

## RESEARCH ARTICLE

# De Novo Assembly and Transcriptome Analysis of Bulb Onion (*Allium cepa* L.) during Cold Acclimation Using Contrasting Genotypes

Jeongsukhyeon Han<sup>1</sup>\*, Senthil Kumar Thamilarasan<sup>1</sup>, Sathishkumar Natarajan<sup>1</sup>, Jong-In Park<sup>1</sup>, Mi-Young Chung<sup>2</sup>, Ill-Sup Nou<sup>1</sup>\*

**1** Department of Horticulture, Sunchon National University, Suncheon, Jeonnam, Republic of Korea,

**2** Department of Agricultural Education, Sunchon National University, Suncheon, Jeonnam, Republic of Korea

\* These authors contributed equally to this work.

\* [nis@sunchon.ac.kr](mailto:nis@sunchon.ac.kr)



## OPEN ACCESS

**Citation:** Han J, Thamilarasan SK, Natarajan S, Park J-I, Chung M-Y, Nou I-S (2016) De Novo Assembly and Transcriptome Analysis of Bulb Onion (*Allium cepa* L.) during Cold Acclimation Using Contrasting Genotypes. PLoS ONE 11(9): e0161987. doi:10.1371/journal.pone.0161987

**Editor:** Xiang Jia Min, Youngstown State University, UNITED STATES

**Received:** April 6, 2016

**Accepted:** August 16, 2016

**Published:** September 14, 2016

**Copyright:** © 2016 Han et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This research was supported by Golden Seed Project (Center for Horticultural Seed Development, No. 213003-04-4-SB110), Ministry of Agriculture, Food and Rural Affairs (MAFRA), Ministry of Oceans and Fisheries (MOF), Rural Development Administration (RDA) and Korea Forest Service (KFS).

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

Bulb onion (*Allium cepa*) is the second most widely cultivated and consumed vegetable crop in the world. During winter, cold injury can limit the production of bulb onion. Genomic resources available for bulb onion are still very limited. To date, no studies on heritably durable cold and freezing tolerance have been carried out in bulb onion genotypes. We applied high-throughput sequencing technology to cold (2°C), freezing (-5 and -15°C), and control (25°C)-treated samples of cold tolerant (CT) and cold susceptible (CS) genotypes of *A. cepa* lines. A total of 452 million paired-end reads were *de novo* assembled into 54,047 genes with an average length of 1,331 bp. Based on similarity searches, these genes were aligned with entries in the public non-redundant (nr) database, as well as KEGG and COG database. Differentially expressed genes (DEGs) were identified using  $\log_{10}$  values with the FPKM method. Among 5,167DEGs, 491 genes were differentially expressed at freezing temperature compared to the control temperature in both CT and CS libraries. The DEG results were validated with qRT-PCR. We performed GO and KEGG pathway enrichment analyses of all DEGs and iPath interactive analysis found 31 pathways including those related to metabolism of carbohydrate, nucleotide, energy, cofactors and vitamins, other amino acids and xenobiotics biodegradation. Furthermore, a large number of molecular markers were identified from the assembled genes, including simple sequence repeats (SSRs) 4,437 and SNP substitutions of transition and transversion types of CT and CS. Our study is the first to provide a transcriptome sequence resource for *Allium* spp. with regard to cold and freezing stress. We identified a large set of genes and determined their DEG profiles under cold and freezing conditions using two different genotypes. These data represent a valuable resource for genetic and genomic studies of *Allium* spp.

## Introduction

Cold and freezing stresses are common environmental abiotic stresses. Plants undergo a variety of physiological and biochemical changes as defense systems for withstanding cold (0–15°C) and freezing (<0°C) conditions. Although not all plants can tolerate cold and freezing temperatures, many plants show tolerance to cold and freezing via a phenomenon termed ‘cold acclimation’ (CA) [1, 2]. During cold acclimation, many genes show differential expression, responding to cold and freezing stress at the transcriptional level [3, 4]. For example, in the model plant *Arabidopsis thaliana*, many transcription factors and thousands of genes are thought to be involved in cold stress responses [4]. In general, plants from temperate regions are considered to be chilling tolerant to varying degrees and their freezing tolerance is increased by previous exposure to cold and freezing temperatures [5]. Recent studies have aimed to analyze the functions of stress-inducible genes, not only to understand the mechanism of cold and freezing stress responses, but also to improve the stress tolerance of other plants by gene transfer. The CBF/DREB (C-repeat [CRT]/dehydration responsive element [DRE] binding factor) transcription factors mediate a cold acclimation pathway that induces transcriptional reprogramming of cold regulated (COR) genes [1, 2]. CBF, an AP2/ERF family transcription factor, binds to and induces a number of COR genes (CBF regulon) that contain a CRT/DRE motif (CCGAC) in their promoters, thereby conferring freezing tolerance [6–11].

Bulb onion (*Allium cepa*) is a major vegetable crop that belongs to the *Amaryllidaceae* family, including 300 species of which only 70 have been cultivated [12]. In onion, the pungency level is related to the content of polyphenols, vitamins and sulphur-containing compounds [13–15]. Those compounds also affect various aspects of tolerance of environmental stresses, including cold and freezing stress. In spite of the status of *Allium* species as a major vegetable crop with both nutritional and medicinal value, genetic research in *Allium* has lagged behind that in other vegetable crops and little genomic information is available, in part because of the enormous size of the *Allium* genome (16.3 Gb) [16, 17]. At present, research examining stress-inducible genes of *A. cepa* has focused on physiology and identification of specific response genes. However, to our knowledge, no study involving individual or large-scale screening of cold and freezing stress-related genes and molecular marker identification has been published to date. Next generation sequencing (NGS) has enabled large-scale transcriptome data analyses that have dramatically improved the efficiency of gene discovery at the genome level without prior knowledge of reference genome sequences. Since the first use of Solexa/Illumina’s Digital Gene Expression (DGE) system to study the zebrafish transcriptome after *Mycobacterium marinum* infection [18], RNA-Seq and DGE technology have been widely used to identify plant genes related to important agronomic traits, including those expressed under stress conditions [19, 20]. Large portions of the transcriptomes of *A. thaliana* and *Brassica rapa* are regulated by abiotic stresses such as cold, salt and drought [4, 21]. In wheat, cold stress induces transcriptome reprogramming, with over 2% of the wheat genome showing a greater than two-fold change in expression [22]. However, transcriptome changes and biochemical functions of cold and freezing stress-regulated genes remain unknown. Molecular markers play a crucial role in applications related to genetic diversity in plant breeding. Many EST-SSR markers have been developed using high-throughput sequencing data, for instance 39,257 EST-SSRs from healthy rubber trees and 4,609 EST-SSRs from date palm EST sequences; in both cases, the primers were validated to have polymorphism in different cultivars [23, 24].

In this study, we obtained transcriptomes from contrasting lines of *Allium cepa* treated with cold (2°C), freezing (-5, -15°C) and control temperature (25°C) using Illumina paired-end sequencing technology. The resulting sequence data were *de novo* assembled and annotated without prior reference genome information. Furthermore, we compared the gene expression

profiles between treated and control plants, and SSRs and SNPs markers were predicted *in silico*. To our knowledge this is the first report on the cold and freezing transcriptome of bulb onion. Hence, this research significantly enhances our understanding of the transcriptional changes underlying the cold and freezing response in *A. cepa*. Our study provides information towards elucidating cold and freezing tolerance mechanisms; in addition, the molecular markers predicted and developed herein should facilitate gene mapping and genetic diversity analysis.

## Materials and Methods

### Plant materials and electrolyte leakage test

Two bulb onion (*Allium cepa*) lines that exhibit contrasting sensitivity to cold stress, cold tolerant (CT) 36122 and cold susceptible (CS) 36101, inbred lines were collected from the Nongwoo Bio Seed Company, Suwon, Korea. The seeds were grown under controlled conditions in a growth chamber for ~4 months at 25°C with 16 h light/8 h dark photo period. For cold and freezing stress treatment 25°C (control) plants were exposed at 24 h to 2°C, -5°C, and -15°C. After treatment, the 4<sup>th</sup> or 5<sup>th</sup> leaf of CT and CS were collected. Immediately after treatment, electrolyte leakage from the cold and freezing-stressed and control plant samples was measured using previously described methods [25, 26], with some modifications. Briefly, 10 leaf discs (1 cm in diameter) excised from fully expanded leaves of 3–4 plants were placed in a glass tube with 10 ml distilled water, and incubated on an orbital shaker at 150 rpm for 30 min at room temperature. The initial conductivity (I) was measured using a CON110 conductivity meter (Oakton Ins. USA). The leaf discs were then kept in a boiling water bath for 10 min and cooled to room temperature before the final conductivity (F) was measured. The relative electrolyte leakage was calculated using the formula I/F X 100. Significance tests were carried out using Least Significant Difference (LSD) at P = 0.05 level was used to compare the means.

### RNA extraction and library preparation

The total RNA was isolated using TRIzol reagent (Invitrogen, USA) according to the manufacturer's instructions. Poly (A)+ mRNA was purified with oligo (dT) beads, and then the mRNA was randomly segmented into small fragments by divalent cations (Fragmentation Buffer, Illumina, Hayward, CA) at 94°C for 5 min. These short fragments were used as templates to synthesize the first-strand cDNA using random hexamer primers. The second-strand cDNA was synthesized using RNaseH and DNA polymerase I. Short cDNA fragments were purified with the QiaQuick PCR extraction kit. After that, the cDNA fragments were connected with sequencing adapters according to Illumina's protocol (San Diego, CA, USA). For each sample, at least ~20 µg total RNAs was used for RNA-sequencing at the Beijing Genomics Institutes (BGI)-Shenzhen, (Shenzhen, China). For the reverse transcription-polymerase chain reaction real-time PCR experiments, the total RNA was prepared following the aforementioned procedures.

### Sequencing, *de novo* assembly and functional annotation

Individual cDNA samples of CT and CS treated at 2°C, -5°C, -15°C and followed by 25°C (control) were prepared for sequencing. The constructed paired-end (PE) sequence libraries of each sample were sequenced by Illumina HiSeq 2000 sequences (2 × 100 bp read length). The generated raw read quality was assessed using FASTX-Toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). To obtain high-quality sequencing data, adaptor sequences and low-quality reads were removed using 'sickle' (<https://github.com/najoshi/sickle>) and 'SeqPrep' (<https://github.com/jstjohn/SeqPrep>), respectively, with the default parameters. We calculated Q20, Q30 and

GC contents for all cleaned data for each sample. Further, poly-N and low-quality reads were removed and high-quality cleaned data were subjected to all downstream calculations. The *de novo* assembly program Trinity (<http://trinityrnaseq.sourceforge.net/>) was used for short read assembly and to calculate the N50 number. We used the pipeline of ‘transcript\_to\_best\_scoring\_ORFs.pl’ pipeline in trinity software to predict the ORFs with default parameters [27]. The raw sequence data obtained by PE RNA-sequencing have been deposited in the NCBI SRA and TSA (National Centre for Biotechnology Information Short Read Archive, <http://www.ncbi.nlm.nih.gov/sra> and Transcript Shotgun Assembly, <http://www.ncbi.nlm.nih.gov/genbank/tsa/>) database under accession number **SRP064878** and **GETF00000000**, respectively.

## Functional annotation and classification

For functional annotation, the assembled genes that might putatively encode proteins were searched against the nr (BLASTX) (NCBI non-redundant protein sequences, <http://www.ncbi.nlm.nih.gov/>), Swiss-Prot (<http://www.expasy.ch/sprot/>), and nt (NCBI non-redundant nucleotide sequences) databases using Blast2GO with minimum E-value of  $<1e^{-10}$  as the threshold [28]. With nr annotations, the Blast2GO program was used to assign GO annotations according to component function, biological process and cellular component categories [29]. In a final step, COG (<http://www.ncbi.nlm.nih.gov/cog/>) [30] and KEGG (<http://www.genome.jp/kegg/>) [31] annotation was performed for all of the genes.

## Mapping and differentially expressed gene analysis

The procedure for constructing the DGE sequencing libraries was the same as that for constructing the transcriptome sequencing libraries. After the raw data were generated and the data-processing steps were completed, the clean reads were mapped to the assembly transcriptome reference sequences using RSEM with default parameters [32]. Gene expression levels were calculated based on the numbers of reads mapped to the reference sequence, using the FPKM [33] method. After calculating gene expression levels, the differentially expressed genes (DEGs) were screened by comparing gene expression levels. Differential expression analysis was performed with software ‘edger’. The resulting P values were adjusted using Benjamini and Hochberg’s approach for controlling the false discovery rate (FDR  $\leq 0.01$ , log<sub>2</sub>FC  $\geq 2$ ). Functional enrichment analysis, including GO and KEGG analyses, was performed using genes of the entire transcriptome. For this analysis, a Bonferroni-corrected P-value  $\leq 0.05$  was used. GO functional enrichment analysis was carried out using Goa tools (<https://github.com/tanghaibao/Goatools>).

## In silico analysis of molecular marker

The detection of SSRs from the total of assembled genes was performed using Msatcommader software [34]. The software accepts FASTA formatted sequence files, the sequence ID, SSR motif, number of repeats (di-, tri-, tetra-, penta- or hexa-nucleotide repeat units), GATK (Genome analysis tool kit) software was used to perform SNP calling with default parameters [35]. Along with the SNP and position, additional information about the called SNPs, such as SNP quality, depth of coverage.

## Quantitative real-time PCR validation

To confirm the DEG results, quantitative real-time reverse transcription PCR (qRT-PCR) was performed. Twenty-five genes were chosen based on FDR  $\leq 0.005$  value to increase the reliability of DEGs for qRT-PCR analysis in the CT vs. CT libraries. Gene-specific primers were

designed based on the sequencing results using Primer 3 (version 6.0) [36] ([S1 Table](#)). The real-time PCR was performed on an Illumina Eco real-time machine (PhileKorea Technology, Korea) using 1 µl first-strand cDNAs mixed with SYBR Premix Ex TaqTM (Toyobo, Japan). The thermal cycling conditions were 30 s at 95°C, followed by 40 cycles of 5 s at 95°C and 31 s at 58°C. All reactions were performed in triplicate. The relative expression levels of each gene were calculated using the  $2^{-\Delta\Delta CT}$  method normalized to the internal control gene, *Allium* ssp. *Actin*.

## Results and Discussion

### Illumina paired-end sequencing and *de novo* assembly

Two well-known contrasting genotypes of Korean onion lines were chosen in this study. Cultivars 36101 and 36122, which are widely adapted to winter and spring/summer seasonal cultivation in South Korea, respectively, were used as cold susceptible (CS) and cold tolerant lines (CT), respectively [37]. To precisely evaluate their tolerance of cold and freezing stress, we compared the electrolyte leakage test of leaf samples after treatment for 24h with different temperature that were further used for transcriptome experiments. Electrolyte leakage is inversely proportion to freezing tolerance [38–40], and there great differences between the two lines during cold (2°C) and freezing stress (-5°C) ([S1 Fig](#)). Comparison of control 25°C and -5°C observations revealed that the tolerant line had adapted to the freezing stress and was able to protect the membrane from further damage. Two way ANOVA statistical analysis showed that observation of electrolyte leakage was significantly different at different temperatures within and between genotypes.

To identify any freezing-tolerance genes induced to create this phenotypic response in onion, total RNA was isolated from plants treated with control temperature (25°C) or 2, -5 and -15°C. Paired-end RNA-sequencing using the Illumina HiSeq 2000 platform was performed with control and treated samples. A total of 452 million paired-end reads were produced for control and treated samples ([Table 1](#)). Clean reads were obtained from paired-end reads by performing high stringency filtering with default parameters to remove adaptor and/or vector read sequences. Ultimately, 450,821,674 (99.6%) pooled high-quality reads were used for *de novo* transcriptome assembly. All clean reads were also deposited in the National Center for Biotechnology Information (NCBI) and can be accessed in the sequence read archive (SRA) under the accession number **SRP064878**. With the Trinity program [41], 93,637 transcripts

**Table 1. Sequencing statistics of the *Allium cepa* L. transcriptomes.**

Summary	Total nucleotides
Total number of reads	452,194,370
Clean reads	450,821,674
Avg. of GC %	42.4
Total no. of isogenes/transcripts	93,637
Avg. length in bp	1,331
Minimum length in bp	500
Maximum length in bp	15,559
N50 length	1,625
Total no. of genes	54,047
Avg. length in bp	807
Minimum length in bp	150
Maximum length in bp	14,970

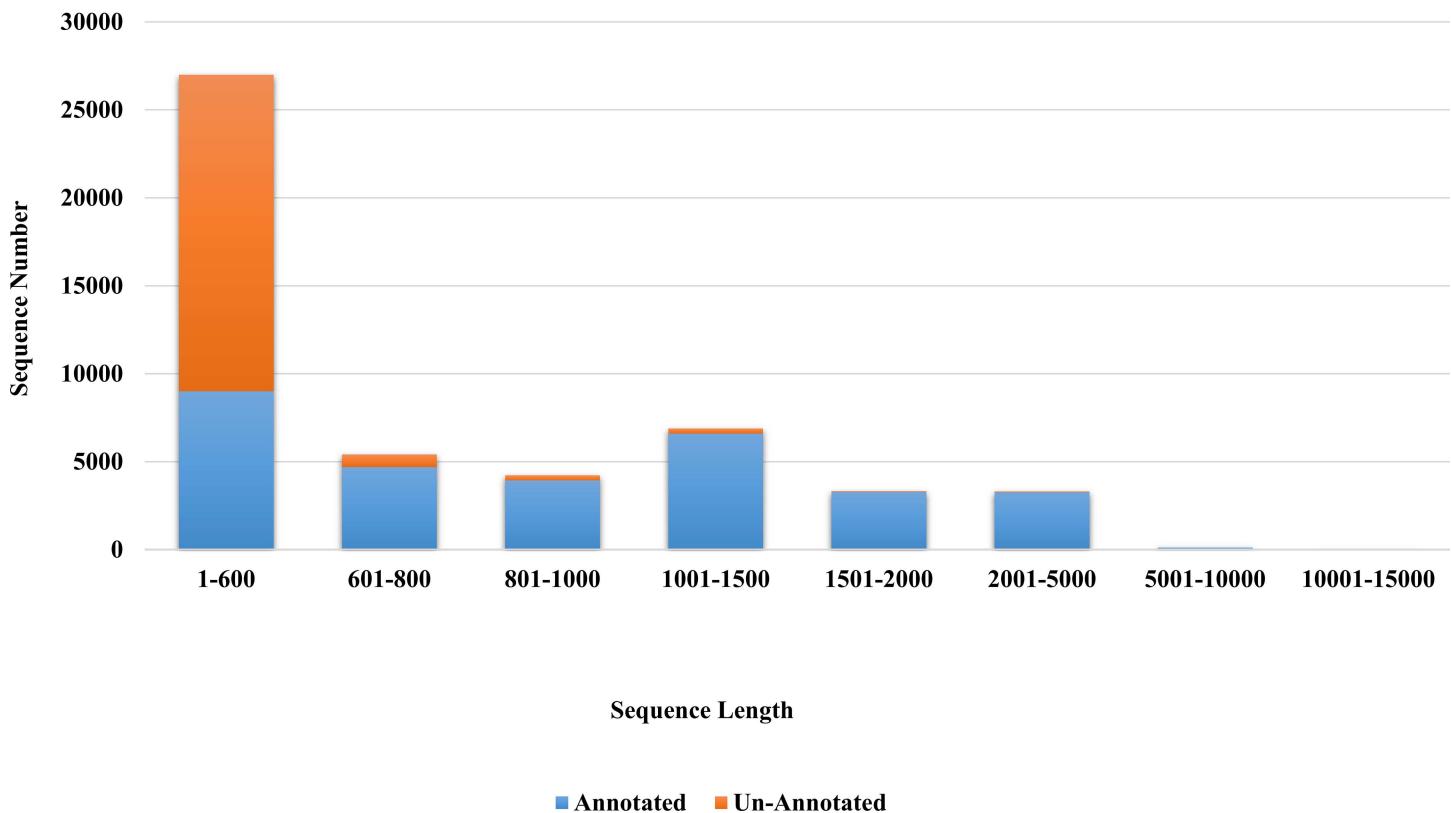
doi:10.1371/journal.pone.0161987.t001

were generated with an average length of 1,331 bp and an N50 length of 1,625 bp; the maximal transcript length was 15,559. These results were comparable to those for other transcriptome analyses in plants, such as *H. ammodendron* (N50 = 1,345 bp, average length = 728bp, [19]) and *Ammopiptanthus mongolicus* (N50 = 1,343 bp, average length = 816 bp, [42]). These data indicate that the assembly was effective in capturing a large portion of the *A. cepa* transcriptome. The assembled transcripts were initially predicted to correspond to 54,047 genes, which were screened for best scoring ORFs predicted using the Trinity program, resulting in 50,280 genes as best candidate genes, were deposited into TSA (GETF00000000). The average length of these genes was 807bp, and their lengths ranged from 150bp to 14,970bp.

### Functional annotation and classification

To identify their putative functions, the assembled genes were screened against public databases such as the NCBI non-redundant (nt, nr), Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Cluster of Orthologous Groups (COG) databases using BLASTX with an E-value threshold of  $1e^{-10}$  (default parameters). This analysis allowed good annotation of genes with more than 1,000bp, whereas 33.3% of sequences less than 600bp could be matched to known proteins (Fig 1). Among all of the genes, 30,926 (61.5%) showed significant similarity to entries in the nr database, similar to values reported for other *de novo*

### Length distribution



**Fig 1. Histogram of sequence length distribution and number of sequences annotated.** Assembled genes were searched in the nr protein database using BLASTX with a cut-off E-value  $1e^{-10}$ .

doi:10.1371/journal.pone.0161987.g001

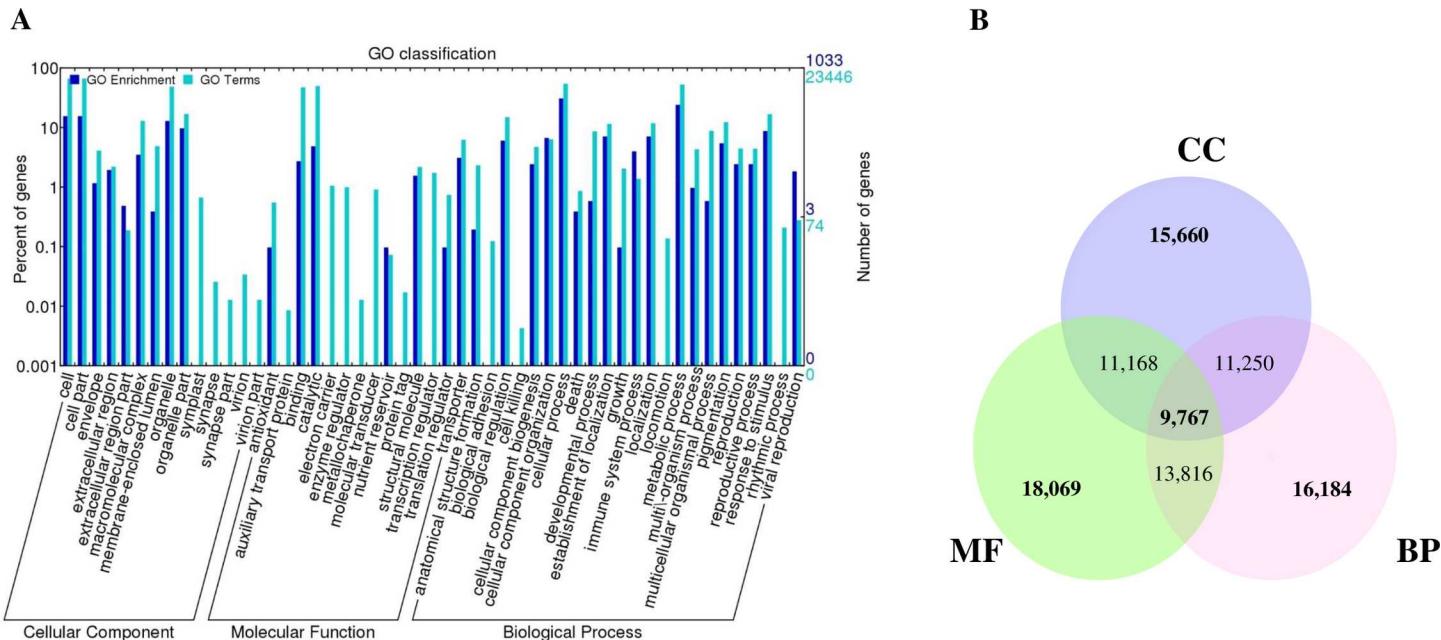
**Table 2. Summary of the functional annotation of assembled genes.**

Public database	Number of genes	Percentage (%)
Annotated in NR	30,926	61.5
Annotated in NT	1,305	3.0
Annotated in GOs	23,446	46.6
Annotated in COG	15,602	31.0
Annotated in KOG	14,993	29.8
Annotated in KO/KEGG	9,656	19.2
Total genes	50,280	100

doi:10.1371/journal.pone.0161987.t002

assemblies [19]. In BLAST searches against the NT database, a smaller percentage of similarities were found Table 2. In total, BLASTX searches against the nr database identified 19,354 accessions that did not match proteins from the public databases, suggesting that our Illumina paired-end sequencing included significant numbers of the cold and freezing stress-related genes in *A. cepa*. Among the BLAST results, the species that provided the most matches was *Vitis vinifera* (5,443) and the next closest species was *Theobroma cacao* (2,420) (S2 Fig).

To classify the functions of the predicted onion cold and freezing transcriptome genes, we used Gene Ontology (GO) analysis. Using the Blast2GO suite, 23,446 (46.6%) sequences were categorized into three different GO terms (Biological Process: BP, Molecular Function: MF and Cellular Component: CC) (Table 2). The three main categories were further classified into 58 functional groups compared with significantly enrichment analysis (Fig 2A and 2B). The most genes were classified under BP (60,578, 42.0%), followed by CC (57,202, 39.6%) and MF (26,316, 18.2%). Under the main three categories, seven groups of sub-categories were



**Fig 2.** (A) Gene ontology classification of the assembled genes. Histogram of the GO annotation and enrichment analysis was generated by the BGI WEGO software. Genes were grouped into three main GO categories: Cellular component (CC), Molecular function (MF), Biological Process (BP). The right Y- axis indicates the number of genes in a category. The left Y-axis indicates the percentage in a specific category. One gene could be assigned with more than one GO term. Color indicates GO enrichment (significantly enriched GO terms with Bonferroni corrected p-value <0.05) and GO terms in all genes. (B) Venn diagram of assembled genes were accounted in three major categories of gene ontology classification.

doi:10.1371/journal.pone.0161987.g002

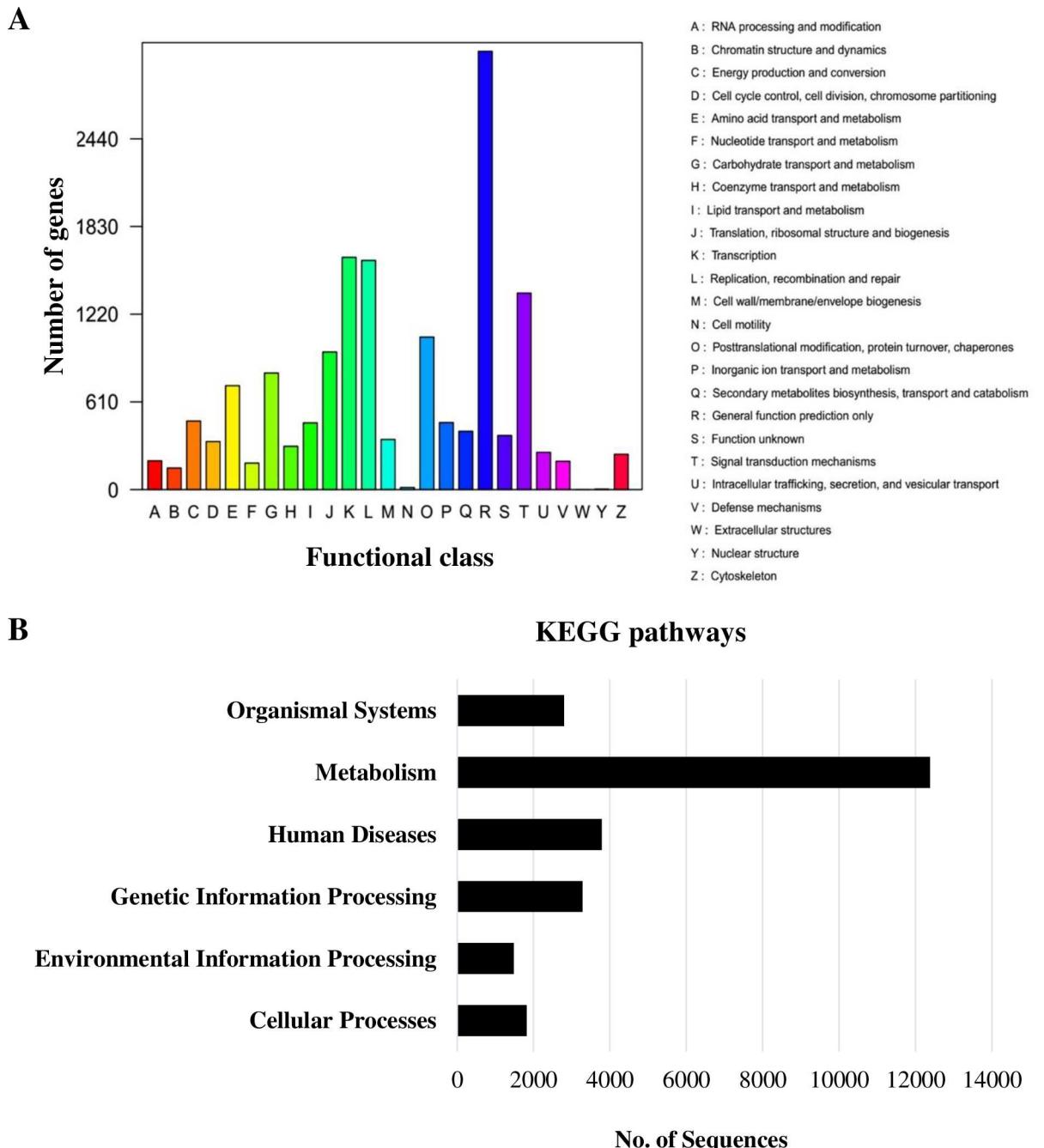
predominant in the GO classification. In the three largest categories of BP, 'cellular process (GO: 0009987)' and 'metabolic process (GO: 0008152)' were highly represented. Under the classification of MF, 'catalytic activity (GO: 0003824)' (11,715, 8.12%) and 'binding (GO: 0005488)' (11,300, 7.8%) were significantly represented, whereas other sub-categories contained only 3,301 genes representing 14.07% of the total genes. For cellular component, three sub-categories 'cell (GO: 0005623)', 'cell parts (GO: 0044464)' and 'organelle (GO: 0043226)' accounted for 68.2% of the genes. Fewer than half of the genes were not annotated in this study, likely due to the sequencing length or depth of coverage, as is common in studies performing *de novo* transcriptome analyses [43]. In addition, some of these genes might be unique to *A. cepa*.

To further investigate their possible functions, we analyzed the assembled genes using the Clusters of Orthologous Groups (COG) database. Of the 15,602 genes with significant similarity to nr database proteins in this study, 10,549 genes were grouped into 25 functional classifications (Fig 3A). The clusters related to 'general function prediction only' (3,049, 19.54%) were the largest group, followed by 'transcription' (1,616, 10.35%), 'replication, recombination and repair' (1,594, 10.21%), and signal transduction mechanisms (1,367, 8.76%).

We further characterized the transcriptome data using pathway analysis with the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. In total, 11,204 (22.28%) sequences were identified in six main categories that included 333 pathways (Fig 3B and S2 Table). 'Metabolism' was the category with the greatest number of genes (12,383, 48.4%), with metabolic pathways and biosynthesis of secondary metabolites as the two pathways represented by the most genes. Overall, these annotation and classification analyses add valuable information to the *A. cepa* transcriptome, with gene functions, annotation and metabolic pathways.

### Differentially expressed gene (DEG) library during cold and freezing stresses

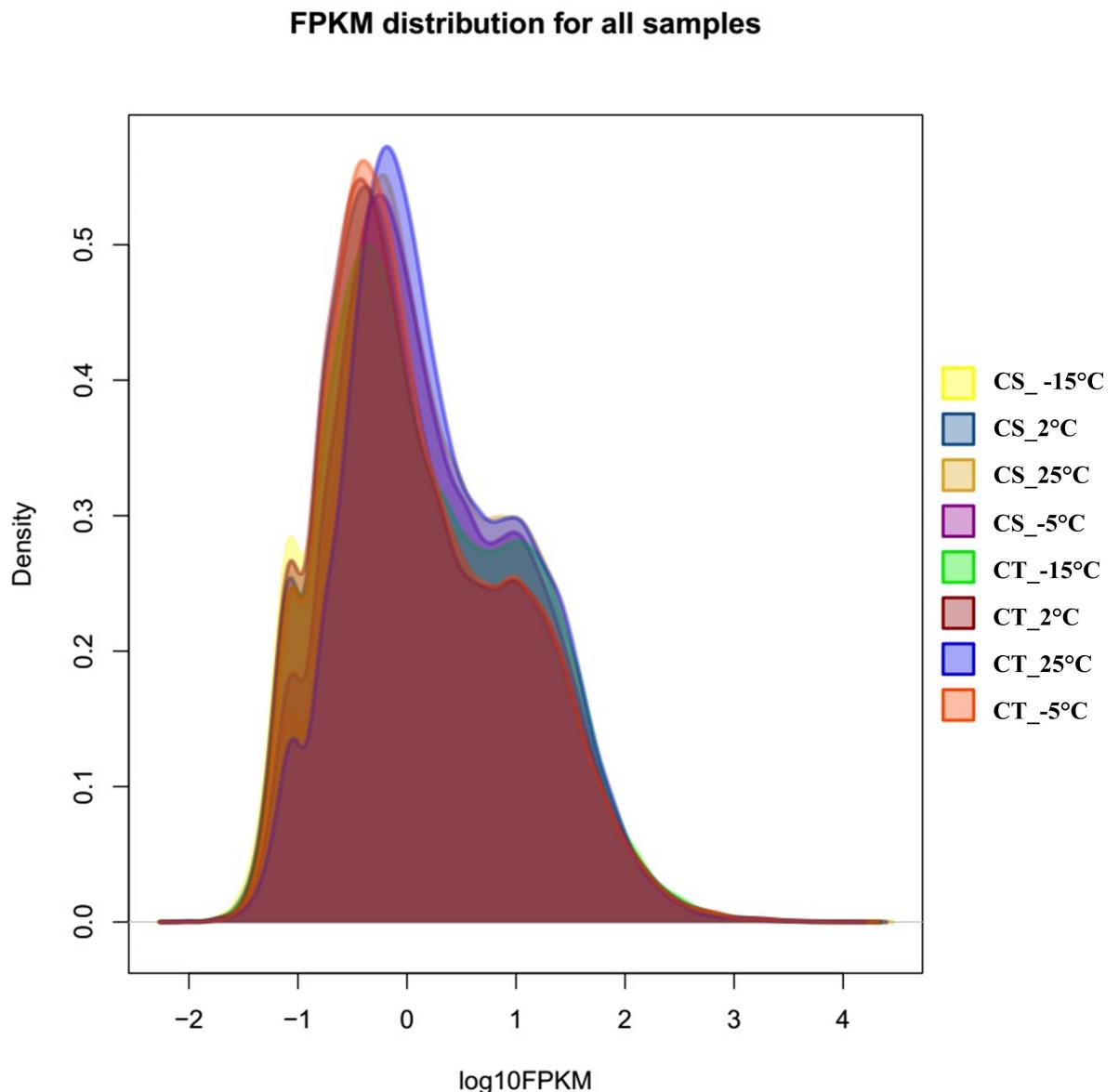
To normalize the datasets for *de novo* transcriptome assemblies, we used RSEM software, which does not require reference genome and provides superior performance over other quantification methods [32]. The FPKM (Fragments per Kilobase per Million), TPM (transcripts per million) and EC (expected count) values for each mapped clean read were calculated (S3 Table). The obtained FPKM values from the libraries ranged from 0 to 15,935, and we visualized the density distribution of expression level in each DEG library based on the  $\log_{10}$  value of FPKM (Fig 4). The distribution patterns of all libraries were overlaid with different peaks values. After normalizing the read density measurement to reveal significant responses to cold and freezing stress, genes were filtered by FDR (value  $\leq 0.01$ ) (False Discovery Rate) and a value of  $\log_2FC \geq 2$  using edgeR software with default parameters. We performed comparative profile analysis using the aligned reads of CS vs. CT, and obtained 2,400 and 4,080 up- and down-regulated genes, respectively, which were commonly expressed in all of the profiles (S3 Fig). The results indicate that a large number of genes were involved in the cold acclimation process. Hence, we compared the two lines CS vs. CT lines during 25°C, 2°C, -5°C and -15°C treatments, finding that a total of 1,822, 2,921, 2,931 and 1,887 were upregulated and 1,786, 2,740, 2,665 and 2,182 were downregulated genes, respectively (S3 Fig). Among all the genes, 705 and 1,203 genes were upregulated and downregulated, respectively, in all the temperature profiles, identifying these 1,208 genes as candidate DEGs between the contrasting lines. However, 165 upregulated and 149 downregulated genes were predominantly expressed throughout the cold treatment at -5°C and also present in temperature profiles 2°C to -15°C between the two lines, whereas 346 upregulated and 334 downregulated genes were only among one of these three profiles. Genes that are commonly regulated across all the temperature treatments



**Fig 3.** (A) Distribution of genes in the transcriptome with Clusters of Orthologous Groups (COG) functional classification. 10,549 genes were grouped into 25 categories. (B) Distribution of assembled genes in each KEGG pathway category. All genes were assigned under six major categories.

doi:10.1371/journal.pone.0161987.g003

between the two lines might be candidate cold-responsive genes that alters plant tolerance during cold stress. Notably, -5°C freezing stress led to the greatest number of DEGs in CS and CT lines. Overall, 2,409 and 2,758 genes were differentially expressed in CS and CT lines, respectively, at 25°C vs. -5°C (S4A and S4B Fig). Of these DEGs, CS (712) and CT (1,258) were successfully annotated using the public databases. To explore the biological functions and



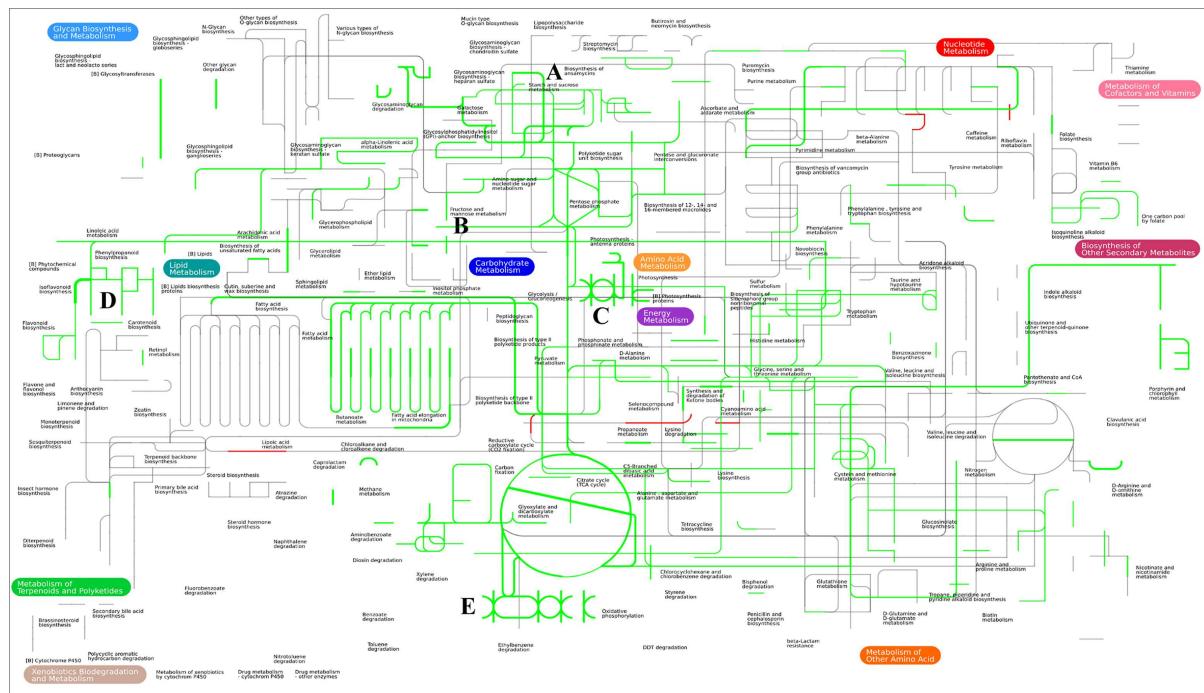
**Fig 4. Density distribution of expression level based on  $\log_{10}$  FPKM, exhibiting the overall differences in expression profile among the libraries.**

doi:10.1371/journal.pone.0161987.g004

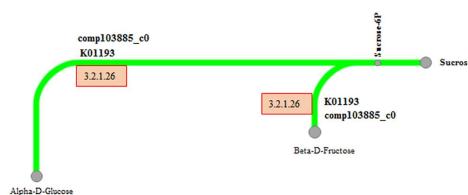
correlation of these DEGs, we performed enrichment analysis of using GO terms of three major categories for the up- and down-regulated genes in the CS and CT profiles. The resultant GO terms enriched in CS and CT, were highly enriched terms are in BP ‘cellular process’ (GO: 0009987), in MF ‘catalytic activity’ (GO: 0003824), and in CC ‘cell’ and ‘cell part’ (GO: 0005623 and GO: 0044464, respectively) ([S5 Fig](#)).

### Metabolic pathway insights related to DEGs

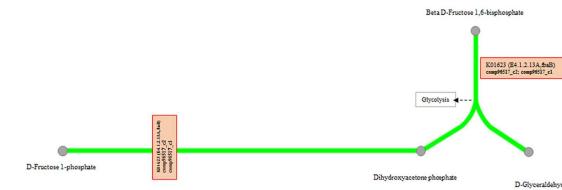
For a global enriched pathway view of freezing metabolism, we used Interactive pathways (Ipath) explorer v2, which integrates 123 KEGG maps from 183 species [44]. Previous studies have effectively used RNA-Seq transcriptomics with interactive pathway analysis [45], for instance for a global view of Chinese bayberry metabolism analysis [46]. In this study, 491



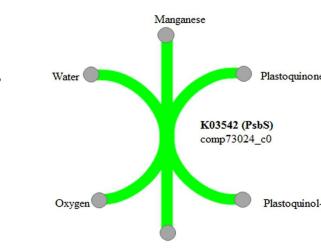
A.



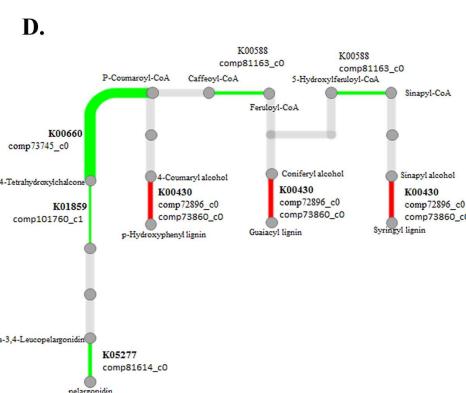
B.



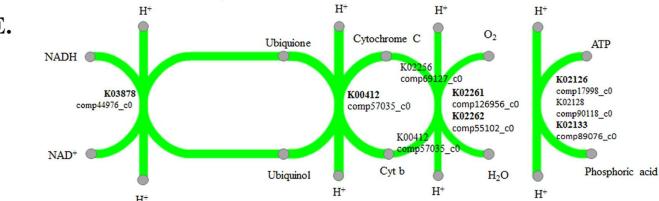
C.



D.



E.



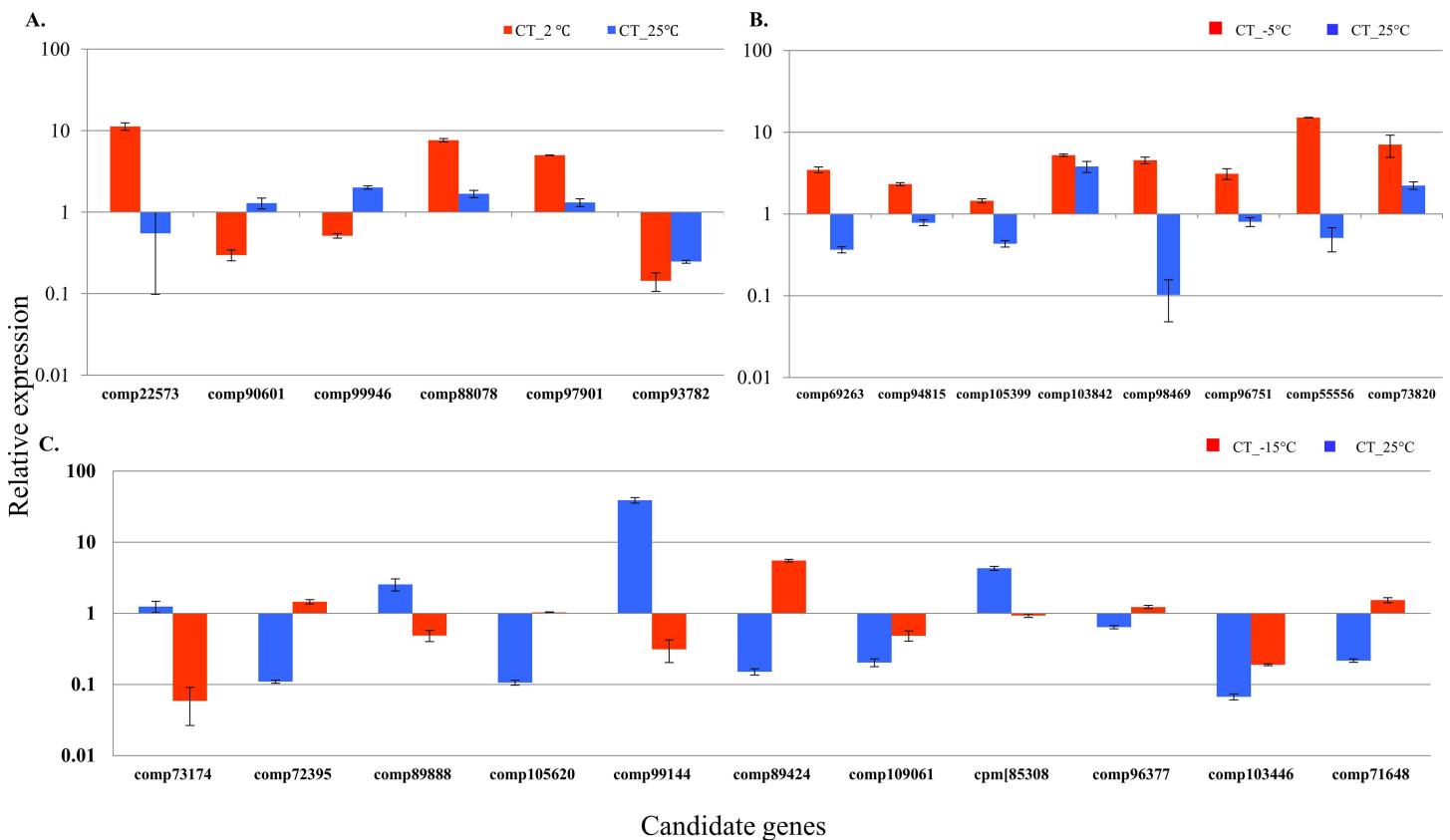
**Fig 5. Metabolic pathways active during freezing temperature in onion.** The red and green indicate up and down-regulated genes with tolerance and susceptibility, respectively. Major pathways of freezing-related metabolism is shown in (A) Starch and sucrose metabolism, (B) Fructose and mannose metabolism, (C) Energy metabolism, (D) Biosynthesis of secondary metabolites, (E) Oxidative phosphorylation.

doi:10.1371/journal.pone.0161987.g005

genes were combined with DEGs of CS and CT profiles at 25 vs. -5°C ([S4A](#) and [S4B](#) Fig, S4 Table). In all, 88 KO (KEGG orthology) ids mapped to 31 metabolic pathways were identified ([Fig 5](#)). At -5°C, expression levels showed up- and down-regulation in the CS and CT profiles, and genes mapped to pathways including metabolism of carbohydrate, biosynthesis of other secondary metabolites, energy, cofactors and vitamins, other amino acids and xenobiotics biodegradation. In some pathways, metabolism of starch and sucrose and fructose and mannose enzymatically induced beta-d-fructose, alpha-d-glucose, and D-glyceraldehyde 3-phosphate and D-fructose 1-phosphate, respectively ([Fig 5A](#) and [5B](#)). More pathways showed differential expression during freezing stress conditions than under other temperatures, in accord with GO comparisons [47]. In energy metabolism, K03542 (Psbs) plastoquinone to plastoquinone-1 may be involved in photosystem II (PSII) intersystem electron transport chain; here, our expression data show that it might be related to attainment of a cold acclimated state and, hence, maximum freezing tolerance ([Fig 5C](#)) [48]. [Fig 5D](#) shows several pathways (K00588, K0060, K01859, K05277 and K00430) involved in biosynthesis of phenylpropanoid, flavonoid and peroxidase may change the composition of aromatic, pungency compounds during freezing stress. In oxidative phosphorylation, K02126, K02262 (O<sub>2</sub> to H<sub>2</sub>O) acts as the primary terminal acceptor from PSI and PSII, and thereby leads to substantial generation of reactive oxygen species (ROS) and may also induce the rapid depletion of cellular damage (NAD/NADH) ([Fig 5E](#)) [49]. Overall, the results revealed interactive metabolic pathways associated with freezing tolerance gene expression during -5°C in CT and CS lines.

### Validation of differential gene expression

To further validate the gene expression profiles, we selected 25 candidate genes from the CT profiles of 25 vs. 2, -5, -15°C, for qRT-PCR ([S1 Table](#)). These genes included cold and freezing-associated genes such as those encoding ZAT12 [50], RAV [9], zinc finger [51], calcium binding protein, CAMK\_KIN1 [52], AP2-like, DREB and UDP-forming [9]. All 25 genes were confirmed by qRT-PCR analysis to be differentially regulated by the respective temperature ([Fig 6](#)), although the fold changes were different from the transcriptome sequencing data. Based on annotation, information on gene function was available for all but two of these genes. The genes significantly induced by cold and freezing stress in *A. cepa* encoded products including 4 zinc finger proteins, 4 hypothetical proteins, UDP forming, CBL-interacting protein kinase, 2 heat shock protein, 2 CMLH and subset of transcription factors, 2 bHLH, AP2 domain containing protein, DREB-2, MYB2, and bZIP. During cold and freezing conditions, drastic changes occur in the transcript levels of genes that encode transcriptional factors, particularly those that function as activators, major TFs involved are DREB, MYB, CBL, bZIP, ZAT, HSPs, bHLH [7, 53, 54]. Apart from these TFs, other genes encode proteins such as chitinase and heat shock proteins that interacts with the anti-freeze related proteins, thus preventing oxidative stress and ice formation [55, 56]. Base on previous reports, the cold-inducible genes were analyzed using qRT-PCR and found to be expressed similarly as in transcriptome data with minor variations. During cold stress (-2°C), upon signal perception by plasma membrane, the signal is transduced into the nucleus by cascades of Ca<sup>2+</sup> dependent kinases such as CIPKs (CBL interacting protein kinase) [57, 58] and they activates transcription factors and other proteins that maintain cell wall integrity, such as UDP proteins [59]. Colder temperatures (-5°C



**Fig 6.** Validation of candidate DEGs during (A) cold, (B) cold & freezing, and (C) freezing stress using qRT-PCR. Error bars indicate standard deviation (SD). X-axis indicates relative expression to the log<sub>2</sub> fold change (log<sub>2</sub>FC), Y-axis indicates selected candidate genes of DEGs.

doi:10.1371/journal.pone.0161987.g006

and -15°C) induces other defense-related genes that are triggered by different types of TFs. For instance, CBF family TFs control the downstream signaling cascade during cold and freezing stress [60]. Further, membrane proteins such as chitinase prevents oxidative stress by fortification of the cell membrane during cold and freezing stress. These membrane binding proteins encode an antifreeze proteins which inhibit the formation of ice in wheat [55]. In *Chenopodium album*, plastid or chloroplast small heat shock proteins confer tolerance during cold stress [56]. In *Arabidopsis*, zinc finger transcription factors increase tolerance by interacting with CBF family proteins [50]. Similarly in *Brassica rapa*, analyses using two cold genotypes they suggest bZIP are involved in cold tolerance [61]. Overall, our expression analysis of the 25 identified genes suggest that these genes are involved during cold and freezing tolerance and may work together to protect *A. cepa* from cold or freezing damage.

**Table 3. Summary of SSRs identified in the assembled genes from transcriptome.**

S. No	Repeats	Min-max length	Most common motif	Genes
1.	Dinucleotide (AA $\geq$ 6)	12–24	AT (447), GT (378)	1,465
2.	Tri nucleotide (AAA $\geq$ 4)	12–24	AAG (305)	3,253
3.	Tetra nucleotide (AAAA $\geq$ 4)	16–24	ATT (63)	370
4.	Penta nucleotide (AAAAA $\geq$ 4)	20–25	AAAAT (12)	59
5.	Hexa nucleotide (AAAAAA $\geq$ 4)	24	AAAAAG, AAAACC, AACCTC (2)	33

doi:10.1371/journal.pone.0161987.t003

### In silico predictions of molecular markers

To enhance the quality of the sequence data assembly and to identify new molecular markers based on SSRs (simple sequence repeats) and SNPs (single nucleotide polymorphisms) in onion, we screened the transcriptome gene dataset for SSRs using Msatcommander software with default parameters. A search for di-, tri-, tetra-, penta-, and hexa- nucleotide repeats yielded 1,258, 2,800, 301, 51 and 27, respectively; mononucleotides are excluded. A total of 5,180 SSRs, with 4,437 potentially unique SSRs, were represented in approximately 85.6% of the genes ([S6 Fig](#)). The minimum and maximum length of the repeats was 12 and 25 respectively, and the average length was 17 nucleotides. The most common motifs are listed in [Table 3](#). A complete list of SSRs and corresponding information are provided in [S5 Table](#). We also mined potential SNPs in CS and CT lines (25°C) using default parameters of software GATK. A massive collection of CS and CT SNPs based on the 25°C libraries included 29,106 and 27,685 genes, respectively, 19,588 were common genes were found with SNPs. Within the detected SNPs, the 63 and 62% of transition (A/G, C/T) type were found in CS and CT lines, respectively. Transversion (A/C, A/T, G/C, G/T) SNP ratios in CS and CT lines (36 and 37%, respectively) can be used to measure the genetic distance between the lines (see details in [S6 Table](#)). Transition polymorphism was more frequent than transversion, consistent with the nature of these changes [62].

## Conclusions

In this study, we used high-throughput sequencing data to characterize the transcriptomes of contrasting genotypes in *Allium cepa* (bulb onion), a species for which little genomic information available. Comparative analysis of expression profiles revealed a large number of cold and freezing responsive differentially expressed genes. Pathway analysis of the 491 enriched DEGs, revealed major metabolism involved during freezing conditions. Re-validation with qRT- PCR demonstrated a high reliability (99%) of candidate cold and freezing-related genes. This study provides a platform for further functional genomic research on *Allium* spp., including genomic resources such as molecular markers (SSRs and SNPs) for cold and freezing stress responsiveness and will support molecular breeding and gene cloning in *Allium* spp.

## Supporting Information

**S1 Fig. Electrolyte leakage analysis of contrasting genotypes in *Allium cepa* L.** Means followed by dissimilar letters are significantly different based on LSD test at P = 0.05 level.  
(TIFF)

**S2 Fig. Species-wise classification based on the nr (BLASTX) hits of the genes identified in the de novo assembled genes.**  
(TIFF)

**S3 Fig. Venn diagram comparison of CS and CT lines showing, DEGs up- and down-regulated in the respective treatments.**  
(TIFF)

**S4 Fig. Volcano plots of CT and CS lines at 25 vs -5°C.** Red and blue dots are indicate as up- and down regulated genes respectively. B. Venn diagram of DEG in CT and CS line at 25 vs -5°C were unique and commonly expressed.  
(TIFF)

**S5 Fig. GO classification of DEGs in CS vs. CS and CT vs. CT.** Histogram of the GO annotation and enrichment analysis was generated by the BGI WEGO software. The genes were grouped into three main GO categories: Cellular component (CC), Molecular function (MF),

Biological Process (BP). The right Y- axis indicates the number of genes in a category. The left Y-axis indicates the percentage in a specific category. One gene could be assigned with more than one GO term. Color indicates GO enrichment (significantly enriched GO terms with Bonferroni corrected p-value <0.05) and GO terms in all genes.  
(TIFF)

**S6 Fig. Venn diagram of SSR motifs (di, tri, tetra, penta and hexa nucleotide) were predicted by *in silico* analysis.**

(TIFF)

**S1 Table. Validation of candidate genes with qPCR analysis using CT libraries of cold and freezing stresses.**

(XLSX)

**S2 Table. Classification of pathways were mapped under KEGG database using assembled genes of transcriptome.**

(XLSX)

**S3 Table. The ID's and FPKM of each gene exhibiting difference among all the libraries.**

(XLSX)

**S4 Table. Combined genes expressed at freezing temperature in CT and CS lines.**

(XLSX)

**S5 Table. Complete list of SSRs marker from assembled genes of transcriptomes.**

(XLSX)

**S6 Table. The number of SNP transition and transversion observed in *A. cepa* L. transcriptomes.**

(XLSX)

## Acknowledgments

This research was supported by Golden Seed Project (Center for Horticultural Seed Development, No. 213003-04-4-SB110), Ministry of Agriculture, Food and Rural Affairs (MAFRA), Ministry of Oceans and Fisheries (MOF), Rural Development Administration (RDA) and Korea Forest Service (KFS).

## Author Contributions

**Conceptualization:** ISN JH SKT.

**Data curation:** SKT SN.

**Supervision:** ISN.

**Validation:** JH MYC.

**Writing – original draft:** SKT.

**Writing – review & editing:** SKT SN JIP.

## References

1. Guy CL. Cold-Acclimation and Freezing Stress Tolerance—Role of Protein-Metabolism. Annual Review of Plant Physiology and Plant Molecular Biology. 1990; 41:187–223. doi: [10.1146/annurev.pp.41.060190.001155](https://doi.org/10.1146/annurev.pp.41.060190.001155) PMID: [WOS:A1990DG17400007](https://pubmed.ncbi.nlm.nih.gov/1690007/).

2. Thomashow MF. Plant Cold Acclimation: Freezing Tolerance Genes and Regulatory Mechanisms. *Annu Rev Plant Physiol Plant Mol Biol*. 1999; 50:571–99. Epub 2004/03/12. doi: [10.1146/annurev.arplant.50.1.571](https://doi.org/10.1146/annurev.arplant.50.1.571) PMID: [15012220](#).
3. Shanker AK, Maheswari M, Yadav SK, Desai S, Bhanu D, Attal NB, et al. Drought stress responses in crops. *Funct Integr Genomics*. 2014; 14(1):11–22. Epub 2014/01/11. doi: [10.1007/s10142-013-0356-x](https://doi.org/10.1007/s10142-013-0356-x) PMID: [24408129](#).
4. Matsui A, Ishida J, Morosawa T, Mochizuki Y, Kaminuma E, Endo TA, et al. Arabidopsis transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. *Plant Cell Physiol*. 2008; 49(8):1135–49. Epub 2008/07/16. doi: [10.1093/pcp/pcn101](https://doi.org/10.1093/pcp/pcn101) PMID: [18625610](#).
5. Sanghera GS, Wani SH, Hussain W, Singh NB. Engineering cold stress tolerance in crop plants. *Curr Genomics*. 2011; 12(1):30–43. Epub 2011/09/03. doi: [10.2174/138920211794520178CG-12-30](https://doi.org/10.2174/138920211794520178CG-12-30) PMID: [21886453](#); PubMed Central PMCID: [PMC3129041](#).
6. Carvalho MA, Pino MT, Jeknic Z, Zou C, Doherty CJ, Shiu SH, et al. A comparison of the low temperature transcriptomes and CBF regulons of three plant species that differ in freezing tolerance: Solanum commersonii, Solanum tuberosum, and Arabidopsis thaliana. *J Exp Bot*. 2011; 62(11):3807–19. Epub 2011/04/23. doi: [10.1093/jxb/err066](https://doi.org/10.1093/jxb/err066). PMID: [21511909](#); PubMed Central PMCID: [PMC3134341](#).
7. Chinnusamy V, Zhu J, Zhu J-K. Cold stress regulation of gene expression in plants. *Trends in Plant Science*. 2007; 12(10):444–51. <http://dx.doi.org/10.1016/j.tplants.2007.07.002>. PMID: [17855156](#)
8. Medina J, Catala R, Salinas J. The CBFs: three arabidopsis transcription factors to cold acclimate. *Plant Sci*. 2011; 180(1):3–11. Epub 2011/03/23. doi: [10.1016/j.plantsci.2010.06.019](https://doi.org/10.1016/j.plantsci.2010.06.019) PMID: [21421341](#).
9. Thamilarasan SK, Park JI, Jung HJ, Nou IS. Genome-wide analysis of the distribution of AP2/ERF transcription factors reveals duplication and CBFs genes elucidate their potential function in Brassica oleracea. *BMC Genomics*. 2014; 15:422. Epub 2014/06/04. doi: [10.1186/1471-2164-15-422](https://doi.org/10.1186/1471-2164-15-422) PMID: [24888752](#); PubMed Central PMCID: [PMC4229850](#).
10. Thomashow MF. Molecular basis of plant cold acclimation: insights gained from studying the CBF cold response pathway. *Plant Physiol*. 2010; 154(2):571–7. Epub 2010/10/06. doi: [10.1104/pp.110.161794](https://doi.org/10.1104/pp.110.161794) PMID: [20921187](#); PubMed Central PMCID: [PMC2948992](#).
11. Zhou MQ, Shen C, Wu LH, Tang KX, Lin J. CBF-dependent signaling pathway: a key responder to low temperature stress in plants. *Crit Rev Biotechnol*. 2011; 31(2):186–92. Epub 2010/10/06. doi: [10.3109/07388551.2010.505910](https://doi.org/10.3109/07388551.2010.505910) PMID: [20919819](#).
12. Brewster JL. Onions and Other Vegetable Alliums. 2nd edition ed: CABI, Oxfordshire, UK 2008. 1–26 p.
13. Bahorun T, Luximon-Ramma A, Crozier A, Aruoma OI. Total phenol, flavonoid, proanthocyanidin and vitamin C levels and antioxidant activities of Mauritian vegetables. *Journal of the Science of Food and Agriculture*. 2004; 84(12):1553–61. doi: [10.1002/jsfa.1820](https://doi.org/10.1002/jsfa.1820)
14. Slimestad R, Fossen T, Vägen IM. Onions: A Source of Unique Dietary Flavonoids. *Journal of Agricultural and Food Chemistry*. 2007; 55(25):10067–80. doi: [10.1021/jf0712503](https://doi.org/10.1021/jf0712503) PMID: [17997520](#)
15. Yang SS, Tu ZJ, Cheung F, Xu WW, Lamb JF, Jung HJ, et al. Using RNA-Seq for gene identification, polymorphism detection and transcript profiling in two alfalfa genotypes with divergent cell wall composition in stems. *BMC Genomics*. 2011; 12:199. Epub 2011/04/21. doi: [10.1186/1471-2164-12-199](https://doi.org/10.1186/1471-2164-12-199) PMID: [21504589](#); PubMed Central PMCID: [PMC3112146](#).
16. Duangjit J, Bohanec B, Chan AP, Town CD, Havey MJ. Transcriptome sequencing to produce SNP-based genetic maps of onion. *Theor Appl Genet*. 2013; 126(8):2093–101. Epub 2013/05/22. doi: [10.1007/s00122-013-2121-x](https://doi.org/10.1007/s00122-013-2121-x) PMID: [23689743](#).
17. McCallum J, Clarke A, Pither-Joyce M, Shaw M, Butler R, Brash D, et al. Genetic mapping of a major gene affecting onion bulb fructan content. *Theoretical and Applied Genetics*. 2006; 112(5):958–67. doi: [10.1007/s00122-005-0199-5](https://doi.org/10.1007/s00122-005-0199-5) PMID: [16404585](#)
18. Hegedus Z, Zakrzewska A, Agoston VC, Ordas A, Racz P, Mink M, et al. Deep sequencing of the zebrafish transcriptome response to mycobacterium infection. *Mol Immunol*. 2009; 46(15):2918–30. Epub 2009/07/28. doi: [10.1016/j.molimm.2009.07.002](https://doi.org/10.1016/j.molimm.2009.07.002) S0161-5890(09)00592-6. PMID: [19631987](#).
19. Tian DQ, Pan XY, Yu YM, Wang WY, Zhang F, Ge YY, et al. De novo characterization of the Anthurium transcriptome and analysis of its digital gene expression under cold stress. *BMC Genomics*. 2013; 14:827. Epub 2013/11/26. doi: [10.1186/1471-2164-14-827](https://doi.org/10.1186/1471-2164-14-827) PMID: [24267953](#); PubMed Central PMCID: [PMC4046746](#).
20. Yates SA, Swain MT, Hegarty MJ, Chernukin I, Lowe M, Allison GG, et al. De novo assembly of red clover transcriptome based on RNA-Seq data provides insight into drought response, gene discovery and marker identification. *BMC Genomics*. 2014; 15:453. Epub 2014/06/11. doi: [10.1186/1471-2164-15-453](https://doi.org/10.1186/1471-2164-15-453) PMID: [24912738](#); PubMed Central PMCID: [PMC4144119](#).

21. Yu S, Zhang F, Yu Y, Zhang D, Zhao X, Wang W. Transcriptome Profiling of Dehydration Stress in the Chinese Cabbage (*Brassica rapa* L. ssp. *pekinensis*) by Tag Sequencing. *Plant Mol Biol Rep.* 2012; 30(1):17–28. doi: [10.1007/s11105-011-0313-6](https://doi.org/10.1007/s11105-011-0313-6)
22. Winfield MO, Lu C, Wilson ID, Coghill JA, Edwards KJ. Plant responses to cold: Transcriptome analysis of wheat. *Plant Biotechnol J.* 2010; 8(7):749–71. Epub 2010/06/22. doi: [10.1111/j.1467-7652.2010.00536.x](https://doi.org/10.1111/j.1467-7652.2010.00536.x) PMID: [20561247](#).
23. Li D, Deng Z, Qin B, Liu X, Men Z. De novo assembly and characterization of bark transcriptome using Illumina sequencing and development of EST-SSR markers in rubber tree (*Hevea brasiliensis* Muell. Arg.). *BMC Genomics.* 2012; 13:192. Epub 2012/05/23. doi: [10.1186/1471-2164-13-192](https://doi.org/10.1186/1471-2164-13-192) PMID: [22607098](#); PubMed Central PMCID: [PMC3431226](#).
24. Zhao Y, Williams R, Prakash CS, He G. Identification and characterization of gene-based SSR markers in date palm (*Phoenix dactylifera* L.). *BMC Plant Biol.* 2012; 12:237. Epub 2012/12/18. doi: [10.1186/1471-2229-12-237](https://doi.org/10.1186/1471-2229-12-237) PMID: [23241238](#); PubMed Central PMCID: [PMC3568718](#).
25. Rapacz M. Regulation of frost resistance during cold de-acclimation and re-acclimation in oilseed rape. A possible role of PSII redox state. *Physiol Plant.* 2002; 115(2):236–43. Epub 2002/06/13. ppl1150209. PMID: [12060241](#).
26. Zhang L, Zhao G, Jia J, Liu X, Kong X. Molecular characterization of 60 isolated wheat MYB genes and analysis of their expression during abiotic stress. *J Exp Bot.* 2012; 63(1):203–14. Epub 2011/09/22. doi: [10.1093/jxb/err264](https://doi.org/10.1093/jxb/err264) PMID: [21934119](#); PubMed Central PMCID: [PMC3245462](#).
27. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 2011; 29(7):644–52. Epub 2011/05/17. doi: [10.1038/nbt.1883](https://doi.org/10.1038/nbt.1883) PMID: [21572440](#); PubMed Central PMCID: [PMC3571712](#).
28. Arasan SK, Park JI, Ahmed NU, Jung HJ, Lee IH, Cho YG, et al. Gene ontology based characterization of expressed sequence tags (ESTs) of *Brassica rapa* cv. Osome. *Indian J Exp Biol.* 2013; 51(7):522–30. Epub 2013/08/01. PMID: [23898551](#).
29. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics.* 2005; 21(18):3674–6. Epub 2005/08/06. doi: [10.1093/bioinformatics/bti610](https://doi.org/10.1093/bioinformatics/bti610) PMID: [16081474](#).
30. Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, et al. The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* 2001; 29(1):22–8. Epub 2000/01/11. PMID: [11125040](#); PubMed Central PMCID: [PMC29819](#).
31. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 2012; 40(Database issue):D109–14. Epub 2011/11/15. doi: [10.1093/nar/gkr988](https://doi.org/10.1093/nar/gkr988) PMID: [22080510](#); PubMed Central PMCID: [PMC3245020](#).
32. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* 2011; 12:323. Epub 2011/08/06. doi: [10.1186/1471-2105-12-323](https://doi.org/10.1186/1471-2105-12-323) PMID: [21816040](#); PubMed Central PMCID: [PMC3163565](#).
33. Harwood. Vector NTI. *Biotechnol Softw J.* 1996; 13(5):22–30.
34. Faircloth BC. Msatcommander: detection of microsatellite repeat arrays and automated, locus specific primer design. *Molecular Ecology Resources.* 2008; 8(1):92–4. doi: [10.1111/j.1471-8286.2007.01884.x](https://doi.org/10.1111/j.1471-8286.2007.01884.x) PMID: [21585724](#).
35. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research.* 2010; 20(9):1297–303. doi: [10.1101/gr.107524.110](https://doi.org/10.1101/gr.107524.110) PMID: [20644199](#).
36. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* 2012; 40(15):e115. Epub 2012/06/26. doi: [10.1093/nar/gks596](https://doi.org/10.1093/nar/gks596) PMID: [22730293](#); PubMed Central PMCID: [PMC3424584](#).
37. Manoharan R, Han J, Vijayakumar H, Subramani B, Thamilarasan S, Park J-I, et al. Molecular and Functional Characterization of FLOWERING LOCUS T Homologs in *Allium cepa*. *Molecules.* 2016; 21(2):217. doi: [10.3390/molecules21020217](https://doi.org/10.3390/molecules21020217).
38. Takagi T, Walker AK, Sawa C, Diehn F, Takase Y, Blackwell TK, et al. The *Caenorhabditis elegans* mRNA 5'-capping enzyme. In vitro and in vivo characterization. *J Biol Chem.* 2003; 278(16):14174–84. Epub 2003/02/11. doi: [10.1074/jbc.M212101200](https://doi.org/10.1074/jbc.M212101200) PMID: [12576476](#).
39. Wallis J, Wang H, Guerra\* D. Expression of a synthetic antifreeze protein in potato reduces electrolyte release at freezing temperatures. *Plant Mol Biol.* 1997; 35(3):323–30. doi: [10.1023/a:1005886210159](https://doi.org/10.1023/a:1005886210159) PMID: [9349256](#)

40. Yu-Jin J, Yong Gu C, III Sup N, and Kwon Kyoo K. Transgenic Tomato Plants Ectopically Expressing BrRZFP1 Gene Encoding C3HC4-type RING Zinc Finger Protein. *Plant Breeding and Biotechnology*. 2014; 2(1):25–34.
41. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotech*. 2011; 29(7):644–52. doi: [10.1038/nbt.1883](https://doi.org/10.1038/nbt.1883)
42. Pang T, Ye CY, Xia X, Yin W. De novo sequencing and transcriptome analysis of the desert shrub, *Ammopiptanthus mongolicus*, during cold acclimation using Illumina/Solexa. *BMC Genomics*. 2013; 14:488. Epub 2013/07/20. doi: [10.1186/1471-2164-14-488](https://doi.org/10.1186/1471-2164-14-488) PMID: [23865740](https://pubmed.ncbi.nlm.nih.gov/23865740/); PubMed Central PMCID: [PMC3728141](https://pubmed.ncbi.nlm.nih.gov/PMC3728141/).
43. Novaes E, Drost DR, Farmerie WG, Pappas GJ Jr., Grattapaglia D, Sederoff RR, et al. High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics*. 2008; 9:312. Epub 2008/07/02. doi: [10.1186/1471-2164-9-312](https://doi.org/10.1186/1471-2164-9-312) PMID: [18590545](https://pubmed.ncbi.nlm.nih.gov/18590545/); PubMed Central PMCID: [PMC2483731](https://pubmed.ncbi.nlm.nih.gov/PMC2483731/).
44. Letunic I, Yamada T, Kanehisa M, Bork P. iPath: interactive exploration of biochemical pathways and networks. *Trends Biochem Sci*. 2008; 33(3):101–3. Epub 2008/02/16. doi: [10.1016/j.tibs.2008.01.001](https://doi.org/10.1016/j.tibs.2008.01.001) PMID: [18276143](https://pubmed.ncbi.nlm.nih.gov/18276143/).
45. Cantacessi C, Jex AR, Hall RS, Young ND, Campbell BE, Joachim A, et al. A practical, bioinformatic workflow system for large data sets generated by next-generation sequencing. *Nucleic Acids Res*. 2010; 38(17):e171. Epub 2010/08/05. doi: [10.1093/nar/gkq667](https://doi.org/10.1093/nar/gkq667) PMID: [20682560](https://pubmed.ncbi.nlm.nih.gov/20682560/); PubMed Central PMCID: [PMC2943614](https://pubmed.ncbi.nlm.nih.gov/PMC2943614/).
46. Feng C, Chen M, Xu CJ, Bai L, Yin XR, Li X, et al. Transcriptomic analysis of Chinese bayberry (*Myrica rubra*) fruit development and ripening using RNA-Seq. *BMC Genomics*. 2012; 13:19. Epub 2012/01/17. doi: [10.1186/1471-2164-13-19](https://doi.org/10.1186/1471-2164-13-19) PMID: [22244270](https://pubmed.ncbi.nlm.nih.gov/22244270/); PubMed Central PMCID: [PMC3398333](https://pubmed.ncbi.nlm.nih.gov/PMC3398333/).
47. Bogdanovic J, Mojovic M, Milosavic N, Mitrovic A, Vucinic Z, Spasojevic I. Role of fructose in the adaptation of plants to cold-induced oxidative stress. *Eur Biophys J*. 2008; 37(7):1241–6. Epub 2008/01/25. doi: [10.1007/s00249-008-0260-9](https://doi.org/10.1007/s00249-008-0260-9) PMID: [18214465](https://pubmed.ncbi.nlm.nih.gov/18214465/).
48. Huner NPA, Öquist G, Sarhan F. Energy balance and acclimation to light and cold. *Trends in Plant Science*. 3(6):224–30. doi: [10.1016/s1360-1385\(98\)01248-5](https://doi.org/10.1016/s1360-1385(98)01248-5).
49. Tian D-Q, Pan X-Y, Yu Y-M, Wang W-Y, Zhang F, Ge Y-Y, et al. De novo characterization of the Anthurium transcriptome and analysis of its digital gene expression under cold stress. *BMC Genomics*. 2013; 14(1):1–14. doi: [10.1186/1471-2164-14-827](https://doi.org/10.1186/1471-2164-14-827).
50. Vogel JT, Zarka DG, Van Buskirk HA, Fowler SG, Thomashow MF. Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of *Arabidopsis*. *The Plant Journal*. 2005; 41(2):195–211. doi: [10.1111/j.1365-313X.2004.02288.x](https://doi.org/10.1111/j.1365-313X.2004.02288.x) PMID: [15634197](https://pubmed.ncbi.nlm.nih.gov/15634197/).
51. Shinozaki K, Yamaguchi-Shinozaki K, Seki M. Regulatory network of gene expression in the drought and cold stress responses. *Current Opinion in Plant Biology*. 2003; 6(5):410–7. doi: [10.1016/S1369-5266\(03\)00092-X](https://doi.org/10.1016/S1369-5266(03)00092-X) PMID: [12972040](https://pubmed.ncbi.nlm.nih.gov/12972040/).
52. Wang H, Zou Z, Wang S, Gong M. Global Analysis of Transcriptome Responses and Gene Expression Profiles to Cold Stress of *Jatropha curcas* L. *PLoS One*. 2013; 8(12):e82817. doi: [10.1371/journal.pone.0082817](https://doi.org/10.1371/journal.pone.0082817) PMID: [PMC3857291](https://pubmed.ncbi.nlm.nih.gov/PMC3857291/).
53. Yamaguchi-Shinozaki K, Shinozaki K. Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu Rev Plant Biol*. 2006; 57:781–803. Epub 2006/05/04. doi: [10.1146/annurev.arplant.57.032905.105444](https://doi.org/10.1146/annurev.arplant.57.032905.105444) PMID: [16669782](https://pubmed.ncbi.nlm.nih.gov/16669782/).
54. Chinnusamy V, Zhu JK, Sunkar R. Gene regulation during cold stress acclimation in plants. *Methods Mol Biol*. 2010; 639:39–55. Epub 2010/04/14. doi: [10.1007/978-1-60761-702-0\\_3](https://doi.org/10.1007/978-1-60761-702-0_3) PMID: [20387039](https://pubmed.ncbi.nlm.nih.gov/20387039/); PubMed Central PMCID: [PMC3064467](https://pubmed.ncbi.nlm.nih.gov/PMC3064467/).
55. Yeh S, Moffatt BA, Griffith M, Xiong F, Yang DS, Wiseman SB, et al. Chitinase genes responsive to cold encode antifreeze proteins in winter cereals. *Plant Physiol*. 2000; 124(3):1251–64. Epub 2000/11/18. PMID: [11080301](https://pubmed.ncbi.nlm.nih.gov/11080301/); PubMed Central PMCID: [PMC59223](https://pubmed.ncbi.nlm.nih.gov/PMC59223/).
56. Haq NU, Ammar M, Bano A, Luthe DS, Heckathorn SA, Shakeel SN. Molecular Characterization of *Chenopodium album* Chloroplast Small Heat Shock Protein and Its Expression in Response to Different Abiotic Stresses. *Plant Mol Biol Rep*. 2013; 31(6):1230–41. doi: [10.1007/s11105-013-0588-x](https://doi.org/10.1007/s11105-013-0588-x)
57. Kolukisaoglu U, Weinl S, Blazevic D, Batistic O, Kudla J. Calcium sensors and their interacting protein kinases: genomics of the *Arabidopsis* and rice CBL-CIPK signaling networks. *Plant Physiol*. 2004; 134(1):43–58. Epub 2004/01/20. doi: [10.1104/pp.103.033068](https://doi.org/10.1104/pp.103.033068) PMID: [14730064](https://pubmed.ncbi.nlm.nih.gov/14730064/); PubMed Central PMCID: [PMC316286](https://pubmed.ncbi.nlm.nih.gov/PMC316286/).
58. Li R, Zhang J, Wei J, Wang H, Wang Y, Ma R. Functions and mechanisms of the CBL–CIPK signaling system in plant response to abiotic stress. *Progress in Natural Science*. 2009; 19(6):667–76. doi: [10.1016/j.pnsc.2008.06.030](https://doi.org/10.1016/j.pnsc.2008.06.030)

59. Ciereszko I, Johansson H, Kleczkowski LA. Sucrose and light regulation of a cold-inducible UDP-glucose pyrophosphorylase gene via a hexokinase-independent and abscisic acid-insensitive pathway in *Arabidopsis*. *Biochemical Journal*. 2001; 354(Pt 1):67–72. PMID: [PMC1221629](#).
60. Agarwal M, Hao Y, Kapoor A, Dong CH, Fujii H, Zheng X, et al. A R2R3 type MYB transcription factor is involved in the cold regulation of CBF genes and in acquired freezing tolerance. *J Biol Chem*. 2006; 281(49):37636–45. Epub 2006/10/04. doi: [10.1074/jbc.M605895200](#) PMID: [17015446](#).
61. Hwang I, Jung HJ, Park JI, Yang TJ, Nou IS. Transcriptome analysis of newly classified bZIP transcription factors of *Brassica rapa* in cold stress response. *Genomics*. 2014; 104(3):194–202. Epub 2014/07/31. doi: [10.1016/j.ygeno.2014.07.008](#) PMID: [25075938](#).
62. Guan X, Nah G, Song Q, Udall JA, Stelly DM, Chen ZJ. Transcriptome analysis of extant cotton progenitors revealed tetraploidization and identified genome-specific single nucleotide polymorphism in diploid and allotetraploid cotton. *BMC Research Notes*. 2014; 7(1):1–10. doi: [10.1186/1756-0500-7-493](#)