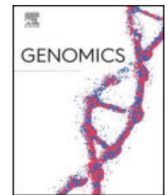
Genomics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno



Identification of genes involved in fruit development/ripening in *Capsicum* and development of functional markers

Meenakshi Dubey[a,b,1], Vandana Jaiswal[a,1], Abdul Rawoof[a], Ajay Kumar[a,c], Mukesh Nitin[a], Sushil Satish Chhapekar[a], Nitin Kumar[a,d], Ilyas Ahmad[a], Khushbu Islam[a], Vijaya Brahma[a,e], Nirala Ramchiary[a,b,*]

[a] Translational and Evolutionary Genomics Lab, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India
[b] Department of Biotechnology, Delhi Technological University, Delhi 110042, India
[c] Department of Plant Science, School of Biological Sciences, Central University of Kerala, Kararagod 671316, India
[d] Department of Bioengineering and Technology, Institute of Science and Technology, Gauhati University, Gopinath Bordoloi Nagar, Guwahati 781014, India
[e] School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi 110067, India

ARTICLE INFO

Keywords:
Capsicum
Fruit development/ripening
Orthologous genes
Real time
Simple sequence repeats

ABSTRACT

The molecular mechanism of the underlying genes involved in the process of fruit ripening in *Capsicum* (family Solanaceae) is not clearly known. In the present study, we identified orthologs of 32 fruit development/ripening genes of tomato in *Capsicum*, and validated their expression in fruit development stages in *C. annuum*, *C. frutescens*, and *C. chinense.* In silico expression analysis using transcriptome data identified a total of 12 out of 32 genes showing differential expression during different stages of fruit development in *Capsicum*. Real time expression identified gene *LOC107847473* (ortholog of *MADS-RIN*) had substantially higher expression (> 500 folds) in breaker and mature fruits, which suggested the non-climacteric ripening behaviour of *Capsicum*. However, differential expression of *Ehtylene receptor 2-like* (*LOC107873245*) gene during fruit maturity supported the climacteric behaviour of only *C. frutescens* (hot pepper). Furthermore, development of 49 gene based simple sequence repeat (SSR) markers would help in selection of identified genes in *Capsicum* breeding.

1. Introduction

The development/ripening of fleshy fruits are a complex phenomenon and affected by plant hormone signalling. On the basis of ethylene signalling and respiration, fruit ripening is divided into two categories- climacteric (ethylene dependent) and non-climacteric (ethylene independent) [17,20]. There are three important components of ethylene biosynthesis viz. *1-aminocyclopropane-1-carboxylic acid* (*ACC*), *ACC synthases* (*ACSs*) and *ACC oxidases* (*ACOs*, [67]). Ethylene is perceived by ethylene receptors (*ETRs*) and then interacts with *Constitutive triple response 1* (*CTR1*) kinase gene to activate *ethylene insensitive 2* (*EIN2*) and starts transcriptional cascade including *EIN3/EIN3-like* (EIL) and ethylene responsive factors. Climacteric fruit ripening, involving ethylene, is well characterized in tomato using several ripening mutants like never-ripe (*nr*), colorless non-ripening (*cnr*), ripening inhibitors (*rin*), and non-ripening (*nor*). However, molecular mechanism of non-climacteric fruit ripening is not well understood. It has been suggested that abscisic acid (ABA) may play important role in non-climacteric fruit ripening. Some key genes of ABA biosynthesis such as *FaNCED1*, *FaBG3*, *FaPYR1/FaABAR* etc. have been shown to be differentially expressed during fruit ripening in strawberry (model for non-climacteric fruit, [19,25,26,42]). A new group of genes called *SEP* gene (*SlMADS1*) was reported to play an important role in ripening of both climacteric and non-climacteric fruits [11–12]. Furthermore, *SlMADS-RIN* (*RIPENING INHIBITOR*, typical SEP gene) is well characterized fruit ripening gene reported in tomato which is responsible for softening, carotenoid accumulation, ethylene production and perception in fruit. *SlMADS-RIN* also interacts with other fruit ripening genes like *Tomato AGAMOUS-LIKE* (*TAGL1*), *FRUITFULL 1* (*FUL1*) and *FRUITFULL 2* (*FUL2*). Ortholog of *SlMADS-RIN* in *Capsicum* has been cloned and its potential role in fruit ripening in *Capsicum* has been reported [12].

Genus *Capsicum* (2n = 2× = 12, genome size = 3.3 GB) belongs to the family Solanaceae and is one of the most widely grown vegetable crops. It is originated from tropical regions around 7500 BCE and perhaps was one of the first cultivated crops [49]. Out of 38 spp., six spp. namely *C. annuum*, *C. chinense*, *C. frutescens*, *C. pubescens*, *C. baccatum*

* Corresponding author at Translational and Evolutionary Genomics Lab, School of Life Sciences, Jawaharlal Nehru University, New Delhi 110067, India.
E-mail address: nrudsc@gmail.com (N. Ramchiary).

[1] These authors contributed equally.

https://doi.org/10.1016/j.ygeno.2019.01.002
Received 15 October 2018; Received in revised form 27 November 2018; Accepted 2 January 2019
0888-7543/ © 2019 Elsevier Inc. All rights reserved.

Please cite this article as: Dubey, M., Genomics, https://doi.org/10.1016/j.ygeno.2019.01.002

ARTICLE IN PRESS

and *C. assamicum* are cultivated [49]. *Capsicum* fruits are rich in antioxidants like ascorbic acid (vitamin C), carotenoids, β-carotene (provitamin A) and phenolic compounds; thus considered good for human health. Besides, anti-cancerous, anti-microbial and anti-inflammatory properties of *Capsicum* have also been suggested [54]. *Capsicum* fruit is typically considered as non-climacteric [1,20,33,47], but climacteric behaviour of fruit ripening of some hot pepper has also been observed [20,48,63]. Till now, molecular mechanism involved in fruit development and ripening in *Capsicum* remains unclear. Therefore, in the present study, we identified genes involved in fruit development/ripening in *Capsicum* using well characterized tomato fruit development/ripening genes to better understand the genetic architecture of this trait in *Capsicum*. The expression of identified genes was studied by both in silico analysis of transcriptome data and qRT-PCR analysis. Furthermore, simple sequence repeat (SSR) markers based on identified genes and SSRs- linked to the genes were also developed in order to utilize these genes in *Capsicum* breeding programs.

2. Materials and methods

2.1. Plant material

A total of 47 *Capsicum* genotypes belonging to *C. chinense*, *C. frutescens* and *C. annuum* which were collected from different parts of India including Assam, Manipur, Jammu, Mizoram, NBPGR, Uttarakhand, Nagaland, and Canada were used in the current study (Table S1). Seeds of each of the 47 accessions were grown in glass house (24-26 °C with 16 h light period) after surface sterilization using 4% sodium hypochlorite. One month old plants were then transferred to Jawaharlal Nehru University Experimental field, New Delhi, India and grown there until fruit setting and maturity. Three genotypes one each from *C. chinense* (NB5), *C. frutescens* (MCM) and *C. annuum* (JH23), used for RNA isolation and qRT-PCR for gene expression validation analysis,were maintained in Jawaharlal Nehru University glasshouse providing standard conditions of light, temperature and humidity.

2.1.1. Selection of tomato genes for fruit development/ripening and identification of chilli orthologs

In *Capsicum*, only few genes have been characterized for fruit development/ripening. However, in tomato, closely related to *Capsicum*, substantial work has been done; pathways involved in fruit development/ripening are well known; and genes involved in the pathways are also well characterized. We have selected genes involved in fruit development/ripening in tomato from previous studies. For each of the selected tomato genes, coding sequences (CDS) were retrieved from NCBI. CDS of these fruit specific tomato genes were BLASTed against *C. annuum* reference genome sequences (GCF_000710875.1_Pepper_Zunla_1_Ref_v1.0_genomic.fa) using BlastN [3] and orthologous genes present in *Capsicum* were identified. For filtration e-value of 1e−10 and ≥80% identity were used as threshold.

2.1.2. Expression analysis of Capsicum fruit developmental/ripening genes using transcriptome data

The available transcriptome sequences in the lab from *C. annuum* (PRJNA505972), *C. chinense* (PRJNA327797) and *C. frutescens* (PRJNA327800) were used for in silico expression analysis of fruit development/ripening genes. The transcriptome sequencing was done using tissues of flower and different developmental stages of fruit i.e. early fruit [20 days post anthesis (DPA)], breaker fruit (30–45 DPA) and mature fruit (45–60 DPA) of each *Capsicum* species. Fruit samples of *C. annuum*, *C. chinense* and *C. frutescens* at above mentioned three developmental stages were collected in liquid nitrogen and stored at −80 °C until RNA extraction. Further, total RNA from each sample was extracted using RNAeasy kit following manufacturer's standard procedure. For each sample, RNA from three biological replicates was pooled together and their integrity was checked using agilent bioanalyser and subjected for RNA sequencing. Total RNA (5–10 μg) was used to construct RNAseq libraries using illumina's TruSeq RNA sample Prep Kits (illumina, San Diego, CA) following manufacturer's protocols and 2 × 100 bp paired-end reads per samples were generated using illumina HiSeq 1000.

Raw RNAseq reads were subjected for quality check using FastQC (v0.11.5) tools and poor quality reads with phred score < 20 along with adapter sequences were trimmed using TrimGalore (v0.4.4). After quality check and filtration, clean reads were mapped to *C. annuum* ref. seq genome (assembly Pepper Zunla 1 Ref_v1.0) using TopHat v2.1.1 [30] with default parameters. Mapped reads were assembled and the transcript abundance were calculated using Cufflinks v.2.2.1. Transcripts having FPKM ≥0.2 were filtered and their differential expression across different stages of fruits was performed using Cuffdiff v.2.2.1 using *p*-value ≤ .01 and FDR ≤ 0.05 [58].

2.2. RNA isolation and qRT-PCR analysis

Total RNA was extracted from fruits at three developmental stages- (i) early fruit (20 DPA), (ii) breaker fruit (30–45 DPA) and (iii) mature fruit (45–60 DPA), leaf and fully opened flower from three genotypes belonging to *C. chinense* (NB5), *C. frutescens* (MCM) and *C annuum* (JH23) using RNA isolation kit of RBC Bioscience following manufacture's protocol. Quality and quantity of isolated RNA were checked in agarose gel (1%) and Nano Drop 1000 (Thermo Scientific). Equal amount (one μg) of RNA was converted into cDNA using SuperScript III first-strand synthesis system (Invitrogen, USA) using manufacturer's instructions. For qRT-PCR, primers were designed using primer express 3.0.1 software. The qRT-PCR was conducted using SYBR Premix Ex Taq (Clontech, USA) on ABI7500 Fast system (Applied Biosystems, USA); thermal profile included initial denaturation at 95 °C for 2 min, followed by 40 cycles of amplification for 15 s at 95 °C and 1 min at 60 °C. For qRT-PCR, two biological and three technical replicates were used. Actin was used as internal control. Relative expression of genes were estimated using $2^{-[\Delta\Delta Ct]}$ method.

2.3. Simple sequence repeats mining and primer designing

Simple sequence repeats (SSRs) was mined using gene sequences along with 5Kb upstream and 5 Kb downstream sequences with online tool "WebSat" [41]. For SSRs mining, stretch with di, tri, tetra, penta and hexa nucleotides with minimum six repeats were selected and mononucleotide repeats were excluded. A cutoff of 100 bp was set as minimum interval distance between two SSRs, and SSRs present within this range were considered as overlapping and thus excluded from further analysis. Using primer 3 software [51] incorporated in WebSat tool, primers were designed using the following criteria - (i) primer size: 18–27 bp with optimum 22 bp, (ii) melting temperature (Tm): 50–65 °C with optimum 60 °C, (iii) GC %: 40–80, (iv) amplicon size: 100–400 bp, (v) max Tm difference between forward and reverse primers: 1 °C, (vi) maximum 3′ stability: 250.

2.4. DNA isolation and SSR amplification

Genomic DNA was extracted from leaf of 47 *Capsicum* accessions using CTAB method [50]. Integrity and quantity were checked on 1% agarose gel and Nanodrop, respectively. For each PCR, 25 ng of DNA, 1× PCR buffer (50 mM KCl, 20 mM Tris–HCl pH 8.4), 1.5 mM $MgCl_2$, 0.125 mM of each dNTPs, 0.5 mM of each primer, and 0.5 U of Taq DNA polymerase (iTaq) were used. PCR mix were put into thermal cycler Eppendorf (Germany) using following profile- denaturation 95 °C for 4 min, 35 cycles of denaturation (94 °C for 2 min), annealing (50–65 °C, as per SSR Tm, for 45 s), elongation (72 °C for 30 s), and final extension at 72 °C for 2 min. PCR product was then resolved using 10 X PAGE on vertical gel electrophoretic system (C.B·S Scientific Co.). Gel was run at 250 V for 6 h. Resolved gel was then visualized using silver staining

ARTICLE IN PRESS

[56]. Genotyping data of each SSRs were recorded as presence/absence and/or length variation. Data was finally converted into allele size (bp) using fragment size calculator.

2.5. SSR polymorphism and diversity

The polymorphism information content (PIC) of SSRs were calculated using formula PIC = 1-Σ xi2, where xi, is relative frequency of ith allele, given by Anderson et al. [4] using PIC calc software. Genetic diversity parameters including number of alleles (na), effective number of alleles (ne), Shannon's information index (I), expected homozygosity, expected heterozygosity, and Nei's gene diversity index were calculated using POPGEN32 v1.32 [68].

3. Results

3.1. Genome wide identification of genes involved in fruit development/ripening

After extensive literature search, a total of 32 genes were selected which are reported to be involved in fruit development/ripening in tomato (Table 1). On the basis of information revealed by published studies, representative network of above mentioned genes has been constructed to understand how the genes are involved in fruit development/ripening(Fig. S1). These 32 genes were distributed on all 12 tomato chromosomes; the maximum of six genes were found on chromosome 10 followed by five genes on chromosome 1, four on chromosomes 3, three on chromosome 6, two each on chromosomes 2, 4, 5, 7, 8 and 9, and one gene each on chromosome 11 and 12, respectively (Fig. S2). Using these 32 tomato fruit development/ripening genes, a total of 41 orthologous genes were identified in *Capsicum* (with ≥80% identity) which are summarized in Table 2. For seven tomato genes i.e.*Tomato constitutive triple response 1 (TCTR1)*, *Lipoxygenase* (*TOMLOX-C*), *RIPENING INHIBITOR* (*MADS-RIN*), *Chalcone synthase* (*CHS-1*), *Glutamate dehydrogense* (*GDH-*1), *DNA demethylase* (*SlDML-*2), and *Cystathionine gamma synthase* (*CGS*), two or more orthologous genes were identified in *Capsicum* genome. Out of above mentioned 41 orthrologous genes, thirty eight genes were mapped on the 12 *Capsicum* genome chromosomes; and the remaining three genes were mapped in scaffolds which are unassigned to any chromosomes. Interestingly, out of 32 tomato genes, > 50% genes (18) were mapped on homologous *Capsicum* chromosomes (Fig. S2). However, 14 genes were mapped on different *Capsicum* chromosomes or scaffolds.

Table 1
List of genes associated with fruit development/ripening in tomato.

Tomato fruit development/ripening genes	Function of Gene	Reference/s
Abscisic stress-ripening protein 1(ASR1)	It is transcriptional regulator of β-hex and thus involved in fruit softening during ripening	[21]
Phytoene synthase 1, chloroplastic (PSY1)	Activates lycopene biosynthesis	[16]
Ethylene-inducible CTR1-like protein kinase (TCTR1)	It is ethylene receptor and involved in ethylene signalling during fruit ripening	[36]
Lipoxygenase (TomloxC)	It is involved in the generation of volatile C6 aldehyde and alcohol flavor compounds during fruit ripening	[18]
MADS-box transcription factor (MADS-RIN)	Associated with ripening time. Shorter ripening time of fruit was observed in SlMADS1-silenced tomatoes	[69]
Polygalacturonase-2a (PG2)	Involved in cell wall metabolism including depolymerization and solubilization and polyuronide degradation during fruit maturation	[55,60]
TAGL1 transcription factor (TAGL1)	Associated with fruit pigmentation. Silencing of the TOMATO AGAMOUS-LIKE 1 (TAGL1) MADS box gene results in altered fruit pigmentation	[24]
NAC domain protein (NAC1)	Functions as a positive regulator of fruit ripening by affecting ethylene synthesis and carotenoid accumulation	[71]
Ehtylene receptor 2-like (ETR4)	It is ethylene receptor and involved in ethylene signalling during fruit ripening	[29]
Pramotor binding protein 1-like CNR(Lespl-CNR)	It is transcription factor that targets LeMADS-RIN, LeHB1, SlAP2a and SlTAGL1; and thus affects fruit pigmentation and ripening	[40], [9]
Chalcone synthase 1 (CHS1)	Involved in flavonoid biosynthesis	[5]
Agamous-like MADS-box protein AGL8 homolog (FUL2)	Involvd in lycopene synthesis, silence line showed altered pigmentation	[8], [65]
MADS-box transcription factor,TDR4 transcription factor (FUL1/TDR)	Involved in production of cuticle components and volatiles, and glutamic acid (Glu) accumulation during fruit development/ripening	[13]
Glutamate dehydrogense (GDH1)	It is involved in anoxia-reoxygenation during fruit ripening	[59]
Golden 2-like protein (GLK2)	It is transcription factor that regulates plastid and chlorophyll levels thus responsible for coloration during fruit ripening	[46]
Protein ros 1-like (SlDML2)	It is responsible for DNA demethylation and taregts CNR and RIN genes, and thus affects fruit development.	[35], [15]
Apatela 2-like protein (AP2)	Transcription factor that regulates fruit ripening via regulation of ethylene biosynthesis and signalling	[27]
Lutescent 2	Associated with ripening duration. Delayed ripening has been observed in mutant gene	[7]
Hydroquinone glucosyltransferase (LOC101244237)	It is involved in fruit aroma an perheps targets flavonoid, flavanols, hydroquinone, xenobiotics and chlorinated pollutants	[39]
Short-chain dehydrogenase-reductase(SlscADH1)	Indirect affect the catabolism of phospholipids and/or integrity of membranes	[6]
Senescence-inducible chloroplast stay-green protein 1 (SGR1)	It is target of RIN and also has physical interaction with phytoene synthase 1 (PSY1) and promotes the biosynthesis of carotenoids in tomato. Besides, also involved in ethylene signalling	[64]
Alternative oxidase 1a (AOX1a)	It is involved in fruit development, ripening, carotenoids, respiration and ethylene production	[66]
Ethylene response factor (ERF6)	It is invoved in carotenoid biosynthesis and ethylene signalling during ripening	[37]
Cystathionine gamma synthase (CGS)	It is involved in methionine synthesis (required for ethylene production) in tomato fruit ripening	[28]
Auxin response factor (ARF4)	Associated with ripening-related fruit quality traits including enhanced fruit density at mature stage, increased firmness, prolonged shelf-life and reduced water (weight) loss at red ripe stage	[52]
Glycosyltransferase (NSGT1)	It converts the cleavable diglycosides of the smoky-related phenylpropanoid volatiles into noncleavable triglycosides, thereby preventing their deglycosylation and release from tomato fruit upon tissue disruption.	[57]
Spermidine synthase (spdsyn)	Associated with reduced shriveling and decay symptom development.	[45]
Alpha-mannosidase (LOC100500729)	It is involved in fruit softenning during ripening	[22]
Glycoalkaloid metabolism 1 (GAME1)	It is involved in steroidal glycoalkaloids (SGAs) metabolism, GAME1 in the glycosylation of SAs and in reducing the toxicity of SA metabolites to the plant cell durind ripening	[23]
Xyloglucan endotransglucosylase-hydrolase (XTH5)	It is involved in fruit softenning and wall-loosening during ripening	[44]
MADS-box protein 1 (LOC543884)	It intract with histone acetyltransferases (HAT) and histone deacetylases (HDAC) and control downstream genes and involved in organ differenciation during fruit development	[14]
Beta-hexosaminidase 1(LOC100529103)	β-D-*N*-acetylhexosaminidase involved in ripening-associated fruit softening	[43]

ARTICLE IN PRESS

Table 2
List of tomato fruit development/ripening genes orthologs identified in *Capsicum* genome.

Tomato gene	Capsicum ortholog	E-value	Score (bits)	Tomato gene	Capsicum ortholog	E-value	Score (bits)
1. *ASR1*	*LOC107867643*	5.34E-174	616	16. *SlDML2*	*LOC107843639*	0	3494
2. *PSY1*	*LOC107868281*	1.59E-121	442		*LOC107843640*	8.51E-62	246
3. *TCTR1*	*LOC107843641*	0	937		*LOC107843637*	0	782
	LOC107843193	2.30E-64	254	17. *AP2*	*LOC107857848*	2.71E-74	285
4. *TomloxC*	*LOC107874197*	0	760	18. *Lutescent 2*	*LOC107854549*	0	839
	LOC107874182	1.50E-158	566	19 *Hydroquinone glucosyltransferase*	*LOC107839289*	0	2593
	LOC107844216	2.16E-82	313	20. *SlscADH1*	*LOC107849564*	0	1496
5. *MADS-RIN*	*LOC107847473*	4.92E-91	340	21. *SGR1*	*LOC107866321*	1.79E-179	634
	LOC107843064	4.95E-86	324	22. *AOX1a*	*LOC107870439*	0	904
6. *PG2*	*LOC107843830*	4.82E-132	477	23. *ERF6*	*LOC107879350*	0	1035
7. *TAGL1*	*LOC107878477*	6.30E-125	453	24. *CGS*	*LOC107858737*	9.62E-104	383
8. *NAC1*	*LOC107870355*	0	854		*LOC107849703*	1.97E-140	505
9. *ETR4*	*LOC107873245*	0	3219	25. *ARF4*	*LOC107847819*	0	1378
10. *LeSPL-CNR*	*LOC107859362*	3.18E-150	536	26. *NSGT1*	*LOC107864612*	0	1197
11. *CHS1*	*LOC107871256*	0	1877	27. spdsyn	*LOC107847831*	2.49E-95	355
	LOC107872666	0	1013	28. *Alpha-mannosidase*	*LOC107875980*	5.53E-109	401
12. *FUL1*	*LOC107855404*	8.69E-92	189	29. *GAME1*	*LOC107878054*	0	1903
13. *FUL2*	*LOC107845304*	8.59E-46	342	30. *XTH5*	*LOC107843860*	0	955
14. *GDH1*	*LOC107843990*	2.87E-130	472	31. *MADS-box protein 1*	*LOC107845303*	2.08E-69	268
	LOC107852706	4.91E-113	414	32. *Beta-hex*	*LOC107843457*	0	1751
15. *GLK2*	*LOC107845460*	2.05E-148	531				

3.2. In silico expression of development/ripening genes in Capsicum spp

During one of our earlier studies we conducted transcriptome analysis of three *Capsicum* genotypes i.e. *C. annuum* (unpublished), *C. chinense* (PRJNA327797) and *C. frutescens* (PRJNA327800) at three fruit developmental stages (early, breaker and mature) and fully opened flower. Out of above mentioned 42 genes, 38 genes were available in transcriptome data. Thus using transcriptome data, in silico expression analysis was conducted for 38 (out of 41) *Capsicum* genes. In silico expression analysis showed that 12 genes were differentially (up or down) expressed in different fruit developmental stages compared to other tissues (Fig. 1). These 12 genes *LOC107847473* (*MADS-RIN*), *LOC107878477* (*TAGL1*), *LOC107873245* (*ETR4*), *LOC107859362* (*Lespl-CNR*), *LOC107845304* (*FUL2*), *LOC107855404* (*FUL1*), *LOC107845460* (*GLK2*), *LOC107839289* (*Hydroquinone glucosyltransferase*), *LOC107866321* (*SGR1*), *LOC107864612* (*NSGT1*), *LOC107843860* (*XTH5*), and *LOC107845303* (*MADS-protein1*) were further selected for real time validation using qRT-PCR. Above mentioned 12 genes are reported to be involved in different metabolic processes during fruit development/ripening which includes ripening time, pigmentation, metabolite accumulation, cuticle synthesis, fruit aroma, fruit softening, etc. (Table 1) in tomato.

3.3. Real time validation of genes involved in fruit development/ripening in Capsicum species

The 12 selected genes were further validated using three genotypes viz. *C. chinense*, *C. frutescens* and *C. annuum*. Altogether five tissues (leaf, flower, early fruit, breaker fruit, and mature fruit) were utilized for qRT-PCR. Primers used for qRT-PCR are given in Table S2. In order to see the relative expression of selected genes among different genotypes belonging to different *Capsicum* spp., leaf tissue of *C. chinense* was used as control. Interestingly, most of the genes were differentially expressed in flower and different stages of fruit development; and almost showed similar expression pattern (except for few cases) across three spp. of *Capsicum*, although the extent of expression varied among different spp. (Fig. 2). On comparison of same tissues in different spp., it has been observed that some genes showed drastic difference in expression level. For example, *FUL2* and *Golden 2 like gene (GLK2)* had the highest expression in *C. annuum* as compared to *C. chinense* and *C. frutescens*. Similarly, *NSGT1* and *MADS-protein1* had the highest expression in *C. chinense* and *C. frutescens*, respectively.

Further, we have compared expression levels in different tissues within genotype using leaf of corresponding genotype as control, and observed that selected genes were differentially expressed in fruit tissues as compared to leaf. For example, *TAGL1* gene showed upto 14.76 fold higher expression in fruit tissues (breaker and mature) as compared to their own leaf in *C. chinense* and *C. frutescens*, and in *C. annuum TAGL1* had highest expression in flower (11.5 fold). *MADS-RIN* showed > 500 folds higher expression in breaker and mature fruits in all the different spp. Further, *ETR4*, *LeSPL*, *Xyloglucan endotransglucosylase-hydrolase (XTHS)*, and *MADS-protein1* had higher expression in flower, early and breaker fruits, however, during maturity its expression had gone down. Other genes like *FUL1*, *FUL2* and *SGR1* were highly expressed in fruit tissues as compared to leaf; however, among three genotypes, their expressions were highest in *C. frutescens*. Contrasting results were observed in case of two genes; *Golden 2 like gene (GLK2)* expression was found to be up-regulated during fruit development in case of *C. annuum*, however in *C. chinense* and *C. frutescens*, *GLK2* was down-regulated during fruit development. Similarly, during fruit development *hydroquinone glucosyltransferase* gene got up-regulated in *C. frutescens*, but was down-regulated in case of *C. chinense* and *C. annuum*. *NSGT1* showed higher expression in fruit tissues in *C. chinense* and *C. frutescens*, however, in *C. annuum* flower had maximum expression for the same gene. Relative expression (in terms of fold change as compared to leaf) of all the 12 genes across five tissues in each of the three genotypes is summarized in Table 3.

3.4. Development of gene based SSR markers

For utilization of fruit development/ripening genes in *Capsicum* breeding, we developed user friendly PCR based SSR markers which were either present within gene or in close vicinity (within 5 Kb regions). Altogether, 49 SSRs were developed, of which, 14 SSRs were from the gene sequence, and 35 SSRs were present in close vicinity (5 Kb upstream and downstream) to the genes (Table 4). Out of 14 SSRs (present within gene sequence), one SSR (SSR_CF-14) was present in exon, however, remaining 13 SSRs were intronic SSRs. Furthermore, we analyzed the presence of SSRs within 1.5 kb upstream of transcripton start site (TSS) of genes (12 SSRs out of 35 SSRs were identified) and for potential promoter or TF motifs using PlantPAN database. Overall a total of 11 TF motifs family along with their putative functions were identified (Table S3).

Out of 49 SSRs, 38, 10 and 1 were di-, tri- and tetra nucleotides,

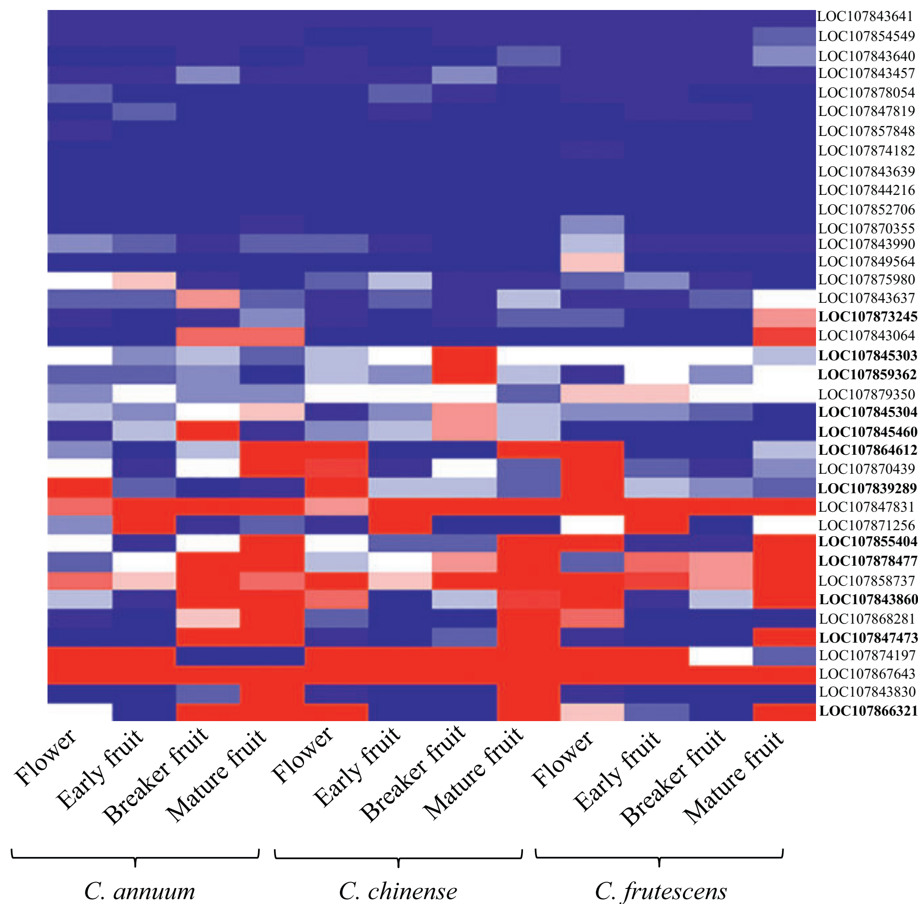


Fig. 1. Expression patterns of *Capsicum* orthologs of tomato fruit development/ripening genes as observed in in silico analysis of transcriptome data. Twelve genes given in bold were differentially expressed during fruit development, and validated using qRT-PCR analysis.

respectively (Fig. 3A). For di-nucleotide SSRs five types of motifs i.e. GA/TC, CA/TG, AT/TA, AC/GT, and AG/CT were identified. Similarly, for tri- and tetra nucleotide SSRs seven and one kind of motifs was identified, respectively (Fig. 3B). Among all the above mentioned 13 types of motifs, motif AT/TA had maximum frequency (31 SSRs). Number of motif repeats and total motif length of 41 SSRs were found up to 54 and 108, respectively (SSR_CF-47). Above mentioned 14 gene based SSRs were from nine genes i.e.*LOC107843641* (SSR_CF-1), *LOC107874197* (SSR_CF-2), *LOC107843830* (SSR_CF-3), *LOC107845304* (SSR_CF-4, SSR_CF-5), *LOC107855404* (SSR_CF-6, SSR_CF-7, SSR_CF-8, SSR_CF-9), *LOC107843990* (SSR_CF-10), *LOC107845460* (SSR_CF-11), *LOC107854549* (SSR_CF-12, SSR_CF-13), and *LOC107875980* (SSR_CF-14). Remaining 35 SSRs were present within 5 Kb up and downstream regions of 23 genes. The primer sequences, melting temperature, amplicon size (bp) and other details of 49 SSRs are provided in Table 4. A representative gel image of SSR profiling in *Capsicum* accessions is given in Fig. 3C.

3.5. Genetic characterization of SSRs and diversity analysis

For SSRs characterization, we have selected 42 SSRs out of 49 SSRs (including all the 14 genic SSRs and 28 non genic SSRs, Table 4). Out of 42 SSRs, 23 SSRs were found to be polymorphic including six genic SSRs. Genetic diversity parameters such as effective number of alleles (ne), Shannon index (I), expected homo- and heterozygosity, Nei gene diversity and polymorphic information content (PIC) for each of the 23 polymorphic SSRs are summarized in Table 5. A total of 50 alleles with 2–3 alleles per locus were scored. Shannon diversity index (I) ranged from 0.51 (SSR_CF-30) – 0.97 (SSR_CF-5) with an average 0.64. Average Nei's gene diversity was 0.42 (ranged 0.26 to 0.58). Maximum PIC was identified for SSR_CF-5 (0.51) followed by SSR_CF-3 (0.44), however, minimum PIC was observed for SSR_CF-30 (0.22).

4. Discussion

The fruit development/ripening of *Capsicum*, despite the main economic harvest of the crop being fruit, has not been well characterized. However, its close relative, tomato, the fruit development and ripening genes have been identified [32]. Therefore, in the present study, we explored the *Capsicum* genome for identification of fruit development/ripening genes through comparative genomics. For this purpose, we used a total of 32 fruit development/ripening genes which are involved both in ethylene -dependent and –independent fruit ripening pathway in tomato (Fig. S1) and could identify a total of 42 orthologous genes in *C. annuum* reference genome. These genes are involved in different fruit development/ripening traits including colouring, cell wall formation, softening, aroma, metabolite accumulation etc. (details are given in Table 1). Interestingly, one or more orthologs were identified for each of the 32 tomato genes in *Capsicum*, which suggest that the large size of *Capsicum* genome may harbour more than one orthologs of tomato genes (Table 2). Furthermore, of the 32 genes, 18 were mapped onto the homologous *Capsicum* chromosomes suggesting these genes are still conserved in the syntenic chromosomal regions.

Although, *Capsicum* is the closest relative to tomato, but regulatory mechanisms (including gene expression) of these two crops differed, and perhaps responsible for diversified features of these two crops [32]. Our in silico expression analysis revealed that out of 32 genes, only 12

ARTICLE IN PRESS

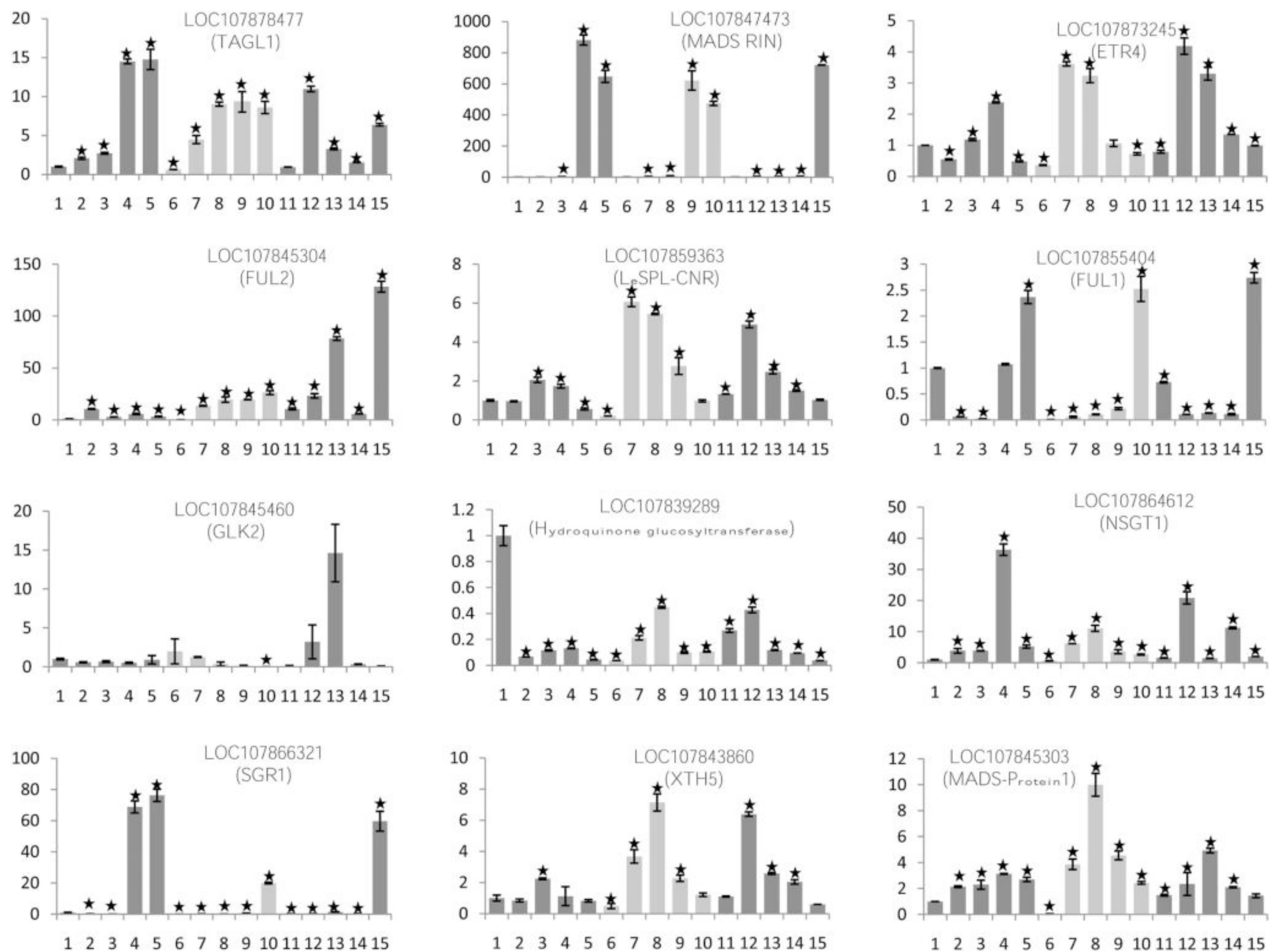


Fig. 2. Real time validation of twelve fruit development/ripening genes using qRT-PCR. Y-axis shows relative expression (as compared to NB5 leaf) in terms of fold change of 15 tissues involving three spp. 1 = NB5_leaf, 2 = NB5_flower, 3 = NB5_early fruit, 4 = NB5_breaker fruit, 5 = NB5_mature fruit, 6 = MCM_leaf, 7 = MCM_flower, 8 = MCM_early fruit, 9 = MCM_breaker fruit, 10 = MCM_mature fruit, 11 = JH23_leaf, 12 = JH23_flower, 13 = JH23_early fruit, 14 = JH23_breaker fruit, 15 = JH23_mature fruit; * represents 0.05 level of significance.

genes were differentially expressed in *Capsicum* during fruit development/ripening; indicating different ripening behaviour of *Capsicum* than that of tomato. Similar results have also been reported where expression pattern of *colorless non ripening* gene (*CNR*), *Golden-2-like gene* (transcription factor that regulates plastid and chlorophyll levels thus responsible for coloration during fruit ripening), *ACO, ACS* and *HB-1* were very distinct in tomato and *Capsicum* [32]. *Capsicum* is typically classified as non-climacteric fruit (unlike tomato which is a

Table 3
Relative expression of selected genes in terms of fold change as compared to leaf of corresponding genotype.

Capsicum gene ID	Tomato gene	NB5 (*C. chinense*)					MCM (*C. frutescence*)					JH23 (*C. annuum*)				
		A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
LOC107878477	*TAGL1*	1	2.17*	2.71*	14.52*	14.76*	1	7.14*	14.41*	14.41*	13.7*	1	11.52*	3.44*	1.63*	6.71*
LOC107847473	*MADS- RIN*	1	1.44	3.27*	881.45*	647.67*	1	5.73*	9.09*	9.09*	476.23*	1	4.42*	5.2*	6.91*	707.54*
LOC107873245	*ETR4*	1	0.55*	1.18*	2.38*	0.48*	1	9.91*	8.87*	8.87*	1.99*	1	5.27*	4.16*	1.70*	1.25
LOC107845304	*FUL2*	1	10.6*	2.86*	5.82*	3.3*	1	423.74*	610.82*	610.82*	821.49*	1	2.19*	7.41*	0.56*	12.13*
LOC107859362	*LeSPL-CNR*	1	0.96	2.05*	1.73*	0.56*	1	30.1*	26.98*	26.98*	4.84*	1	3.69*	1.86*	1.13	0.78*
LOC107855404	*FUL1*	1	0.06*	0.02*	1.07	2.37*	1	9.29*	17.66*	17.66*	420.11*	1	0.15*	0.18*	0.15	3.77*
LOC107845460	*GLK2*	1	0.59	0.67	0.51	0.91	1	0.63	0.18	0.18	0.002*	1	32.81*	149.68*	3.45	0.48
LOC107839289	*Hydroquinone glucosyltransferase*	1	0.07*	0.12*	0.13*	0.05*	1	5.55*	11.71*	11.71*	2.77*	1	1.59*	0.44*	0.35*	0.14*
LOC107864612	*NSGT1*	1	3.84*	3.84*	36.31*	5.21*	1	9.8*	17.63*	17.63*	4.23*	1	13.52*	0.94	7.27*	1.32
LOC107866321	*SGR1*	1	0.15*	0.02*	68.82*	76.25*	1	171.58*	72.69*	72.69*	20,160.1*	1	1.26	6.12*	1.72	255.47*
LOC107843860	*XTH5*	1	0.85	2.24*	1.13	0.83	1	7.48*	14.53*	14.53*	2.47*	1	5.79*	2.34*	1.85*	0.54*
LOC107845303	*MADS-Protein1*	1	2.14*	2.29*	3.13*	2.69*	1	86.34*	223.09*	223.09*	54.26*	1	1.58*	3.32*	1.41	0.97

A = Leaf, B = flower, C = Early fruit, D = Breaker fruit, E = Mature fruit.
* Represents 0.05 level of significance.

Table 4
Summary of SSRs developed from fruit development/ripening genes in *Capsicum*.

SSR[@]	Capsicum gene[$]	Motif	Forward primer	Length	TM	Reverse primer	Length	TM	Amplicon size
SSR_CF-1*	*LOC107843641*	(TA)14	CCAACTTATGTGATCCTCTTCT	22	55	GGTCAATGCACAAGGTAGTT	20	55	385
SSR_CF-2*	*LOC107874197*	(TCT)6	AAGACACTAAGAGAAAGGATGC	22	55	AGTACCCGAGGTTATGGAAT	20	55	384
SSR_CF-3*	*LOC107843830*	(TA)9	GGGGAACTATCAATGGAAAT	20	55	CGTATCGAGTAAGTCAGAGTGAA	23	56	394
SSR_CF-4*	*LOC107845304*	(CAA)6	AGTCTCTTCTTTGTGAGTGTTG	22	54	GACAGCTAACTAGACAGGTTTGT	23	55	383
SSR_CF-5*	*LOC107845304*	(TA)14	CGCACAAACTTCACAACTAA	20	55	CTTTCATACCTCTCAAGAATCC	22	55	354
SSR_CF-6*	*LOC107855404*	(TA)6	AACCGTAGATGGAAGTCCTAC	21	55	TAAGTGTGCTGCTCCTCTTC	20	56	389
SSR_CF-7*	*LOC107855404*	(AAC)6	GTAGTCTCCATCTCCATACCTG	22	55	CGGGTGTAATCAACTCTCTTA	21	55	134
SSR_CF-8*	*LOC107855404*	(AT)6	CGTGAGATTAACATCGTATTCC	22	56	AGTCCTCGTGTTTGTGTGTT	20	56	202
SSR_CF-9*	*LOC107855404*	(AT)11	ATAGTGAACACACAAACACGAG	22	55	GACGTAGCAGGAGTTTTACTTT	22	55	283
SSR_CF-10*	*LOC107843990*	(GT)6	ATAATCTTCAGTATCAGGCTCG	22	55	AACAACCTCTGGATGGTATCT	21	55	253
SSR_CF-11*	*LOC107845460*	(AT)6	TTGAAAGGAGGGAGTATCTT	20	53	CCATCATACATTTGCTTCC	19	54	297
SSR_CF-12*	*LOC107854549*	(AAC)6	GTGATACACCCTTATATGACCC	22	55	GTCGTACTTCTTGTCGTAGGTAT	23	54	270
SSR_CF-13*	*LOC107854549*	(TATC)6	GATTCCCAGAAGTGTGAAAA	20	55	CTACTGTCATTGTTTGGTTGAC	22	55	362
SSR_CF-14*	*LOC107875980*	(GA)6	AGTTCAAGTGAGCGAACATAC	21	55	ACAGTTTAGTGGTAGAACCCTT	22	54	235
SSR_CF-15*	*LOC107843641-up*	(TA)6	TCCTGGTTTGATTATATGGC	20	55	GGTAGTTCGTGGATTACTTTTC	22	55	398
SSR_CF-16*	*LOC107847473-up*	(TA)6	TTCCTACACATGCTCATCACT	21	56	TAAGAATGGAGAAGTGGCAA	20	56	359
SSR_CF-17*	*LOC107878477-up*	(AT)12	CTTTCCTTCAATACCACTTACC	22	55	TTTACTTCGCCATTGCTACT	20	55	316
SSR_CF-18	*LOC107845304-up*	(TA)10	AAAGATAGAGACTTCAGTTGCC	22	55	GGGTTATTTCGGTGTTTTG	19	55	390
SSR_CF-19*	*LOC107845304-up*	(ATT)8	TCAACTCTTACGACTGACATGA	22	56	AACACCACCATTAAAGACCTC	21	56	326
SSR_CF-20	*LOC107845304-up*	(AT)6	CCCCTTCGATTAGTTTGTATAG	22	55	GGTTACTGTTGGAGTCGTTAGT	22	55	270
SSR_CF-21*	*LOC107843990-up*	(AT)10	TAGTGGCCTATTCATGTTG	19	52	GCACGTCTTAATTCTTATCC	20	52	184
SSR_CF-22*	*LOC107845460-up*	(AGA)6	GATACTTTACTGGATGGTTGCT	22	55	TGTTCTACACTCGTATTTGGG	21	55	269
SSR_CF-23*	*LOC107854549-up*	(AT)7	TACTATAACAGCAATTACCGCC	22	56	CCGCTTCGATTGTATATGAA	20	56	187
SSR_CF-24*	*LOC107839289-up*	(TA)12	AACAGTATAAGAACGTGGTGTG	22	54	GACATCGCAGTCAGTAATAAAC	22	54	322
SSR_CF-25*	*LOC107866321-up*	(TA)7	TAGAAGGTTCATCCAAACTCTC	22	55	GAAGTAGTTGCGTATTGGGTAT	22	55	375
SSR_CF-26	*LOC107866321-up*	(CA)6	TAACCTTCACTAACACCTCACA	22	55	GGCAAAGAGAATGAGTAGAAAC	22	55	278
SSR_CF-27*	*LOC107879350-up*	(TA)7	GAGATATTTATGGGGTAAGTCG	22	55	AAGGTGTGTTGTAGGGGTTA	20	55	395
SSR_CF-28*	*LOC107858737-up*	(TA)10	TACTTGACTGCTTGATTCCTAC	22	54	AAGCCTAACCAAAGCTAAAG	20	53	310
SSR_CF-29	*LOC107858737-up*	(AG)8	CTAAACTAGCACTTATCCCGAC	22	55	CCTCTTTTGTTACCTCTTTGAC	22	55	250
SSR_CF-30*	*LOC107847819-up*	(TAC)8	AGGGGTAGTAGTGGAAATTGTT	22	56	ACAGGTGAAGTAGAGGAAGATG	22	55	116
SSR_CF-31*	*LOC107847831-up*	(AT)9	TGTAACTTGTAACCCAACACG	21	56	GATCTCAAGCTCTTCTTTCTTG	22	55	389
SSR_CF-32*	*LOC107845303-up*	(TA)9	GTTTGCATGTGAGTTATGTAGG	22	55	AGAATCACTGGGCTATTCAAC	21	56	371
SSR_CF-33	*LOC107843457-up*	(AAT)6	CTACTTCCGGTTGAAGATTGT	21	56	CAGGATGTTTATCTGTTGCAC	21	56	393
SSR_CF-34*	*LOC107843457-up*	(AT)21	AAATACCCTCAAATCCTGTG	20	54	ATTCAACAATGGAGTCAACC	20	55	188
SSR_CF-35	*LOC107843457-up*	(AT)6	CACACGACATAGTCATAGGAAG	22	55	TATATTGAGGGGTCATTTGG	20	55	382
SSR_CF-36*	*LOC107874197-down*	(AT)6	GAACTTTAGTCCCCAAGCAT	20	56	GTGAACCTCAAACTCTACCATT	22	55	399
SSR_CF-37*	*LOC107847473-down*	(TA)13	CAACAAGTGTCTGGCAATAAC	21	56	CACGAGTGAAGTACGTGTAGC	21	56	305
SSR_CF-38*	*LOC107870355-down*	(AT)10	AGAGATCGACGGCAACTTA	19	56	TGACGTGGCGTATGAAAT	18	56	391
SSR_CF-39*	*LOC107873245-down*	(AT)9	GCACCCGATAATGTAAGAA	19	54	TTTGCATCTCTCATAGCACT	20	54	246
SSR_CF-40*	*LOC107871256-down*	(TA)9	AAGCCCGTACTAGATTGTTAAG	22	55	ATACTGAAGAAGGATACAACCG	22	55	355
SSR_CF-41*	*LOC107845460-down*	(AT)14	GATTTTCCCTGTGAGTGGTA	20	55	TGTCCTATTCTCTGTGTGTTTG	22	55	282
SSR_CF-42*	*LOC107843639-down*	(TA)6	TGGAAGCATCTATTGGAGAA	20	56	GTCACATAAGTTGCGATAAAGC	22	57	383
SSR_CF-43*	*LOC107866321-down*	(AT)9	ATCTATGGAGTCATTTGGTGAG	22	56	GTTCTTGGGTCATACTTCTTTG	22	56	380
SSR_CF-44	*LOC107866321-down*	(TC)7	AGTAACCATGTGTGCTGACTAA	22	55	GTGTTTGAGTAGGATTGGAGAT	22	55	282
SSR_CF-45*	*LOC107879350-down*	(ATC)8	AACAACATATCAGCCTCTGC	20	55	AGGAGTTAGAACAAGAATGCTC	22	55	358
SSR_CF-46*	*LOC107847831-down*	(CT)6	GGCGATTGCTACTAATAACTCT	22	55	CAAGCTATGTATCGTCAGTCAG	22	55	256
SSR_CF-47*	*LOC107875980-down*	(AC)54	CGTCAAGTCTACGAGTAAGGA	21	55	CCTCCTGTTTGGTTGTAAGTAG	22	56	315
SSR_CF-48*	*LOC107878054-down*	(AGA)8	TGGAGAGTTTAGTAGTTTCGTG	22	54	GAGTATGAAGATGAGCGTTAGA	22	54	396
SSR_CF-49*	*LOC107843860-down*	(TA)14	TGTTAAGAGAGCAGTGTGGTTA	22	55	ACTACCATTACCTTCTCCGAAT	22	56	349

[@] SSRs marked with star sign were synthesized for genotyping; SSRs in bold represented polymorphism.
[$] Up and down represent 5 Kb upstream and 5 downstream sequences of corresponding gene.

climacteric fruit), and this may be reason that out of the 12 differentially expressed genes, 11 (excluding *Ethylene receptor like protein, ETR*) belonged to the ethylene independent pathway. In the present study, *ETR* (*Capsicum* gene ID; *LOC107873245*) found to be up-regulated during fruit maturity in only *C. frutescens* supported the climacteric behaviour of hot pepper as suggested in earlier studies [20,48].

Although, key regulators of the ethylene independent fruit ripening pathway are still not known, but *LeMADSRIN* (*RIN*), a MADS-box transcription factor, has been considered as one of the major regulators [17,38]. In the present study, interestingly *RIN* (*Capsicum* gene ID; *LOC107847473*) unexpectedly up-regulated (> 500 folds higher expression as compared to leaf) in breaker and mature fruits in all the three spp.; suggesting that *RIN* is a potential regulator of ethylene independent fruit ripening pathway in non-climacteric fruit. Furthermore, interacting partner genes of *RIN*, like *FUL1* (*LOC107855404*), *FUL2* (*LOC107845304*), *TAGL1* (*LOC107878477*) showed maximum expression during fruit maturation as expected, although level of expression differed among three spp.; homologs of *Stay green 1* (*SGR1*,*Capsicum* gene ID; *LOC107866321*), target of *RIN* and involved in chlorophyll degradation [64], also found to be up regulated during breaker and mature fruits stages suggested involvement of *SGR1* gene in non climacteric fruit ripening. Unlike, a few genes showed different expression pattern across spp., for example, *Golden 2 like gene* (*GLK2*, *Capsicum* gene ID; *LOC107845460*), which is known for regulating plastid and chlorophyll levels (coloration) during fruit ripening was up-regulated in *C. annuum* in contrast to *C. chinense* and *C. frutescens* where *it was* down-regulated during fruit maturation. On the basis of above results present study suggested that fruit development/ripening pathway may vary somewhat across three spp., although most of the pathway may be similar across spp.

Moreover, to utilize the fruit development and ripening genes identified in the present study in breeding program, we also developed 49 SSR markers which are present either within or nearby to genes. As expected lesser number of SSRs (only 14, out of 49 SSRs) were present within the genes as compared to nearby regions, and only six were polymorphic (out of 14) suggesting the high conservation of genic

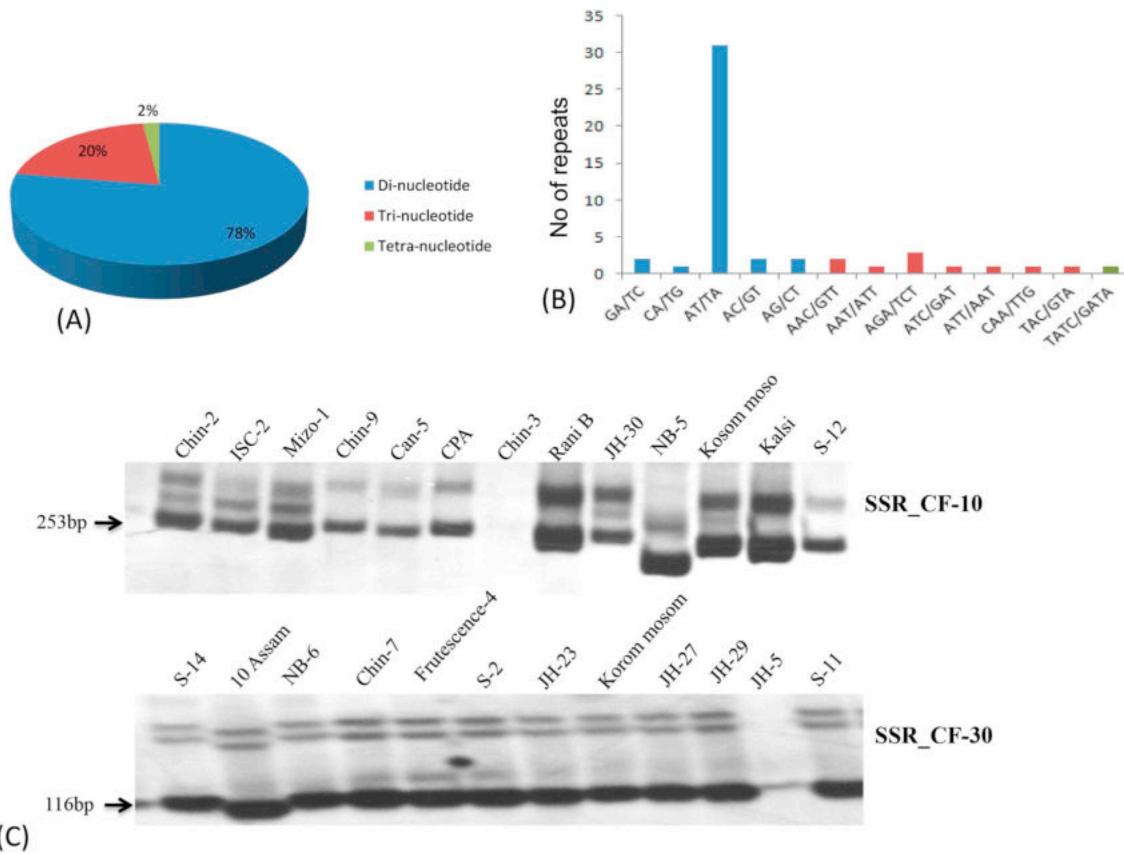


Fig. 3. Frequency distribution of different SSRs motifs (A, B) and representative gel image of SSR profiling (two SSRs) in *Capsicum* germplasm.

regions. From earlier studies it was suggested that the high conservation is maintained among the genes in different species as any mutation in this region may lead to drastic change [31,70]. In plants, di and tri-nucleotides repeat motifs have been found most abundant and varies from species to species [34,53,61]. In our present study, di-nucleotide (AT/TA) SSRs was of maximum frequency followed by tri-nucleotide SSRs. In order to see the utilization of SSRs developed in population genetics study, we genotyped 47 *Capsicum* accessions using 42 SSRs. Out of 42 SSRs, 23 (~55%) SSRs were found to be polymorphic which was comparable to earlier studies [2,10,62]; and out of 23 SSRs, six including SSR_CF-1 to SSR_CF-5 and SSR_CF-11 were present within the gene orthologs of *TCTR1, GLK, TOMLOX, PG2 and FUL2* (Tables 5 and

Table 5
Diversity parameters observed in 47 *Capsicum* accessions using 23 polymorphic SSR markers.

SSR[#]	No of observed allele	Effective number of allele	Shannon index	Expected homozygosity	Expected heterozygosity	Nei gene diversity	PIC[@]
SSR_CF-1	2	1.873	0.659	0.529	0.471	0.466	0.357
SSR_CF-11	2	1.719	0.609	0.577	0.423	0.418	0.331
SSR_CF-15	2	1.555	0.542	0.639	0.361	0.357	0.293
SSR_CF-17	2	1.614	0.568	0.616	0.384	0.380	0.308
SSR_CF-19	2	1.858	0.654	0.533	0.467	0.462	0.355
SSR_CF-2	2	1.719	0.609	0.577	0.423	0.418	0.331
SSR_CF-21	3	1.680	0.712	0.591	0.409	0.405	0.359
SSR_CF-22	2	1.572	0.550	0.632	0.368	0.364	0.298
SSR_CF-25	2	1.719	0.609	0.577	0.423	0.418	0.331
SSR_CF-27	2	1.941	0.678	0.510	0.490	0.485	0.367
SSR_CF-3	3	2.036	0.852	0.486	0.514	0.509	0.440
SSR_CF-30	3	1.356	0.510	0.735	0.265	0.263	0.220
SSR_CF-31	2	1.873	0.659	0.529	0.471	0.466	0.357
SSR_CF-32	2	1.815	0.641	0.546	0.454	0.449	0.348
SSR_CF-37	2	1.957	0.682	0.506	0.494	0.489	0.369
SSR_CF-38	2	1.516	0.524	0.656	0.344	0.340	0.282
SSR_CF-4	2	1.614	0.568	0.616	0.384	0.380	0.308
SSR_CF-42	2	1.896	0.666	0.522	0.478	0.473	0.361
SSR_CF-43	2	1.504	0.518	0.661	0.339	0.335	0.279
SSR_CF-45	2	1.815	0.641	0.546	0.454	0.449	0.348
SSR_CF-46	2	1.978	0.688	0.500	0.500	0.494	0.372
SSR_CF-47	2	1.815	0.641	0.546	0.454	0.449	0.348
SSR_CF-5	3	2.399	0.972	0.411	0.589	0.583	0.511

[#] SSRs in bold represent genic SSRs.
[@] PIC represents polymorphic information content.

2), respectively. Unexpectedly, we observed higher PIC for genic SSRs (0.31 to 0.51) than that of nearby SSRs (0.22 to 0.37). Highly polymorphic genic SSRs may serve as perfect markers in fruit breeding in *Capsicum* for the selection of associated genes; although non-genic SSRs also prove as potential flag marks for genes since they are tightly linked to genes (within 5 Kb). Wide range of other diversity parameters like heterozygosity (0.27–0.59), effective number of allele (1.36–2.4), gene diversity (0.26–0.58), and diversity index (0.51–0.97) also suggested that SSRs developed in our present study may prove useful for population genetics study.

5. Conclusion

Present study identified genes involved in fruit development/ripening in *Capsicum* through comparative genomic approach using tomato genes. Extensive expression of *MADS-RIN* ortholog in *Capsicum* fruits suggested that *Capsicum* shows non-climacteric ripening behaviour. Across three spp. most of the genes showing similar expression pattern suggested fruit development/ripening in three spp. is almost similar except for few exceptions, like *ETR* gene was up-regulated during fruit maturation in *C. frutescens* only, which suggested involvement of ethylene in fruit ripening in *C. frutescens*. User-friendly SSR markers developed in this study could be used in population genetics studies, QTL mapping and fruit breeding program in *Capsicum* species.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2019.01.002.

Acknowledgement

This work was supported by the research grants from the Department of Biotechnology, Government of India, in the form of Ramalingaswami *Re*-Entry Fellowship and Research Grant, and Science and Engineering Research Board, Department of Science and Technology, Govemnmet of India (SB/EMEQ-086/2014) to Dr. Nirala Ramchiary. MD acknowledges WOS-A (SR/WOS-A/LS/1339/2015), VJ acknowledges the DST-INSPIRE faculty (DST/INSPIRE/04/2016/001189) award received from the Department of Science and Technology, Ministry of Science and Technology, Government of India for supporting grant to carry out research work. VB acknowledges University Grants Commission for providing Postdoctoral Fellowship for Women.

Author's contribution

NR and VJ conceived and designed research. MD, VJ, IA and KI conducted field and lab experiments. AR, AK, MN, SC, NK, VB performed transcriptome analysis. AR and MN conducted in silico analysis. VJ developed SSRs and analyzed data. VJ and MD wrote the manuscript. All authors read and approved the manuscript.

Conflict of interest

Authors declare no conflict of interest.

References

[1] W.M. Aizat, J.A. Able, J.C. Stangoulis, A.J. Able, Characterisation of ethylene pathway components in non-climacteric capsicum, BMC Plant Biol. 3 (2013) 191.

[2] L.W. Al-Qadumii, T.S. Monther, M. Hussein, Assessment of in silico BAC-based simple sequence repeat (SSR) marker development for tomato (*Solanum lycopersicum* L.), African J. Biotech. 11 (2012) 13938–13946.

[3] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.

[4] P.A. Anderson, G.J. Lawrence, B.C. Morrish, M.A. Ayliffe, E.J. Finnegan, J.G. Ellis, Inactivation of the flax rust resistance gene M associated with loss of a repeated unit within the leucine rich repeat coding region, Plant Cell 9 (1997) 641–651.

[5] T. Ariizumi, Y. Shinozaki, H. Ezura, Genes that influence yield in tomato, Breed. Sci. 63 (2013) 3–13.

[6] V. Baldassarre, Spadafora N.D. CabassiVG, A. Aprile, C.T. Müller, H.J. Rogers, A. Ferrante, Wounding tomato fruit elicits ripening-stage specific changes in gene expression and production of volatile compounds, J. Exp. Bot. 66 (2015) 1511–1526.

[7] C.S. Barry, G.M. Aldridge, G. Herzog, Q. Ma, R.P. McQuinn, J. Hirschberg, J.J. Giovannoni, Altered chloroplast development and delayed fruit ripening caused by mutations in a zinc metalloprotease at the lutescent2 locus of tomato, Plant Physiol. 159 (2012) 1086–1098.

[8] M. Bemer, R. Karlova, A.R. Ballester, Bovy A.G. TikunovYM, M. Wolters-Arts, Rossetto P. de Barros, G.C. Angenent, R.A. de Maagd, The Tomato FRUITFULL homologs *TDR4/FUL1* and *MBP7/FUL2* regulate ethylene-independent aspects of fruit ripening, Plant Cell 24 (2012) 4437–4451.

[9] W. Chen, J. Kong, T. Lai, K. Manning, C. Wu, Y. Wang, C. Qin, B. Li, Z. Yu, X. Zhang, M. He, P. Zhang, M. Gu, X. Yang, A. Mahammed, C. Li, T. Osman, N. Shi, Wang, S. Jackson, Y. Liu, P. Gallusci, Y. Hong, Tuning LeSPL-CNR expression by SlymiR157 affects tomato fruit ripening, Sci. Rep. 5 (2015) 7852.

[10] M. Chunyang, W. Xiaochun, Z. Yanyan, Y. Yuxiang, Y. Shuangjuan, W. Zhiyong, Z. Xiaowei, S. Jiuwen, Z. Xiaolan, Y. Qiuju, Z. Qiang, Genetic diversity analysis of *Capsicum* genus by SSR markers, Mol. Plant Breed 8 (2017) 70–78.

[11] T. Dong, Z. Hu, L. Deng, Y. Wang, M. Zhu, J. Zhang, G. Chen, A Tomato MADS-Box transcription factor, SlMADS1, acts as a negative regulator of fruit ripening, Plant Physiol. 163 (2013) 1026–1036.

[12] T. Dong, G. Chen, S. Tian, Q. Xie, W. Yin, Y. Zhang, et al., A non-climacteric fruit gene *CaMADS-RIN* regulates fruit ripening and ethylene biosynthesis in climacteric fruit, PLoS One 9 (4) (2014) e95559.

[13] M. Fujisawa, Y. Shima, H. Nakagawa, M. Kitagawa, J. Kimbara, T. Nakano, T. Kasumi, Y. Itoa, Transcriptional regulation of fruit ripening by Tomato *FRUITFULL* homologs and associated MADS Box proteins, Plant Cell 26 (2014) 89–101.

[14] J. Gaffe, C. Lemercier, J.P. Alcaraz, M. Kuntz, Identification of three tomato flower and fruit MADS-box proteins with a putative histone deacetylase binding domain, Gene 471 (2011) 19–26.

[15] P. Gallusci, C. Hodgman, E. Teyssier, G.B. Seymour, DNA methylation and chromatin regulation during fleshy fruit development and ripening, Fron. Plant Sci. (2016), https://doi.org/10.3389/fpls.2016.00807.

[16] G. Giorio, A.L. Stigliani, C.D. Ambrosio, 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 (2008) 527–535.

[17] J.J. Giovannoni, Genetic regulation of fruit development and ripening, Plant Cell 16 (2004) 170–180.

[18] C. Guoping, H. Rachel, W. David, T. Andy, L. Zhefeng, G. Donald, Identification of a specific isoform of tomato lipoxygenase (*tomloxc*) involved in the generation of fatty acid-derived flavor compounds, Plant Physiol. 136 (2004) 2641–2651.

[19] Y. Han, R. Dang, J. Li, J. Jiang, N. Zhang, M. Jia, L. Wei, Z. Li, B. Li, W. Jia, Sucrose *nonfermenting1-related protein kinase2.6*, an ortholog of open *stomata1*, is a negative regulator of strawberry fruit development and ripening, Plant Physiol. 167 (2015) 915–930.

[20] B.-Z. Hou, C.-L. Li, Y.-Y. Han, Y.-Y. Shen, Characterization of the hot pepper (Capsicum frutescens) fruit ripening regulated by ethylene and ABA, BMC Plant Biol. 18 (2018) 162.

[21] M. Irfan, S. Ghosh, V. Kumar, N. Chakraborty, S. Chakraborty, A. Datta, Insights into transcriptional regulation of β-D-nacetylhexosaminidase, an N-glycan-processing enzyme involved in ripening-associated fruit softening, Exp. Bot. 65 (2014) 5835–5848.

[22] M. Irfan, S. Ghosh, V.S. Meli, A. Kumar, V. Kumar, V. Chakraborty, S. Chakraborty, A. Datta, Fruit ripening regulation of αmannosidase expression by the MADS Box transcription factor *ripening inhibitor* and ethylene, Front. Plant Sci. 7 (2016) 10.

[23] M. Itkin, I. Rogachev, N. Alkan, T. Rosenberg, S. Malitsky, L. Masini, S. Meir, Y. Iijima, K. Aoki, R.D. Vos, D. Prusky, S. Burdman, J. Beekwilder, A. Aharonia, *Glycoalkaloid metabolism1* is required for steroidal alkaloid glycosylation and prevention of phytotoxicity in tomato, Plant Cell 23 (2011) 4507–4525.

[24] M. Itkin, M. Seybold, D. Breitel, I. Rogachev, S. Meir, A. Aharoni, *tomato agamous-like1* is a component of the fruit ripening regulatory network, Plant J. 60 (2009) 1081–1095.

[25] H.F. Jia, D. Lu, J.H. Sun, C.L. Li, Y. Xing, L. Qin, Y.Y. Shen, Type 2C protein phosphatase ABI1 is a negative regulator of strawberry fruit ripening, J. Exp. Bot. 64 (2013) 1677–1687.

[26] I.Y. Kadomura, K. Miyawaki, A. Takahashi, T. Masuda, S. Noji, Light and abscisic acid independently regulated *FaMYB10* in Fragaria × ananassa fruit, Planta 241 (2015) 953–965.

[27] R. Karlova, F.M. Rosin, J. Busscher-Lange, V. Parapunova, P.T. Do, A.R. Fernie, P.D. Fraser, C. Baxter, G.C. Angenent, R.A. deMaagd, Transcriptome and metabolite profiling show that apetala2a is a major regulator of tomato fruit ripening, Plant Cell 23 (2011) 923–941.

[28] Y.S. Katz, G. Galili, R. Amir, Regulatory role of *cystathionine-c-synthase* and de novo synthesis of methionine in ethylene production during tomato fruit ripening, Plant Mol. Biol. 61 (2006) 255–268.

[29] B.M. Kevany, D.M. Tieman, M.G. Taylor, V.D. Cin, H.J. Klee, Ethylene receptor degradation controls the timing of ripening in tomato fruit, Plant J. 51 (2007) 458–467.

[30] D. Kim, G. Pertea, C. Tranell, H. Pimentel, R. Kelley, S.L. Saizberg, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions, Genome Biol. 14 (2013) R36.

[31] K.S. Kim, S.T. Ratcliffe, B.W. French, L. Liu, T.W. Sappington, Utility of EST-derived SSRs as population genetics markers in a beetle, J. Heredity 99 (2008) 112–124.

[32] S. Kim, M. Park, S.I. Yeom, Y.M. Kim, J.M. Lee, H.A. Lee, E. Seo, J. Choi, K. Cheong,

ARTICLE IN PRESS

et al., Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species, Nat. Genet. 46 (2014) 270–278.
[33] S. Klie, S. Osorio, T. Tohge, M.F. Drincovich, A. Fait, J.J. Giovannoni, A.R. Fernie, Conserved changes in the dynamics of metabolic processes during fruit development and ripening across species, Plant Physiol. 164 (2014) 55–68.
[34] S.P. Kumpatla, S. Mukhopadhyay, Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species, Genome 48 (2005) 985–998.
[35] Z. Lang, Y. Wang, K. Tang, D. Tang, T. Datsenka, J. Cheng, Y. Zhang, A.K. Handa, J.K. Zhua, Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) E4511–E4519.
[36] J. Leclercq, L.C. Adams-Phillips, H. Zegzouti, B. Jones, A. Latche, J.J. Giovannoni, J.C. Pech, M. Bouzayen, *LeCTR1*, a Tomato *CTR1*-like gene, demonstrates ethylene signaling ability in arabidopsis and novel expression patterns in tomato, Plant Physiol. 130 (2002) 1132–1142.
[37] J.M. Lee, J.G. Joung, R. McQuinn, M.Y. Chung, Z. Fei, D. Tieman, H. Klee, J. Giovannoni, Combined transcriptome, genetic diversity and metabolite profiling in tomato fruit reveals that the ethylene response factor SlERF6 plays an important role in ripening and carotenoid accumulation, Plant J. 70 (2012) 191–204.
[38] Z.F. Lin, S.L. Zhong, D. Grierson, Recent advances in ethylene research, J. Exp. Bot. 60 (2009) 3311–3336.
[39] T. Louveau, C. Leitao, S. Green, C. Hamiaux, B.V. Rest, O. Dechy-Cabaret, R.G. Atkinson, C. Chervin, Predicting the substrate specificity of a glycosyltransferase implicated in the production of phenolic volatiles in tomato fruit, FEBS J. 278 (2011) 390–400.
[40] K. Manning, M. Tor, M. Poole, Y. Hong, A.J. Thompson, G.J. King, J.J. Giovannoni, G.B. Seymouret, A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening, Nat. Genet. 38 (2006) 948–952.
[41] W.S. Martins, D.C.S. Lucas, K.F.S. Neves, D.J. Bertioli, WebSat - a web software for microsatellite marker development, Bioinformation 3 (2009) 282–283.
[42] L. Medina-Puche, R. Blanco-Portales, F.J. Molina-Hidalgo, G. Cumplido-Laso, N. García-Caparrós, E. Moyano-Cañete, J.L. Caballero-Repullo, Extensive transcriptomic studies on the roles played by abscisic acid and auxins in the development and ripening of strawberry fruits, Funct. Integr. Genomics 16 (2016) 671–692.
[43] V.S. Meli, S. Ghosh, T.N. Prabha, N. Chakraborty, S. Chakraborty, A. Datta, Enhancement of fruit shelf life by suppressing N-glycan processing enzymes, PNAS 107 (2010) 2413–2418.
[44] E. Miedes, K. Herbers, U. Sonnewald, E.P. Lorences, Overexpression of a cell wall enzyme reduces xyloglucan depolymerization and softening of transgenic tomato fruits, J. Agric. Food Chem. 58 (2010) 5708–5713.
[45] S. Nambeesan, T. Datsenka, M.G. Ferruzzi, A. Malladi, A.K. Mattoo, A.K. Handa, Overexpression of yeast spermidine synthase impacts ripening, senescence and decay symptoms in tomato, Plant J. 63 (2010) 836–847.
[46] C.V. Nguyen, J.T. Vrebalov, N.E. Gapper, Y. Zheng, S. Zhong, Z. Fei, J.J. Giovannonib, Tomato GOLDEN2-LIKE transcription factors reveal molecular gradients that function during fruit development and ripening, Plant Cell 26 (2014) 585–601.
[47] J.M. Palma, F.J. Corpas, L.A. Del Río, López-HE, Pepper fruit as a model to study the metabolism of antioxidants, ROS and RNS, Free Radic. Biol. Med. 5 (2014) 39.
[48] V. Paul, R. Pandey, G.C. Srivastava, The fading distinctions between classical patterns of ripening in climacteric and non-climacteric fruit and the ubiquity of ethylene-an overview, J. Food Sci. Technol. 49 (2012) 1–21.
[49] N. Ramchiary, M. Kehie, V. Brahma, S. Kumaria, P. Tandon, Application of genetics and genomics towards Capsicum translational research, Plant Biotechnol. Rep. 8 (2013) 101–123.
[50] S.O. Rogers, A.J. Bendich, Extraction of DNA from milligram amounts of fresh herbarium and mummified plant tissues, Plant Mol. Biol. 5 (1985) 69–76.
[51] S. Rozen, H. Skaletsky, Primer3 on the WWW for general users and for biologist programmers, Methods Mol. Biol. 132 (2000) 365–386.
[52] M. Sagar, C. Chervin, I. Mila, Y. Hao, J.P. Roustan, M. Benichou, Y. Gibon, B. Biais, P. Maury, A. Latche, J.C. Pech, M. Bouzayen, M. Zouine, Sl-ARF4,an Auxin Response Factor involved in the control of sugar metabolism during tomato fruit development, Plant Physiol. 161 (2013) 1362–1374.
[53] D. Saha, R.S. Rana, S. Chakraborty, S. Datta, A.A. Kumar, A.K. Chakraborty, P.G. Karmakar, Development of a set of SSR markers for genetic polymorphism detection and interspecific hybrid jute breeding, Crop J. 5 (2017) 416–429.
[54] M. Sarpras, R. Gaur, V. Sharma, S.S. Chhapekar, J. Das, A. Kumar, et al., Comparative analysis of fruit metabolites and pungency candidate genes expression between bhut jolokia and other Capsicum species, PLoS One 11 (2016) e0167791.
[55] Y. Sitrit, A.B. Bennett, Regulation of tomato fruit polygalacturonase mRNA accumulation by ethylene: a re-examination, Plant Physiol. 116 (1998) 1145–1150.
[56] H. Tagelstrom, Detection of mitochondrial DNA fragments, in: A.R. Hoelzel (Ed.), Molecular Genetic Analysis of Populations: A Practical Approach, IRL Press, Oxford, 1992, pp. 89–114.
[57] Y.M. Tikunov, J. Molthoff, R.C.H. de Vos, J. Beekwilder, A.V. Houwelingen, J.J.V. Hooft, M. Nijenhuis-de Vries, C.W. Labrie, W. Verkerke, H.V. de Geest, M.V. Zamora, S. Presa, J.L. Rambla, A. Granell, R.D. Hall, A.G. Bovya, *NON-SMOKY GLYCOSYL TRANSFERASE1* prevents the release of smoky aroma from tomato fruit, Plant Cell 25 (2013) 3067–3078.
[58] C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D.R. Kelley, H. Pimentel, S.L. Salzberg, J.L. Rinn, L. Pachter, Differential gene and transcript expression analysis of RNA-Seq experiments with TopHat and Cufflinks, Nat. Protoc. 7 (2012) 562–578.
[59] K.J. Tsai, C.Y. Lin, C.Y. Ting, M.C. Shih, Ethylene-regulated glutamate dehydrogenase fine-tunes metabolism during anoxia-reoxygenation, Plant Physiol. 172 (2016) 1548–1562.
[60] G.A. Tucker, Genetic manipulation of fruit ripening, Biotec Genet. Eng. Rev. 8 (1990) 133–160.
[61] R.K. Varshney, T. Thiel, N. Stein, P. Langridge, A. Graner, *In silico* analysis on frequency and distribution of microsatellites in ESTs of some cereal species, Cell. Mol. Biol. Lett. 7 (2002) 537–546.
[62] P.R. Ved, K. Rajesh, K. Sanjay, R. Ashutosh, K. Sanjeet, S. Major, P.S. Sheo, B.R. Awadesh, P. Rajneesh, Genetic diversity in Capsicum germplasm based on microsatellite and random amplified microsatellite polymorphism markers, Physiol. Mol. Biol. Plants 19 (2013) 575–586.
[63] L. Villavicencio, S. Blankenship, D. Sanders, W. Swallow, Ethylene and carbon dioxide production in detached fruit of selected pepper cultivars, J. Am. Soc. Hortic. Sci. 124 (1999) 402–406.
[64] R.H. Wang, X.Y. Yuan, L.H. Meng, B.Z. Zhu, H.L. Zhu, Y.B. Luo, D.Q. Fu, Transcriptome analysis provides a preliminary regulation route of the ethylene signal transduction component, *slein2*, during tomato ripening, PLoS One 11 (2016) e0168287.
[65] S. Wang, G. Lu, Z. Hou, Z. Luo, T. Wang, H. Li, J. Zhang, Z. Ye, Members of the tomato FRUITFULL MADS-box family regulate style abscission and fruit ripening, J. Exp. Bot. 65 (2014) 3005–3014.
[66] F. Xu, S. Yuan, D.W. Zhang, X. Lv, H.H. Lin, The role of alternative oxidase in tomato fruit ripening and its regulatory interaction with ethylene, J. Exp. Bot. 63 (2012) 5705–5716.
[67] S.F. Yang, N.E. Hoffman, Ethylene biosynthesis and its regulation in higher plants, Annu. Rev. Plant Physiol. 35 (1984) 155–189.
[68] F.C. Yeh, R.C. Yang, T.B.J. Boyel, Z.H. Ye, J.X. Mao, POPGENE, the User-Friendly Shareware for Population Genetic Analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada, 1997.
[69] J. Zhang, Z. Hu, Q. Yao, X. Guo, V. Nguyen, F. Li, G. Chen, A tomato MADS-box protein, SlCMB1, regulates ethylene biosynthesis and carotenoid accumulation during fruit ripening, Sci. Rep. 8 (2018) 3413.
[70] M. Zhang, W. Mao, G. Zhang, F. Wu, Development and characterization of polymorphic EST-SSR and genomic SSR markers for tibetan annual wild barley, PLoS One 9 (2014) 948.
[71] M. Zhu, G. Chen, S. Zhou, Y. Tu, Y. Wang, T. Dong, Z. Hu, A new tomato NAC (NAM/ATAF1/2/CUC2) transcription factor, SlNAC4, functions as a positive regulator of fruit ripening and carotenoid accumulation, Plant Cell Physiol. 55 (2014) 119–135.