carotenoid biosynthesis pathway, steroid pathway, kinases, transcription factors and several corresponding to proteins of unknown function.Gene annotation analysis showed that many of the genes were associated with carotenoid biosynthesis processes. The main carotenoid biosynthesis pathway genes phytoene synthase, phytoene desaturase,zeta carotene isomerase, zetacarotene desaturase, carotene isomerase, lycopene cyclase, zeaxnthin epoxidase, vialoxathin epoxidase, capsanthin capsorubin synthase, neoxanthion epoxidase, as well as some transcription factors and transferase/transport proteins

were densely connected (Nisar *et al.,* 2015). Among the pathway genes, phytoene synthase had the highest number of interactive proteins.

Elucidation of pepper carotenoid metabolic pathway revealed that metabolic pathway contained seventeen main pathway enzymes. The public domain contains twenty-nine full coding genes. In this study, it is clear that *Ccs* gene is responsible for colour formation. Red colour in chillies is due to the presence of some compounds called capsanthin and capsorubin. The carotenoids are produced in chillies as a result of the combination of products which are formed in the carotenoid biosynthesis pathway. GGPP (geranyl geranyl pyrophosphate) is the precursor for carotenoid pathway. *Capsanthin capsorubin synthase* enzyme produces carotenoid end products by catalyzing antheraxanthin into capsanthin and violaxanthin into capsorubin in chili pepper fruits (Camara, 1980).

In the present experiment, primers Ccs cds and Ccs promoter amplified the Ccs gene and among this, Ccs cds, primer is more specific and gave single, clear, intact band in all varieties including yellow colour accession. From this it is concluded that *Ccs* gene is responsible for the production of major colour compounds capsanthin and capsorubin. The sequence variations in the *Ccs* gene can be a reason for orange and yellow colour in chilli.

Elucidation of carotenoid metabolic pathway in *Capsicum annuum* revealed that carotenoid biosynthesis pathway having many genes which are associated with the carotenoid biosynthesis pathway. Among the pathway genes phytoene synthase enzyme had highest number of interacting proteins. *Psy*, *pds*, and *zds*, which catalyze the first three committed step in carotenoid biosynthesis, were showed to be the major key regulators of carotenoid biosynthesis in plants (Matthews, 2003; Toledo-Ortiz, 2010; Tuan, 2011). This finding suggests the existence of *Psy* isoforms in chilli.

Ccs variants could provide candidate molecular markers for selection of pepper lines with high dark red color and yellow carotenids.

Our study demonstrates that SNPs relating to alteration of Lysine at 425th position of capsanthin capsorubin synthase polypeptide by Glutamic acid is correlated to change in the colour in Byadagi Kaddi and Byadagi Dabbi. This study could also help to further breeding programmes by SNP phenotyping and marker assisted selection.

The most significant finding from the present investigations is the presence of SNP at 425th position of Ccs polypeptide chain that could distinguish Byadagi cultivars from other genotypes. In Byadagi chilli cultivars, at 425th position of CCS polypeptide chain, Lysine was found replaced by Glutamic acid. In rest of the genotypes studied Lysine was present at 425th position.

Byadagi Kaddi is different from Byadagi Dabbi in one SNP. At 43rd position, Tyrosine was found replaced by phenyl alanine in Byadagi Kaddi. In rest of *Capsicum annuum* genotypes at 43rd position Tyrosine was present but in *Capsicum frutescens* and *Capsicum chinense*, serine was found instead of Tyrosine at 43rd position. Vellayani Samrudhi, an orange coloured variety has deletions at 278, 348 and 358th position. In the yellow coloured accession CC8-1, premature Stop Codon (UAG) was seen at 200th position which stopped further transcription and translation of Capsanthin-capsorubin synthase gene. Vellayani Thejus and CC8-1 share the same sequence except for the stop codon at 200th position. The varieties Ujwala and Anugraha share the same amino acid through out the CCs gene sequence. Anugraha is a near isogenic line of the cross between Ujawala and Pusa Jwala and that may be the reason for the same amino acid sequence.

The identified SNPs (15 Nos.) in the present study have to be further characterized and validated. Many SNP annotation tools are available for analysis of human genome sequence data. But such tools are lacking for the analysis of plant genome sequence data. But the identified SNPs when matched with the reported SNPs in Capcisum sp., it was found that the SNPs identified in the present

investigations were new. However, transcriptome analysis of Ccs gene in the different genotypes, homology modeling the Ccs enzyme and prediction of active sites could derive more information on the identified SNPs

Summary

6. SUMMARY

The study entitled "Analysis of *Capsanthin-capsorubin synthase* gene in Byadagi chilli (*Capsicum annuum* L.) and elucidation of carotenoid metabolic pathway" was undertaken at the CPBMB, College of Horticulture during the period 2014 - 2016.

The present study was conducted to analyse *Capsanthin-capsorubin synthase* gene (*Ccs*) in Byadagi chilli and to elucidate the carotenoid metabolic pathway for production of capsanthin and capsorubin. The studies were focused on seven genetically distinct chilli varieties/accessions of three different *Capsicum* spp. based on colour at fully ripe fruit stage. The accessions selected were Byadagi Kaddi, Byadagi Dabbi, Ujwala, Anugraha (*Capsicum annuum*), Vellayani Samrudhi (*Capsicum frutescens*), Vellayani Thejus and CC8-1 (*Capsicum chinense*).

Genomic DNA was isolated from tender leaves of one month old plants by CTAB method. Two chilli Ccs gene specific SSR primers *viz.* Ccs Cds and Ccs promoter were used to amplify the *Ccs* gene. The amplified PCR products obtained with *Ccs* Cds and *Ccs* promoter were sequenced by outsourcing and sequence data were analyzed using bioinformatics tools.

The salient findings of the study were as follows:-

- The protocol suggested by Rogers and Bendich (1994) was found good for DNA isolation from young and immature leaves of chilli genotypes. The RNA contamination was completely removed through RNase treatment.
- The quality and quantity of DNA was analysed by NanoDropR ND-1000 spectrophotometer. The absorbance OD 260/280 ratio ranged from 1.8-1.85 which indicated good quality DNA for *Ccs* gene analysis.

- The *Ccs* gene was found amplified in all the genotypes including the yellow fruited accession CC8-1. Size of amplified product was 1.5 kb with *Ccs* Cds primer in all the genotypes.
- *Ccs* promoter amplified product was 920 bp in Byadagi Kaddi, Byadagi Dabbi, Ujwala, Anugraha, Vellayani Samrudhi and 1200bp in Vellayani Thejus and CC8-1.
- BLASTn analysis of the *Ccs* gene amplified with *Ccs* Cds primer showed 99-100 per cent similarity with the reference nucleotide sequence in all the genotypes.
- BLASTx analysis of *Ccs* gene sequence amplified with *Ccs* Cds primer showed 99-100 per cent similarity with the reference amino acid sequence in the seven genotypes studied.
- Analysis of conserved domains revealed that lycopene beta cyclase was the conserved domain in *Capsicum annuum* and *C. chinense* genotypes while in *C. frutescens,* NADB super family protein was the conserved domain.
- The number of ORFs in the *Ccs* sequence amplified with *Ccs* Cds primer ranged from six to seven in the genotypes studied and the number of amino acids coded ranged from 463-469 in *C. annuum,* 298 in *C. frutescens* and 217 in *C. chinense*.
- Multiple sequence alignment of the sequences revealed SNP variations in the genotypes studied and SNP variation caused change in amino acid coded.
- SNP variations were observed in five genotypes viz. Byadagi Kaddi, Byadagi Dabbi, Vellayani Samrudhi, Vellayani Thejus and CC8-1 while no SNP variations were seen in the Ujwala and Anugraha varieties.
- Byadagi Kaddi had two SNPs leading to change in amino acids at 43rd and 425th position of *Capsanthin capsorubin synthase* peptide. Tyrosine (Y) was found replaced by Phenyl alanine (F) in the 43rd position and Lysine (K) was found replaced by Glutamic acid (E) in the 425th position.

160

- Byadagi Dabbi also had the same amino acid change at 425th position, Lysine (K) was replaced by Glutamic acid (E).
- Premature Stop codon UAG was observed in yellow fruited variety CC8-1 at 200th position.
- BLASTn analysis of Ccs gene sequence amplified with *Ccs* promoter primer in seven genotypes showed 90-99 per cent similarity with the reference nucleotide sequence.
- Multiple sequence alignment of the promoter region showed structural changes in the sequences.
- Several SNPs in the sequences, a tandem repeat structure, insertion, deletions and various cis regulatory elements like heat stress related cis-elements (HSE), Myb binding site (MYBPZM) and light responsive elements, TATA box, and CAAT box were observed in the promoter region.
- *Ccs* gene was located in chromosome six of *Capsicum annuum* and in the genome map of chilli it was seen in between 9497216 – 9500911kb.
- *Ccs* Cds gene specific primer was seen to bind 18 bp downstream region of the sequence. The *Ccs* promoter was seen upstream of the protein coding region.
- Elucidation of carotenoid metabolic pathway in *Capsicum annuum* revealed that 17 enzymes were present in the carotenoid biosynthesis pathway.
- Gene regulatory network analysis using Cytoscape showed that network contained 94 nodes and many of the genes were associated with carotenoid biosynthesis processes.
- The main 17 carotenoid metabolic pathway genes, some transcription factors and transferase/transport proteins were densely connected.
- Among the pathway genes, phytoene synthase had the highest number (30 No.) of interactive proteins.

References

References

Aoki, K., Ogata, Y. and Shibata, D. 2007. Approaches for extracting practical information from gene co-expression networks in plant biology. *Pl. Cell Physiol.* 48: 381–390.

Arathi, B. P., Sowmyaa, P. R., Vijay, K., Baskaran, V. and Lakshminarayana, R. 2015. Metabolomics of carotenoids: The challenges and prospects - A review. *Trends. Food Sci. Technol.* 45: 105-117.

Aza-Gonzalez, C., Herrrera-Isidron, L., Nunez-Palenius, H. G., Martinez de la Vega, O. and Ochoa-Alejo, N.2013. Anthocyanin accumulation and expression analysis of biosynthesis related genes during chili pepper fruit development-*Biol. Pl.* 57: 49-55.

Babu, T. P. H. 2000. RAPD analysis to assess the genetic stability in tissue culture derived black pepper (*Piper nigrum* L.) plant. M.Sc. (Ag.) Thesis, Kerala Agricultural University, Thrissur, 83p.

Baerenfaller, K., Massonnet, C. and Walsh, S. 2012. Systems-based analysis of Arabidopsis leaf growth reveals adaptation to water deficit. *Mol. Syst. Biol.* 8: 606.

Baryshnikova, A. 2011. Exploratory analysis of biological networks through visualization, clustering, and functional annotation in cytoscape. *CSHL press* 8: 110-120

Barsan, C., Sanchez-Bel, P., Rombaldi, C., Egea, I., Rossignol, M., Kuntz, M., Zouine, M., Latche, A., Bouzayen, M. and Pech, J. C. 2010.Characteristics of the tomato chromoplast revealed by proteomicanalysis. *J. Exp. Bot.* 61: 2413 - 2431.

Bauer, A. M. and Bryant, S. H. 2004. CD-Search: protein domain annotations on the fly. *Nucl. Acid Res.* 32: 327-331

Berke, T. G. and Shieh, S. C. 2001. *Capsicum*, chillies, paparika, bird's eye chilli In: Peter K V (Ed.) Handbook of Herbs and Spices. *Wood head publishing limited*, Cambridge. 111-122.

Borovsky, Y., Tadmor, Y., Bar, E., Meir, A., Lewinsohn, E. and Paran, I. 2012. Induced mutation in β- carotene hydroxylase results in accumulation of β-carotene and conversion of red to orange color in pepper fruit. *Theor. Appl. Genet.* 126:557–565

Bosland, P. W. 1996. *Capsicums*: Innovative uses of an ancient crop. In: J. Janick (ed.) *Progress in new crops. ASHS press*, Arlington. 479 - 487.

Bosland, P. W. and Votava, E. J. 2000. *Peppers, vegetable and spice Capsicums*. CABI Publishing, 204p.

Bouvier, F., Hugueney, P., Harlingue, A., Kuntz, M. and Camara, B. 1994. Xanthophyll biosynthesis in chromoplasts: isolation and molecular cloning of an enzyme catalyzing the conversion of 5,6-epoxycarotenoid into ketocarotenoid. *Pl. J.* 6: 45 - 54.

Brangeon, J. 1981. Carotenoid metabolism during chloroplast to chromoplast transformation in *Capsicum annuum* fruit. *Planta*.151: 359-364.

Camara, B. 1980. Biosynthesis of keto-carotenoids in *Capsicum annuum* fruits. *FEBS Lett.* 118:315–318.

Camara, B. and Brangeon, J. 1981. Carotenoid metabolism during chloroplast to chromoplast transformation in *Capsicum annuum* fruit. *Plant* 151: 359-364.

Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C. A., Keseler, I. M., Kothari, A., Krummenacker, M., Latendresse, M., Mueller, L. A., Ong, Q., Paley, S., Subhraveti, P., Weaver, D. S. and Karp, P. D. 2015. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucl. Acid Res.* 44: 471-480

Castillo, E. G., Jaime, R. F., Cortes, A. F., Garfias, A. E. J., Gloria, E. L., Martinez, O., Alejo, N. O. and Bustamante, R. 2012. The *Capsicum* transcriptome DB: a hot tool for genomic research. *Bioinforma*. 8(1): 43-47.

Cazzonelli, C. I. and Pogson, B. J. 2010. Source to sink: regulation of carotenoid biosynthesis in plants. *Trends Pl. Sci.* 15: 266-274.

Costa, L. V., Lopes R., Lopes, M. T. G., Figueiedo A. F., Barros, W. S., and Alves, S. R. M. 2009. Cross compatibility of domesticated hot pepper and cultivated sweet pepper. *Crop Breed. Appl. Bio.* 9: 37-44.

Cunningham, F. X. and Gantt, E. 1998. Genes and enzymes of carotenoid biosynthesis in plants. *Annu. Rev. Pl. Physiol. Pl. Mol. Biol.* 49: 557–583.

Cunningham, F. X., Chamovitz, D., Misawa, N., Gantt, E. and Hirschberg, J. 1993. Cloning and functional expression in *Escherichia coli* of a cyanobacterial gene for lycopene cyclase, the enzyme that catalyzes the biosynthesis of β-carotene. *FEBS Lett.* 328:130-138.

Cunningham, F. X., Pogson, B., Sun, Z. R., Mcdonald, K. A., Dellapenna, D. and Gantt, E. 1996. Functional analysis of the beta and epsilon lycopene cyclase enzymes of Arabidopsis reveals a mechanism for control of cyclic carotenoid formation. *Pl. Cell* 8: 1613-1626.

Delgado-Vargas, F.and Peredes-Lopez, O. 2003. Natural Colorants for Food and Nutraceutical Uses. *CRC Press.*327p.

Deli, J., Matus, Z. and Toth, G. 1996. Carotenoid composition in the fruits of *Capsicum annuum* cv. Szentesi kosszarvu during ripening. *J. Agric. Food Chem.* 44: 711-716.

Deli, J., Molnar, P., Matus, Z., and Toth, G. 2001. Carotenoid composition in the fruits of red paprika (*Capsicum annuum* var. lycopersiciforme rubrum) during ripening; biosynthesis of carotenoids in red parika. *J. Agric. Food Chem.* 49: 1517- 1523.

Eisenreich, W., Rohdich, F. and Bacher, A. 2001. Deoxyxylulose phosphate pathway to terpenoids. *Trends. Pl. Sci.* 6: 78-84.

Finger, F. L., Rego, M. M., Cruz, C. D. and Casali, V. W. D. 2009. A diallel study of yield components and fruit quality in chilli pepper (*Capsicum baccatum*). *Euphytica* 168: 275-287.

Fraser, P. D. and Bramley. P. M. 2004. The biosynthesis and nutritional uses of carotenoids. *Progr. Lipid Res.*43: 228–265.

Gallego, F. J. and Martinez, I. 1996. Molecular typing of rose cultivars using RAPD. *J. Hort. Sci.* 71(6): 901-908.

Ghosh, S. and Mandi, S. S. 2015. SNP in chalcone synthase gene is associated with variation of 6-gingerol content in contrasting landraces of *Zingiber officinale.* Roscoe. *Gene.* 5p.

Giorio, G., Stigliani, A. L. and Ambrosio, C. 2007. Phytoene synthase genes in tomato (*Solanum lycopersicum* L.) - new data on the structures, the deduced amino acid sequences and the expression patterns. *FEBS J.* 275: 527-535.

Goodwin, T. W. 1980. The Biochemistry of the Carotenoids, (2nd ed.) London, UK: Chapman and Hall. 224pp.

Govindarajan, V. S. 1985. Capsicum - Production, technology, chemistry and quality. Part II: processed products, standards, world production and trade. *CRC Crit. Rev. Food Sci. Nut.* 23 (3): 207- 88.

Guerin, C., Joet, T., Serret, J., Lashermes, P., Vaissayre, V., Agbessi, M. D. T., Beule, T., Severac, D., Amblard, P., Tregear, J., Gasselin, T. D., Morcillo, F. and Dussert, S. 2016. Gene coexpression network analysis of oil biosynthesis in an interspecific backcross of oil palm. *The Pl. J.* 87: 423-441.

Guzman, I., Hamby, S., Romero, J., Bosland P. W. O. and Connell, M. A. 2010. Variability of carotenoid biosynthesis in orange colored *Capsicum* spp. *Pl. Sci.* 179: 49-59.

Ha, S. H., Kim, J. B., Park, J. S., Lee, S. W. and Cho, K. J. 2007. A comparison of the carotenoid accumulation in Capsicum varieties that show different ripening colours: deletion of the capsanthin-capsorubin synthase gene is not a pre requisite for the formation of a yellow pepper. *J. Exp. Bot.* 58: 3135–3144.

Hall, L. 2011. BioEdit: an important software for molecular biology. *GERF Bull. Bio sci.* 2(1): 60-61.

Hecker, M., Lambecka, S., Toepfer, S., Someren, E. V. and Guthke, R. 2008. Gene regulatory network inference: Data integration in dynamic models-a review. *BioSyst.* 96: 86-103.

Higashi, Y. and Saito, K. 2013. Network analysis for gene discovery in plant-specialized metabolism. *Pl. Cell. Environ.* 36: 1597-1606.

Hirschberg, J. 2001. Carotenoid biosynthesis in flowering plants. *Curr. Opinion Pl. Biol.* 4: 210-218

Hurtado-Hernandez, H. and Smith, P. G. 1985. Inheritance of mature fruit colour in (*Capsicum* annuum L.). *J. Hered.* 76: 211–213.

Isaacson, T., Ohad, I., Beyer, P. and Hirschberg, J. 2004. Analysis in vitro of the enzyme CRTISO establishes a poly-cis-carotenoid biosynthesis pathway in plants. *Pl. Physiol.* 136: 4246-4255.

Jeknic, Z., Morre, T. J., Jeknic, S., Jevremovic, S., Subotic, A. and Chen, T. H. H. 2012. Cloning and functional charactarization of a gene for capsanthin - capsorubin from tiger lilly. *Pl. Cell. Physiol.* 53(11): 1898 - 1912.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. and Tanabe, M. 2011. KEGG for integration and interpretation of large-scale molecular data sets. *Nucl. Acid Res.* 1-6

Kim, H. J., Baek, K. H., Lee, S. W., Kim, J. E., Lee, B. W., Cho, H. S., Kim, W. T., Choi, D. and Hur, C. G. 2008. Pepper EST database: comprehensive in silico tool for analyzing the chili pepper (*Capsicum annuum*) transcriptome. *BMC Pl. Biol.* 8:101

Kuntz, M., Romer, S., Suire, C., Hugueney, P., Weil, J. H., Schantz, R. and Camara, B. 1998. Identification of a cDNA for the plastid-located geranylgeranyl pyrophosphate synthase from *Capsicum annuum*: correlative increase in enzyme activity and transcript level during fruit ripening. *The Pl. J.* 2: 25-34

Lang, Y. Q., Yanagava, S., Sasanuma, T. and Sasakuma., T. 2004. Orange fruit colour in *Capsicum* is due to deletion in the *Capsanthin capsorubin synthase* gene. *Breed. Biosci.* 54:33-39.

Lefebvre, V., Kuntz, M., Camara, B. and Palloix, A. 1998.The *Capsanthin capsorubin synthase* gene: a candidate gene for the *y* locus controlling the red fruit colour in pepper. *Pl. Mol. Biol.*36: 785-789.

Li, Z., Wang, S., Gui, X. L., Chang, X. B. and Gong, Z. H. 2013. A further analysis of the relationship between yellow ripe-fruit color and the *Capsanthin capsorubin synthase* gene in pepper (*Capsicum* sp.) indicated a new mutant variant in *C. annuum* and a tandem repeat structure in promoter region. *PLoS ONE* 8(4): e61996.

Lu, F. H., Yoona, M. Y., Cho,Y., Chung, J. W., Kim, K. T., Cho, M. C., Cheong, S. R. and Park, Y. J. 2011. Transcriptome analysis and SNP/SSR marker information of red pepper variety YCM334 and Taean. *Sci. Hortic.* 129: 38-45.

Manju, P. R. and Sreelathakumary, I. 2002. Genetic cataloging of hot chilli (*Capsicum annuum* Jacq.) types of Kerala. *J. Trop. Agric.* 40: 42-44.

Maoka, T. 2009. Recent progress in structural studies of carotenoids in animals and plants. *Arch. Biochem. Biophys.* 483: 191–195.

Maoka, T., Mochida, K. and Kozuka, M. 2001. Cancer chemopreventive activity of carotenoids in the fruits of red paprika *Capsicum annuum* L. *Cancer Lett.* 172: 103-109.

Matsufuji, H., Nakamura, H., Chino, M. and Takeda, M. 1998. Antioxidant activity of capsanthin and the fatty acid esters in paprika (*Capsicum annuum*). *J. Agric. Food Chem.* 46: 3468-3472.

Matthews, P. D., Luo, R. and Wurtzel, E. T. 2003. Maize phytoene desaturase and ζ-carotene desaturase catalyse a poly-z desaturation pathway: Implications for genetic engineering of carotenoid content among cereal crops. *J. Exp. Bot.* 54: 2215-2230.

Mialoundama, A. S., Heintz, D., Jadid, N., Nkeng, P., Rahier, A., Deli, J., Camara, B. and Bouvier, F. 2010. Characterization of plant carotenoid cyclases as members of the flavoprotein family functioning with no net redox Change. *Pl. Physiol.* 153: 970-979

Moehs, C. P., Tian, L. Osteryoung, K. W. and Dellapenna, D. 2001. Analysis of carotenoids biosynthetic gene expression during marigold petal development. *Pl. Mol. Biol.* 45: 281-293.

Nisar, N., Li, L., Lu, S., Khin, N. C. and Pogson, B. J. 2015. Carotenoid metabolism in plants. *Mol. Pl.* 8: 68–82.

Oltvai, Z. N. and Barabasi, A. L. 2002. Life's Complexity Pyramid. *Sci.* 298: 763-764.

Pan, Y., Bradley, G., Pyke, K., Ball., G., Lu, C., Fray, R., Marshall, A., Jayasuta, S., Baxter, C., Wijk, V. R., Boyden, L., Cade, R., Chapman, N. H., Fraser, P. D., Hodgman, C. and Seymour G. B. 2013. Network inference analysis identifies an APRR2-Like gene linked to pigment accumulation in tomato and pepper fruits. *Pl. Physiol.* 161: 1476-1485.

Park, H., Kreunen, S. S., Cuttriss, A. J., Dellapenna, D. and Pogson, B. J. 2002. Identification of the carotenoid isomerase provides insight into carotenoid biosynthesis, prolamellar body formation, and photomorphogenesis. *Pl. Cell.* 14: 321-332.

Pecker, I., Gabbay, R., Cunningham, F. X. J. and Hirschberg, J. 1996. Cloning and characterization of the cDNA for lycopene β-cyclase from tomato reveals decrease in its expression during fruit ripening. *Pl. Mol. Biol.* 30: 807-819.

Perrin, B. E., Ralaivola, L., Mazurie, A., Bottani, S., Mallet, J. and Buc, F. A. 2003. Gene networks inference using dynamic bayesian networks. *Bioinfo.*19(2): 138-148

Placzek, S., Schomburg, I., Chang, A., Jeske, L., Ulbrich, M., Tillack, J. and Schomburg, D. BRENDA in 2016: new perspectives and new tools in BRENDA. *Nucl. Acid Res.* 45: 380-388

Popovsky, S. and Paran, I. 2000. Molecular genetics of the *y* locus in pepper: its relation to capsanthin-capsorubin synthase and to fruit color. *Theor. Appl. Genet.* 101: 86-89.

Reddy, M. K., Naresh, P., Reddy, K. M. and Ravishankar, K. V. 2012. Allelic variation at capsanthin capsorubin synthase gene for ripening fruit color in chilli (*Capsicum annuum* L.) *Indian J. Genet.* 72: 72-78

Rensink, W. A., Lee, Y., Liu, J., Lobst, S., Ouyang, S. and Buell, C. R. 2005. Comparative analysis of six solanaceous transcriptomes reveal a high degree of sequence conservation and species-specific transcripts. *BMC Genomics.* 6: 124p.

Renuka, S. K. 2014. Morphological and molecular characterization of Byadagi chilli (*Capsicum annuum* L.) M.Sc. (Ag.) thesis, Kerala Agricultural University, Thrissur, 102p.

Rodriguez-Burruezo, A., Gonzalez-Mas, M.C. and Nuez, F. 2010. Carotenoid composition and vitamin A value in ají (*Capsicum baccatum* L.) and rocoto (*C. pubescens* R. & P.), two pepper species from the Andean region. *J. Food Sci.* 75: 446-453.

Rodriguez-Burruezo, A., Prohens, J., Raigon, M. D., and Nuez, F. 2009. Variations for bioactive compounds in aji (*Capsicum baccatum* L.) and rocoto (*C. pubescens* R. & P.) and implications for breeding. *Euphytica.* 170: 169-181.

Rodriguez-Concepcion, M. and Boronat, A. 2002. Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Pl. Physiol.* 130: 1079-1089.

Rodriguez-Uribe, L., Guzman, I., Rajapakse, W., Richins, R. D., and Connell, M. A. 2012. Carotenoid accumulation in orange-pigmented *Capsicum annuum* fruit, regulated at multiple levels. *J. Exp. Bot.* 63(1): 517–526.

Rogers, S. O. and Bendich, A. J. 1994. *Plant Molecular Biology Manual* (2nd Ed.). *Springer* Publishing, Netherlands. 183-190 p.

Ronen, G., Cohen, M., Zamir, D. and Hirschberg, J. 1999. Regulation of carotenoid biosynthesis during tomato fruit development: expression of the gene for lycopene epsilon-cyclase is down-regulated during ripening and is elevated in the mutant delta. *Pl. J.* 17: 341-351.

Salmeron, P. and Garrido, J. M. 1976. Standards criteria and values in processing paparika. *Ion.* 36: 531-540.

Sandeep, 2007. Genetic variability, correlation, morphological and molecular diversity in Byadagi Kaddi and Byadagi Dabbi chillies (*Capsicum annuum* L.) Accessions. M.Sc. Thesis. University of Agricultural Sciences, Dharwad, Karnataka. 100p.

Sandmann, G., Albrecht, M., Schnurr, G., Knorzer, O. and Boger, P. 1999. The biotechnological potential and design of novel carotenoids by gene combination in *Escherichia coli. Trends Biotechnol.* 17: 233-236.

Send, R. and Sundholm, D. 2007. The role of the β-ionone ring in the photochemical reaction of rhodopsin. *J. Phys. Chem.* 111: 27-33.

Sghaier, Z. A. and Mohammed, C. 2005. Genomic DNA extraction method from pearl millet (*Pennisetum glaucum*) leaves. *Afr. J. Biotechnol.* 4(8): 862-866.

Shah, S. N. M., Tian, S. L., Gong, Z. H. and Arisha, M. H. 2014. Studies on metabolism of capsanthin and its regulation under different conditions in pepper fruits (*Capsicum* spp.). *Annu. Rev. Res. Biol.* 4: 1106-1120.

SPI (Spice Board India). 2015. Spice Wise Area and Production (online) http:/www.indian spices.com/pdf/spicewisearprd.xls. 20 Dec 2016

Sreenath, H. L., Muniswami, B. and Naidu, M. M. 1992. Embryo culture of three interspecific crosses in coffee. *J. Pl. Crops* 20: 243-247.

Sugawara, R. C. H., Koike, T., Lopez, R., Gibson, T. J., Desmond, G. Higgins, D. G. and Thompson, J. D. 2003. Multiple sequence alignment with the Clustal series of programs. *Nucl. Acid. Res.* 31: 3497-3500.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. 1994. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acid Res.* Vol. 25: 4876-4882

Thorup, T. A., Tanyolac, B., Livingstone, K. D., Popovsky, S., Paran, I. and Jahn M. 2000. Candidate gene analysis of organ pigmentation loci in the *Solanaceae*. *Proc. nat. acad. Sci.* 97: 11192-11197.

Thul, S. T., Darrokar, M. P., Shaseny, A. K. A. and Khanuja, S. P. S. 2011. Molecular profiling of genetic variability in *Capsicum* species based on ISSR and RAPD markers. *Mol. Biotechnol.* 22: 429-435.

Tian, S. L., Li, L., Shah, S. N. M. and Gong, Z. H. 2015. The relationship between red fruit colour formation and key genes of capsanthin biosynthesis pathway in *Capsicum annuum*. *Biol. Pl.* 59(3): 507-513.

Tian, S. L., Li, L., Chai, W. G., Shah, S. N. M. and Gong, Z. H. 2014. Effects of silencing key genes in the capsanthin biosynthetic pathway on fruit color of detached pepper fruits. *BMC Pl. Biol.* 14:314p.

Toledo-Ortiz, G., Huq, E. and Rodríguez-Concepcion, M. 2010. Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. *Proc. Nature. Acad. Sci.* USA 107: 11626-11631.

Tuan, P. A., Kim, J. K., Park, N. I., Lee, S. Y. and Park, S. U. 2011. Carotenoid content and expression of phytoene synthase and phytoene desaturase genes in bitter melon (*Momordica charantia*). *Food Chem.*126: 1686-1692.

Vijesh, N., Chakrabarti, S. K. and Sreekumar, J. 2013. Modeling of gene regulatory networks: A review. *J. Biomed. Sci. Eng.* 6: 223-231

Watts, D. J. and Strogatz, S. H. 1998. Collective dynamics of small-world networks. *Nature.* 393: 440-442.

Wei, P., Chen, C., Yu, Q., Gady, A., Yu, Y., Liang, G. and Gimitter, F. G. 2014. Novel expression patterns of carotenoid pathway related genes in citrus leaves and maturing fruits. *Tree Genet. Genom.* 10: 439-448.

Welsch, R., Arango, J., Bar, C., Salazar, B., Al-Babili, S., Beltran, J., Chavarriaga, R., Ceballos, H., Tohme, J. and Beyer, P. 2010. Provitamin A accumulation in cassava (*Manihott esculenta*) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. *The Pl. Cell.* 22: 3348-3356.

Wettasinghe, R. and Peffley, E. B. 1998. A rapid and efficient extraction method for onion DNA. *Pl. Breed.* 117: 588-589.

Yan, J, Kandianis, C. B. and Harjes, C. E. 2010. Rare genetic variation at *Zea mays* crtRB1 increases β-carotene in maize grain. *Nat. Genet.* 42: 322-327.

Zeb and Mahmood, 2004. Carotenoids contents from various sourses and their potential health applicatios. *J. Nutri*, 3: 199-204.

Zhu, C., Naqvi, S., Capell, T. and Christou, P. 2009. Metabolic engineering of ketocarotenoid biosynthesis in higher plants. *Arch. Biochem. Biophys.* 483: 182-190.

Annexures

ANNEXURE I

Details of laboratory equipments used for the study

1. High speed refrigerated centrifuge : Kubota 6550, Japan
2. Horizontal electrophoresis system : BIO-RAD, USA
3. Thermal cycler : Proflex thermal cycler
4. NanoDrop^{R} ND-1000 spectrophotometer : NanoDrop^{R} Technologies Inc. USA
5. Gel documentation : Gel Documentation System, BIORAD, USA
6. Water purification system : Millipore, Germany
7. Ice flaking machine : F100 Compact, Ice matics
8. Laminar Air Flow : HML-104, Thermadyne

ANNEXURE II

Reagents for DNA isolation by CTAB method as per Rogers and Bendich (1994)

Reagents:

A. 2X CTAB Buffer:

- 2% CTAB (w/v)
- 100 mM Tris (pH 8)
- 0.5 M EDTA (pH 8)
- 1.4 M NaCl

B. 10% CTAB solution:

- 10 % CTAB (w/v)
- 0.7 M NaCl

C. TE buffer:

- 10 mM Tris (pH 8)
- 1 mM EDTA (pH 8)

ANNEXURE III

Composition of buffers and dyes used for agarose gel electrophoresis

1. **50X TAE buffer (pH 8):**
    - 242g Tris base
    - 57.1 ml glacial acetic acid
    - 100 ml 0.5 mM EDTA (pH 8)

2. **Tracking/ loading dye (6X):**
    - 0.25% Bromophenol blue
    - 0.25% Xylene cyanol
    - 30% Glycerol in water

3. **Ethidium bromide:**
    - The dye was prepared as a stock solution of 10 mg/ml in water and was stored at room temperature in dark bottle.

ANALYSIS OF *CAPSANTHIN CAPSORUBIN SYNTHASE* GENE IN BYADAGI CHILLI (*CAPSICUM ANNUUM* L.) AND ELUCIDATION OF CAROTENOID METABOLIC PATHWAY

By

NARESH S.

(2014-11-107)

Abstract of Thesis

Submitted in partial fulfillment of the

requirements for the degree of

Master of Science in Agriculture

(Plant Biotechnology)

Faculty of Agriculture
Kerala Agricultural University, Thrissur

Centre for Plant Biotechnology and Molecular Biology

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR – 680 656

KERALA, INDIA

2017

Abstract

Byadagi chilli is famous for its deep red colour and negligible or zero pungency. Demand for Byadagi chilli has increased enormously as a source of natural red colour in food industry, confectionaries, cosmetics, beverages, pharmaceuticals and even as a dye in textile industries. Byadagi chilli is mainly exported as oleoresin which serves as a substitute for paprika oleoresin. The red color of chilli fruits is due to several related carotenoid pigments. The most important pigments are capsanthin and its isomer capsorubin.

The present study was conducted to analyze *Capsanthin-capsorubin synthase* gene (*Ccs*) in Byadagi chilli and to elucidate the carotenoid metabolic pathway for production of capsanthin and capsorubin. The studies were focused on seven genetically distinct chilli varieties /accessions of three different *Capsicum* spp. based on colour at fully ripe fruit stage. The accessions selected were Byadagi Kaddi, Byadagi Dabbi, Ujwala, Anugraha (*Capsicum annuum*), Vellayani Samrudhi (*Capsicum frutescens),* Vellayani Thejus and CC8-1 (*Capsicum chinense*).

Genomic DNA was isolated from tender leaves of one month old plants by CTAB method. Two chilli Ccs gene specific SSR primers *viz.* Ccs Cds and Ccs promoter were used to amplify the Ccs gene. The amplified PCR products obtained with Ccs Cds and Ccs promoter were sequenced by outsourcing and sequence data analyzed using bioinformatics tools.

The Ccs gene was found amplified in all the genotypes including the yellow fruited accession CC8-1. Size of amplified product was 1.5kb with Ccs cds primer in all the genotypes. For Ccs promoter, amplified product was 920bp in Byadagi Kaddi, Byadagi Dabbi, Ujwala, Anugraha, Vellayani Samrudhi and 1200bp in Vellayani Thejus and CC8-1

BLASTn analysis of the *Ccs* gene amplified with Ccs cds primer showed 99-100 per cent similarity with the reference nucleotide sequence in all the genotypes. BLASTx analysis of Ccs gene sequence amplified with Ccs cds primer showed 99-

100 per cent similarity with the reference amino acid sequence in the seven genotypes studied. Analysis of conserved domains revealed that lycopene beta cyclase was the conserved domain in *Capsicum annuum* and *C. chinense* genotypes while in *C. frutescens* NADB super family protein was the conserved domain. The number of ORFs in the Ccs sequence amplified with Ccs cds primer ranged from six to seven in the genotypes studied and the number of amino acids coded ranged from 463-469 in *C. annuum,* 298 in *C. frutescens* and 217 in *C. chinense.*

Multiple sequence alignment of the sequences revealed SNP variations in the genotypes studied and SNP variation caused change in amino acid coded. SNP variations were observed in five genotypes *viz.* Byadagi Kaddi, Byadagi Dabbi, Vellayani Samrudhi, Vellayani Thejus and CC8-1 while no SNP variations were seen in the varities Ujwala and Anugraha

Byadgi kaddi had two SNPs leading to change in amino acids at 43rd and 425th position of Capsanthin capsorubin synthase peptide. Tyrosine (Y) was found replaced by Phenyl alanine (F) in the 43rd position and Lysine (K) was found replaced by Glutamic acid (E) in the 425th position. Byadagi dabbi also had the same amino acid change at 425th position, Lysine (K) was replaced by Glutamic acid (E). Premature Stop codon UAG was observed in yellow fruited variety CC8-1 at 200th position

BLASTn analysis of Ccs gene sequence amplified with Ccs promoter primer in seven genotypes showed 90-99 per cent similarity with the reference nucleotide sequence. Multiple sequence alignment of the promoter region could see structural changes in the sequences. Several SNPs in the sequences, a tandem repeat structure, insertion, deletions and various cis regulatory elements like heat stress related cis-elements (HSE), Myb binding site (MYBPZM) and light responsive elements, TATA box, and CAAT box could be observed in the promoter region.

Ccs gene was located in Chromosome six of *Capsicum annuum* and in the genome map of chilli it was seen in between 9497216 – 9500911kb. Ccs cds gene

specific primer was seen to bind 18bp downstream region of the sequence. The Ccs promoter was seen upstream of the protein coding region.

Elucidation of carotenoid metabolic pathway in *Capsicum annuum* revealed that 17 enzymes were present in the carotenoid biosynthesis pathway. Gene regulatory network analysis, using cytoscape showed that network contained 94 nodes and many of the genes were associated with carotenoid biosynthesis processes. The main seventeen carotenoid metabolic pathway genes, some transcription factors and transferase/transport proteins were densely connected. Among the pathway genes, Phytoene synthase had the highest number (30 No.) of interactive proteins.

The identified SNPs in the present study have to be further characterized and validated. Transcriptome analysis of Ccs gene in the different genotypes, homology modeling the Ccs enzyme and prediction of active sites could derive more information on the identified SNPs