Genomics 114 (2022) 110273

Contents lists available at ScienceDirect

Genomics

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

Fruit transcriptional profiling of the contrasting genotypes for shelf life reveals the key candidate genes and molecular pathways regulating post-harvest biology in cucumber

N. Pradeepkumara [a], Parva Kumar Sharma [d], A.D. Munshi [a], T.K. Behera [a], Reeta Bhatia [b], Khushboo Kumari [a], Jogendra Singh [a], Sarika Jaiswal [d], Mir Asif Iquebal [d], Ajay Arora [c], Anil Rai [d], Dinesh Kumar [d], R.C. Bhattacharya [e], S.S. Dey [a,*]

[a] Division of Vegetable Science, ICAR-Indian Agricultural Research Institute, New Delhi, India
[b] Division of Floriculture and Landscaping, ICAR-Indian Agricultural Research Institute, New Delhi, India
[c] Division of Plant Physiology, ICAR-Indian Agricultural Research Institute, New Delhi, India
[d] Centre for Agricultural Bioinformatics, ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India
[e] ICAR-National Institute of Plant Biotechnology, New Delhi, India

ARTICLE INFO

Keywords:
Cucumber
Developmental stages
Shelf life
RNA-seq
GO analysis
KEGG pathways
Variant identification
Gene regulatory networks
Web resource

ABSTRACT

Cucumber fruits are perishable in nature and become unfit for market within 2–3 days of harvesting. A natural variant, DC-48 with exceptionally high shelf life was developed and used to dissect the genetic architecture and molecular mechanism for extended shelf life through RNA-seq for first time. A total of 1364 DEGs were identified and cell wall degradation, chlorophyll *a*nd ethylene metabolism related genes played key role. Polygalacturunase (PG), Expansin (EXP) and xyloglucan were down regulated determining fruit firmness and retention of fresh green colour was mainly attributed to the low expression level of the chlorophyll catalytic enzymes (CCEs). Gene regulatory networks revealed the hub genes and cross-talk associated with wide variety of the biological processes. Large number of SSRs (21524), SNPs (545173) and InDels (126252) identified will be instrumental in cucumber improvement. A web genomic resource, CsExSLDb developed will provide a platform for future investigation on cucumber post-harvest biology.

1. Introduction

Reduction of food loss after harvest through different post-harvest technologies is being practiced widely in developed countries. These technologies are rarely used in developing countries because of increased cost of the final product [1]. Around 30% fruits are lost after harvest as they are perishable with limited shelf life. Therefore, reducing this loss through different cost-effective approach is a global mission and need to be taken in war footing [2]. Among the available options, genetic mechanism to enhance the shelf life is considered as the most economical for their adoption even in the low-income economies which could be an important tool to reduce food loss and hunger in large. Therefore, breeding and engineering of food crops with extended shelf life can make important contribution in ensuring food security, improving farm profitability [1] and to achieve the sustainable development goals (SDGs) of United Nations.

Cucumber (*Cucumis sativus* L.) is an important vegetable crop grown worldwide and cultivated as a commercial crop in around 150 countries with world production of 87.8 million metric tonnes [3]. Cucumber fruits are highly perishable and become unfit for marketing 2–3 days after harvesting. The principal reasons for post-harvest loss are mainly associated with fruit discoloration due to loss of chlorophyll pigments, shrivelling or wilting caused by loss of moisture which affects the firmness and physiological deterioration due to undesirable temperature [4]. Besides, the cucumber fruits start losing the attractive fresh green colour after 8–10 days of pollination even when they are kept in the plants for harvesting or stored under different controlled storage conditions after harvesting. Consumers do not prefer to purchase fruits with yellowish green colour. Therefore, retention of attractive green colour with minimal shrinkage in the pedicel after harvesting are important traits determine shelf life and marketability of the fruits. Cucumber is native to India and tremendous variation for different traits is available

* Corresponding author at: Division of Vegetable Science, ICAR-Indian Agricultural Research Institute, New Delhi, India.
E-mail address: shyam.iari@gmail.com (S.S. Dey).

https://doi.org/10.1016/j.ygeno.2022.110273
Received 11 September 2021; Received in revised form 17 January 2022; Accepted 21 January 2022
Available online 31 January 2022

in the local accessions and wild feral forms distributed in different parts of India [5]. However, this crop is now distributed all over the world and its dispersal from the Indian sub-continent to Occident is well documented [6]. Variation in shape, size, colour, surface texture, wartiness and spininess in cucumber has been reported since medieval times [6,7]. Besides, its use as salad, the fruit extract and derived particles including nanoparticles are widely used in medicinal, therapeutic and cosmetic purpose [8–11].

One natural variant, DC-48 was isolated at Division of Vegetable Science, ICAR-IARI, New Delhi with exceptional shelf life in year 2012. Initially only one plant was identified with fruits completely different from rest of the plants in the population. This plant was maintained through selfing and single plant selection was practiced for another 8 generations to develop homozygous and homogeneous genotype with the desirable variation. The fruits of the DC-48 remain fresh without losing green colour and firmness up to 10–15 days after harvesting when stored at room condition. In contrast, most of the cucumber genotypes start losing the attractive green colour within 2–5 days after harvesting from plants when kept in room condition. Besides, the genotype can retain its epicarp firmness and texture without shrinkage for 10–15 days after harvesting unlike most of the cultivated genotypes. Therefore, this genotype holds huge promise in developing cucumber genotypes with extended shelf life. However, the underlying physiological and molecular basis for the extraordinary shelf life of DC-48 is unknown. Till now, no information is available reporting the molecular mechanisms and physiological factors responsible for extended shelf life in cucumber. Under this circumstance, RNA-seq approach to analyse the transcripts, understanding the expression level of the key genes associated with retention of green colour, fruit firmness and their profiling would be instrumental to get insight about this extremely useful trait in cucumber.

In the recent past, high-throughput RNA sequencing technology (RNA-Seq) used widely as an efficient tool for transcriptomic study of less understood trait [12]. RNA-seq is established as a key tool to discover, profile and analyse the RNA transcripts for several plant species for number of traits [12]. This technique has the unique advantage of microarray and EST-sequencing which allows single-base resolution with minimal background noise. At the same time, it is high throughput with high sensitivity. RNA-seq allows quantification of gene expression, transcribed region mapping, distinguishes the different isoforms and allelic expression [13–16]. RNA-Seq have been used widely in cucumber for studying the sex expression [17–19], resistance to biotic stresses [20,21] fruit morphology [22] and domestication process [23]. Molecular mechanisms regulating fruit ripening and post-harvest quality of different fruits crops were studied through comparative transcriptomics using RNA-seq in wide variety of crops [24–26]. In melons, comparative transcriptome study among climacteric and non-climacteric fruits revealed role of ethylene biosynthesis and signalling and chlorophyll degradation through microarray analysis [27].

The molecular mechanism of post-harvest biology with extended shelf life is poorly understood in cucumber. Cucumber fruit follow non-climacteric pattern of ripening [28]. Pre-harvest factors play important role in determining post-harvest behaviour of the fruits and understanding the molecular networks and pathways in the developmental stages of the fruits before harvesting are critical to determine the ripening pattern [27]. However, several physiological traits before and after harvest varies among genotypes and ethylene production like climacteric fruits has been reported in different genotypes [29]. In contrast to the ripening respiratory pattern, ethylene release and production in cucumber is also associated with fruit decay in terms of softening and pale colour instead of physical changes associated with ripening [30]. Differential pattern and pace in change in the colour and turning the fruits epidermis to yellowish from the attractive green was earlier recorded in cucumber genotypes [31,32]. The present genotype under investigation (DC-48) retains not only the green colour but it had intact fruit firmness and fruits decay was extremely slow and not visible even after 15 days of harvesting under room temperature (27-30 °C). In number of crops the role of the cuticle layer, structure and integrity of the cell wall, biosynthesis and rate of ethylene release and ripening related upstream components are reported to be associated with fruit softening. The genes encoding enzymes of cell-wall degradation, ethylene biosynthesis, and upstream transcription factors are potential candidate genes for extended shelf life in cucumber. It has been reported that cell wall, tissue turgor and cuticle play pivotal role in retention of fruit texture and firmness after harvest in number of crops [33–36] although not reported in cucumber. Fruit firmness is largely determined by the rigid cell structure of the cell wall and reduced rigidity of the cell wall is directly associated with fruit softening [1]. Transpiration water loss from the fruit surface and shelf life is influenced by the cuticles, mainly due to its role in water loss [35,37]. Studies on the mechanisms associated with fruit ripening and cell wall degradation metabolism can enhance the understanding about the post-harvest biology and extended shelf life in cucumber natural variant, DC-48.

Role of the different enzymes associated with the genes of the cell wall stability is not known. However, some of the key genes encoding enzymes associated with fruit texture are pectin β-methylesterase (PME), polygalacturonase (PG), β -galactosidase (β-GAL), β-and pectate lyase (PL), or of matrix glycans, such as *endo*-β-glucanase (EGase), xyloglucan endo-transglycosylase-hydrolases (XTHs), and expansins (EXPs) [36,38]. Ethylene biosynthesis and response related pathways have studied in different crops although their role in cucumber post-harvest biology is not reported extensively. ACC synthase (ACS) and ACC oxidase (ACO) are two major families of genes encode key enzymes of ethylene biosynthesis [39]. Chlorophyll catabolic enzymes (CCEs) associated with degradation of chlorophyll (Chl), results in loss of green colour that typically occurs during leaf senescence [40]. STAYGREEN (SGR) phenotype with delayed leaf senescence results from mutation in CCE genes [41,42]. The role of the cucumber SGR genes (*CsSGR*) in leaf senescence and resistance to foliar fungal disease is being established [40]. However, it is not known if *CsSGR* is associated with retention of fruit epicarp colour after harvesting. The genetics and inheritance pattern of the extended shelf life characterised by reduced degradation of Chl and intact firmness of the fruit epicarp is not reported.

In cucumber, there is no specific information about the molecular network responsible for extended shelf life. The identified natural variant, DC-48 with exceptional shelf life is the ideal candidate genotype to get insight about the post-harvest life of cucumber fruits. Comparative transcriptome study with contrasting genotypes for shelf life would be very effective at this stage to understand this extremely useful trait and its further use in cucumber improvement programme. Therefore, the present investigation was conducted to understand the molecular network associated with unique post-harvest biology and extended shelf life in cucumber using the natural variant, DC-48 through RNA-seq. Mining of putative DNA markers (SSRs) and variants (SNPs and InDels) was also attempted to develop reliable molecular markers for future investigation in molecular mapping and marker assisted breeding of this trait in cucumber. Database development was undertaken for a ready reference to the molecular mechanisms and DNA based markers for future improvement programme.

2. Materials and methods

2.1. Plant material and sampling

The two contrasting cucumber genotypes, namely, DC-48 (extended shelf life) and DC-83 (poor shelf life) were grown under insect proof nethouse at the research field of Division of Vegetable Science, ICAR-Indian Agricultural Research Institute, New Delhi (28.6377° N, 77.1571° E) during the spring-summer season of 2019. The genotype, DC-48 and DC-83 were slicing type cucumber and belong to the same species, *C. sativus* L. The genotype, DC-83 is early in maturity and first harvesting of fruits 45–50 days after sowing whereas DC-48 is comparatively late maturity in nature with first harvesting at 55–60 days after sowing. Both the

genotypes have medium to dark green leaves. Individual fruit weight of DC-48 and DC-83 at marketable maturity are 300 g and 350 g, respectively. Fruits of both the genotypes are light green in colour. These genotypes are suitable for cultivation under open field and they are non-parthenocarpic in nature. The sowing of the seeds was done in the last fortnight of February. Both the genotypes were monoecious in nature and female flowers of the individual plants were pollinated with the pollens collected from the same plant for fertilization and fruit development as there was no natural insect pollinators inside the net-house. All the cultural practices were followed for successful raising of healthy cucumber plants under insect proof net-house as per ICAR-IARI, New Delhi standard cultivation protocol. The fully developed fruits at different developmental stages from the healthy plants were used for determining the shelf life and whole genome RNA-sequencing.

2.2. Determining the shelf life after harvesting

The fruits of DC-48 and DC-83 were harvested at 10 days after pollination for determining shelf life. Retention of fruits beyond this stage results loss of fresh green colour makes the fruits less attractive to consumers besides oversized fruits [43]. Ten fruits of each genotype were harvested at 7 days interval for three times for determining shelf life. The fruits were stored under room temperature condition (27-30 °C) without use of any wrapping materials/film. Scoring of the fruits was done based on the following criteria:

2.2.1. Index-I (based on loss of colour)

Stage-1: Good colour (green).
Stage-2: Acceptable colour (somewhat lighter green).
Stage-3 Unacceptable colour (first occurrence of yellowing).
Stage-4: Unacceptable colour (yellowing).
Stage-5: Unacceptable colour (extreme yellowing).

The fruits were characterised by counting the days when the fruits reach the Stage - 3.

2.2.2. Index-II (based on moisture loss)

Stage-1: No moisture loss (no shrivelling).

Stage-2: Acceptable moisture loss (Shrivelling initiated in the pedicel end).

Stage-3 Unacceptable moisture loss (Clear shrivelling visible in the pedicel end and initiation of shrivelling in the main body of the fruits).

Stage-4 Unacceptable moisture loss (Prominent shrivelling in the entire fruits).

Stage-5 Unacceptable moisture loss (Shrivelled fruits with appearance of wide ridge in the fruits).

Like the Index-I, the fruits were characterised by counting the days when reach stage-3.

2.3. Determination of chlorophyll content, moisture loss and fruit firmness at different time intervals

The estimation of total chlorophyll was based on a standard spectrophotometric method [44]. The upper surface of the fruit epicarp (2 mm) was peeled with a sharp peeler and the same peeler was used for all the samples to keep the sample uniform in the contrasting genotypes for estimation of chlorophylls. The samples were collected at 5DAP, 10DAP and at 7 days after harvesting (DAH). One fruit each from 5 randomly selected plants were collected and equal amount of samples from each fruits were collected and homogenised. The samples were refrigerated immediately after collection till the chlorophyll estimation. Chlorophylls were also extracted in acetone solution and quantified in spectrophotometer at OD of 645 nm and 663 nm.

Similarly, one fruit each from five randomly selected plants of each genotype were harvested at 10 DAP. Weight of the fruits were taken immediately after harvesting and thereafter stored under room temperature (27-30 °C) to determine moisture loss at different intervals after harvest. The wight of the fruits were recorded at 5 DAH and 10 DAH after storage. Average of five fruits were taken both for total chlorophyll content and fruit weight to compare the genotype, DC-48 and DC-83.

Fruit firmness (lb/in.^{2}) was also measured from five individual fruits at 7 days after harvest for the genotypes DC-48 and DC-83. Firmness is measured with the help of penetrometer at different regions of fruits (pedicel, middle and distal end) and average of three different locations were taken for data analysis.

2.4. RNA extraction and sequencing

Total RNA was extracted from the fruits of two contrasting genotypes for transcriptome analysis in triplicate. The surface of the fruit epicarp (2 mm) was peeled with a sharp peeler and the same peeler was used for all the samples to keep the sample uniform. Fruit epicarp was taken for transcriptome analysis as texture, structure and composition of epicarp are the principal factors determining post-harvest biology and shelf life of the cucumber fruits. These samples were immediately immersed into liquid nitrogen and later stored in -80 °C. The RNA extracted from the epicarp of the fruits was analysed separately. Total RNA was extracted using the TRIzol reagent (Invitrogen, Waltham, MA, USA). The concentration (optical density 260 nm/280 nm ratio) and quality (optical density 260 nm/230 nm ratio) of the extracted RNA were determined using Bio-analyzer (Agilent, UK). The final concentration of RNA was adjusted to 100 ng μL^{-1} for whole genome RNA sequencing using Illumina HiSeq X10. For the quantitative real-time polymerase chain reaction (qPCR) analysis, RNA was treated with DNaseI from the DNA-free kit (Promega Corporation, USA) and then checked by PCR to ensure that there was no contaminating DNA. First-strand cDNA was synthesized using the High-Capacity cDNA Reverse Transcription Kit as per the manufacturer's protocol (Promega Corporation, USA).

2.5. Pre-processing of data and transcriptome assembly

The paired-end HiSeq Illumina reads of 2*151 bp, generated from two extreme varieties of cucumber, i.e., DC-48 and DC-83, (labelled as 1 and 2, respectively) at two developmental stages viz., 5 DAP (Days after pollination) and 10 DAP (labelled as A and B, respectively) in triplicates were subjected for data cleaning and pre-processing. The pre-processing of raw reads was done using *Trimmomatic* v0.39 tool [45] to remove low quality reads and adapter sequences. We used filters like discarding the reads with read length $\leq$ 36 bp, phred score $\leq$ 3 (i.e., bases having values 3 and above are retained) and HEADCROP:10 (i.e., removed 10 bases from the beginning of the reads). The quality visualization of reads was done through FASTQC-0.11.8 tool [46].

After pre-processing, the high-quality reads were used for the construction of de novo transcriptome assembly using *Trinity* assembler v2.5.1 [47] for identification of differentially expressed genes, transcription factors, variants and gene regulatory networks. The analyses were done with four different combinations, *ie.,*DC-48 at 5 DAP vs. DC-48 at 10 DAP (1A:1B), DC-48 at 5 DAP vs. DC-83 vs. 5 DAP (1A:2A), DC-48 at 10 DAP vs. DC-83 at 10 DAP (1B:2B) and DC-83 at 5 DAP vs. DC-83 at 10 DAP (2A:2B).

2.6. Transcriptome assembly and identification of differentially expressed genes

The clean reads for each of the samples were mapped onto the cucumber reference genome using read aligner HISAT2 2.2.0 for reference-based assembly [48]. The transcripts were reconstructed for each sample using StringTie program v2.1.4 [49] over the cucumber genome annotation file (http://ftp.gramene.org/CURRENT_RELEASE/gtf/cucumis_sativus/). The resulting transcripts of all the samples were merged using Stringtie with "-merge" parameter to generate a global, unified set of transcripts across multiple samples.

For the identification of differentially expressed genes, paired-end reads of each sample were mapped individually onto the (i) generated reference-based cucumber transcriptome assembly for reference-based assembly and (ii) generated de novo cucumber transcriptome assembly for de novo assembly using Bowtie2 [50]. Transcript quantification was performed using RSEM (RNA-Seq by expectation maximization) [51] to calculate the expression. After this edgeR (Empirical analysis of Digital Gene Expression in R) [52] at stringent parameters (False Discovery Rate (FDR) < 0.001, p-value <0.05 and $\log_2$fold change $= \pm 2$) was employed for identification of significant differentially expressed genes (DEGs). This was performed for four possible comparisons, *namely,* 1A:1B, 1A:2A, 1B:2B and 2A:2B to obtain DEGs and their related plots by running utility PERL scripts.

2.7. Homology search, annotation, functional characterization and transcription factor identification

Sequences of the differentially expressed transcripts were extracted from the transcript assembly file using the tool *seqtk* (https://github.com/lh3/seqtk). The homology search was performed against NCBI nr database (ftp://ftp.ncbi.nlm.nih.gov/blast/db/) using standalone local ncbi-blast-2.2.31+ at threshold *E*-value of 10^{-3} [53]. Blast2GO Pro ver. 3.1 was used for the gene ontology to categorise transcripts into molecular functions, biological processes and cellular components, KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways analysis, enzyme classes, domains and families [54].

Blastx search at $1^{e\text{-}03}$ was used for identification of transcription factors (TFs) against PlantTFDB 4.0 (Plant Transcription Factors Database) version 5.0 for all the differential expressed genes obtained from four combinations of datasets [55].

2.8. Variant identification

Mining of putative genic region DNA markers (Simple Sequence Repeat; SSRs) was performed from the de novo assembled transcripts using MISA (MIcroSAtellite identification tool) to get the mono-, di-, tri, tetra-, penta-, hexanucleotide and compound repeats [56]. For these markers, primers were generated using Primer3 tool [57] with parameters (annealing temperature- min: 57 °C, optimal: 60 °C, maximum: 63 °C, primer size- min: 15, optimal: 18, maximum: 28 oligonucleotides).

Single Nucleotide Polymorphism (SNPs) and Insertion and Deletion (InDels) calling was performed using the bcftools pipeline where the reads of each of the samples 1A and 2A were individually mapped onto the reference cucumber genome (http://ftp.gramene.org/CURRENT_RELEASE/fasta/cucumis_sativus/dna/) using Bowtie2 tool [58]. SAMtools was used to pre-process the alignment/map files (SAM/BAM) for sorting, duplicate removal, read group addition and build BAM index for the BAM file [59]. These indexed BAM files were further used in bcftools for calling SNPs and INDELs. In order to obtain the significant variants, stringent parameters like read depth ≥ 15 and quality >30 were used. For generation of circular map of variants to visualize the relative distribution of SNPs and InDels over 7 chromosomes, Circos tool was used [60].

2.9. Gene regulatory networks

For construction of gene regulatory networks (GRNs), we filtered the "House Keeping Genes", "Uncharacteristic proteins" and "Similar genes" from the top up- and down- regulated DEGs of all the four sets (1A:1B, 1A:2A, 1B:2B and 2A:2B). These DEGs were selected based on the log fold change value for gene network analysis. The networks were visualized and analysis was carried out using *Cytoscape version 3.7.2*, where gene correlation was computed using the Pearson Correlation Coefficient using the normalized expression values for each group [61]. Network Analyzer plug-in was used for the network centrality and topology. Hub genes were identified on the basis of degree and betweenness.

2.10. Real time PCR validation

Eleven differentially expressed genes were selected for validation using qRT-PCR. Besides, 5 primers based on the genes associated with cell wall stability (Expansin, β-galactosidase and Xyloglucan endotransglucosylase) and chlorophyll break down (Pheophytinase and 7-hydroxymethyl chlorophyll *a* reductase) were also used for their expression analysis at three different time intervals after harvest viz. 1DAH, 3DAH and 7DAH. Gene sequences for primer designing were taken from cucurbits genomic database (http://cucurbitgenomics.org). The primers were designed using IDT primer quest software (https://www.idtdna.com/). cDNA was synthesized from DNase treated total RNA (2 μg) using Go Script™reverse transcription system kit (Promega, USA) as per the manufacturer's instructions and diluted 20 times with nuclease free water. The qRT-PCR was performed on light-cycler 96 system Real-Time PCR (Roche, USA) in a final volume of 10 μL containing 1 μL diluted cDNA (200 ng), 5 μL 2xSYBR Green (Go Taq qPCR system, USA), 0.4 μL each of forward and reverse primer (10 μM), and 3.2 μL RNase-free water as per the manufacturer's instructions. The thermal cycling conditions were as follows: 95 °C for 1 min followed by 40 repeated cycles of 95 °C for 10 s, 58 °C for 30 s, and 72 °C for 30 s. Relative gene expression was determined using $2 - {}^{\Delta\Delta}$CT method by normalizing to the *Actin* gene expression. The primers used for qRT-PCR validation and expression analysis of cell wall stability and chlorophyll degradation related genes along with description were listed in Supplementary Materials, Table S1.

2.11. Development of cucumber web-genomic resource

CsExSLDb (*C. sativus* Extended Shelf Life Database) is an online relational database of cucumber (*C. sativus*) transcriptome based on "*three-tier architecture*" having, client-, middle- and database tier, that catalogues the information pertaining to assembled transcripts, differentially expressed genes (transcripts) and the pathways they are involved, transcription factors, putative SSR markers, their primers and variants (SNPs and InDels). All these information are in the form of tables in MySQL in the database. For the user queries, fetching and execution, scripting in PHP has been done in the middle tier. For database browsing, web pages are developed in client tier. This resource is freely available for academic use at http://backlin.cabgrid.res.in/csexsldb/.

3. Results

3.1. Determination of shelf life of the harvested fruits

Shelf life of the cucumber fruits of contrasting genotypes, DC-48 and DC-83 was determined using two different indices based on colour change from green to yellow and extent of shrivelling. Besides, two additional parameters, total chlorophyll content and moisture loss as determined by weight of fruits at different time intervals were also conducted. Appearance of fruits at 5DAP (Fig. 1A), 10DAP (Fig. 1B) and 15 days after storage in room condition is presented in Fig. 1C. Days taken for changing the colour to stage-3 with unacceptable fruits were significantly different in the genotypes, DC-48 and DC-83 at three different intervals of harvesting. The genotype, DC-48 took 12.7–13.8 days as compared to 2.2–2.3 days in DC-83 to reach the stage-3 with first appearance of yellowing in the epicarp. Similarly, the shrivelling in the epicarp was delayed significantly in the genotype, DC-48. The appearance of the fruit cross section representing shrinkage 15 days after harvesting is presented in Fig. 1D. The days taken to the stage-3 characterised by clear shrivelling visible in the pedicel end and initiation of shrivelling in the main body of the fruits was 14.2–14.5 days in DC-48

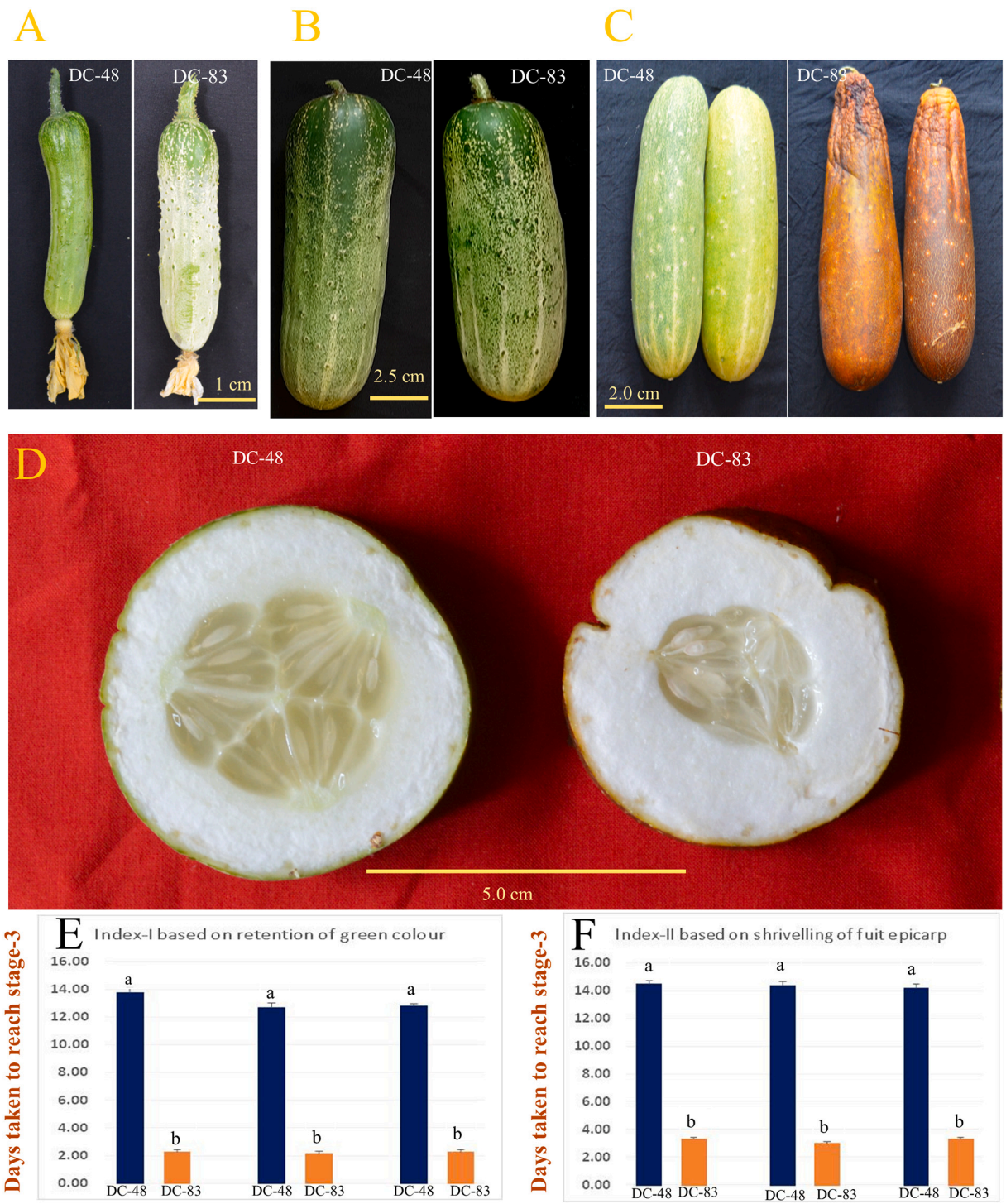

Fig. 1. Appearance and behaviour of fruits of cucumber genotypes, DC-48 and DC-83. (A) Appearance of fresh fruits at 5 days after pollination (5DAP) (B) Appearance of fresh fruits at 10 days after pollination (10 DAP) (C) Appearance of the fruits 15 days after harvesting stored under room condition (D) Cross section of the mid portion of the fruits with predominant shrinkage in DC-83 at 15 days after harvesting stored under ambient condition (E) shelf life of the fruits of DC-48 and DC-83 based on retention of fresh green colour (F) shelf life of fruits of DC-48 and DC-83 based on appearance of shrinkage on the fruit epidermis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

while it was 3.0–3.3 days in DC-83. Shelf life of both genotypes were obtained by using retention of fresh green colour (Fig. 1E) and appearance of shrinkage on the fruit epidermis (Fig. 1F).

Fruits of DC-48 and DC-83 were also analysed for chlorophyll content and moisture loss at different time intervals Figure Supplementary Fig. S1. Average chlorophyll content in the fruit epicarp of DC-48 was 7.46 mg, 6.72 mg and 5.62 mg whereas it was 6.19 mg, 4.97 mg and 1.52 mg in DC-83 at 5DAP, 10DAP and 7 DAH (days after harvesting), respectively. Similarly, different pattern of moisture loss from the harvested fruits of the DC-48 and DC-83 was observed. Average weight of the DC-48 was 298.8 g at harvesting (10 DAP). After storage under room condition, the weight of the fruits reduced to 279.4 and 266.3 g at 5 and 10 days after harvesting (DAH). In DC-83 weight of the fruits changed from 354.2 g to 306.5 g and 271.2 g at 5 and 10 DAH. (Supplementary

Materials, Fig. S1). Analysis of fruit firmness in these contrasting genotypes at 7DAH also revealed significant difference. Firmness was 13.72 lb./in.^{2} in the genotype, DC-48 whereas it was 8.21 lb./in.^{2} in the genotype, DC-83 with poor shelf life indicating its role in determining post-harvest behaviour of the fruits.

3.2. Pre-processing of data and transcriptome assembly

The paired-end Illumina reads were generated from the two cucumber varieties, namely, DC-48 (extended shelf life) and DC-83 (poor shelf life) had 71,959,228 (DC-48, 5 days), 83,176,537 (DC-48, 10 days), 69,580,245 (DC-83, 5 days) and 87,084,805 (DC-83, 10 days) reads. Summary of reads before and after trimming is presented in Table 1. After reads' pre-processing and contamination removal, 147,629,766 high-quality reads from all the samples were pooled for de novo transcriptome assembly. *Trinity* assembler generated a total of 186,184 transcript with N50 value 2919 bp. The average contig length was found to be 1734.86 base pairs with GC content of 38.89% (Table 2). In order to rule out errors (spurious transcripts) in the assembly, reference assembly was used. The assembled transcripts were mapped on *C. sativus* genome (http://ftp.gramene.org/CURRENT_RELEASE/fasta/cucumis_sativus/dna/). Out of a total of 186,184 transcripts, 170,819 transcripts mapped on the available cucumber genome making the mapping percentage as high as 91.75%. We used the successfully mapped transcripts for the downstream analysis to get DEGs and compute GRN. Pooled transcriptome assembly was used to maximize number of DEGs in all dataset. This helps in retaining any novel genes/ transcripts which are not in the reference assembly.

3.3. Transcriptome assembly and identification of differentially expressed genes

A total of 830 (722 unique) and 1866 (1514 unique) DEGs were identified from all the four comparison sets (1A:1B, 1A:2A, 1B:2B and 2A:2B) in response to the extended shelf life with defined parameters (FDR and P-value <0.05) using the EdgeR for reference based and de novo transcriptome assembly, respectively. All the DEGs identified in reference-based assembly were compared with the DEGs identified through the de novo assembly. Out of the 722 reference-based DEGs, 685 (94.87%) DEGs showed significant similarity with the DEGs identified by de novo assembly and are essentially the same transcript. While the rest of the 37 transcripts showed no similarity with any of the DEGs of de novo assembly Figure Supplementary Fig. S2. Believing that the larger number of DEGs identified by de novo approach may provide extra information, we proceeded with these DEGs for further analysis.

Out of these differentially expressed genes, a total of 622, 716, 518 and 10 transcripts were from sets 1A:1B, 1A:2A, 1B:2B and 2A:2B, respectively (Table 3, Supplementary Materials, Table S2). Heatmap showing the hierarchically clustered Spearman correlation matrix resulting from comparing the transcript expression values (TMM-normalized FPKM) for each sample is represented in Fig. 2A. Hierarchical clustering of the DEGs as well as samples is shown as the heatmap where clustering is based on transcripts abundance in the three biological replicates per sample (Figure Supplementary Fig. S3). The shared and unique differentially expressed genes are represented in the form of Venn diagram in Fig. 2B. The MA plot and volcano plot of the identified DEGs are represented in Figure Supplementary Fig. S4. Highest number of transcripts (716) were recorded when DC-48 at 5 DAP was compared with DC-83 at 5 DAP followed by DC-48 at two different developmental stages, with 622 number of transcripts. A total of 526, 435, 5 and 201 DEGs were unique to the comparison sets 1A-1B, 1A-2A, 2A-2B and 1B-2B, respectively as seen in the Venn diagram (Fig. 2B). None of the genes were found common in all the 4 sets of comparison.

Table 1
Summary of Reads' Statistics of the genotypes DC-48 and DC-83 at different fruit developmental stages.

*Sample	Input Read Pairs	Both Surviving	Forward Only Surviving	Reverse Only Surviving	Dropped
1A	71,959,228	31,173,063	40,752,590	10,445	23,130
1B	83,176,537	36,399,704	46,739,130	12,055	25,648
2A	69,580,245	35,861,319	33,686,152	11,825	20,949
2B	87,084,805	44,195,680	42,846,863	14,469	27,793

(*1 is DC-48 and 2 is DC-83; A is 5 DAP and B is 10 DAP)

Table 2
Summary of de novo assembly statistics of *Cucumis sativus*.

Total assembled reads	147,629,766
Total trinity transcripts	186,184
Average contig length (kb)	1734.86
Contig N50 (kb)	2919
Percent GC (%)	38.89

Table 3
Number of Differentially Expressed Genes and Transcription Factors in various comparison sets.

Comparison	Total DEGs	Up-regulated	Down-regulated	Transcription factors
1A:1B	622	90	532	432
1A:2A	716	354	362	437
1B:2B	518	306	212	319
2A:2B	10	4	6	6

3.4. Homology search, annotation, functional characterization and transcription factor identification

Annotation of the four sets, (1A:1B, 1A:2A, 1B:2B, 2A:2B) revealed maximum similarity with *C. sativus* followed by *Cucumis melo* var. *makuwa, C. melo and Cucurbit pepo* subsp. *pepo* (Figure Supplementary Fig. S5) which may be due to their phylogenetic similarity. Gene ontology classification of *Cucumis* gene list was carried out. For 1A:1B, 1A:2A, 1B:2B and 2A:2B, a total of 557, 589, 459 and 9 GO terms, respectively were assigned (Supplementary Materials, Table S3). It was observed that under biological processes 'biosynthetic process' was assigned the highest number of sequences followed by "cellular nitrogen compound metabolic process". In case of cellular, "cellular components" were observed to have highest number of sequences involved. "Ion binding" was observed to be the highest in molecular functions followed by "Oxidoreductase activity" and "DNA binding" (Fig. 3).

A total of 1866 differentially expressed genes were found to control 165 pathways which belonged to five major classes, *namely*, Purine metabolism, Phenylalanine metabolism, Starch and sucrose metabolism, Methane metabolism and Glycosphingolipid biosynthesis in decreasing order. Maximum differentially expressed pathway of Purine metabolism was found on the day 5 of contrasting varieties. Maximum DEGs were found in major pathway of Phenylalanine metabolism, Starch and sucrose metabolism, Methane metabolism when we compared two timelines of good keeping varieties. Glycosphingolipid biosynthesis pathway was found to be most differentiated on day 5 between two contrasting varieties. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed that from a total of 622, 716, 518 and 10 DEGs obtained, 119, 115, 95 and 2 transcripts were assigned under 1A:1B, 1A:2A, 1B:2B and 2A:2B, respectively. It was observed that purine metabolism, and phenylalanine metabolism were common in all four sets with different numbers of transcripts involved (Supplementary Materials, Table 4). Table 4 represents the major KEGG pathways from the four comparison sets.

In comparing the genotype, DC-48 with extended shelf life in two different developmental stages (1A:1B), it was evident that besides the

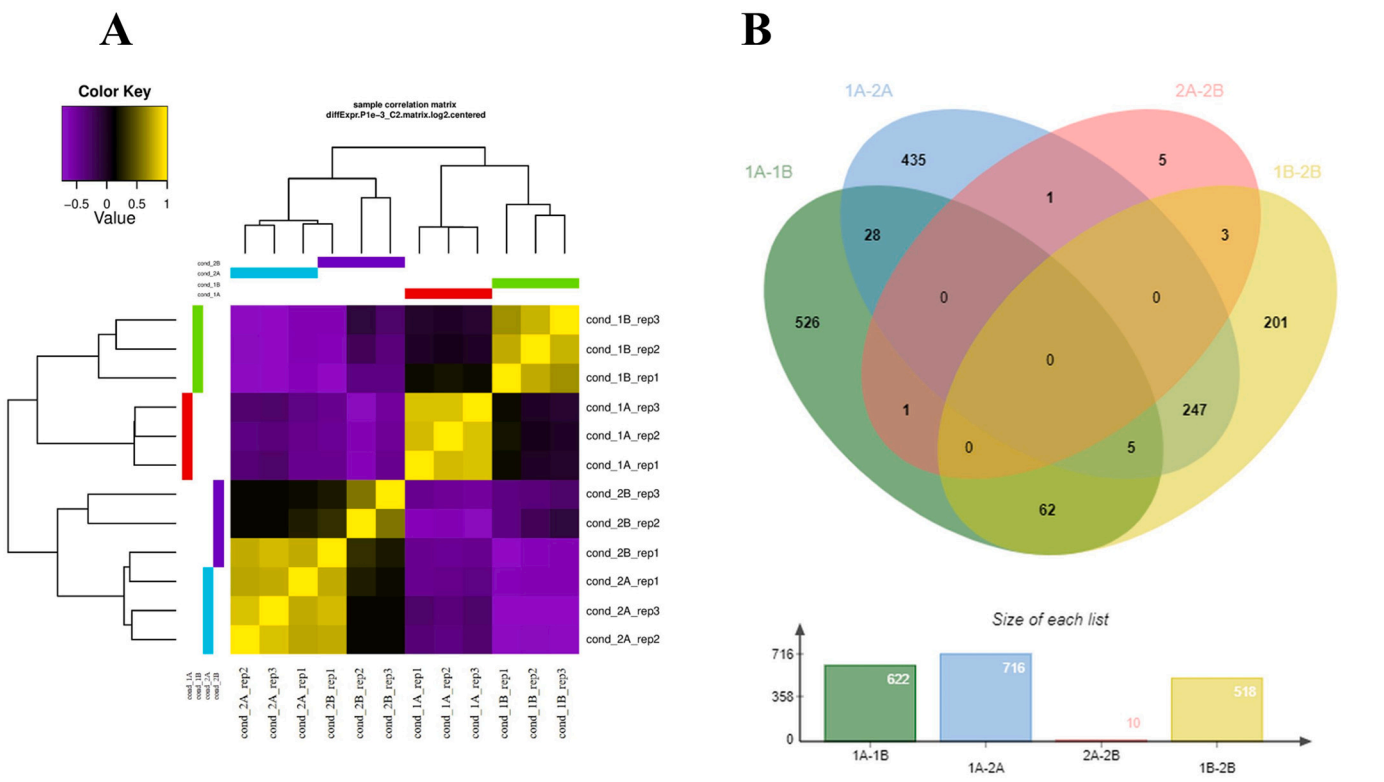

Fig. 2. (A) Heatmap of each sample dataset representing the hierarchically clustered Spearman correlation matrix (B) Venn diagram representing the unique and shared DEGs from different comparison sets.

metabolism associated with secondary metabolites, glyoxalate, dicarboxylate, glycolysis, porphyrin, chlorophyll *a*nd carotenoid metabolism associated with photosynthesis and chlorophyll degradation/synthesis played important role. However, none of the transcripts of the genotype with very less shelf life with rapid loss of green colour at two different developmental stages (2A:2B) were related either to photosynthesis or cell wall degradation metabolism. The transcripts related to the glycolysis, glyoxylate, dicarboxylate, porphyrin, chlorophyll metabolism, pentose and glucuronate interconversions were the major metabolisms in case of 1A:2A. Glycolysis, porphyrin and chlorophyll, glyoxylate and dicarboxylate, oxidative phosphorylation and photosynthesis, glycerolipid metabolism, glycerophospholipid metabolism, carotenoid biosynthesis, cutin, suberine and wax biosynthesis, and thiamine biosynthesis were the major pathways revealed in comparing 1B:2B.

We identified a total of 432, 437, 319 and 6 transcriptions factors (TFs) in 622, 716, 518 and 10 DEGs obtained from four sets of comparison, namely, 1A:1B, 1A:2A, 1B:2B and 2A:2B. ERF family protein, NAC family protein, bHLH family protein and MYB family proteins were most abundant TFs identified in 1A:2A and 1B:2B. Similarly, these TFs were also detected repeatedly in case of 1A:1B (Supplementary Materials, Table S5).

3.5. Differential expression of genes associated with cell wall stability and fruit firmness

Cell wall metabolism is crucial in determining the fruit texture, quality attributes and softening of the fruits. The genes associated with cell wall degradation and structure were expressed differently in the genotypes, DC-48 and DC-83. Polygalacturonase QRT3-like, and Expansin-A4 are crucial genes in softening of the fruits and were significantly down regulated in the genotype, DC-48 at 10DAP (1B:2B) at the stage of commercial maturity when compared with the poor keeping genotype, DC-83. The log2FC values for these two crucial genes were − 10.58 and − 3.89, respectively in the combination DC-48 vs. DC-83 at 10DAP. Similarly, Polygalacturonase At1g48100 was down-regulated in the genotypes DC-48 (−3.21) at 5 DAP also when compared with DC-83 at the same developmental stage. Whereas, Glycosyltransferase family protein, Probable xyloglucan Glycosyltransferase 12 (−2.09), Polygalacturonase inhibitor-like (−3.30), Pectinesterase/pectinesterase inhibitor PPE8B-like (−2.06), probable xyloglucan endotransglucosylase/hydrolase protein 23 (−5.28), xyloglucan endotransglucosylase/hydrolase protein 22-like (−3.44), xyloglucan endotransglucosylase/hydrolase 2-like (−4.52) and probable xyloglucan endotransglucosylase/hydrolase protein 23 (−4.51) were down regulated in the genotype, DC-48 at 5DAP (1A:1B). Polygalacturonase At1g48100 (−3.21) and Pectinesterase (2.54) were also down regulated in the genotype, DC-48 when compared with DC-83 at 5 DAP (1A:2A) (Supplementary Materials, Table S6A).

3.6. Differential expression of genes associated with chlorophyll metabolism

While studying the expression pattern of the chlorophyll metabolism related genes, it was evident that key genes related to the chlorophyll metabolism were differentially expressed in two contrasting genotypes at different developmental stages (Supplementary Materials, Table S6B). The transcripts glutamyl-tRNA reductase 1 (3.13), magnesium-chelatase subunit ChlH (4.01, 6.74), magnesium-protoporphyrin IX monomethyl ester (5.29), chlorophyllide a oxygenase (3.60), magnesium protoporphyrin IX methyltransferase (3.22), chlorophyll *a*-b binding protein 6A (4.44), chlorophyll *a*-b binding protein 8 (6.12), chlorophyll *a*-b binding protein CP24 10A (6.22), chlorophyll *a*-b binding protein CP29.1 (4.29), chlorophyll *a*-b binding protein of LHCII type 1-like (7.44, 5.08), chlorophyll *a*-b binding protein of LHCII type III (5.43), chlorophyll *a*-b binding protein P4 (7.08), photosystem II CP47 chlorophyll *a*poprotein (3.37) and photosystem II P680 chlorophyll A apoprotein (3.87) were significantly upregulated in the genotype DC-48 as compared DC-83 at

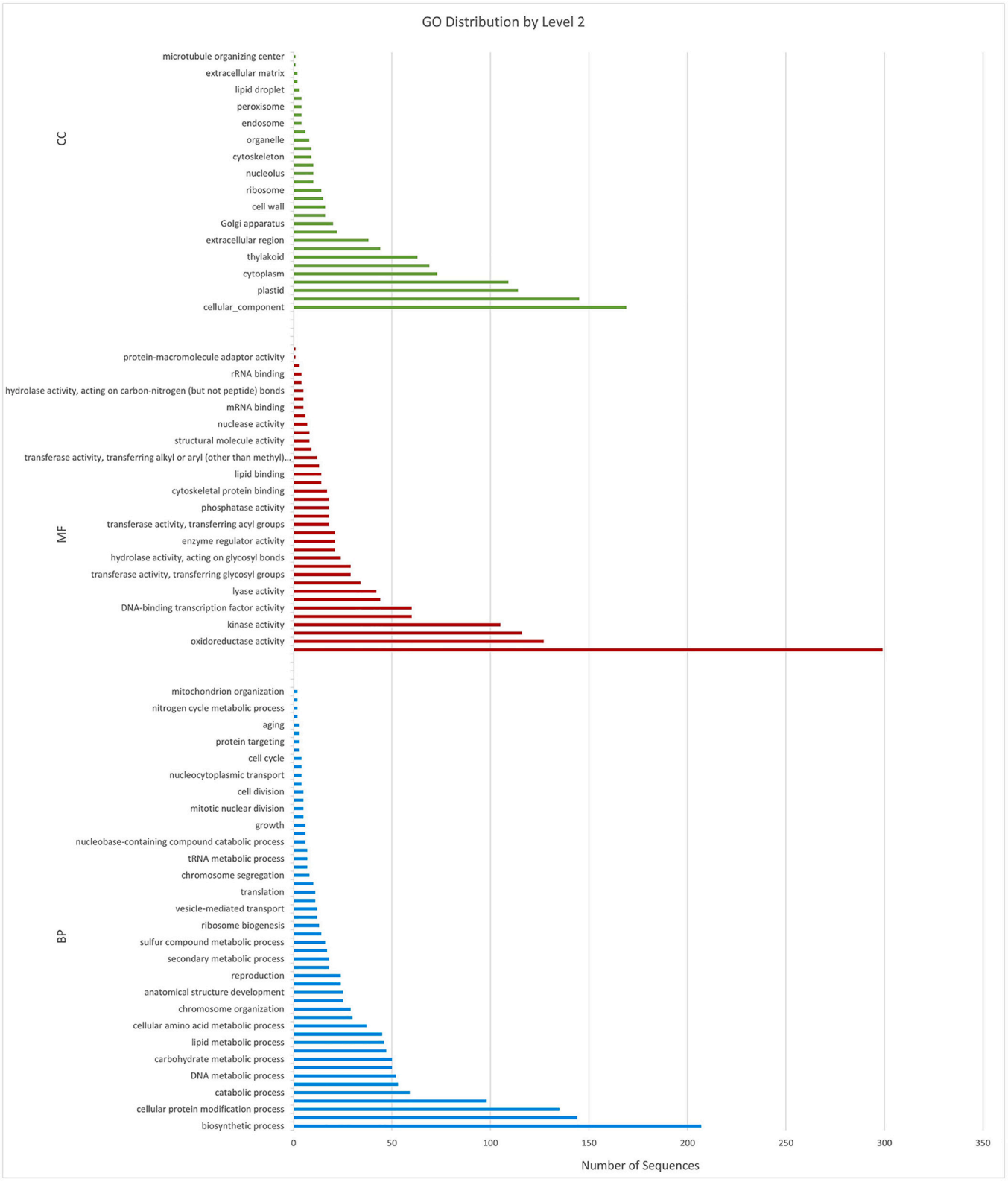

Fig. 3. Gene Ontology counts for Biological processes, Cellular Components and Molecular functions.

10DAP (1B:2B). Similarly, the transcripts magnesium-protoporphyrin IX monomethyl ester [oxidative] cyclase (2.46), chlorophyll *a*-b binding protein 8 (2.08), chlorophyll *a*-b binding protein CP24 10A (2.34) and chlorophyll *a*-b binding protein of LHCII type 1-like (3.16) were upregulated in the genotype, DC-48 at 5DAP too (1A:2A). In studying the genotype, DC-48 at two different developmental stages (1A:1B), it was evident that the key transcripts magnesium-chelatase subunit ChlH (−3.80, 4.23), magnesium-protoporphyrin IX monomethyl ester [oxidative] cyclase (−2.52), chlorophyllide a oxygenase (−2.79), glutamyl-tRNA reductase 2 (−2.99), delta-aminolevulinic acid

Table 4
Major KEGG pathway analyses of the DEGs from the various comparison sets.

Pathways	1A:1B	1A:2A	1B:2B	2A:2B
Purine metabolism	24	43	28	1
Phenylalanine metabolism	18	1	2	1
Starch and sucrose metabolism	9	9	9	0
Methane metabolism	6	0	6	0
Glycosphingolipid biosynthesis	0	7	1	0

dehydratase (-2.19), chlorophyll *a*-b binding protein 8 (−2.27), chlorophyll *a*-b binding protein CP24 10A (−2.45), chlorophyll *a*-b binding protein CP29.1 (−2.18), chlorophyll *a*-b binding protein of LHCII type 1-like (−2.10), chlorophyll *a*-b binding protein of LHCII type III (−2.32), chlorophyll *a*-b binding protein P4 (−3.03) and photosystem I chlorophyll a/b-binding protein 6 (−8.21) were significantly downregulated at 5DAP as compared 10DAP. However, it was notable that there was no significant difference in the expression pattern of key genes in the genotype DC-83 at two different developmental stages (2A:2B).

3.7. Differential expression of genes associated with ethylene biosynthesis

The key genes for ethylene metabolism, 1-aminocyclopropane-1-carboxylate synthase (ACS) (−7.16) and 1-aminocyclopropane-1-carboxylate oxidase (ACO) (−4.20) and 1-aminocyclopropane-1-carboxylate oxidase homolog 1-like (−6.64) were expressed differently in the genotype DC-48 at two different developmental stages (1A:1B). Along with ACO and ACS, several ethylene-responsive transcription factors (ETFs) were down regulated at 5DAP in the genotype, DC-48. However, no significant difference in expression pattern of these genes and ETFs were recorded in the genotype, DC-83 (2A:2B) and among DC-48 and DC-83 (1A:2A; 1B:2B) (Supplementary Materials, Table S6C).

3.8. Variant identification

Based on the de novo assembled transcripts, a total of 21,524 putative SSRs were identified. Out of this, mononucleotides were 12,251, dinucleotide were 3824, trinucleotide were 3075, tetranucleotide were 278, pentanucleotide were 86 and hexanucleotide were 81 in number. Maximum abundance was observed in mono-type (56.91%), followed by di- type (17.76%), and tri- type (14.28%). The primers were designed for 21,477 SSRs while for 47 SSRs primers could not be designed from the genic region SSRs mined from the de novo assembly. Supplementary Materials, Table S7 gives the details of all the identified SSR markers and their primers.

SNPs and InDels were identified in all the four samples (1A, 1B, 2A, 2B). A total of 671,425 SNPs and InDels were identified. Out of this 545,173 were SNPs and 126,252 were InDels. Individual distribution of SNPs and InDels from all the samples is provided in Table 5. However, many SNPs and InDels were shared among these four samples. The Venn diagram shows the shared and unique SNPs and InDels across all the four samples (Fig. 4A). A total of 183,816 variants were unique to a particular sample while 107,468 variants were shared by two samples. A total of 33,311 variants were shared by three samples and 41,485 variants were shared by all the four samples. We also compared the SNP and InDel distribution between the two contrasting cucumber genotypes. Circos plot (Circular map) was generated to depict chromosome-wise SNP and InDel distribution between the two cucumber varieties (Fig. 4B).

Table 5
Distribution of SNPs and INDELs in the samples 1A and 2A.

Condition	SNP	INDEL
1A	129,532	28,769
2A	150,420	36,215

3.9. Gene regulatory networks

For construction of Gene regulatory networks (GRN), DEGs were used [62] to get the idea of the co-expressional network and the cross talk between key genes that are associated with the trait under study. The SNPs of genes used in the GRN are reported to be involved in regulating trait [63]. We constructed GRNs for 1A:1B, 1A:2A, 1B:2B and 2A:2B with 48, 49, 40 and 10 DEGs, respectively after applying the filtering criteria (Supplementary Materials, Table S8). Genes represented in green are up-regulated while those in red are down-regulated. We found more down-regulated hub genes for the comparison set 1A:1B (Fig. 5A; Fig. 5B), while number of up-regulated hub genes were more than down regulated for comparison sets 1A:2A (Fig. 6A; Fig. 6B) and 1B:2B (Fig. 7A; Fig. 7B). Fig. 8 represents the heatmap for the genes involved in GRN construction for comparison set 2A:2B.

3.10. Real time PCR validation

To authenticate the RNA seq data for transcriptomic study, a set of 11 randomly differentially expressed genes including seven down-regulated transcripts and four up-regulated and four down-regulated transcripts for comparison of DC-48 and DC-83 at different developmental stages were selected. There was significantly high correlation (0.98) in expression pattern of the selected DEGs in transcriptomic with qRT-PCR data. Similar pattern of up- and down-regulation of all the selected transcriptomes confirmed the reliability and quality of the transcriptome data and the estimation of gene expressions based on FPKM values (Supplementary Materials, Fig. S6A and Fig. S6B).

Real time PCR was also conducted for five key genes associated with cell wall stability and chlorophyll degradation in the two contrasting genotypes at 3 different time intervals starting from 1 day to 7 days after harvest. We have not taken any genes for ethylene biosynthesis after getting an idea about their limited role because of non-climacteric nature of the cucumber fruits. Relative expression of cell wall stability and chlorophyll degradation related genes revealed that these key genes were highly expressed in the genotype, DC-83 when compared with DC-48 in all three different time intervals. Besides, the all the five genes were up-regulated in both the genotypes with the increasing intervals after harvest though the extent of upregulation was significantly higher in case of DC-83 with poor keeping quality (Supplementary Materials Fig. S7) The expression pattern were similar to the results obtained through RNA-seq at two different time intervals before harvesting the fruits.

3.11. Development of cucumber web-genomic resource

Transcriptome based cucumber specific web-resource, CsExSLDb (*C. sativus* Extended Shelf Life Database) is freely available at http://backlin.cabgrid.res.in/csexsldb/ for academic use. It has 6 tabs, *namely,* “Home”, “Genes”, “Markers”, “Help”, “Download” and “Contact” It catalogues the information related to assembled transcripts along with length, putative function, locations on reference chromosome, BLAST details and the pathways associated. In addition to these, the candidate genes with their expression values, transcription factor and the TF Protein families are also provided. The genic region putative markers like SSRs, their type (mono, di, tri, compound etc.), location on the transcripts along with five possible primers are provided. In case of variants, information for both SNP and InDel has been provided which includes the chromosome number with location, reference and alternate allele, the quality and the read depth. CsExSLDb houses 186,184 transcripts, 1866 DEGs, 1194 Transcription Factors, 21,524 SSR, 545173 SNPs and 126,252 InDels. The entire database occupies ~102 Mb of data. CsExSLDb is designed to be easily navigable and allows users to browse the data easily. The ‘Help’ page will assist the users how to use this database. The “Download” tab provides links to the original data submitted to NCBI for download. Fig. 9 represents the interface for the

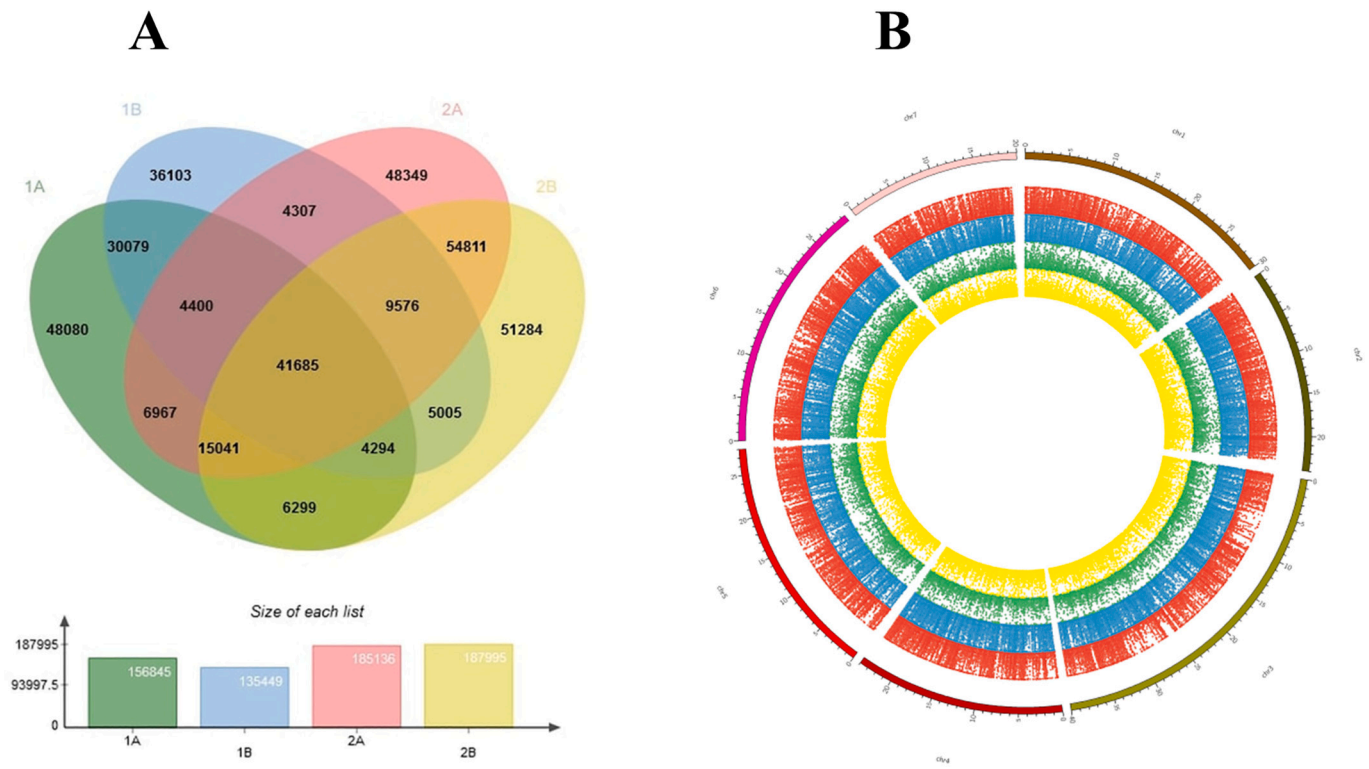

Fig. 4. Venn diagram and Circos plot representing the identified variants (A) Venn diagram representing the unique and shared number of SNPs in each sample dataset (B) Circos plot representing the distribution of variants in the Cucumber transcriptome assembly as compared to reference Cucumber Genome having 7 chromosomes (Red and blue dots represent SNPs under conditions 1A and 2A, respectively while green and yellow dots represent INDELs under conditions 1A and 2A, respectively). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

usage of CsExSLDb.

4. Discussion

Post-harvest biology, shelf life and losses due to decay of fruits are inter-linked with each other. Shelf life of fruits after harvest are determined by an array of physiological, biochemical and developmental changes in a genetically regulated and coordinated manner [64–67]. Based on the respiratory pattern, cucumber fruits are non-climacteric in nature [28]. We report for the first time the detailed molecular networks associated with degradation of the fruits and post-harvest biology in cucumber based on the transcriptional analysis of two contrasting genotypes through RNA-seq at two different developmental stages. The evaluation of the natural variant, DC-48 with extended shelf life revealed that extended shelf life was mainly determined by retention of fruit firmness, flesh texture and attractive green colour of the fruit epidermis without any shrinkage even after several days of harvesting under room condition. In contrast, the genotype, DC-83 become unfit for consumption because of visible shrinkage initiated from pedicel end, softening of fruits and yellowish flesh colour within 2–3 days of harvesting. While investigating chlorophyll content and moisture loss from the fruits at different time interval, it was evident that degradation of chlorophyll was slower in the genotype with better keeping quality. Similarly, moisture loss was more prominent in the poor keeping genotype, DC-83. Moisture loss in the genotype, DC-83 was 13.5% and 23.5% as compared to 6.4% and 10.8% of the original fruit weight in DC-48 at 5DAH and 10 DAH, respectively. Therefore, reduced water loss from the fruits in the genotype, DC-48 was probably responsible for delayed shrinkage of the fruits thus better keeping quality. While comparing fruit firmness at 7DAH, it was evident that the genotype DC-48 had much higher firmness as compared to the poor keeping genotype, DC-83. Therefore, this trait was one of the most important key determinant of shelf life of the cucumber genotypes. The findings of this study regarding differential expression of the key genes associated with chlorophyll, ethylene and cell wall stability at two different developmental stages provided insight about post-harvest behaviour of the fruits.

The total number of DEGs from de novo transcriptome assembly (1514) is more than those from reference-based (722). Since the draft genome assembly of cucumber had total length of 243.5 Mb which is about 30% smaller than the estimated genome size, we can believe the loss of information from the missing parts. Also, we got partial transcripts in de novo transcriptome assembly where each gene comprises of more than one transcript, hence inflating the number of transcripts. Further, genes were predicted having an average of 4.39 exons per gene leading to the increased number of transcripts [68] https://academic.oup.com/nar/article/30/1/191/1332222].

We found the number of transcripts (186184) to be more than the number of genes reported in cucumber. Since in most of the eukaryotes, average number of exons is 5.4 per gene [68]: https://academic.oup.com/nar/article/30/1/191/1332222] and if 3 and 3` UTR regions are added further for every gene, the number of transcripts will be inflated, which is observed here. Limited number of DEGs (only 10) were observed within poor shelf life cucumber (DC-83) across timeline of 5 and 10 days (2A:2B) which indicates that major physiological pathways are already set into motion within/up to 5 days in this genotype and limited number of genes are involved at later stage. To simplify, DC-83 variety is poor in initial 5 days itself in terms of their genetic ability of shelf life and retention of green colour. Interestingly, major differences between genotypes due to their contrasting genetic back-up in terms of shelf life and de-greening was found. We found highest DEGs (518) when we compare two genotypes at their respective time-line (5 and 10 days) which reflects physiological changes between 5 and 10 days. On comparison within genotype, it was found that extended shelf life (good keeping line, i.e., DC-48) had higher number of DEGs (622) in their timeline. Maximum DEGs (>716) were found when we compared two

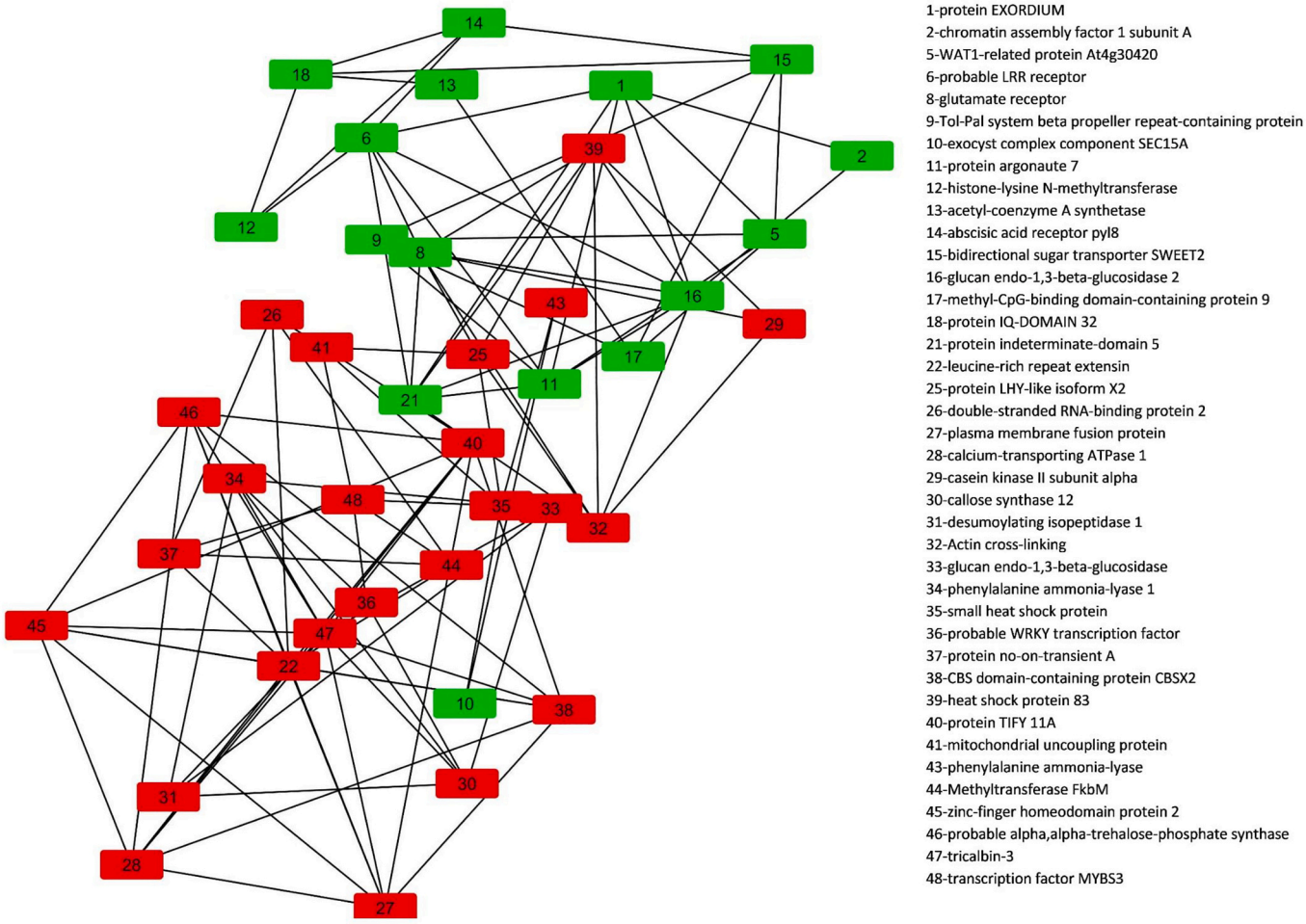

Fig. 5. Gene regulatory network (A) and heat map (B) of the top 48 genes of cucumber transcriptome associated with shelf life in the comparison set 1A-1B.

contrasting genotypes at the earliest data point of 5 days, enabling us to delineate the genetic contrast / genetic backup of two contrasting genotypes decoding their key candidate genes and molecular pathways. DC-48 line is good in the first 5 days itself due to its inherent genetic makeup.

Gene regulatory networks in two genotypes at different time point also establishes the role of key hub genes associated with exceptionally high shelf life and slow degradation of DC-48 fruits. In the poor keeping genotype (DC-83) across timeline of 5 and 10 days (2A:2B) the fold difference was very high (>8) which depicts that limited set of pathways are operating in this poor keeping lines at later stage. The top 10 DEGs used in GRN provided insight about its poor shelf life. These limited genes pertain to pathways of fruit ripening and post-harvest life. The role of the up-regulated transcripts, mediator-associated protein 1, VQ motif-containing protein and serine/threonine-protein kinase SMG1-like in ripening and shelf life has been reported earlier [69–71] in different fruit crops. Prostatic spermine-binding protein and WRKY transcription factor genes are reported to play role in fruit ripening. Prostatic spermine-binding protein genes binds polyamines which act as 'surrogate messengers' and nudge other signalling molecules like plant hormones and NO activating the genetic network to regulate late developmental stages of fruit ripening and senescence [72]. The ethylene-responsive WRKY genes show up-regulation during fruit ripening. They are known to activate the genes associated with development of fruit colour in tomato [22]. Phenylalanine ammonia-lyase known to be involved in prolong postharvest shelf life of fruits like melons. It lowers rate of O_2/CO_2 transmission coefficient, limiting food-borne pathogenic bacterial growth, higher antioxidant activities, and also higher reactive oxygen species (ROS) scavenging and anti-browning activities of related enzymes in the tomatoes [73]. In fig (*Ficus carica*) fruit, pleiotropic drug resistance protein 1 has been found associated with fruit size, colour and post-harvest transportability and shelf life [74]. Up-regulation of these key hub genes in 1A:1B explained the molecular network responsible for poor shelf life of DC-83.

The key hub genes up-regulated in 1A: 2A were 1, 3-isopropylmalate dehydratase small subunit 3, Callose synthase, Photosystem II CP43, Protein Ycf2 and MYC2 TF gene. In melon 1, 3-isopropylmalate dehydratase small subunit 3 has been found to be associated with high fruit flesh firmness during ripening. It catalyzes the stereospecific isomerization in 2-isopropylmaleate in biosynthesis of the amino acid leucine which is a potential precursor of volatile organic compound like 2-ethylhexan-1-ol, nonanol or 4-isopropylhexanol involved in fruit ripening and shelf life [75]. Similarly, callose synthase reported to play important role in softening of cell wall in the process of ripening [76]. Protein Ycf2 gene is involved in plastid protein import, regulation, and involvement in plastid differentiation in fruit ripening [77] and MYC2 gene is a key regulatory factor in the JA pathway and SlFSR expression level associated with shortening of fruit shelf life in tomato [78]. Similarly, Argininosuccinate synthase, Light-dependent short hypocotyls protein gene, Flavin-containing monooxygenase, Nitrate regulatory gene2 (NRG2) and 1, 3-oxoacyl-[acyl-carrier-protein] reductase 4 were down regulated in 1A when compared with 2A. Argininosuccinate synthase and Light-dependent short hypocotyls protein gene are important in synthesis of polyamine and carotenoid biosynthesis, respectively and associated

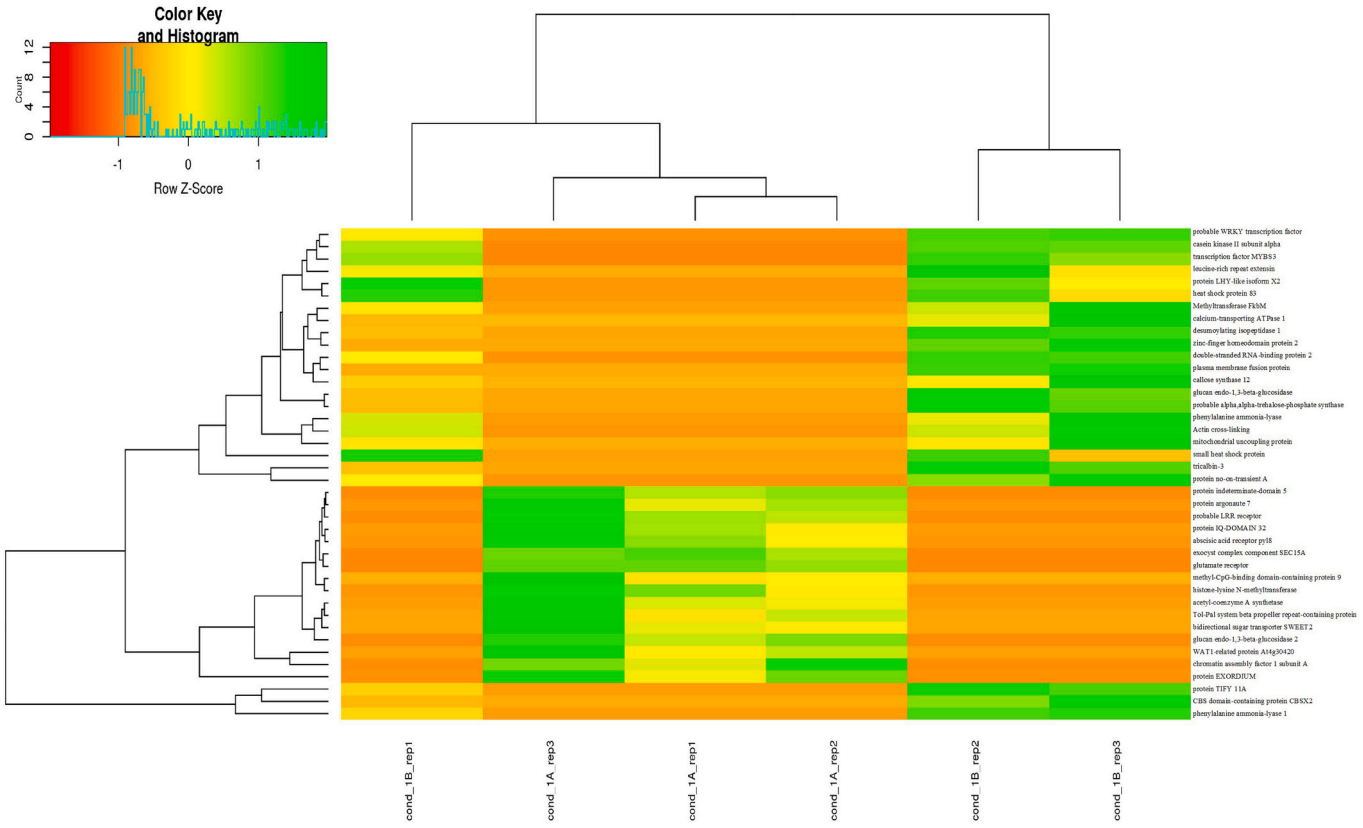

Fig. 5. (*continued*).

with fruit ripening [79,80]. Besides, Flavin-containing monooxygenase and Nitrate regulatory gene2 (NRG2) are key hub genes associated with auxin-mediated transcription contributing faster ripening by earlier initiation of ethylene response (auxin-ethylene crosstalk) [81] nitrate signalling controlling nitric oxide which counters ethylene effects on ripening fruits [82], respectively. 1, 3-oxoacyl-[acyl-carrier-protein] reductase 4 is another key gene associated with fruit ripening in banana [83].

The top transcripts in GRN analysis of 1B:2B were instrumental to dissect the mediation of change in colour and shelf life in the contrasting genotypes. The up-regulated transcripts were Fasciclin-like arabinogalactan protein, HMG genes, ATP-dependent (*S*)-NAD(*P*)H-hydrate dehydratase, Argininosuccinate synthase and Nitrate regulatory gene2 protein-like. Fasciclin-like arabinogalactan protein plays important role in development of cell wall component and fruit metabolism and fruit ripening [84]. HMG genes acts as fruit size regulators with pleiotropic effect on seeds and flowers with coordinated regulation. ATP-dependent (*S*)-NAD(*P*)H-hydrate dehydratase, Argininosuccinate synthase and Nitrate regulatory gene2 protein-like genes were up regulated in HMG genes early developmental stage (1A:2A) also. Therefore, these hub genes play critical role starting from early developmental stage to determine the shelf life in cucumber. Zinc finger CCCH domain-containing protein (TF), cytochrome P450, tyrosine sulfotransferase, UDP-glucuronic acid decarboxylase 2, Phenylalanine ammonia-lyase were key down regulated transcripts in 1B:2B. Zinc finger CCCH domain-containing protein mediates synergistically control persimmon fruit deastringency [85]. Cytochrome P450mediates the synthesis of pigments, odorants, flavours and order−/genus-specific secondary metabolites [86]. Tyrosine sulfotransferase is known for regulating carotenoid pathways and also regulates betalain accumulation of pitaya fruit [87]. The role of UDP-glucuronic aciddecarboxylase 2 in cuticle biogenesis has also been reported [88].

Besides, the top transcripts used in formation of GRN, we have also investigated the role of key transcripts associated with cell wall stability, chlorophyll and ethylene metabolism associated with extraordinary shelf lifekeeping quality of the genotype, DC-48. Identified DEGs and associated pathways in two contrasting genotypes at different time interval revealed that the genes associated with cell wall degradation and softening of fruit surface, loss of green colour and ripening mediated degradation played pivotal role in determining the shelf life of cucumber fruits. Softening of fruits are mainly due to modification of the rigid cell wall which is primary contributor of firmness [1,36,89]. The genes responsible for cell wall metabolism and ripening of the fruits are directly related with shelf life of fruits [33,34,36,90,91]. It was revealed that the genes related to the polygalacturonase (PG) and expansin (EXP) were significantly down regulated in the genotype, DC-48 at the stage of marketing maturity (1B:2B). Besides, PG was also down regulated at 5DAP in DC-48 indicating its role in determining the cell wall integrity and firmness of the fruits even after harvest. Therefore, the retention of fruit firmness in this genotype can be attributed to the PG and EXP metabolism besides other hub genes as indicated earlier. The role of the genes encoding the PG and EXP degradation enzymes in ripening and flesh softening have been established in different fruits [36,38]. In the early stage of fruit development at 5DAP, the genes related to the cell wall degradation like pectinor of matrix glycans, xyloglucan endotransglycosylase-hydrolases (XTHs) were down regulated in the genotype, DC-48 as compared to the harvesting stage at 10DAP, indicating their role in better keeping quality of the fruits after harvest. Therefore, the key mechanisms and networks associated with better keeping quality of DC-48 must have come into action much before harvest at immature stage itself. Role of the early developmental stages in determining the ripening pattern and as a consequence the keeping quality has already been established in melon [27]. Down-regulation of some families of the PG and pectinesterase in early developmental stage (1A:2A) also confirmed their larger role even after harvesting. In tomato and bell pepper the role of PG in fruit softening have been reported by

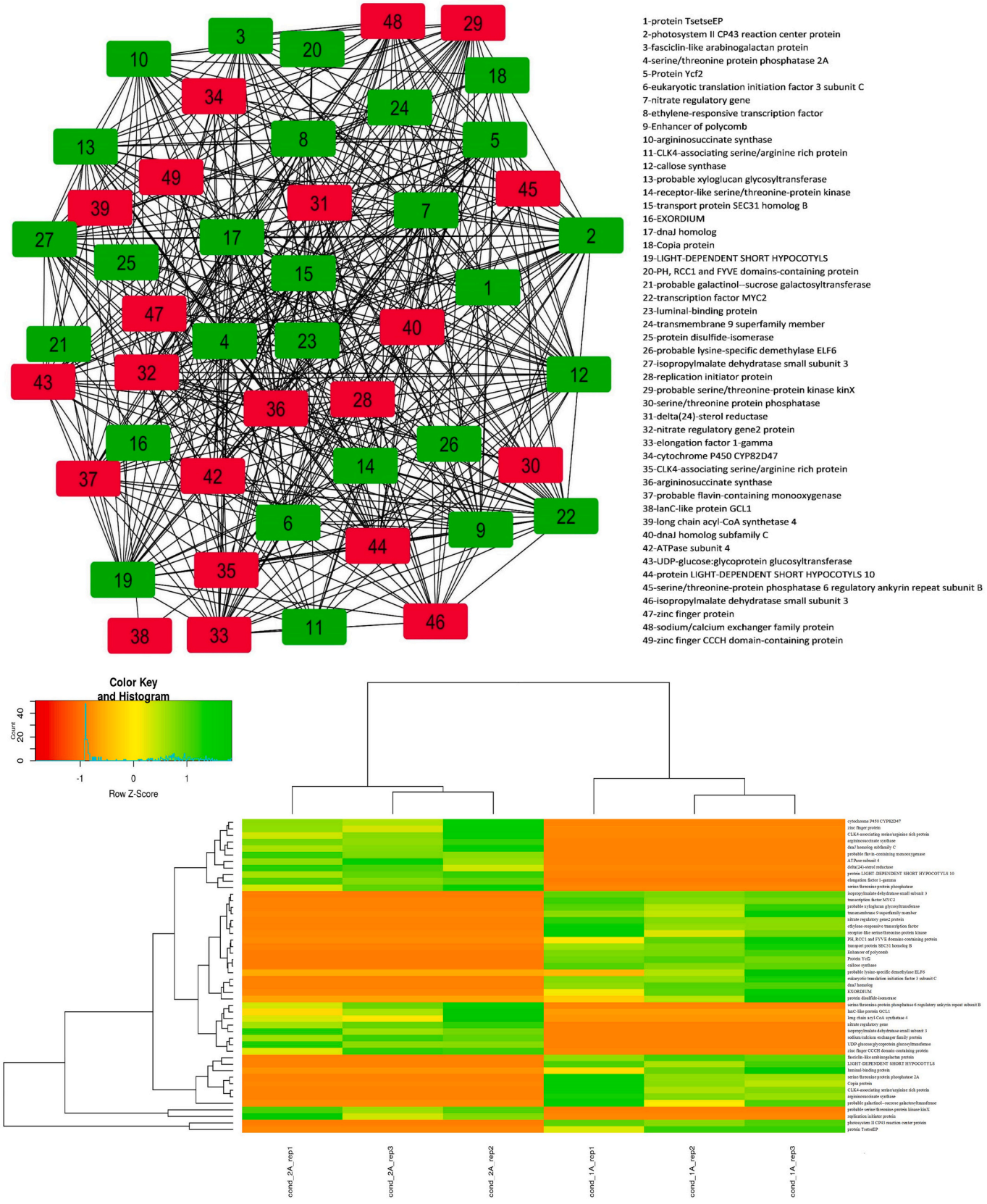

Fig. 6. Gene regulatory network (A) and heat map (B) of the top 49 genes of cucumber transcriptome associated with shelf life in the comparison set 1A-2A.

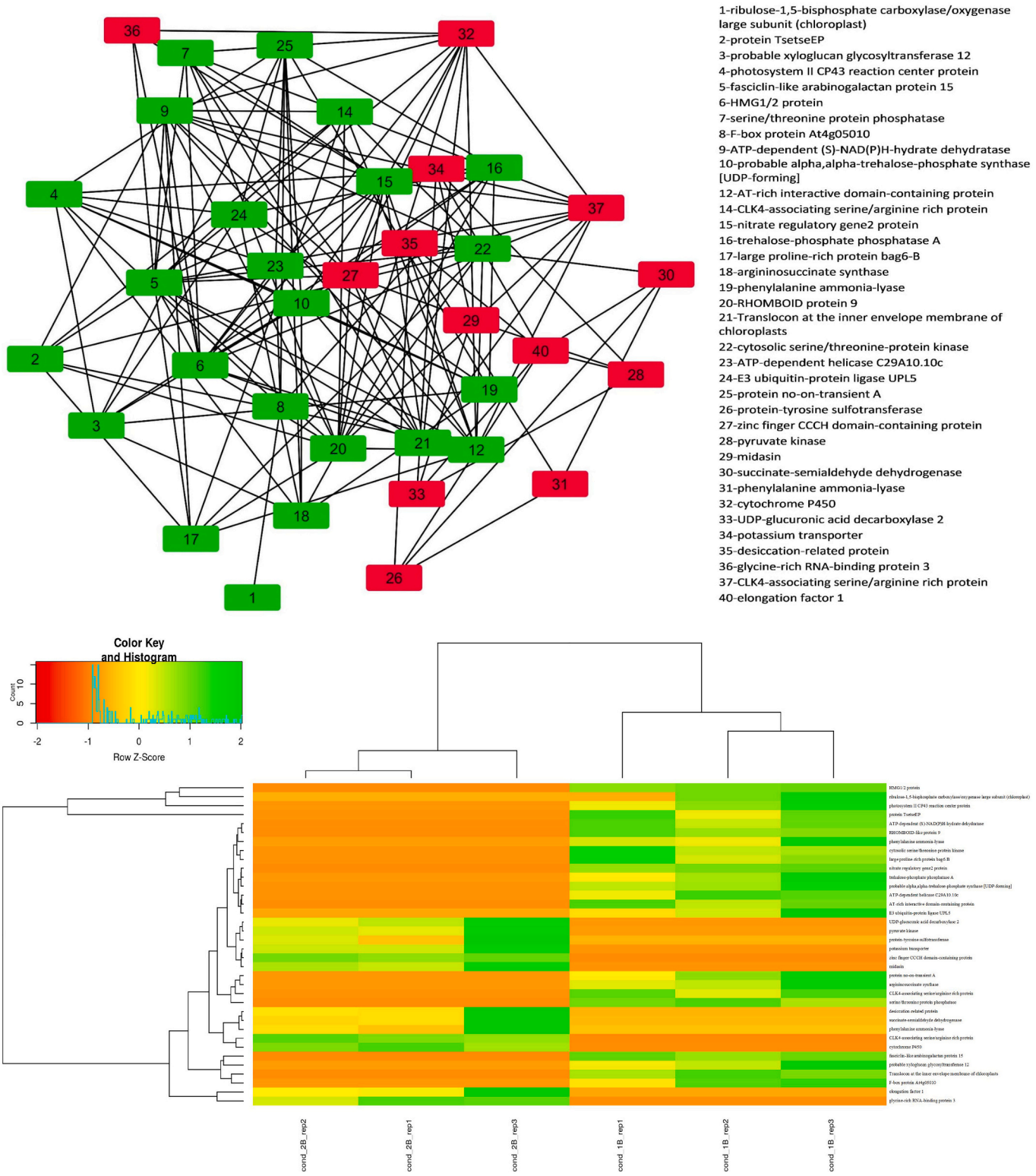

Fig. 7. Gene regulatory network (A) and heat map (B) of the top 40 genes of cucumber transcriptome associated with shelf life in the comparison set 1B-2B.

several workers [33,36,92,93]. These findings established the role of the cell wall degradation metabolisms in enhancing the shelf life in cucumber through retention of fruit firmness.

Besides the stability of the cell wall and maintenance of fruit firmness, ability of the genotype DC-48 to retain fresh green colour even after several days of harvesting was crucial determinant of keeping quality and marketability of the fruits. In cucumber, external appearance and colour of the fruit surface is an important criterion of consumer preference [94]. Epicarp colour formation and retention is largely determined by chlorophyll and carotenoid pigments and their proportion [95]. Green, orange, yellow, light green and white colour in different cucurbit fruits are because of different composition and concentration of chlorophyll and carotenoid [20,96,97]. The appearance of the cucumber fruits after harvest is determined by retention of the green

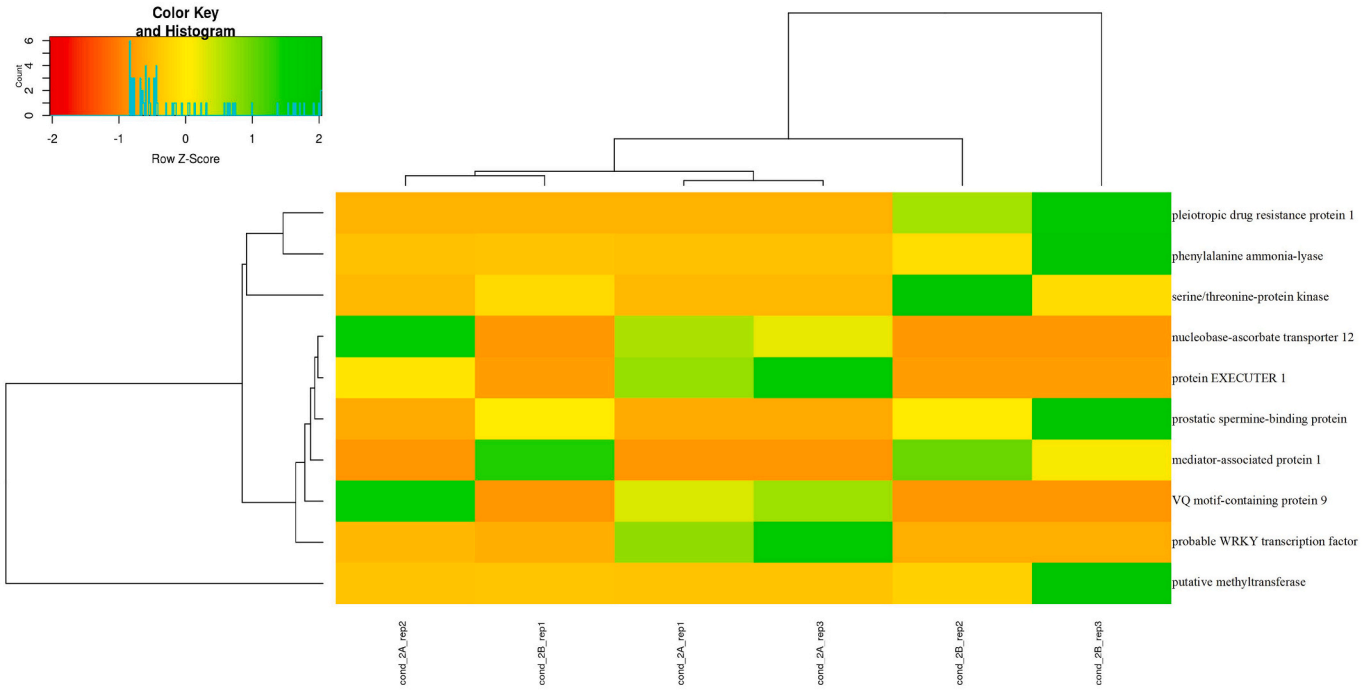

Fig. 8. Heat map of the top 10 genes of cucumber transcriptome associated with shelf life in the comparison set 2A-2B.

flesh colour. Loss of green colour is because of degradation of chlorophyll (Chl) by Chl catabolic enzymes (CCEs) which occurs mainly during the leaf senescence. Stay green phenotype in cucumber is due to mutations in CCE genes resulting delayed leaf senescence [41,42]. For example, the STAYGREEN (SGR) gene encoding the magnesium dechelatase is a key regulator in the Chl degradation pathway [98]. A large number of transcripts encoding magnesium chelatase, magnesium-protoporphyrin, glutamyl-tRNA reductase, chlorophyllide a oxygenase, chlorophyll a-b binding protein were up-regulated in the genotype, DC-48 at 10DAP indicating their role in retention of fresh green colour after harvesting. Lower expression level of most of these genes at 5DAP could be because of their greater role at maturity rather than at immature stage. Therefore, the retention of epicarp green colour of the fruits after harvesting in the genotype, DC-48 was because of down-regulation of the CCEs transcripts. Findings of the cell wall related genes and chlorophyll metabolism establishes that the motion for cell wall related genes were set much before the harvest as we observed their differential expression even at 5DAP whereas the role of the chlorophyll metabolism related genes became greater when the fruits approaching the harvesting maturity. Real time PCR of selected genes related to the cell wall stability and chlorophyll degradation at different time intervals after harvest followed the similar pattern with the results obtained through RNA-seq. Higher expression of the key genes like Expansin, β-galactosidase and Xyloglucan endotransglucosylase in the genotype DC-83 was associated with cell wall degradation and reduced firmness after harvest. These genes were advocated as the key genes associated with fruit firmness in wide variety of fruit crops [36,38]. Besides, the genes associated with chlorophyll break down (Pheophytinase and 7-hydroxymethyl chlorophyll a reductase) also exhibited similar expression pattern with much lower expression in the genotype, DC-48 with higher shelf life and capacity to retain the fresh green colour after harvest for a much longer time. Similar expression pattern before and after harvest of the fruits indicated that the underlying molecular network involved with extended shelf life of cucumber must have initiated much before harvest and the findings through RNA-seq at different fruit development stages before harvest should be applicable for behaviour of the fruits even after harvest also. Correlation between expression pattern of genes at different fruit developmental stages before harvest and ripening pattern and as a result the post-harvest life of the fruits has been established in melon [27].

Besides respiration, shelf life of most of the fruits and decay after harvest is found to be directly correlated with the level of ethylene released [99,100]. While comparing the contrasting genotypes it was found that there was no significant difference in the expression pattern of the key ethylene biosynthesis genes among DC-48 and DC-83 either at 5DAP or 10DAP indicating limited role of the ethylene in determining the post-harvest biology and extended shelf life of the cucumber fruits. Limited role of ethylene in determining shelf life and post-harvestpost-harvest biology of cucumber is attributed to the non-climacteric nature of the cucumber fruits as ripening process is generally non-responsive to ethylene.

Purine, phenylalanine, and phenylpropanoid metabolisms are important for wide variety of functions including their role in fruit ripening, membrane structure and senescence [101–103]. Chlorophyll metabolism related to photosynthesis, glycerolipid metabolism, glycerophospholipid metabolism, carotenoid biosynthesis, cutin, suberine and wax biosynthesis, and thiamine biosynthesis were the major pathways identified among the two contrasting genotypes through KEGG analysis. Role of chlorophyll metabolism and carotenoid biosynthesis in retention of fresh green epicarp colour and turning of green to yellow, brown, or orange in well known [104–107]. Therefore, altered chlorophyll metabolism for slower degradation in the genotype with extended shelf life was the principal factor for retention of fresh green colour of the fruits even after several days of harvesting in the natural variant, DC-48. Retention of firmness phenotype is very complex but the glycerolipid metabolism, glycerophospholipid metabolism, cutin, suberine and wax biosynthesis, revealed through KEGG pathway analysis are reported to be directly or indirectly involved with structure of the cell wall, transpiration loss and fruit firmness [108].

Gene expression value is spatio-temporal in function and we compared DEGs at two different cardinal datapoints of genetic and physiological response of the genes. Genetic response of the trait of degreening and shrivelling can easily be dissected by comparing contrasting varieties at two different timelines (5 DAP and 10 DAP). Physiological response of the trait can be seen within the genotype itself across timeline depicting the pathways involved. This clearly

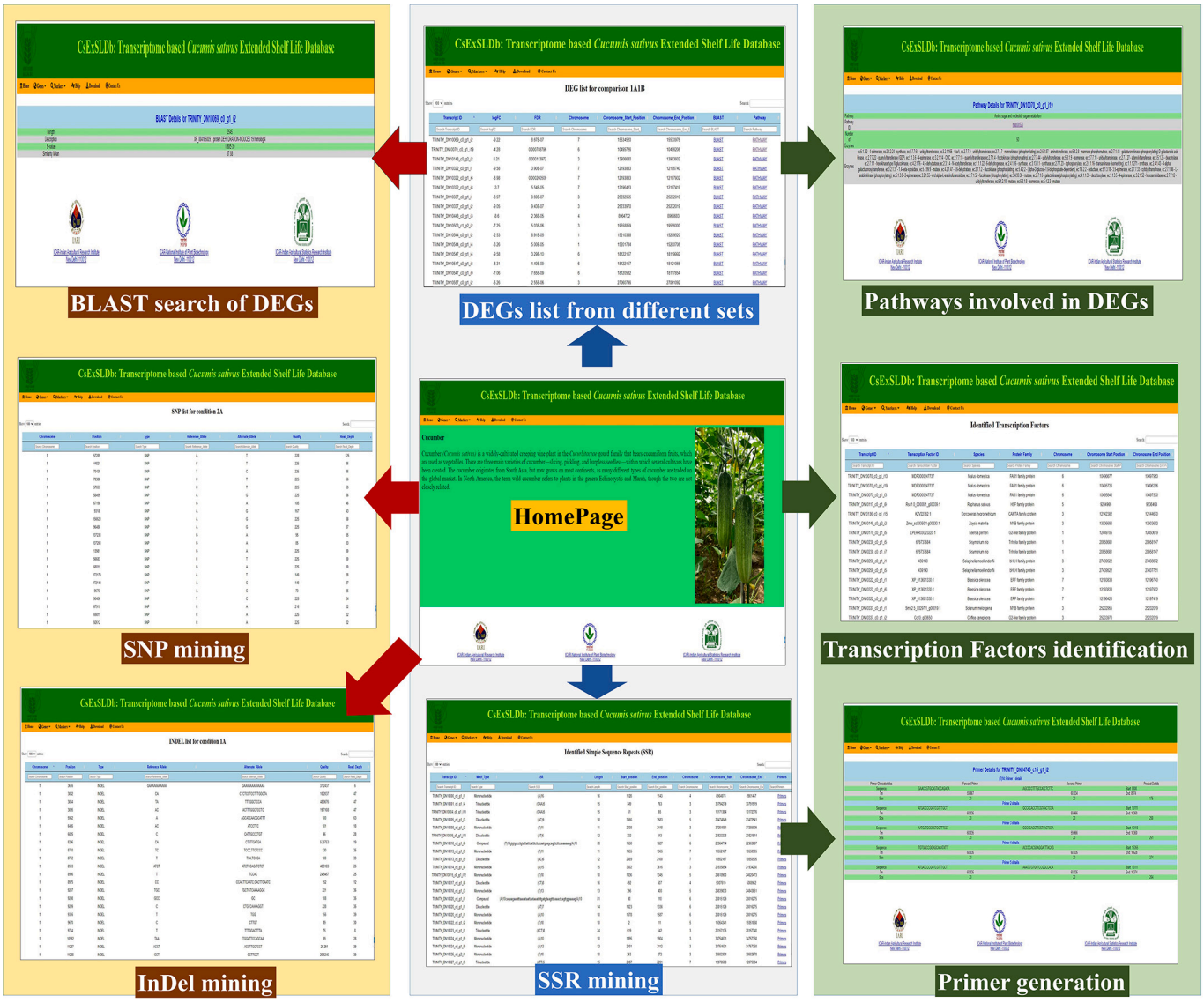

Fig. 9. Web interface of database (CsExSLDb) showing various search option facilities.

demonstrates different lines can delineate the genetic background operating behind these traits. By comparing different lines, the GRN can not only enlist the candidate genes but can also prioritize certain genes like hub genes explaining the pathways. Such candidate gene discovery can be used in future marker discovery program also.

Significantly high correlation and similar expression pattern of the transcripts based on RNA-seq and qRT-PCR confirmed the high quality, reliability and authenticity of the RNA-seq data conducted for comparative transcriptase analysis. Understanding the molecular network responsible for the extended shelf life of the genotype, DC-48 through RNA-seq provided an insight into this trait which can be applied in cucumber improvement programme in breeding genotypes with higher keeping quality. The analysis of the differentially expressed genes revealed complex molecular network responsible for exceptional shelf life of the DC-48 through gene regulatory networks analysis. GRN among the genotypes at different developmental stage indicated the role of the important metabolism and hub genes associated with wide variety of biological process besides chlorophyll, ethylene and cell wall degradation metabolism in extended shelf life of cucumber fruits.

Variants identified in the study can be used as a potential source of putative functional domain markers as they are derived from genic region. Variants in the forms of SSRs, InDels and SNPs can be used as a polymorphic marker by mapping them over reference sequence of cucumber to design genotyping assay. The large number of the identified variants distributed throughout the cucumber genome will be instrumental in future association studies leading to fine mapping of the keeping quality (shelf life). Identification of DNA based putative markers like SSRs and variants like SNPs and InDels are extremely useful in development of linkage map and mapping of the economically important traits in cucumber. The identified SSRs, SNPs and InDels can be used for further enriching the linkage map in cucumber, identification of DNA based markers for economically important traits and fine mapping in cucumber. Besides, these findings will be useful for future research in association studies and marker assisted breeding which is long term requirement in varietal improvement. The web-genomic database, CsExSLDb would be available freely to the researchers worldwide for academic purpose. This genomic database will provide a platform to undertake genetic studies, genomic analysis, marker development and linkage map construction in cucumber improvement programme.

5. Conclusion

In summary, comparative transcriptome analysis for extended shelf

life in two contrasting genotypes of cucumber revealed the key physiological, molecular mechanisms and pathways associated with exceptionally high shelf life in the developed natural variant, DC-48. qRT-PCR with the randomly selected primers established authenticity of the RNA-seq data. At the different developmental stages, 1364 DEGs were identified and most them were associated with purine, phenylalanine, starch and sucrose and methane metabolism and glycophospholipid biosynthesis. Besides, cell wall degradation and chlorophyll metabolism played important role. Polygalacturunase (PG), Expansin (EXP) and xyloglucan endotransglucosylase associated with cell wall degradation played key role in maintaining fruit firmness even after several days of harvest. Retention of fresh green colour of the fruit epidermis was because of low expression level of the chlorophyll catalytic enzymes (CCEs). Mining of large number of SSRs, SNPs and InDels will be instrumental in future cucumber improvement programme. The findings will facilitate development of cucumber genotypes with better shelf life, reduce the post-harvest loss and would be instrumental to reduce hunger at large as per United Nations Sustainable Development Goals. The developed web data base (CsExSLDb) freely available will provide a readymade platform in understanding and utilizing genomic resources developed for shelf life.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2022.110273.

Authors contributions

Conceived theme of the study and designed experiment: S.S. Dey; **Data curation:** S. S. Dey, Mir Asif Iquebal, Sarika Jaiswal; **Computational analysis and development of web-resources:** Parva Kumar Sharma, Mir Asif Iquebal, Sarika Jaiswal, S. S. Dey; **Investigation:** Pradeep Kumara N., S. S. Dey, Khushboo Kumari; **Resources:** S.S. Dey, T.K. Behera, A. D. Munshi; **Supervision:** S. S. Dey, Reeta Bhatia, A. D. Munshi, R. C. Bhattacharya; **Visualization:** S.S. Dey, T.K. Behera, A. D. Munshi, R. C. Bhattacharya; **Writing original draft:** Pradeep Kumara N., S. S. Dey, Parva Kumar Sharma, Mir Asif Iquebal, A. D. Munshi, Sarika Jaiswal and Dinesh Kumar; **Review and editing:** S.S. Dey, Reeta Bhatia, T. K. Behera, Ajay Arora, Dinesh Kumar and Anil Rai; All authors read and approved the final manuscript.

Data availability

The datasets supporting the conclusions of this article are available in the NCBI's Sequence Read Archive with accessions (BioProject ID: PRJNA702645, BioSample accession: SAMN17976848 to SAMN17976859).

Declaration of Competing Interest

The authors declare no competing interests.

Acknowledgements

Authors are thankful to the NAHEP/CAAST (2019-20) research grant by Indian Council of Agricultural Research (ICAR) for financial support. Authors are also thankful to Indian Council of Agricultural Research, Ministry of Agriculture and Farmers' Welfare, Govt. of India for providing financial assistance in the form of CABin grant (F. no. Agril. Edn.4–1/2013-A&P) as well as Advanced Super Computing Hub for Omics Knowledge in Agriculture (ASHOKA) facility at ICAR-IASRI, New Delhi, India.).

References

[1] W.R. Friedman, B.S. Halpern, E. McLeod, M.W. Beck, C.M. Duarte, C.V. Kappel, A. Levine, R.D. Sluka, S. Adler, C.C. O'Hara, E.J. Sterling, Research priorities for achieving healthy marine ecosystems and human communities in a changing climate, Front. Mar. Sci. 7 (2020) 5.

[2] R. Porat, A. Lichter, L.A. Terry, R. Harker, J. Buzby, Postharvest losses of fruit and vegetables during retail and in consumers' homes: quantifications, causes, and means ofprevention, Postharvest Biol. Technol. 139 (2018) 135–149.

[3] Food and Agriculture Organization of the United Nations, FAOSTAT Database, FAO, Rome, Italy, 2019. Retrieved from, http://www.fao.org/faostat/en/#data/QC.

[4] M. Manjunatha, R.K. Anurag, Effect of modified atmosphere packaging and storage conditions on quality characteristics of cucumber, J. Food Sci. Technol. 51 (2014) 3470–3475.

[5] P. Sebastian, H. Schaefer, I.R.H. Telford, S.S. Renner, Cucumber (*Cucumis sativus* L.) and melon (*C. melo*) have numerous wild relatives in Asia and Australia, and the sister species of melon is from Australia, Proc. Natl. Acad. Sci. 107 (2010) 14269–14273.

[6] H.S. Paris, M.C. Daunay, J. Janick, Occidental diffusion of cucumber (*Cucumis sativus* L.) 500-1300 CE: two routes to Europe, Ann. Bot. 109 (2011) 117–126.

[7] J. Janick, H.S. Paris, D.C. Parrish, The cucurbits of Mediterranean antiquity: identification of taxa from ancient images and descriptions, Ann. Bot. 100 (2007) 1441–1457.

[8] M. Vats, S. Bhardwaj, A. Chhabra, Green synthesis of copper oxide nanoparticles using *Cucumis sativus* L. (Cucumber) extracts and their bio-physical and biochemical characterization for cosmetic and dermatologic applications, Endocr metab immune disord drug targets 21 (2021) 726–733, https://doi.org/10.2174/1871530320666200705212107.

[9] M.M. Fiume, W.F. Bergfeld, D.V. Belsito, R.A. Hill, C.D. Klaassen, D.C. Liebler, J. G. Marks, R.C. Shank, T.J. Slaga, P.W. Snyder, F.A. Andersen, Safety assessment of *Cucumis sativus* L. (cucumber)-derived ingredients as used in cosmetics, Int. J. Toxicol. 33 (2014) 47–64, https://doi.org/10.1177/1091581814526892.

[10] A.H. Atta, S.A. Saad, S.A. Atta, S.M. Mouneir, S.M. Nasr, H.M. Desouky, H. M. Shaker, *Cucumis sativus* L. and *Cucurbita maxima* L. extract attenuate diabetes-induced hepatic and pancreatic injury in a rat model, J. Physiol. Pharmacol. 71 (2020).

[11] B. Chandrasekhar, B. Mukherjee, S.K. Mukherjee, Blood sugar lowering potentiality of selected Cucurbitaceae plants of Indian origin, Indian J. Med. Res. 90 (1989) 300–305.

[12] Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics, Nat. Rev. Genet. 10 (2009) 57–63.

[13] A. Mortazavi, B.A. Williams, K. McCue, L. Schaeffer, B. Wold, Mapping and quantifying mammalian transcriptomes by RNA-Seq, Nat. Methods 5 (2008) 621–628.

[14] J.C. Vera, C.W. Wheat, H.W. Fescemyer, M.J. Frilander, D.L. Crawford, I. Hanski, J.H. Marden, Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing, Mol. Ecol. 17 (2008) 1636–1647.

[15] B.T. Wilhelm, J.R. Landry, RNA-Seq - quantitative measurement of expression through massively parallel RNA-sequencing, Methods. 48 (2009) 249–257.

[16] F. Ozsolak, P.M. Milos, RNA sequencing: advances, challenges and opportunities, Nat. Rev. Genet. 12 (2011) 87–98.

[17] S. Guo, Y. Zheng, J.G. Joung, S. Liu, Z. Zhang, O.R. Crasta, B.W. Sobral, Y. Xu, S. Huang, Z. Fei, Transcriptome sequencing and comparative analysis of cucumber flowers with different sex types, BMC Genomics 11 (2010) 1–13, https://doi.org/10.1186/1471-2164-11-384.

[18] J. Song, Y. Zhang, S. Song, W. Su, R. Chen, G. Sun, Y. Hao, H. Liu, Comparative RNA-Seq analysis on the regulation of cucumber sex differentiation under different ratios of blue and red light, Bot. Stud. 59 (2018) 1–18, https://doi.org/10.1186/s40529-018-0237-7.

[19] M. Pawełkowicz, L. Pryszcz, A. Skarzyńska, R.K. Wóycicki, K. Posyniak, J. Rymuszka, Z. Przybecki, W. Pląder, Comparative transcriptome analysis reveals new molecular pathways for cucumber genes related to sex determination, Plant Reprod. 32 (2019) 193–216.

[20] L. Zhang, X.K. Feng, Y.K. Ng, S.C. Li, Reconstructing directed gene regulatory network by only gene expression data, BMC Genomics 17 (2016) 341–353.

[21] X. Wang, C. Cheng, K. Zhang, Z. Tian, J. Xu, S. Yang, Q. Lou, J. Li, J.F. Chen, Comparative transcriptomics reveals suppressed expression of genes related to auxin and the cell cycle contributes to the resistance of cucumber against *Meloidogyne incognita*, BMC Genomics 19 (2018) 1–14, https://doi.org/10.1186/s12864-018-4979-0.

[22] L. Wang, C. Cao, S. Zheng, H. Zhang, P. Liu, Q. Ge, Z. Ren, Transcriptomic analysis of short-fruit 1 (*sf1*) reveals new insights into the variation of fruit-related traits in *Cucumis sativus* L, Sci. Rep. 7 (2017) 1–13, https://doi.org/10.1038/s41598-017-02932-5.

[23] E.M. Abdel-Salam, F. Mohammad, A.A. Alatar, S. Quaiser, A.A. Hend, Comparative analysis between wild and cultivated cucumbers reveals transcriptional changes during domestication process plants, Plants. 9 (2020) 63.

[24] C. Bonghi, M. Begheldo, F. Ziliotto, A. Rasori, L. Trainotti, A. Tadiello, R. Tosetti, P. Tonutti, Transcriptome analyses and postharvest physiology of peaches and nectarines, Acta Hortic. 877 (2010) 69–73.

[25] D. Sanhueza, P. Vizoso, I. Balic, R. Campos-Vargas, C. Meneses, Transcriptomic analysis of fruit stored under cold conditions using controlled atmosphere in *Prunus persica* cv. "red pearl", front, Plant Sci. 26 (2015) 788.

[26] Z. Yun, H. Qu, H. Wang, F. Zhu, Z. Zhang, X. Duan, B. Yang, Y. Cheng, Y. Jiang, Comparative transcriptome and metabolome provides new insights into the regulatory mechanisms of accelerated senescence in litchi fruit after cold storage, Sci. Rep. 6 (2016) 1–16.

[27] M. Saladié, J. Cañizares, M.A. Phillips, M. Rodriguez-Concepcion, C. Larrigaudière, Y. Gibon, M. Stitt, J.E. Lunn, J. Garcia-Mas, Comparative transcriptional profiling analysis of developing melon (*Cucumis melo* L.) fruit from climacteric and non-climacteric varieties, BMC Genomics 16 (2015) 1–20.

[28] J.B. Biale, R.E. Young, Respiration and ripening in fruit retrospect and prospect, in: J. Friend, M.J.C. Rhodes (Eds.), Recent Advances in the Biochemistry of Fruit and Vegetables, Academic, 1981, pp. 1–39.

[29] 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.

[30] M.E. Saltveit, R.F. McFeeters, Polygalacturonase activity and ethylene synthesis during cucumber fruit development and maturation, Plant Physiol. 66 (1980) 1019–1023.

[31] L.C.O. Lima, B.M. Hurr, D.J. Huber, Deterioration of beit alpha and slicing cucumbers (*Cucumis sativus* L.) during storage in ethylene or air: responses to suppression of ethylene perception and parallels to natural senescence, Postharvest Biol. Technol. 37 (2005) 265–276.

[32] T. Nilsson, Effects of ethylene and 1-MCP on ripening and senescence of European seedless cucumbers, Postharvest Biol. Technol. 36 (2005) 113–125.

[33] D.A. Brummell, Cell wall disassembly in ripening fruit, Funct. Plant Biol. 33 (2006) 103–119.

[34] A.J. Matas, N.E. Gapper, M.Y. Chung, J.J. Giovannoni, J.K. Rose, Biology and genetic engineering of fruit maturation for enhanced quality and shelf life, Curr. Opin. Biotechnol. 20 (2009) 197–203.

[35] I. Lara, B. Belge, L.F. Goulao, The fruit cuticle as a modulator of postharvest quality, Postharvest Biol. Technol. 87 (2014) 103–112.

[36] J.A. Mercado, F. Pliego-Alfaro, M.A. Quesada, Fruit shelf life and potential for its genetic improvement, Breeding Fruit Quality. (2011) 81–104.

[37] C. Paniagua, R. Blanco-Portales, M. Barceló-Muñoz, J.A. García-Gago, K. W. Waldron, M.A. Quesada, J. Muñoz-Blanco, J.A. Mercado, Antisense down regulation of the strawberry β galactosidase gene FaβGal4 increases cell wall galactose levels and reduces fruit softening, J. Exp. Bot. 67 (2016) 619–631.

[38] Y. Han, S.S. Korban, Transgenic approaches to improve fruit quality, in: Breeding for Fruit Quality, John Wiley and Sons, Inc, 2001, pp. 151–172.

[39] H.J. Klee, J.J. Giovannoni, Genetics and control of tomato fruit ripening and quality attributes, Annu. Rev. Genet. 45 (2011) 41–59.

[40] Y. Wang, J. Tan, Z. Wu, K. VandenLangenberg, T.C. Wehner, C. Wen, X. Zheng, K. Owens, A. Thornton, H.H. Bang, E. Hoeft, P.A.G. Kraan, J. Suelmann, J. Pan, Y. Weng, Staygreen, stay healthy: a loss-of-susceptibility mutation in the staygreen gene provides durable, broad-spectrum disease resistances for over 50 years of US cucumber production, New Phytol. 221 (2019) 415–430.

[41] H. Thomas, H. Ougham, The stay-green trait, J. Exp. Bot. 65 (2014) 3889–3900.

[42] B. Kuai, J. Chen, S. Hörtensteiner, The biochemistry and molecular biology of chlorophyll breakdown, J. Exp. Bot. 69 (2018) 751–767.

[43] R.W. Robinson, D.S. Decker-Walters, Cucurbits, CAB International, New York, 1999.

[44] S. Ranganna, Manual of Analysis of Fruits and Vegetable Products, Tata McGraw Hill Book Co., New Delhi, 1979.

[45] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics. 30 (2014) 2114–2120.

[46] S. Andrews, FastQC: a quality control tool for high throughput sequence data, Available online at, http://www.bioinformatics.babraham.ac.uk/projects/fastqc, 2010.

[47] B.J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P.D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M.D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C.N. Dewey, R. Henschel, A. Regev, De novo transcript sequence reconstruction from RNA-seq using the trinity platform for reference generation and analysis, Nat. Protoc. 8 (2013) 1494–1512, https://doi.org/10.1038/nprot.2013.084.

[48] B. Langmead, C. Trapnell, M. Pop, S.L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome, Genome Biol. 10 (2009) 1–10.

[49] B. Li, C.N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome, BMC Bioinformatics. 12 (2011) 1–16, https://doi.org/10.1186/1471-2105-12-323.

[50] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics (Oxford, England). 26 (2010) 139–140, https://doi.org/10.1093/bioinformatics/btp616.

[51] 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.

[52] A. Conesa, S. Götz, Blast2GO: a comprehensive suite for functional analysis in plant genomics, International Journal of Plant Genomics. 2008 (2008), 619832, https://doi.org/10.1155/2008/619832.

[53] J. Jin, F. Tian, D.C. Yang, Y.Q. Meng, L. Kong, J. Luo, G. Gao, PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants, Nucleic Acids Res. 982 (2016).

[54] T. Thiel, W. Michalek, R. Varshney, A. Graner, Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.), Theor. Appl. Genet. 106 (2003) 411–422.

[55] A. Untergasser, I. Cutcutache, T. Koressaar, J. Ye, B.C. Faircloth, M. Remm, S. G. Rozen, Primer3 new capabilities and interfaces, Nucleic Acids Res. 40 (2012) 115.

[56] B. Langmead, S. Salzberg, Fast gapped-read alignment with bowtie 2, Nat. Methods 9 (2012) 357–359, https://doi.org/10.1038/nmeth.1923.

[57] B. Li, C.N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome, BMC Bioinformatics. 12 (2011) 1–16, https://doi.org/10.1186/1471-2105-12-323.

[58] M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, M.A. Marra, Circos: an information aesthetic for comparative genomics, Genome Res. 19 (2009) 1639–1645.

[59] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (2003) 2498–2504.

[60] L. Zhang, Z. Zhang, T. Zheng, W. Wei, Y. Zhu, Y. Gao, X. Yang, S. Lin, Characterization of carotenoid accumulation and carotenogenic gene expression during fruit development in yellow and white loquat fruit, horticultural, Plant J. 2 (2016) 9–15.

[61] D.C. Kim, C. Liu, X. Wu, B. Zhang, J. Gao, Inference of gene regulatory networks by integrating gene expressions and genetic perturbations, in: IEEE International Conference on Bioinformatics and Biomedicine, 2013, pp. 182–187.

[62] A.N. Stepanova, J.M. Alonso, Ethylene signaling and response pathway a unique signaling cascade with a multitude of inputs and outputs, Physiol. Plant. 123 (2005) 195–206.

[63] N. Etheridge, B.P. Hall, G.E. Schaller, Progress report: ethylene signaling and responses, Planta. 223 (2006) 387–391.

[64] C.S. Barry, J.J. Giovannoni, Ethylene and fruit ripening, J. Plant Growth Regul. 26 (2007) 143–159.

[65] M. Sakharkar, F. Passetti, J.E. de Souza, M. Long, S.J. de Souza, ExInt: an exon intron database, Nucleic Acids Res. 30 (2002) 191–194.

[66] Y. Gao, N. Zhu, X. Zhu, M. Wu, C.Z. Jiang, D. Grierson, G. Qu, Diversity and Redundancy of the Ripening Regulatory Networks Revealed by the fruitENCODE and the new CRISPR/Cas9 CNR and NOR Mutants, Horticulture Research 6, 2019, pp. 1–10.

[67] Y.J. Ye, Y.Y. Xiao, Y.C. Han, W. Shan, Z.Q. Fan, Q.G. Xu, J.F. Kuang, W.J. Lu, P. Lakshmanan, J.Y. Chen, Banana fruit VQ motif-containing protein5 represses cold-responsive transcription factor MaWRKY26 involved in the regulation of JA biosynthetic genes, Sci. Rep. 6 (2016) 1–14.

[68] T. Taybi, S. Patil, R. Chollet, J.C. Cushman, A minimal serine/threonine protein kinase circadianly regulates phosphoenolpyruvate carboxylase activity in crassulacean acid metabolism-induced leaves of the common ice plant, Plant Physiol. 123 (2000) 1471–1482, https://doi.org/10.1104/pp.123.4.1471.

[69] A.K. Mattoo, A.K. Handa, Higher polyamines restore and enhance metabolic memory in ripening fruit, Plant Sci. 174 (2008) 386–393, https://doi.org/10.1016/j.plantsci.2008.01.011.

[70] L. Wang, X.L. Zhang, L. Wang, Y. Tian, N. Jia, S. Chen, N.B. Shi, X. Huang, C. Zhou, Y. Yu, Z.Q. Zhang, X.Q. Pang, Regulation of ethylene-responsive SlWRKYs involved in color change during tomato fruit ripening, Sci. Rep. 7 (2017) 16674, https://doi.org/10.1038/s41598-017-16851-y.

[71] Y. Zhu, D. Li, T. Belwal, L. Li, H. Chen, T. Xu, Z. Luo, Effect of Nano-SiOx/chitosan complex coating on the physicochemical characteristics and preservation performance of green tomato, Molecules. 24 (2019) 4552, https://doi.org/10.3390/molecules24244552.

[72] Y. Cui, Z. Wang, S. Chen, A. Vainstein, H. Ma, Proteome and transcriptome analyses reveal key molecular differences between quality parameters of commercial-ripe and tree-ripe fig (*Ficus carica* L.), BMC Plant Biol. 19 (2019) 146, https://doi.org/10.1186/s12870-019-1742-x.

[73] M. Zarid, V. García-Carpintero, C. Esteras, J. Esteva, M.C. Bueso, J. Cañizares, M. B. Picó, A.J. Monforte, J.P. Fernández-Trujillo, Transcriptomic analysis of a near-isogenic line of melon with high fruit flesh firmness during ripening, J. Sci. Food Agric. 101 (2020) 754–777, https://doi.org/10.1002/jsfa.10688.

[74] S. Amsbury, P. Kirk, Y. Benitez-Alfonso, Emerging models on the regulation of intercellular transport by plasmodesmata-associated callose, J. Exp. Bot. 69 (2017) 105–115, https://doi.org/10.1093/jxb/erx337 (PMID: 29040641).

[75] N.M. Sadali, R.G. Sowden, Q. Ling, R.P. Jarvis, Differentiation of chromoplasts and other plastids in plants, Plant Cell Rep. 38 (2019) 803–818, https://doi.org/10.1007/s00299-019-02420-2.

[76] M. Liu, Z. Zhang, Z. Xu, L. Wang, C. Chen, Z. Ren, Overexpression of SlMYB75 enhances resistance to Botrytis cinerea and prolongs fruit storage life in tomato, Plant Cell Rep. 40 (2021) 43–58, https://doi.org/10.1007/s00299-020-02609-w.

[77] G. Winter, C.D. Todd, M. Trovato, G. Forlani, D. Funck, Physiological implications of arginine metabolism in plants, Front. Plant Sci. 6 (2015) 534, https://doi.org/10.3389/fpls.2015.00534.

[78] B. Llorente, L. D'Andrea, M.A. Ruiz-Sola, E. Botterweg, P. Pulido, J. Andilla, P. Loza-Alvarez, M. Rodriguez-Concepcion, Tomato fruit carotenoid biosynthesis is adjusted to actual ripening progression by a light-dependent mechanism, Plant J. 85 (2016) 107–119, https://doi.org/10.1111/tpj.13094.

[79] G. Khaksar, W. Sangchay, P. Pinsorn, L. Sangpong, S. Sirikantaramas, Genome-wide analysis of the Dof gene family in durian reveals fruit ripening-associated and cultivar-dependent Dof transcription factors, Sci. Rep. 9 (2019) 1–13, https://doi.org/10.1038/s41598-019-48601-7.

[80] N. Xu, R. Wang, L. Zhao, C. Zhang, Z. Li, Z. Lei, F. Liu, P. Guan, Z. Chu, N. M. Crawford, Y. Wang, The Arabidopsis NRG2 protein mediates nitrate Signalling and interacts with and regulates key nitrate regulators, Plant Cell 28 (2016) 485–504, https://doi.org/10.1105/tpc.15.00567.

[81] M.H. Asif, D. Lakhwani, S. Pathak, P. Gupta, S.K. Bag, P. Nath, P.K. Trivedi, Transcriptome analysis of ripe and unripe fruit tissue of banana identifies major metabolic networks involved in fruit ripening process, BMC Plant Biol. 14 (2014) (2014) 316, https://doi.org/10.1186/s12870-014-0316-1.

[82] A. Leszczuk, P. Kalaitzis, K.N. Blazakis, A. Zdunek, The role of arabinogalactan proteins (AGPs) in fruit ripening– a review, Hortic Res. 7 (2020) 1–12, https://doi.org/10.1038/s41438-020-00397-8.

[83] W. Jamil, W. Wu, H. Gong, J.W. Huang, M. Ahmad, Q.G. Zhu, R. Jin, X.F. Liu, X. R. Yin, C2H2-type zinc finger proteins (DkZF1/2) synergistically control persimmon fruit Deastringency, Int. J. Mol. Sci. 20 (2019) 5611, https://doi.org/10.3390/ijms20225611.
[84] D.R. Nelson, R. Ming, M. Alam, M.A. Schuler, Comparison of cytochrome P450 genes from six plant genomes, Trop. Plant Biol. 1 (2008) 216–235, https://doi.org/10.1007/s12042-008-9022-1.
[85] C. Chen, F. Xie, Q. Hua, N. Zur, L. Zhang, Z. Zhang, R. Zhang, J. Zhao, G.Y. Qin, Integrated sRNAome and RNA-Seq analysis reveals miRNA effects on betalain biosynthesis in pitaya, BMC Plant Biol. 22 (20) (2020) 437, https://doi.org/10.1186/s12870-020-02622-x.
[86] Z. Yang, T. Zhang, T. Lang, G. Li, G. Chen, E. Nevo, Transcriptome comparative profiling of barley eibi1 mutant reveals pleiotropic effects of HvABCG31 gene on cuticle biogenesis and stress responsive pathways, Int. J. Mol. Sci. 14 (2013) 20478–20491, https://doi.org/10.3390/ijms141020478.
[87] L.F. Goulao, C.M. Oliveira, Cell wall modifications during fruit ripening: when a fruit is not the fruit. Trend, Food Sci. Technol. 19 (2008) 4–25.
[88] M. Saladié, A.J. Matas, T. Isaacson, M.A. Jenks, S.M. Goodwin, K.J. Niklas, R. Xiaolin, J.M. Labavitch, K.A. Shackel, A.R. Fernie, A. Lytovchenko, A re-evaluation of the key factors that influence tomato fruit softening and integrity, Plant Physiol. 144 (2007) 1012–1028.
[89] A.R. Vicente, M. Saladie, J.K. Rose, J.M. Labavitch, The linkage between cell wall metabolism and fruit softening: looking to the future, J. Sci. Food Agric. 87 (2007) 1435–1448.
[90] D.A. Brummell, J.M. Labavitch, Effect of antisense suppression of endopolygalacturonase activity on polyuronide molecular weight in ripening tomato fruit and in fruit homogenates, Plant Physiol. 115 (1997) 717–725.
[91] G. Rao, I. Paran, Polygalacturonase: a candidate gene for the soft flesh and deciduous fruit mutation in Capsicum, Plant Mol. Biol. 51 (2003) 135–141.
[92] M. Adami, P.D. Franceschi, F. Brandi, A. Liverani, D. Giovannini, C. Rosati, Identifying a carotenoid cleavage dioxygenase (ccd4) gene controlling yellow/white fruit flesh color of peach, Plant Mol. Biol. Report. 31 (2013) 1166–1175.
[93] L. Li, H. Yuan, Chromoplast biogenesis and carotenoid accumulation, Arch. Biochem. Biophys. 539 (2013) 102–109.
[94] W.R. Henderson, G.H. Scott, T.C. Wehner, Interaction of flesh color genes in watermelon, J Hered. 89 (1998) 50–53.
[95] Y. Burger, H. Paris, R. Cohen, N. Katzir, Y. Tadmor, E. Lewinsohn, A.A. Schaffer, Genetic diversity of *Cucumis melo* L, Hortic Rev. Am Soc Hortic Sci. 36 (2009) 165–198.
[96] Y. Shimoda, H. Ito, A. Tanaka, Arabidopsis *STAY-GREEN*, Mendel's green cotyledon gene, encodes magnesium-dechelatase, Plant Cell 28 (2016) 2147–2160.
[97] C.D. Gussman, J.C. Goffredz, T.J. Gianfagna, Ethylene production and fruit-softening rates in several apple fruit ripening variants, Hortic. Sci. 28 (1993) 135–137.
[98] X.Y. Zheng, D.W. Wolff, Ethylene production, shelf life and evidence of RFLP polymorphisms linked to ethylene genes in melon (*Cucumis melo* L.), Theor. Appl. Genet. 101 (2000) 613–624.
[99] C. Stasolla, R. Katahira, T.A. Thorpe, H. Ashihara, Purine and pyrimidine nucleotide metabolism in higher plants, J. Plant Physiol. 160 (2003) 1271–1295.
[100] C.M. Fraser, C. Chapple, The phenylpropanoid pathway in Arabidopsis, in: Arabidopsis Book 9, 2011, e0152, https://doi.org/10.1199/tab. 0152.
[101] M.B. Pascual, J. El-Azaz, F.N. de la Torre, R.A. Cañas, C. Avila, F.M. Cánovas, Biosynthesis and metabolic fate of phenylalanine in conifers, Front. Plant Sci. 7 (2016) 1030, https://doi.org/10.3389/fpls.2016.01030.
[102] S.B. Lee, J.E. Kim, H.T. Kim, G. Lee, B. Kim, J.M. Lee, G.Theor. Appl. Genet. 133 (2020) 1897–1910.
[103] J. Xie, S. Yao, J. Ming, L. Deng, K. Zeng, Variations in chlorophyll and carotenoid contents and expression of genes involved in pigment metabolism response to oleocellosis in citrus fruits, Food Chem. 272 (2019) 49–57.
[104] L. Meng, Z. Fan, Q. Zhang, C. Wang, Y. Gao, Y. Deng, B. Zhu, H. Zhu, J. Chen, W. Shan, X. Yin, BEL1-LIKE HOMEODOMAIN 11 regulates chloroplast development and chlorophyll synthesis in tomato fruit, Plant J. 94 (2018) 1126–1140.
[105] B. Lai, B. Hu, Y. Qin, J. Zhao, H. Wang, G. Hu, Transcriptomic analysis of Litchi chinensis pericarp during maturation with a focus on chlorophyll degradation and flavonoid biosynthesis, BMC Genomics 16 (2015) 225.
[106] J.X. Shi, A. Adato, N. Alkan, Y. He, J. Lashbrooke, A.J. Matas, S. Meir, S. Malitsky, T. Isaacson, D. Prusky, D. Leshkowitz, The tomato SlSHINE3 transcription factor regulates fruit cuticle formation and epidermal patterning, New Phytol. 197 (2013) 468–480.
[107] P.F. Cavagnaro, D.A. Senalik, L. Yang, P.W. Simon, T.T. Harkins, C.D. Kodira, S. Huang, Y. Weng, Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.), BMC Genomics 11 (2010) 1–18.
[108] W.Y. Zhu, L. Huang, L. Chen, J.T. Yang, J.N. Wu, M.L. Qu, D.Q. Yao, C.L. Guo, H. L. Lian, H.L. He, J.S. Pan, A high-density genetic linkage map for cucumber (*Cucumis sativus* L.): based on specific length amplified fragment (SLAF) sequencing and QTL analysis of fruit traits in cucumber, Front. Plant Sci. 7 (2016) 437.