Indian J. Genet. Plant Breed., (2024); 84(1): 114-120
https://doi.org/10.31742/ISGPB.84.1.11
ISSN: 0975-6906
www.isgpb.org

RESEARCH ARTICLE

Transcriptomic analysis of almond (*Prunus dulcis*) cultivars with differential flowering time

Javed Iqbal Mir*, Susheel Kumar Raina[1*], Om Chand Sharma, Wasim Hassan Raja, Sajad-Un-Nabi, Salwee Yasmeen and Sheikh M. Sultan[1]

Abstract

Almond is one of the oldest and major tree nut crops consumed for its nutritious kernels. Since transcriptome analysis is crucial to identifying genes controlling important agronomic, an attempt was made for comparative transcriptomic analysis to identify the genetic differences in almond cultivars Waris and Ferralise, varying in blooming dates. Sequencing data from Illumina TruSeq libraries was processed using the FASTQC-Trimmomatic-HISAT2-Stringtie-Stringtie merge-DESeq2 pipeline. Overall alignment rates for different libraries ranged from 82.04-91.33%. DESeq2 analysis identified 3707 differentially expressed genes, out of which 1558 genes were up-regulated while the 2149 genes were down-regulated in Waris compared to Ferralise. Differential regulation of genes associated with the auxin signaling pathway, ubiquitin ligases, serine/threonine-protein phosphatases, proteasomes, translation initiation factors, phosphatidylinositol signaling and pentose phosphate pathway provides new insights into the molecular regulation of blooming in almonds. These findings will pave the way for molecularly differentiating early and late flowering almond cultivars and have implications in breeding late-blooming almond cultivars.

Keywords: Almond, transcriptome, flowering, sequences, genes.

Introduction

Almond (*Prunus dulcis*) is an important nut crop grown for its nutritious kernels. There are reports of consumption of almonds for medical purposes by ancient Greeks, Persians, Chinese and Indian habitants (Albala 2009). Almonds are considered the most nutritious of all nut crops, possessing phytonutrients of considerable health value like protein, fiber, vitamin E, Omega-3 fatty acids and minerals like potassium, magnesium, phosphorus, calcium and iron (Ahmed and Verma 2009). Daily 30–50 g intake is recommended as part of a healthy diet (Barreca et al. 2020). It is a very high-value crop and its nuts have a very good shelf life. Although a native of the Mediterranian region, its cultivation has spread over a range of countries, including the USA, Spain, Seria, Italy, Iran, Morocco, Algeria, Tunisia, Greece and Turkey (Ahmed and Verma 2009). The USA is the largest almond producer, followed by Spain and Australia (Barreca et al. 2020). In India, the crop is mainly cultivated in Kashmir, where it was introduced during 16th century by Persian settlers (Ahmed and Verma 2009). However, there has been a sharp decline in almond production in the Valley in last few years due to inclement weather from March and April that drastically affected the production of almonds in Kashmir as the majority of growers use an indigenous variety that blooms early, leading to drastic fall in production. As per official data available till 2011, land used for almond production has reduced to 7,107 hectares from 16,418 hectares and production has come down to 6360 metric tons from 16, 537 metric tons. Hence,

ICAR-Central Institute of Temperate Horticulture, KD Farm, Old Air Field, Rangreth, Srinagar 191 132, Jammu and Kashmir, India.

[1]ICAR-National Bureau of Plant Genetic Resources, Regional Station Srinagar, KD Farm, Old Air Field, Rangreth, Srinagar 190 001, Jammu and Kashmir, India.

***Corresponding Author:** Javed Iqbal Mir, ICAR-Central Institute of Temperate Horticulture, KD Farm, Old Air Field, Rangreth, Srinagar 191 132, Jammu and Kashmir, India, E-Mail: javidiqbal1234@gmail.com; Susheel.Raina@icar.gov.in

How to cite this article: Mir J.I., Raina S.K., Sharma O.C., Raja W.H., Un-Nabi S., Yasmeen S. and Sultan S.M. 2024. Transcriptomic analysis of almond cultivars with differential flowering time. Indian J. Genet. Plant Breed., **84**(1): 114-120.

Source of support: ICAR, Govt of India.

Conflict of interest: None.

Received: Jan. 2023 **Revised:** Dec. 2023 **Accepted:** Jan. 2024

the growers have been shifting towards cultivating apples (https://indianexpress.com/article/business/commodities/almond-production-on-decline-as-growers-blame-jammu-kashmir-valley-weather-4744288/).

In almonds, flowering time, self-compatibility and kernel sweetness are important traits while developing new breeding programs (Prudencio et al. 2021). This nut crop is among the earliest to bloom temperate fruit crops. Blooming is often associated with early spring frosts (Kumawat et al. 2023). These frosts cause heavy damage to the flowers and discourage activity of pollinators. This leads to a heavy reduction in fruit set, thereby posing serious economic loss to the growers. Late-blooming almond cultivars escape early spring frosts, and the identification/development of late-blooming almond cultivars is an important trait that breeders focus on.

With the advent of recent genomics technologies and sequencing of fruit crop genomes, the search for genes regulating traits of economic importance are being explored as potential candidates for recent genome editing technologies, including CRISPER-CAS9 and others. Although several genes have been identified in almond genome to be associated with late blooming, the clear-cut picture of gene regulation involved in late blooming is still missing in this nut crop. A single gene, *Late blooming* (*Lb*), determining very late flowering, has been identified in a mutant line, Tardy Nonpareil. However, the regulation of gene expressions has not been studied in detail, although several transcription factors that play a role in flowering time have been described (Sánchez-Pérez R et al. 2014). The advent of technologies such as RNA sequencing utilizes high-throughput next-generation sequencing to understand gene regulation. RNA sequencing has potential applications in deciphering the dynamics of functional genes, genome-wide differences, and identification of genes associated with fruit quality (Zhang and Hao 2020). Since transcriptomic analysis is emerging as a powerful tool in understanding the gene regulation underlying specific developmental stages, the present investigation was undertaken to study the gene regulation underlying late blooming in almonds.

Materials and methods

The experiment was conducted on the almond plants growing in the orchard of ICAR-Central Institute of Temperate Horticulture, Srinagar (34.05' N, 74°50' E; elevation - 1640 m a.s.l.). The area is characterized by Karewa soils having pH of 6.86 and EC of 0.38 dSm^{-1}. The climate of this region is temperate, with cold conditions from November to February. The almond cultivars used in the study were Waris and Ferralise, both budded on bitter almond rootstock. Waris is a local almond cultivar known for its soft-shelled and good-quality kernels and is categorized as early bloomer. Ferralise is an exotic cultivar introduced from Australia and is categorized as late bloomer, flowering 7 to 10 days later than Waris. However, nut quality of Ferralise is considered not to be of good standard.

RNA samples were extracted from the leaves as well as flowers of almond cultivars at the pop-corn stage of flowers. Total RNA was isolated from samples using the RNeasy Plant Mini Kit. The quality of the isolated RNA was checked on 1% Formaldehyde Denaturing Agarose gel and quantified using Qubit® 2.0 Fluorometer. Total RNA samples were prepared by pooling equal amounts of RNA from the two samples (flowers and leaves) and two biological replicates for each cultivar were outsourced to Xcelris Company for further processing.

Illumina 2 x 150 PE library preparation: The libraries were prepared from samples with input total RNA ~1µg using Illumina TruSeq Stranded mRNA Library Preparation Kit as per the manufacturer's protocol. Briefly, the total RNA was subjected to Oligo dT beads to enrich mRNA fragments, which were then subjected to purification, fragmentation, and priming for cDNA synthesis. The fragmented mRNA was converted into first-strand cDNA, followed by second-strand cDNA synthesis, A-tailing, adapter-index ligation and finally amplified by the recommended number of PCR cycles. Library quality and quantity check was performed using Agilent DNA High Sensitivity Assay Kit. The amplified libraries were analyzed on Bioanalyzer 2100 (Agilent Technologies) using a high-sensitivity (HS) DNA chip as per the manufacturer's instructions.

After obtaining the Qubit concentration for the library and the mean peak size from the Bioanalyzer profile, library was loaded into the Illumina platform for cluster generation and sequencing. Paired-end sequencing allows the template fragments to be sequenced in both the forward and reverse directions. The library molecules bind to complementary adapter oligos on the paired-end flow cell. The adapters were designed to allow selective cleavage of the forward strands after re-synthesis of the reverse strand during sequencing. The copied reverse strand was then used to sequence from the opposite end of the fragment.

Total RNA was subjected to pair-end library preparation with Illumina TruSeq stranded mRNA Library Preparation Kit. The mean size of the libraries was between 378bp and 402bp. The libraries were sequenced using 2x150bp PE chemistry on Illumina platform for generating ~4 GB data per sample.

The FASTQC files of the transcriptome for each biological replicate were processed for RNAseq analysis using GALAXY Europe platform (https://galaxyproject.eu). The pipeline used for analysis consisted of FASTQC-Trimmomatic-HISAT2-Stringtie-Stringtie merge-DESeq2 as per the protocol described in the link (https://www.youtube.com/watch?v=KVh98S89yUU&t=272s).

Briefly, quality of the FASTQ files was checked by FASTQC tool and these files were later subjected to Trimmomatic

Table 1. Description of the raw and trimmed reads generated from each library used in study

	Raw read pairs			Read-pairs after trimming			% trimmed
	Number	Sequence length	%GC	Number	Sequence length	%GC	
Waris R1	12855553	35–151	50	11178042	3–151	50	13.05
Waris R2	15615082	35–151	48	13554786	3–151	48	13.19
Ferralise R1	11416648	35–151	47.5	9858792	3–151	47	13.65
Ferralise R2	15901151	35–151	48	13817457	3–151	48	13.1

using sliding window option and quality threshold of 30. The files were processed for further analysis only after having passed the requisite threshold. The paired files for each sample obtained from Trimmomatic as output files were used for sequence alignment using HISAT2 tool. Employing a reference-based transcriptomic approach, the almond reference genome for alignment was the 'Normalzedfasta' form of GCF_902201215.1_AlmondV2_genomic.fna. The HISAT2 output files were subjected to Stringtie analysis using 'short reads' as input option and selecting all the HISAT2 files without using any reference genome used to guide assembly. The Stringtie output files were analyzed by Stringtie merge tool using GCF_902201215.1_ALMONDv2_genomic.gff as reference annotation. Output of the Stringtie merge was subjected to another round of Stringtie, although using GTF/GFF3 dataset of reference genome to guide the assembly. The count files for two cultivars obtained from second round of Stringtie were processed using DESeq2 tool. Only the genes with a probability (*q*-value) of differential expression over 0.95, equivalent to selecting below 0.05 by FDR, and a fold change (FC) higher than or equal to |2|, were considered as statistically significant.

BlastX analyzed the differentially expressed genes against *Prunus dulcis*_v2.0 protein database on GDR Rosaceae (Jung et al. 2019)/ Ensembl Plants (https://plants.ensembl.org/index.html). However, for differentially expressed genes with MSTRG tags, these genes' coordinates (Chromosome No., start and end) were used to create a BED file. The BED file was uploaded on Galaxy Europe platform and used as input file for tool "BEDtools get fasta". The almond genome file viz. "GCF_902201215.1_Almond v2_genomic.fna" was used as a source for fasta file. The fasta sequences obtained were analyzed by BlastX 2.12.0 against the *Prunus dulcis*_v2.0 protein database on the GDR Rosaceae website. Wherever required, accession to ID conversion was achieved by using PubChem upload tools/ DAVID (Huang et al. 2009; Sherman et al. 2021). Functional enrichment of differentially expressed genes was achieved using gProfiler (Kolberg et al. 2023) with g:SCS threshold method and a threshold value of 0.05. DAVID platform (https://david.ncifcrf.gov) was also used for the functional enrichment of up-regulated and down-regulated genes as well as KEGG pathway gene analysis (Huang et al. 2009; Sherman et al. 2021).

Results

Total RNA from two biological replicates of almond cultivars 'Waris' and 'Ferralise' was isolated and used for tconstructing RNA-Seq libraries. The FASTQC analysis of the raw reads revealed that all of them passed the quality check. The number of reads from each library is presented in Table 1. The sequence length of the reads was 35–151 with a GC content ranging from 47.5 to 50% (Table 1). When subjected to Trimmomatic, the raw reads resulted in reduced sequences for each library (Table 1). The sequence length among the trimmed reads was 3–151 with a GC content ranging from 47 to 50%.

The trimmed reads were aligned with the almond genome GCF_902201215.1_AlmondV2_genomic.fna using HISAT2. The overall alignment rate for different libraries is shown in table (Table 2). The uniquely mapped reads among the libraries ranged from 56.33 to 71.85% while the multi-mapped reads ranged from 11.99 to 30.66%.

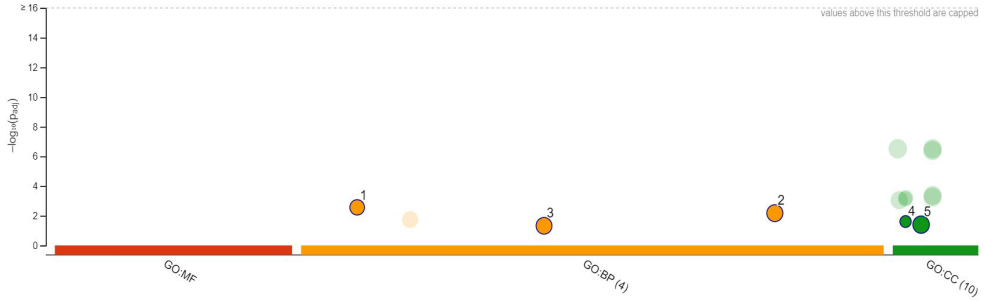
Genome reconstruction was achieved with Stringtie using all the HISAT2 output files. DESeq2 mediated differential gene expression with a *p*-value of ≤ 0.05% and a fold change (FC) higher than or equal to |2|, resulting in 3707 differentially expressed genes.

Of the total 3707 differentially expressed genes, 1558 genes were up-regulated genes while 2149 genes were down-regulated in Waris compared to Ferralise. However, out of 3707 differentially expressed genes, stable gene IDs could be obtained for 3364 genes (1375: up-regulated and 1989: down-regulated) only due to lack of significant hits or other reasons. The list of up-regulated and down-regulated genes, their description, fold changes, and Padj- (Data Availability at NCBI Sequence Read Archive under the accession PRJNA898899).

Differentially expressed genes might perform biological functions crucial to late blooming. Hence, Gene Ontology (GO) enrichment analysis was performed to investigate the functions of the differentially induced/repressed genes using g:Profiler. As presented in Fig. 1, we detected significantly enriched terms, including organelle organization (GO:0006996), organonitrogen compound biosynthetic process (GO:1901566), small molecule metabolic process (GO:0044281), photosystem II (GO:0009523) and protein-containing complex (GO:0032991). Together, these results suggest that diverse factors, including intracellular organization of organelles, biosynthesis of organonitrogen

Table 2. The overall alignment rate for different libraries

	Waris R1	Waris R2	Ferralise R1	Ferralise R2
Total read pairs	11178042	13554786	9858792	13817457
Unmapped read pairs	1453810	3025752	1592768	2192069
Uniquely mapped read pairs	6296746	8410059	7083781	9768971
Multi-mapped read pairs	3427486	2118975	1182243	1856417
Improperly aligned read pairs	47860	59699	41992	61862
Overall alignment rate (%)	91.33	82.04	88.09	88.59



ID	Source	Term ID	Term Name	p_{adj} (query_1)
1	GO:BP	GO:0006996	organelle organization	2.719×10^{-3}
2	GO:BP	GO:1901566	organonitrogen compound biosynthetic process	6.686×10^{-3}
3	GO:BP	GO:0044281	small molecule metabolic process	4.738×10^{-2}
4	GO:CC	GO:0009523	photosystem II	2.452×10^{-2}
5	GO:CC	GO:0032991	protein-containing complex	3.958×10^{-2}

Fig. 1. Figure above shows a graph depicting different Gene Ontology (GO) categories along X-axis and log10(P_{adj}) values of significant GO terms along Y-axis. The table below displays source, GO term IDs and term names alongwith P_{adj}- values of significant GO terms identified by g: Profiler, using g:SCS threshold method and a threshold value of 0.05.

compounds and protein complex photosystem II which executes the initial reaction of photosynthesis in higher plants, dominate in regulating the blooming in almonds.

DAVID was also used to identify the key terms associated with differentially-regulated genes. Among the up-regulated genes, key terms associated with the cellular component (CC) category included Secreted, Apoplast, and Mitochondrion inner membrane. Molecular function categories included Actin-binding, Hydrolase, Peroxidase, Ligase. The biological process (BP) category included Cell shape, Mitosis, tRNA processing, Auxin signaling pathway and Cell wall biogenesis/degradation. Among the down-regulated genes, key terms associated with the cellular component (CC) category included Plastid, Chloroplast, Proteasome, CF (1), Photosystem II, Mitochondrion and Cytoskeleton. Molecular function (MF) included Ribosomal protein, Ribonucleoprotein, RNA-binding, tRNA-binding, Lyase Isomerase and Translocase.

The key auxin signaling genes up-regulated were Prudul26B018413, Prudul26B017395, Prudul26B015310, Prudul26B026112, Prudul26B009102, Prudul26B002838, Prudul26B019178, Prudul26B023459 and Prudul26B012627. The key ligase up-regulated included ubiquitin ligases like Prudul26B001738, Prudul26B012009, Prudul26B000017, Prudul26B025427, Prudul26B011246, Prudul26B020058, Prudul26B000018, Prudul26B020171 and several others including 4-coumarate--CoA ligases. Important up-regulated genes under the term secreted included epidermal patterning factors (Prudul26B022647, Prudul26B011156, dirigent proteins (Prudul26B028046, Prudul26B028940), laccases like Prudul26B002680, peroxidases (Prudul26B009097, Prudul26B035878 and others), serine carboxypeptidase-like and others like Prudul26B031356 and Prudul26B019449.

Key genes down-regulated under translocase term included Prudul26B000346, Prudul26B012649, Prudul26B012195, Prudul26B023898, Prudul26B001685, Prudul26B032108, Prudul26B020831 and others. Important genes under the term Isomerase included DNA topoisomerases (Prudul26B028436, Prudul26B031676), peptidyl-prolyl cis-trans isomerases and others. Key genes down-regulated under the term ribonucleoprotein included ribosomal proteins, Serine/threonine-protein phosphatase Prudul26B012400, elongation factor 2 Prudul26B036019 and others. Important down-regulated genes under the term proteasome included 26S proteasome subunit (Prudul26B011537, Prudul26B000335, Prudul26B023604, Prudul26B004942, Prudul26B021054 and others), eukaryotic translation initiation factor 3 subunit M-like

(Prudul26B016193), COP9 signalosome complex subunits (Prudul26B018846 and Prudul26B030983) and others. Key downregulated genes under the term photosystem II included Chlorophyll a-b binding proteins (Prudul26B033007, Prudul26B030364, Prudul26B012874, Prudul26B029006, Prudul26B030999), Prudul26B005976 and oxygen-evolving enhancer protein Prudul26B022911. The key genes down-regulated under the term CF(1) included ATP synthases Prudul26B000346, Prudul26B023590, Prudul26B011120, Prudul26B030071 and Prudul26B012649.

Among the up-regulated genes, DAVID identified three significant terms associated with KEGG pathway viz. "Pyrimidine metabolism", "Purine metabolism," and Seleno-compound metabolism". Eleven genes associated with Pyrimidine metabolism included Prudul26B005033, Prudul26B019171, Prudul26B019863, Prudul26B023061, Prudul26B006864, Prudul26B000239, Prudul26B002976, Prudul26B031560, Prudul26B027565, Prudul26B009647 and Prudul26B027851. Ten genes associated with Purine metabolism included Prudul26B020988, Prudul26B024676, Prudul26B025442, Prudul26B011836, Prudul26B001861, Prudul26B007801, Prudul26B004967, Prudul26B018243, Prudul26B026229 and Prudul26B003809. Four genes associated with Selenocompound metabolism included Prudul26B026100, Prudul26B016594, Prudul26B013509, Prudul26B001779.

Among the down-regulated genes, DAVID identified some significant terms like the Pentose phosphate pathway, Carbon metabolism, phosphatidylinositol signaling system, carbon fixation in photosynthetic organisms, photosynthesis, ribosome and porphyrin metabolism. Key genes associated with pentose phosphate pathway included Prudul26B005893, Prudul26B027583, Prudul26B026776, Prudul26B026252, Prudul26B017739, Prudul26B022769, Prudul26B004881, Prudul26B011408, Prudul26B029241. Key genes associated with phosphatidylinositol signaling system included Prudul26B026001, Prudul26B026197, Prudul26B008531, Prudul26B021755, Prudul26B002429, Prudul26B016492, Prudul26B007113, Prudul26B011781, Prudul26B014107, Prudul26B008592, Prudul26B031681. Key genes associated with porphyrin metabolism included Prudul26B022335, Prudul26B031162, Prudul26B002741, Prudul26B024350, Prudul26B008142, Prudul26B020315, Prudul26B006121, Prudul26B002846, Prudul26B026773, Prudul26B001844, Prudul26B016580.

Discussion

RNA sequencing technology allowed us to study differences in gene expression patterns in cultivars differing in blooming dates. In the present study, we observed the up-regulation of genes associated with the auxin signaling pathway. It is pertinent to mention here that auxin is known to control the initiation of flower primordial (Kaur et al. 2021). Application of auxin in Arabidopsis increase H3K9ac (Antihistone 3 lysine 9 acetylation) at LFY and FIL (FILAMENTOUS FLOWER) loci which leads to mRNA accumulation of these genes and promotes floral primordium initiation. In the absence of auxin treatment, the transcription of LFY and FIL is inhibited by transcriptional co-repressor TOPLESS (TPL) and the histone deacetylase (HDA19) that occupies the MP (MONOPTEROS)-bound site of these genes.

DAVID also revealed the up-regulation of ubiquitin ligase genes in Waris. The ubiquitination pathway plays a role in embryogenesis, photo-morphogenesis, hormone regulation, senescence and defense. Overexpression of *CaRING1* in Arabidopsis (*Arabidopsis thaliana*) leads to enhanced resistance against *Pseudomonas syringae* pv *tomato* and *Hyaloperonospora arabidopsidis* infections (Lee et al. 2011). Similarly, ACD6 (Accelerated Cell Death 6) protein has an important role in plant response to biotic and abiotic stresses through SA signaling (Jasinski et al. 2021). Down-regulation of putative E3 ubiquitin-protein ligase *RING1a* (LOC117629512) and other ACCELERATED CELL DEATH 6-like (LOC117612820) genes observed in Waris probably indicates differential responses of the two cultivars towards biotic stresses including pathogens.

Among the key down-regulated genes included Serine/threonine-protein phosphatase. The serine/threonine phosphatase PP2A family plays an important role in regulating protein dephosphorylation in a highly coordinated manner in association with other regulatory subunits (Chawla et al. 2020). Arabidopsis plants with mutations in Protein phosphatase 2A exhibited late flowering compared to wild type (Kataya et al. 2015). In line with these findings, the down-regulation of Serine/threonine phosphatases in early blooming Waris indicates genetic differences for these loci when compared to the late bloomer Ferralise.

Our study also observed down-regulation of elongation factors and eukaryotic translation initiation factors. Translation initiation factors eIF3f, eIF3h and eIF3e have been associated with pollen germination, embryogenesis and pollen growth and germination in *Arabidopsis* (Xia et al. 2010; Roy et al. 2011). These findings reveal the probable differential regulation of translation initiation factors in almond cultivars varying in bloom dates.

We also report the down-regulation of genes associated with phosphatidylinositol signaling. In Arabidopsis, disruption of biosynthesis/hydrolysis of PIP2 interferes with vesicle trafficking and impairs the growth of pollen tube, supporting a role for PIP2 in the regulation of pollen tube growth (Rodas-Junco et al., 2020). Given these findings, differential regulation of genes associated with phosphatidylinositol signaling in almond cultivars probably plays a role in reproductive processes.

Plants have developed intricate signaling mechanisms in order to regulate their response to external signals and

protein kinases are an important component of these signaling cascades (Raina et al. 2012). Present studies reveal differential regulation of serine/threonine protein kinase genes. Recently, the involvement of serine/threonine protein kinases was reported in the recognition of self-incompatibility in almonds (Xu et al. 2022). Hence, the differential regulation of the protein kinase genes observed in our studies probably indicates differential incompatibility reactions in the two cultivars under study.

F-box proteins regulate an array of developmental processes, including photo-morphogenesis, circadian clock regulation, self-incompatibility, and floral meristem and floral organ identity determination (Jain et al. 2007). F-box genes show an increasing trend with pollen tube development and play an important role in self-incompatibility (SI) reactions (Yan et al. 2019). The differential regulation of F-box proteins probably points towards the differential SI reactions among the two cultivars under the present study.

An interesting observation was that in the initial stage of data analysis, we got two kinds of gene IDs; some with prefix Prudul26A and others with prefix Prudul26B. When we submitted these gene IDs to g:Profiler for functional profiling, the results were generated using only gene IDs with prefix Prudul26B as input. The gene IDs with prefix Prudul26A were not recognized by g:Profiler. In order to improve the results, we extracted the fasta sequences of gene IDs with prefix Prudul26A and blasted these sequences against *P. dulcis* genome using BlastX in Ensembl plants (https://plants.ensembl.org/index.html). The hits with maximum score were considered as equivalent conversions for the concerned query gene ID. Most of the genes with prefix Prudul26A retained the suffix when processed using the above procedure. For example, Prudul26A017255 when was blasted, the resulting hit was Prudul26B017255. Here only the prefix changed but there was no change in suffix i. e 017255. However, there were some exceptions also. For example, the fasta sequence of Prudul26A019447 when blasted against *P. dulcis* genome revealed several hits in results and a hit with maximum score of 388 i. e Prudul26B019254, was considered as most appropriate gene ID conversion. However, the results had several other hits, including one with gene ID Prudul26B019447 (with same suffix as query). However, Prudul26B019447 was a hit with a low score (196) compared to Prudul26B019254 (388). Hence we considered Prudul26B019254 as a genuine conversion of Prudul26A019447 and not the gene ID Prudul26B019447.

Data Availability

The sequencing data from this study have been submitted to the NCBI Sequence Read Archive under the accession PRJNA898899.

Authors' Contribution

Conceptualization of research (JIM, SKR); Designing of the experiments (JIM, SKR); Contribution of experimental materials (JIM, SKR, OCS); Execution of field/lab experiments and data collection (JIM, SKR, WHR, SUN, SY); Analysis of data and interpretation (JIM, SKR); Preparation of the manuscript (JIM, SKR, OCS, SMS).

Acknowledgment

The authors thank the Indian Council of Agricultural Research (ICAR) for financial assistance.

References

Albala K. 2009. Almonds along the Silk Road: The exchange and adaptation of ideas from West to East. Petits Propos. Culin., **88**: 19–34.

Barreca D., Nabavi S. M., Sureda A., Rasekhian M., Raciti R, Silva A. S., Annunziata G., Arnone A., Tenore G. C., Süntar İ. and Mandalari G. 2020. Almonds (*Prunus Dulcis* Mill. D. A. Webb): A Source of Nutrients and Health-Promoting Compounds. Nutrients, **12**: 672. doi: 10.3390/nu12030672.

Chawla S., Marothia D. and Pati P. K. 2020. Role of Serine/Threonine Phosphatase PP2A Class and Its Regulators in Salinity Stress Tolerance in Plants. In: Pandey, G.K. (eds) Protein Phosphatases and Stress Management in Plants. Springer, Cham. https://doi.org/10.1007/978-3-030-48733-1_4

Huang D. W., Sherman B.T., Lempicki R. A. 2009. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protoc., **4**: 44-57.

Jain M., Nijhawan A., Arora R., Agarwal P., Ray S., Sharma P., Kapoor S., Tyagi A. K. and Khurana J. P. 2007. F-box proteins in rice. Genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. Plant Physiol., **143**: 1467-83. doi: 10.1104/pp.106.091900

Jasinski S., Fabrissin I., Masson A., Marmagne A., Lécureuil A., Bill L. and Chardon F. 2021. ACCELERATED CELL DEATH 6 Acts on Natural Leaf Senescence and Nitrogen Fluxes in Arabidopsis. Front Plant Sci., **11**: 611170. doi: 10.3389/fpls.2020.611170.

Jung S. et al. 2019. 15 years of GDR: New data and functionality in the Genome Database for Rosaceae. Nucleic Acids Res. 47: D1137-D1145

Kataya A. R., Heidari B. and Lillo C. 2015. Protein phosphatase 2A regulatory subunits affecting plant innate immunity, energy metabolism, and flowering time--joint functions among B'η subfamily members. Plant Signal Behav., **10**: e1026024. doi: 10.1080/15592324.2015.1026024.

Kaur A., Maness N., Ferguson L., Deng W. and Zhang L. 2021. Role of plant hormones in flowering and exogenous hormone application in fruit/nut trees: a review of pecans. Fruit Research, **1**: 15 https://doi.org/10.48130/FruRes-2021-0015

Kolberg L., Raudvere U., Kuzmin I., Adler P., Vilo J. and Peterson H. 2023. g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res., **51**(W1): W207-W212. doi: 10.1093/nar/gkad347

Kumawat K.L., Raina S.K., Kumar D. et al. 2023. Association of Reproductive Phenology with Air Temperature in Almond (*Prunus dulcis* [Mill.] D.A. Webb) Cultivars Under Northwestern Himalayan Conditions. Erwerbs-Obstbau. https://doi.org/10.1007/s10341-023-00991-9

Lai Z., Schluttenhofer C., Bhide K. et al. 2014. MED18 interaction with distinct transcription factors regulates multiple plant functions. Nat Commun., **5**: 3064 https://doi.org/10.1038/ncomms4064

Lee D. H., Choi H. W. and Hwang B. K. 2011. The pepper E3 ubiquitin ligase RING1 gene, CaRING1, is required for cell death and the salicylic acid-dependent defense response. Plant Physiol., **156**: 2011-25. doi: 10.1104/pp.111.177568

Prudencio Á. S., Hoeberichts F. A., Dicenta F., Martínez-Gómez P. and Sánchez-Pérez R. 2021. Identification of early and late flowering time candidate genes in endodormant and ecodormant almond flower buds. Tree Physiol., **41**: 589-605. doi: 10.1093/treephys/tpaa151.

Raina S. K., Wankhede D. P., Jaggi M., Singh P., Jalmi S. K., Raghuram B., Sheikh A. H. and Sinha A. K. 2012. CrMPK3, a mitogen activated protein kinase from *Catharanthus roseus* and its possible role in stress induced biosynthesis of monoterpenoid indole alkaloids. BMC Plant Biol., **12**: 134. doi: 10.1186/1471-2229-12-134.

Rodas-Junco B. A., Nic-Can G. I., Muñoz-Sánchez A. and Hernández-Sotomayor S. M. T. 2020. Phospholipid Signaling Is a Component of the Salicylic Acid Response in Plant Cell Suspension Cultures. Int. J. Mol. Sci., **21**: 5285. doi: 10.3390/ijms21155285.

Roy B., Copenhaver G.P. and Von Arnim A.G. 2011. Fluorescence tagged transgenic lines reveal genet.ic defects in pollen growth-application to the eIF3 complex. - PLoS One **6**: e17640.

Sánchez-Pérez R., Del Cueto J., Dicenta F. and Martínez-Gómez P. 2014. Recent advancements to study flowering time in almond and other *Prunus* species. Front Plant Sci., **5**: 334. doi: 10.3389/fpls.2014.00334.

Sherman B.T., M. Hao, J. Qiu, X. Jiao, M.W. Baseler, H.C. Lane, T. Imamichi and W. Chang. 2022. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res., doi:10.1093/nar/gkac194.

The Galaxy Community. 2022. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update Nucleic Acids Research, gkac247 doi:10.1093/nar/gkac247

Xia C., Wang Y.J., Li W.Q., Chen Y.R., Deng Y., Zhang X.Q., Chen L.Q. and Ye D. The Arabidopsis eukaryotic translation initiation factor 3, subunit F (AteIF3f), is required for pollen germination and embryogenesis. - Plant J., **63**: 189-202, 2010.

Xu Y., Zhang Q., Zhang X., Wang J., Ayup M., Yang B., Guo C., Gong P. and Dong W. 2022. The proteome reveals the involvement of serine/threonine kinase in the recognition of self- incompatibility in almond. J. Proteom., **256**: 104505. doi: 10.1016/j.jprot.2022.104505.

Yan F., Zhou H., Yue M., Yang G., Li H., Zhang S. and Zhao P. 2019. Genome-Wide Identification and Transcriptional Expression Profiles of the F-box Gene Family in Common Walnut (*Juglans regia* L.). Forests, **10**: 275. https://doi.org/10.3390/f10030275

Zhang C. and Hao Y. J. 2020. Advances in Genomic, Transcriptomic, and Metabolomic Analyses of Fruit Quality in Fruit Crops, Hort. Plant J., **6**: 361-371, ISSN 2468-0141. https://doi.org/10.1016/j.hpj.2020.11.001.