De novo Assembly and Transcriptome Analysis of Cultivated Strawberry (Fragaria × Ananassa) Receptacles under Freezing Stress

Xiaodong Chen¹, Mizhen Zhao¹*, Fuhua Pang¹ and Jing Wang¹
¹Institute of Pomology, Jiangsu Academy of Agricultural Sciences/Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement, Jiangsu, P.R. China
*For correspondence: njzhaomz@163.com

Abstract

The octoploid strawberry (Fragaria × ananassa) is an important species cultivated and consumed around the world for its excellent flavour and nutritional value. Strawberry floral organs are very sensitive to low-temperature stress, which often leads to economic loss. In the current study RNA-Seq method was used to understand the strawberry receptacle response to freezing stress. The results indicated that out of 356,807,838 clean reads obtained from six libraries, 42,365 unigenes were generated through assembling them in a de novo method. These unigenes were further annotated in Pfam, KEGG, Swissprot, String, and NR database. A total of 528 differentially expressed genes were identified in our research. Among these genes, 428 genes were upregulated and 100 genes were downregulated by freezing stress. The GO classification and KEGG enrichment analysis showed that the most highly represented pathways were environmental adaptation and carbohydrate metabolism. Moreover, we identified 3 CBF/DREB transcription factor genes and 1 COR gene in upregulated DEGs. This analysis reveals the global transcriptomic response of strawberry receptacles to freezing stress and may help to understand the molecular mechanisms involved in freezing stress tolerance in strawberry receptacles. © 2018 Friends Science Publishers

Keywords: Freezing stress; Low temperatures; Receptacle tissue; Strawberry (fragaria × ananassa); Transcriptome

Introduction

The cultivated strawberry is an important horticultural crop for the production of fresh fruit and juice (Kang et al., 2013; Chen et al., 2016). In the current growing areas, strawberries usually blossom during the winter and early spring. Low temperatures are common during these times (Ariza et al., 2015) and can injure the strawberry flowers (Ki and Warmund, 1992; Ariza et al., 2015). Fruit quantity and weight are substantially reduced when flowers are injured.

Low-temperature stress, which includes chilling (0-15°C) and freezing temperatures (≤0°C), has a big impact on the growth, development, productivity and distribution of plants (Chinnusamy et al., 2007; Zhan et al., 2016). Responding to low-temperature stress, plants exhibit many physiological and biochemical changes, mainly through the C-repeat (CRT)-binding factor/dehydration-responsive element (DRE) binding factor (CBF/DREB)-mediated signal transduction pathway (Chinnusamy et al., 2007; Shi et al., 2015; Zhan et al., 2016). CBF1 (DREB1b), CBF2 (DREB1c), CBF3 (DREB1a) are rapidly induced by low temperatures (Fowler and Thomasaw, 2002; Medina et al., 2011). These CBFs can activate the expression of many cold responsive (COR) genes, such as COR48, COR47, COR15 and KIN1, to protect plants against low-temperature stress (Chinnusamy et al., 2003; Shi et al., 2015; Zhan et al., 2016). Many biological processes are regulated by CBFs, such as ROS detoxification, phosphoinositide metabolism, membrane transport, hormone metabolism and the production of cryoprotective molecules (Chinnusamy et al., 2007; Koehler et al., 2012). The CBF signalling pathway has been well studied in model plant (Arabidopsis) and has been reported to have a function in many crops (Chinnusamy et al., 2007; Koehler et al., 2015). F. × ananassa CBF1 (FaCBF1) was induced by chilling stress in leaf tissue, but not in receptacle tissue. In addition, overexpression of FaCBF1 enhanced the freezing tolerance in leaves in F. × ananassa ‘Honeoye’, but did not improve the tolerance in receptacles (Owens et al., 2002).

Several studies have identified genes involved in the cold response in strawberry leaves (NDong et al., 1997; Koehler et al., 2015). Furthermore, potential protein markers for cold-tolerant octoploid strawberries have been identified using proteomics (Koehler et al., 2012). Recently, integrative “omics” approaches were used to study low-temperature responses in strawberry leaf and root tissue (Koehler et al., 2015). However, it is still unclear how strawberry receptacles respond to freezing temperatures as well as if other CBF genes are involved in freezing stress tolerance.

To cite this paper: Chen, X., M. Zhao, F. Pang and J. Wang, 201x. De novo assembly and transcriptome analysis of cultivated strawberry (fragaria × ananassa) receptacles under freezing stress. Int. J. Agric. Biol. 00: 000-000
In this work, we conducted a de novo transcriptome assembly of *F. × ananassa* and analyzed the differentially expressed genes between controls and freezing treated samples to study how strawberry receptacles respond to early freezing stress. Among our transcriptome data, 428 unigenes and 100 unigenes were found to be upregulated or down regulated by freezing treatment. We summarized strawberry receptacles’ major transcriptomic changes in response to early freezing stress, which included alterations in signal transduction and metabolic pathways. We also identified several CBF/DREB genes, which might be involved in regulating low temperature response in strawberry receptacle tissue.

**Materials and Methods**

**Plant Materials and Freezing Treatment**

Rooted runners of ‘Honeoye’ *F. × ananassa* were grown in plastic pots (10-cm) at 18-22°C under natural light and day length conditions (32:04 N, 118:78 E; from 2016.09-2016.12). Plants with open flowers (Fig. 1) were selected for freezing treatment (-2°C, 6 h) in darkness. After freezing treatment, receptacles with attached pistils were collected for RNA-seq sequencing. Control samples were collected before transferring to the cold chamber. Six strawberry receptacles were harvested from three plants per replicate in freezing-treated samples and control samples. All of the collected strawberry receptacles were frozen in liquid nitrogen and stored at -80°C immediately.

**RNA Sequencing**

Total RNA was extracted from receptacles of control and freezing-treated (-2°C for 6 h) plants using RNA prep Pure Plant Kit for Polysaccharides and Polyphenolics-rich (Tian Gen, Beijing, China). Genomic DNA was removed from total RNA with Thermo Scientific DNaseI following the manufacturer’s instructions. The purity and concentration of RNA were measured by Nano Drop 2000. The integrity of RNA was evaluated by gel electrophoresis. The mRNA was specifically separated from total RNA by using magnetic beads coupling with Oligo (dT) (Invitrogen). Then, these mRNAs were randomly fragmented to around 300 bp. The first-strand cDNA was obtained by using reverse transcriptase with random primers, the second-strand cDNA was subsequently synthesized, forming the stable double strand structure. The double-strand cDNA was end repaired by using End Repair Mix. Then a-tailed and adapters were added. The library was amplified by using PCR (15 cycles). Purpose fragments were separated by 2% certified low range ultra-agarose. Sequencing libraries were sequenced by 2×150 bp paired-end sequencing on an Illumina Hiseq 2500 platform.

**Sequencing Data Assembly and Functional Annotation**

Seq Prep was used to trim dirty reads (adapter sequences, reads without insertions, and reads with a length less than 20 bp) and low-quality reads (reads with more than 10% bases “N”). Then, the clean reads were assembled in a de novo way by using Trinity (Grabherr *et al*., 2011) with default parameters to generate singletons and contigs. The open reading frames (ORFs) of assembled unigenes were predicted by using the ORF prediction process offered by Trinity. The unigenes were predicted and annotated with Basic Local Alignment Search Tool (BLASTX) programs against non-redundant, String, Swissprot protein, and the Kyoto Encyclopedia of Genes and Genomes (KEGG) databases with an E-value of less than \(1 \times 10^{-5}\). Blast2GO software (Conesa *et al*., 2005) was used to perform GO annotations. Unigenes were classified assigned to biological process, molecular function and cellular component, and the number of assigned unigenes were calculated in each category. The assembled unigenes were further compared to the COG database to carry out functional annotation, classification, and protein evolution analysis.

**Identification and Functional Analysis of Differentially Expressed Genes (DEGs)**

Sequencing reads were mapped back to reference transcriptome and by using the software RSEM (Li and Dewey, 2011). The FPKM (Fragments Per Kilobase of exon model per Million mapped reads) values were calculated by normalizing the read counts using RSEM (Li and Dewey, 2011). EdgeR package (Robinson *et al*., 2010) were used to identify DEGs between controls and freezing-treated samples. Genes with |log2 fold change| ≥ 1 and FDR < 0.05 were considered as DEGs and were submitted to a functional analysis. The Goatools and KOBAS software (Xie *et al*., 2011) were used to analyze the GO enrichment and KEGG pathways.

**Gene Expression Analysis**

Total RNA samples were reverse-transcribed into cDNA using the BioTeke supermoIII RT Kit. The SYBR reaction kit (TakaRa) was used to perform the qRT-PCR experiment. And the experiment was carried out on Applied Biosystems 7500 Real-Time PCR System. The relative gene expression level was calculated against the GAPDH (Gu *et al*., 2016). The qRT-PCR primers used in this study are listed in Supplementary Table S1.

**Results**

**RNA Sequence Data Analysis**

Strawberry receptacle tissue is very sensitive to freezing stress, a short period of low temperature stress will lead to the damage of strawberry receptacle tissue (Fig. 1). To better understand the strawberry receptacle’s early response to freezing stress, RNA samples were collected from receptacles before (controls) and after freezing treatment (-2°C, 6 h), in triplicate.
RNA-Seq Analysis of Strawberry / Int. J. Agric. Biol., Vol. 00, No. 0, 201x

Table S1: List of primers used in this study

<table>
<thead>
<tr>
<th>Primers</th>
<th>Primer Sequence 5'-3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>c41210_g1-QRT-F2</td>
<td>GGAGTACGTCGACAGGCTGTGCTG</td>
</tr>
<tr>
<td>c41210_g1-QRT-R2</td>
<td>CAGCTGGCAAGACGGATTCA</td>
</tr>
<tr>
<td>c12042_g1-QRT-F1</td>
<td>CGATTGCTAGGTGATGTAATG</td>
</tr>
<tr>
<td>c12042_g1-QRT-R1</td>
<td>GATGTTGACAGATGTAATG</td>
</tr>
<tr>
<td>c43922_g2-QRT-F1</td>
<td>GATGTTGACAGATGTAATG</td>
</tr>
<tr>
<td>c43922_g2-QRT-R1</td>
<td>GATGTTGACAGATGTAATG</td>
</tr>
<tr>
<td>c63129_QRT-F1</td>
<td>CTCAGGCTACATTCTCCTG</td>
</tr>
<tr>
<td>c63129_QRT-R1</td>
<td>CTGAGCTGAGACACCCGACG</td>
</tr>
<tr>
<td>c28135_QRT-F1</td>
<td>GGTGATTGTAGCAGGCTGTG</td>
</tr>
<tr>
<td>c28135_QRT-R1</td>
<td>CTCATTCCGATAACATG</td>
</tr>
<tr>
<td>ACTIN-QRT-F2</td>
<td>GGTTCAAGGTCTGGTATTGCTG</td>
</tr>
<tr>
<td>ACTIN-QRT-R2</td>
<td>GGAGTAGCTCCAGAGGCTGTG</td>
</tr>
</tbody>
</table>

Table S2: Statistics of raw sequencing data

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Total raw reads</th>
<th>Total raw bases</th>
<th>GC%</th>
<th>Q30%</th>
</tr>
</thead>
<tbody>
<tr>
<td>LD_1</td>
<td>64152698</td>
<td>9686966798</td>
<td>46.74</td>
<td>91.37</td>
</tr>
<tr>
<td>LD_2</td>
<td>59049929</td>
<td>905238392</td>
<td>46.96</td>
<td>90.52</td>
</tr>
<tr>
<td>LD_3</td>
<td>60368504</td>
<td>9115644104</td>
<td>47.13</td>
<td>90.61</td>
</tr>
<tr>
<td>BCS_1</td>
<td>63268188</td>
<td>9553496388</td>
<td>47.02</td>
<td>90.58</td>
</tr>
<tr>
<td>BCS_2</td>
<td>60240240</td>
<td>9096276240</td>
<td>47.66</td>
<td>90.94</td>
</tr>
<tr>
<td>BCS_3</td>
<td>56820888</td>
<td>8579893688</td>
<td>48.62</td>
<td>91.01</td>
</tr>
</tbody>
</table>

BSC_1, BCS_2 and BCS_3 represent three replications for control samples. LD_1, LD_2 and LD_3 represent three replications for cold-treated samples.

Table S3: Statistics of clean sequencing data

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Total clean reads</th>
<th>Clean reads ratio (%)</th>
<th>Total clean bases GC%</th>
<th>Q30%</th>
</tr>
</thead>
<tbody>
<tr>
<td>LD_1</td>
<td>62904736</td>
<td>98.04</td>
<td>9277423438</td>
<td>46.59</td>
</tr>
<tr>
<td>LD_2</td>
<td>58531860</td>
<td>95.12</td>
<td>8610411785</td>
<td>46.79</td>
</tr>
<tr>
<td>LD_3</td>
<td>59002880</td>
<td>95.34</td>
<td>8690854344</td>
<td>46.96</td>
</tr>
<tr>
<td>BCS_1</td>
<td>61844334</td>
<td>97.75</td>
<td>9103752976</td>
<td>46.85</td>
</tr>
<tr>
<td>BCS_2</td>
<td>58933164</td>
<td>97.83</td>
<td>8684021911</td>
<td>46.66</td>
</tr>
<tr>
<td>BCS_3</td>
<td>55600864</td>
<td>97.85</td>
<td>8190893688</td>
<td>46.66</td>
</tr>
</tbody>
</table>

BSC_1, BCS_2 and BCS_3 represent three replications for control samples. LD_1, LD_2 and LD_3 represent three replications for cold-treated samples.

Fig. 1: A representative picture of receptacle tissue before (controls) and after freezing treatment.

We obtained 184,470,194 raw reads from the freezing-treated samples and 180,328,916 from the control samples, with an average GC content 46.94% and 46.87%, respectively. The average Q30 percentage (sequencing error rate < 0.1%) for the two RNA samples were 90.83% and 90.84%, respectively (Table 1 and Supplementary Table S2). These statistics qualified our RNA sequencing results. Then, we removed low-quality reads and obtained a total of 180,429,476 and 176,378,362 clean reads for freezing-treated and control samples, respectively. The two samples had similar mean clean lengths (147 nt) (Table 1 and Supplementary Table S3). Because the genome map of F. × ananassa was incomplete, the cleaned reads were de novo assembled with Trinity software. A total of 116,731 transcripts with a N50 of 1,939 were obtained, containing 42,365 unigenes with a N50 of 2,004. The mean length for transcripts and unigenes was 1,444.71 and 1,297.63, respectively. More than half of the unigenes were between 401 nt and 2000 nt. This is indicative of a high-quality assembly. The percentage of sequences annotated in Pfam, KEGG, Swissprot, String, and NR were 40.57%, 27.78%, 43.85%, 31.92% and 69.91%, respectively (Table 2). Most of the annotated unigenes were similar to gene sequences from the diploid woodland strawberry (Fragaria virginiana) (Fig. 2). Furthermore, we calculated correlations between all samples. The results showed a high correlation between 3 biological replicates among freezing-treated samples and controls (Fig. 3).

GO/COG/KEGG Analyses of the Assembled Unigenes

To obtain more information from the RNA-Seq results, GO classifications COG categories and KEGG pathway analysis were performed. GO classifications can provide enrichment analysis on gene sets. Totally 13,561 unigenes were annotated in more than one GO classifications. These items were further divided into 60 functional subgroups, including 17 molecular functions, 19 cellular components, 24 biological process. Among molecular functions, the most two abundant GO terms were binding and catalytic activity. Among cellular component, cell part, membrane were the most enriched GO terms. Among biological process, the two represented GO terms were metabolic process and cellular process (Fig. 4). The COG (Clusters of Orthologous Groups of proteins) database was generated to predict the functions of individual proteins or protein sets (Tatusov et al., 2000). 7437 unigenes were distributed among 25 categories. Among these categories, R category (General function prediction only; 12.18%) was the largest group, then followed by the T category (Signal transduction mechanisms; 11.98%). Three other categories containing more than 6% unigenes were O (Posttranslational modification, protein turnover, chaperones; 9.28%), J (Translation, ribosomal structure and biogenesis; 9.46%), and G (Carbohydrate transport and metabolism; 6.90%) (Fig. 5).

To gain a better understanding of the metabolism changes activated in strawberry. The assembled unigenes were assigned to KEGG database. A total of 11,772 assembled unigenes were annotated in KEGG database. The largest pathway group was carbon metabolism, which contained 401 unigenes. This was followed by ribosome (374) and biosynthesis of amino acids (350) (Fig. 6). These annotations provide a valuable foundation for further researching on the processes and pathways involved in strawberry freezing stress response.
To analyse the effect of freezing treatment on gene expression, we examined the DEGs between controls and freezing-treated samples. Totally, 528 DEGs were identified between controls and freezing-treated samples (FDR < 0.05 and |log2FC| ≥ 1). Of these DEGs, 428 were upregulated and 100 were downregulated. These DEGs were used for cluster analysis. Overall, the upregulated DEGs and downregulated DEGs were separated into two major clusters (Supplementary Fig. S1). To better understand the DEGs’ functions, we performed a GO enrichment and KEGG pathway analysis. A total of 530 upregulated genes and 88 downregulated genes were assigned to GO annotation. Among upregulated DEGs, most of the DEGs (63 DEGs) were annotated with metabolic process, followed by cellular process (59 DEGs) and binding (51 DEGs) (Fig. 7). Among downregulated DEGs, metabolic process (13 DEGs) was the largest enriched group, then followed by the GO term “catalytic activity” (11 DEGs) and “binding” (11 DEGs) (Fig. 7). In addition, the GO term enrichment result was analysed by directed acyclic graph (Supplementary Fig. S2). The result showed that DEGs among these GO terms: “oxygen oxidoreductase activity”, “phenylpropanoid catabolic process”, “lignin catabolic process”, “external encapsulating structure”, “cell wall”, “extracellular region”, and “apoplast” changed significantly in response to freezing stress.
To gain a better understanding of the metabolism pathways influenced by the freezing stress, these genes were analyzed in KEGG database. Fifty-three pathways were enriched, which contained 185 DEGs. Among these pathways, “Environmental adaptation” and “Carbohydrate metabolism” were the most highly represented, with ten and seven genes in these two pathways, respectively.

**Potential Genes Involved in Cold Signal Transduction Pathway in Strawberry Receptacle Tissue**

The CBF signal transduction pathway have a crucial role in regulating many biological processes in response to low temperature stress. It was reported that the *F. × ananassa* CBF1 (*FaCBF1*) was not induced by chilling stress in receptacle tissue, and the overexpressing of *FaCBF1* did not improve the freezing tolerance ability in strawberry receptacle tissue (Owens et al., 2002). These results imply that there might exist other cold response signaling pathways involved in regulating low temperature response in strawberry receptacle tissue. Among our RNA-seq data, we identified 3 CBF/DREB transcription factor genes and 1 COR gene in up-regulated DEGs. And the expression results were validated by qRT-PCR. The expression level of *DREB1F*-like (c41210), *DREB2C*-like (c43922) and *COR47*-like (c28135) were largely induced by freezing stress, while the expression of *DREB1D*-like (c36129) was slightly induced by freezing stress (Fig. 8). These results indicated that *DREB1F*-like (c41210), *DREB2C*-like (c43922) and *COR47*-like (c28135) might be involved in low-temperature signal transduction in strawberry receptacle tissue.
In this study, we revealed the effect of freezing stress on strawberry receptacles at the transcriptome level. We obtained 42,365 unigenes in total, of which 69.91% were annotated in the NR database. The majority of annotated unigenes were similar to genes from the diploid woodland strawberry (*F. vesca*). Using bioinformatics, 528 DEGs were identified between controls and freezing-treated samples.

**Fig. 7:** Gene ontology classification of the DEGs in strawberry. A total of 530 up-regulated genes and 88 down-regulated genes were assigned to three main GO categories: Biological Process, Cellular Component, Molecular Function.

**Fig. S1:** Hierarchical clustering analysis of DEGs under freezing treatment in strawberry receptacle tissue.

**Discussion**

In this study, we revealed the effect of freezing stress on strawberry receptacles at the transcriptome level. We obtained 42,365 unigenes in total, of which 69.91% were annotated in the NR database. The majority of annotated unigenes were similar to genes from the diploid woodland strawberry (*F. vesca*). Using bioinformatics, 528 DEGs were identified between controls and freezing-treated samples.
Among these DEGs, we found the upregulation of 428 genes and downregulation of 100 genes. We further analyzed the DEGs in terms of signalling pathways, metabolism and hormone signal transduction. The careful analysis of DEGs allowed us to understand the mechanisms of freezing stress on strawberry receptacles.

Low temperature stress is a big threat to crop yields. The physiological changes and mechanisms of low temperature tolerance have been revealed by studying the model species (Chinnusamy et al., 2007). The RNA-seq method has been widely used to identify DEGs between different cold treatments or different cultivars in many non-model plants, such as in Ocimum americanum var. pilosum (Zhan et al., 2016), Spartina pectinata (Nah et al., 2016), and Mangifera indica (Shankalyani et al., 2016).

It is known that low-temperature stress could cause the overproduction of ROS (Reactive oxygen species), induce the expression of of secondary metabolism genes (Janska et al., 2010), trigger the accumulation of low-molecular weight nitrogenous compounds, soluble sugars and sugar alcohols (Janska et al., 2010). It is shown from the results of the GO and KEGG analyses that many low temperature stress-related categories were identified, including: “oxygen oxidoreductase activity”, “phenylpropanoid catabolic process”, “lignin catabolic process”, “external encapsulating structure”, “cell wall”, “extracellular region”, “apoplast”, “Environmental adaptation” and “Carbohydrate metabolism”. This discovery implicates that genes which identified in our RNA-Seq data might have crucial roles in regulating freezing tolerance in strawberry receptacle tissue.

Cold shock causes a transient increase of cytosolic calcium levels. These signals are amplified through secondary messengers, such as calcium binding proteins (Jeon and Kim, 2013). Many calcium binding proteins were found to be upregulated after freezing treatment. These signals might be further transferred to the nucleus to activate the expression of stress-related TFs (Jeon and Kim, 2013; Nah et al., 2016). Arabidopsis CBFs (CBF1/DREB1B, CBF2/DREB1C and CBF3/DREB1A) were reported to have an important role in regulating the response of cold acclimation (Jeon and Kim, 2013). Constitutive overexpression of any CBF (CBF1, CBF2, and CBF3) genes in Arabidopsis induces the increase of freezing tolerance ability (Jaglo-Ottosen et al., 1998; Kasuga et al., 1999; Gilmour et al., 2004). The expression and function of the strawberry FaCBF1 gene has been characterized. Data showed that FaCBF1 was not induced by cold stress in receptacles and had no role in increasing freezing tolerance in receptacles (Owens et al., 2002). This result was further confirmed by our RNA-seq data: no orthologues of any CBF (CBF1, CBF2 or CBF3) genes were identified among the DEGs. However, we found the upregulation of several...
other DREB genes, such as DREB1F-like (c41210) and DREB2C-like (c43922). DREB1F-like (c41210) is the ortholog of AtDREB1D. In Arabidopsis, DREB1D is also known as CBF4. Although the expression of CBF4 was not induced by cold stress in Arabidopsis, the CBF4/DREB1D-overexpressed plants exhibited an increase in drought and freezing tolerance (Chinnusamy et al., 2007). These data indicated that FaDREB1D, which identified in our DEGs, might have an important role in regulating the freezing stress response in strawberry receptacle tissue.

Conclusion

In conclusion, we conducted a de novo transcriptome analysis of the freezing-stress response in the receptacles of the cultivated strawberry. By analyzing the DEGs in detail, we revealed the effects of low temperature on strawberry receptacles at the transcriptome level. Notably, we identified 3 CBF/DREB transcription factor genes and 1 COR gene, which might be involved in cold signal transduction in strawberry receptacle tissue. Furthermore, our transcriptome data will provide a valuable foundation for further characterizing the molecular mechanisms of freezing stress in strawberry growth and development.

Acknowledgements

This work was supported by the Jiangsu Planned Projects for Postdoctoral Research Funds (No. 1601103B); Jiangsu Provincial Natural Science Foundation of China (No. BK20130722) and National Natural Science Foundation of China (No. 31301764).

References


(Received 23 June 2017; Accepted 30 December 2017)