Comparison of Gene Expression in the Red Imported Fire Ant, Solenopsis Invicta, in Different Temperature Conditions

Youngjin Park (happy2pyj@gmail.com)
Animal and Plant Quarantine Agency
Mohammad Vatanparast
Animal and Plant Quarantine Agency
Robert Puckett
Texas A&M University
Deuk-Soo Choi
Animal and Plant Quarantine Agency

Research Article

Keywords: Solenopsis invicta, RIFA, transcriptome, RNA-seq, KEGG pathway, differentially expressed gene

Posted Date: March 1st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-255031/v1

Version of Record: A version of this preprint was published at Scientific Reports on August 13th, 2021.
See the published version at https://doi.org/10.1038/s41598-021-95779-w.

Abstract

The red imported fire ant (RIFA), *Solenopsis invicta* Buren is native to South America and is known as a global problematic invasive species. This study focused on the molecular response of RIFA by comparing gene expression profiles after exposing ants to low (10℃) and high (40℃) temperature stress and comparing to untreated controls (30℃). A total of 99,085 unigenes were obtained, of which 19,154 were annotated with gene descriptions, gene ontology terms, and metabolic pathways. 86 gene ontology (GO) functional sub-groups and 23 EggNOG terms resulted. Differentially expressed genes (DEGs) with $\log_2$FC ≥ 10 were screened and were compared at different temperatures. We found 203, 48, and 66 specific DEGs co-regulated at 10, 20, and 40℃. Comparing transcriptome profiles for differential gene expression resulted in various DE proteins and genes, including cytochrome P450, NADH dehydrogenase subunit 1, cuticle protein and heat shock protein (HSP) which have previously been reported to be involved in cold and high temperature resistance. GO analysis revealed that antioxidant activity up-regulated under high temperature stress. We verified the RNA-seq data by qPCR on 20 up and down-regulated DEGs. These finding provide a basis for the future understanding of adaptation mechanism of RIFA and molecular mechanism underlying the response to low and high temperatures.

Introduction

The red imported fire ant (RIFA), *Solenopsis invicta* Buren (Hymenoptera: Formicidae), is a global invasive and aggressive species native to South America that is considered to be one of the 100 most impactful invasive pests around the world[1,2]. RIFA was first reported from the US in the 1930's and since has spread to other temperate areas around the globe 2011[1]. RIFA populations are now established in the United States, Mexico, Australia, New Zealand, China, Malaysia, Japan, Singapore, and the West Indies[3,4]. In Korea, RIFA is listed as a quarantine pest. RIFA is well-documented to drive negative impacts human health, public safety, ecosystems, agriculture and native biodiversity in their invasive range[5]. RIFA invasions potentially threaten 41 species on the China National List of Protected Wildlife, including 22 birds, one amphibian and 18 reptiles and also it has been predicted that they will likely create resource limitations in arthropod communities[6]. As a result of the deleterious effects of RIFA invasions mentioned above, RIFA is of great concern to the Animal and Plant Quarantine Agency of Korea.

Temperature is one of the critical abiotic factors that determine the distribution and life history of insects[7,8]. Population dynamics and geographical distribution of insects are affected by temperatures trough interfering with their metabolic processes such as alimentation, digestion, detoxification, mating, and development[9–13]. Extreme temperatures are potential hazards for the stability of insect populations and can be detrimental to their development. In warm regions, RIFA demonstrates ecological adaptability to extreme high temperatures, but its geographical distribution is directly impacted by cold temperatures[14]. It has been well documented that temperature indices provide a useful predictive tool for predicting the potential distribution of RIFA in newly invaded systems[15]. RIFA is of great concern in China, where studies have been conducted to determine its tolerance to extreme temperatures in order to predict

its potential range expansion[14]. Recently, scientists using 'omic' technologies have determined which pathways are important for allowing a species of beetle to cope in temperature stress[16]. Transcriptomics and the fast development of novel high-throughput sequencing technologies, such as RNA-Seq, has provided an opportunity to investigate signaling-associated genes and triggered putative function(s) and pathway(s) at low and high temperature stress conditions, in insects[17,18].

RNA-seq technologies in a New Zealand alpine stick insect demonstrated upregulation of cuticle genes following cuticle modification in response to low temperature were the results of first used[19]. Since 2014, transcriptome analysis using RNA-seq has been used to investigate gene expression changes when coping with thermal stress in several species of insects (*Drosophila virilis*[20], *Cryptolaemus montrouzieri*[18], *Microdera punctipennis*[21], *Nilaparvata lugens, Sogatella furcifera, Laodelphax striatellus*[22], *Galeruca daurica*[16], and *Monochamus alternatus*[7]). The findings of such studies demonstrate that cold stress can change the expression levels of hundreds of genes associated with transcription, metabolism, and cuticular organization, especially enzyme-related genes responsible for the upregulation of encoding cytochrome P450s (P450), antioxidative enzymes, and aldehyde dehydrogenase [18,23,24].

In this study, we used RNA-Seq and *de novo* transcriptome assemblies to generate transcriptomes and examine the changes in the regulation of transcription associated with cold and heat treatment in *S. invicta*. Detailed differential expression analysis revealed a number of candidate genes that are potentially related to the cold and heat tolerance of RIFA. We performed qRT-PCR to validate the RNA-seq data. We aimed to develop a basis for the adaptive mechanism and a rich resource for the discovery and identification of novel genes involved in the cold and heat stress response in *S. invicta.*

Results

Sequencing, RNA-Seq Assembly, and Functional Annotation

To investigate the transcriptome responses to heat and cold stress in *S. invicta*, quality filtering for Illumina raw data (Table S2) was completed. In total, 44.53 Gb of clean data passed the Illumina quality filter after transcriptome sequencing of four cDNA samples with Q30 > 94% (Table S3). To perform the *de novo* transcriptome assembly, all high-quality reads (Table S3) were pooled. Using paired-end joining and clustering according to the similarity of contigs, these contigs were further assembled into 107,264 transcripts with a mean length of 757.72 bp and an N50 of 1,504 bp, and 99,085 unigenes with a mean length of 615.38 bp and an N50 of 1,051 bp (Table S4 and S5). The length distribution of unigenes closely followed the length distribution of transcripts. This indicates a high-quality assembly, providing a sequence basis for future studies.

Annotation of predicted proteins

For validating and annotating, the assembled unigenes were searched against five public databases (NR, NT, UniProt, Pfam, GO, EggNOG, and KEGG) using BLASTX with a cut off *E*-value of 10^{-5}. After

annotation, genes with a significant blast hit to arthropods were identified. In total, 19,154 (19.33%) unigenes were found in at least one public database (UniProt). The NT database (41,925 annotated unigenes, 42.31%) had the most matches, followed by the NR database (21,232, 37.28%) (Fig. 1, Table 1). Overall, most of the unigenes either could not be annotated or their descriptions were uninformative (e.g., putative, unknown, hypothetical, or unnamed protein). Overall, the unigene sequences were most similar to gene sequences from *S. invicta* (56.80%) and more than 70% showed similarity with ant genera (*Solenopsis* sp, *Trachymyrmex* sp, *Acromyrmex* sp, *Atta* sp, *Camponotus* sp, and *Cyphomyrmex* sp), as observed via BLASTX matches in NR database.

Table 1
Statistics of annotation analysis of unigenes.

Database	Unigene	300 ≤ length < 1000	Length ≥ 1000
NT_Annotation	41,925	5,363	36,441
NR_Annotation	36,937	19,637	6,067
Pfam_Annotation	22,252	4,461	36
EggNOG_Annotation	31,092	17,091	5,823
KO_EUK_Annotation	33,907	19,081	6,224
GO_Annotation	22,091	12,752	4,536
UniProt_Annotation	19,154	11,116	3,858

ORF prediction for unigenes was performed using the TransDecoder program. ORFs of at least 100 amino acids in length were extracted. 14.86% (14,721) of total predicted unigenes (99,085) included at least one ORF and 49.3% showed a full open reading frame (Table 2).

Table 2
Statistics of Open Reading Frame (ORF) prediction.

Assembly	Total unigene	ORF predicted unigene	Single ORF predicted unigene	Multiple ORF predicted unigene	
merge	99,085	14,721 (14.86%)	13,325 (90.52%)	1,396 (9.48%)	
Assembly	**# of ORF**	**Complete**	**Internal**	**5' partial**	**3' partial**
merge	16,235	8,004 (49.3%)	3,733 (22.99%)	3,462 (21.32%)	1,036 (6.38%)

GO and EggNOG analysis for global functional classification

For functional annotation of the unigenes, the Gene Ontology (GO) database and EggNOG database were used to classify the annotated unigenes using BLASTX. As one unigene can have different functional annotations, a total of 22,091 genes have been annotated to GO terms inferred from BLAST results. 86

GO functional sub-groups could be obtained according to the three main GO groups ‘biological process’, ‘cell component’, and ‘molecular function’ in the GO database. 8,646 unigenes belonged to the biological process group, 6,676 unigenes fell in the cellular component group, and 6,769 unigenes were categorized in the molecular function group (Fig. 2A). The most frequent GO terms were ‘metabolic process (9,546 unigenes)’, ‘cellular process (12,843 unigenes)’, ‘cell part (14,118 unigenes)’, ‘catalytic activity (8,895 unigenes)’ and ‘binding (9,186 unigenes)’.

We used the EggNOG database to reveal functional and biological classification. In total, 31,092 unigenes were assigned to 23 EggNOG terms (Fig. 2B) that belonged to three functional classes including ‘information storage and processing’, ‘cellular processes and signaling’, and ‘metabolism’. The largest number of unigenes were classified as ‘translation, ribosomal structure, and biogenesis (1,399 unigenes)’, ‘transcription (1,264)’, ‘replication, recombination, and repair (1275)’, ‘intracellular trafficking, secretion, and vesicular transport (1,528)’ and ‘post-translational modification, protein turnover, chaperones (2,057)’ (Fig. 2B).

Differential gene expression under different temperature

When compared to untreated controls (T30), up-regulated DEGs under various temperature treatments were identified. 4,596, 2,953, and 4,068 unigenes were DEGs for T10, T20 and T40, respectively, with a criterion of p-value < 0.05 and $|\log_2 FC| \geq 2$ (Fig. 3A).

To identify the most probable temperature-reactive specific genes, a volcano plot with a criterion of p-value < 0.05 and $\log_2 FC \geq 5$ was constructed for each treatment temperature, in comparison as the control temperature (Fig. 3B-D). The largest number of expression changes in unigenes were observed in response to T10 (Fig. 3B). GO analysis revealed that most unigenes in the category of ‘biological process’ belong to the ‘cellular process (12,843)’ sub-group (Table S6), and ‘cellular part (14,118)’ and ‘binding (9,186)’ included the largest number of unigenes from the ‘cellular component’ and ‘molecular function’ classifications, respectively (Table S6).

To explore the more specific and exclusive genes involved in cold and high temperatures stress conditions, a Venn diagram was plotted for T10, T20, and T40 in comparison with the T30 as a control group (p-value < 0.05). As shown in Fig. 4A, 203, 48, and 66 differentially expressed genes (DEGs) were identified when comparing the control (30°C) and stressor temperature 10, 20, and 40°C, respectively. There were 13 common DEGs that were consistently up-regulated in all three groups. The same 51 DEGs were up-regulated ($FC \geq 10$) between T10 and T20, when compared with T30. There were 29 and 5 DEGs up-regulated ($FC \geq 10$) between T10 and T40; T20 and T40, respectively as compared to the T30 control. We believe two general groups of unigenes are related to temperatures fluctuations. The first group includes the unigenes that are presented in all (T10, T20, and T40 (13)) or two (T10 and T40 (42)) treatment temperatures, and the second group are the unigenes that specifically expressed more than 10 times at one temperature (T10 (203), T20 (48), and T40 (66)) (Fig. 4A). To understand the comparative distribution of unigenes in the first group, a heatmap was constructed (Fig. 4B and C). ‘Venom carboxylase-6-like’, ‘cGMP-dependent protein kinase’ and ‘growth hormone regulated TBC protein 1-A’

showed high expression levels when the ants were incubated at 10°C in comparison with 40°C and to 30°C controls. Interestingly, 'histone' unigenes (histone H3-like centromeric protein, histone H2A, histone H4, and histone H2B-like) showed high degrees of fold change at 40°C in comparison with 10°C (Fig. 4B). Among 13 matching unigenes 'aromatic-L-amino-acid decarboxylase' and 'homeobox protein orthopedia-like' showed lower expression levels at T10 and T40 when compared to those at T20 (Fig. 4C).

Specific unigenes ($log_2FC \geq 10$) for each treatment were the second group of unigenes that were investigated. To confirm functional unigenes, KEGG and GO analysis was performed (Figs. 5 and 6, Table S8, 9, and 10). KEGG pathway enrichment analysis revealed the primary DEG pathways (Fig. 5A, B, and C). Pathway enrichment was observed among all groups, although the T20 group showed lesser effects on pathway enrichment than T10 and T40 (Fig. 5B). A large number of co-regulated DEGs under cold and high temperature stresses were significantly enriched in the 'Metabolic pathway' which in T10 is dominated by 'pyruvate carboxylase', 'diacylglycerol kinase 1', 'hexaprenyldihydroxybenzoate methyltransferase', 'sphingomyelin phosphodiesterase 1-like', 'adenylate cyclase type 6', 'ADP-dependent glucokinase', and 'porphobilinogen deaminase-like'. Another identified enriched-pathway for T10 group included 'Lysosome' which includes 'CD63 antigen-like' and 'sphingomyelin phosphodiesterase 1-like' unigenes. 'RNA transport pathway' and 'carbon metabolism pathway' are another pathway enriched in the T10 group (Fig. 5A). Three proteins presented in the 'Metabolic pathway' in T20 were 'muscle M-line assembly protein unc-89', 'histone-lysine N-methyltransferase SETMAR-like', and 'acetyl-coenzyme A transporter 1'(Fig. 5B). Additionally, 'fatty acid synthase', 'phosphodiesterase 8A', 'procollagen-lysine,2-oxoglutarate 5-dioxygenase 1', 'fatty acid synthase-like', 'fructose-bisphosphate aldolase-like', 'histone-lysine N-methyltransferase SETMAR-like', 'eye-specific diacylglycerol kinase', 'tyrosine aminotransferase', 'beta-1,3-galactosyltransferase 5', and 'S-adenosylmethionine decarboxylase proenzyme' were the transformed components of the metabolic pathway of T40 group. 'Protein processing in endoplasmic reticulum', 'Biosynthesis of amino acids', 'Lysine degradation', 'Fatty acid metabolism', and 'Fatty acid synthase' are other enriched pathways in the T40 group (Fig. 5C).

GO annotation was used to clarify the functions of DEGs that were significantly different ($FC \geq 10$) between treatments (Fig. 6). Between T10 and T30 control, the most significant items were 'Biological Process: multicellular organismal process', 'Cellular Component: organelle' and 'Molecular Function: catalytic activity and binding' (Fig. 6A). When comparing T20 and T30 control, unigenes were assigned to 'Biological Process: developmental process', 'Cellular component: membrane part' and 'Molecular Function: binding and molecular function regulator' (Fig. 6B). Comparison between T40 and T30 control revealed the following significant increase of unigene percentages: 'Biological Process: response to stimulus', 'Cellular component: organelle' and 'Molecular Function: catalytic activity' (Fig. 6C).

Validation of Gene Expression profiles by q-PCR

qPCR and gel electrophoresis of twenty common DEGs identified in the RNA sequence data was performed to confirm the accuracy and reproducibility of the Illumina RNASEq. 10 up-regulated DEGs were included; hsp70 like-protein (h70; c412839_g5_i2), synapsin (Syn; c408336_g1_i2), cytochrome P450 (P450; c407395_g2_i1), serine protease (SerP; c391510_g1_i1), fatty acid synthase like (FaSyn;

c412971_g1_i1), glycerol-3-phosphate dehydrogenase (GPDH; c391490_g1_i2), cuticle protein (Cut; c385485_g1_i2), trypsin-2-like (Try; c395664_g1_i3), lipase-3-like (Lip; c412461_g4_i1), and chymotrypsin (Chy; c414255_g1_i1). Another 10 DEGs that showed down-regulation at T10 in comparison to T30 controls were validated by qPCR including; general odorant-binding protein 72 (Odo; c410048_g3_i1), small G protein signaling modulator 3 homolog (SGP; c422133_g1_i1), scavenger (Scv; c412512_g2_i1), RNA binding protein 33 like (RBi; c375277_g1_i1), monocarboxylate transporter 1-like (MCa; c411974_g1_i2), dipeptidase 1-like (Dip; c412614_g1_i6), calmodulin-like protein 4 (Cal; c400490_g2_i1), transmembrane channel-like protein 2 (Tra; c412553_g2_i3), anoctamin-4 (Ano; c409055_g3_i1), and fibrinogen silencer-binding protein like (Fib; c411675_g1_i1). The results of qPCR and Illumina FPKM ratio were plotted in Fig. 7. These data demonstrated that expression changes were in the same direction for the qPCR. The Illumina sequencing data were consistent with qPCR data, verifying the reliability and accuracy of the transcriptome analysis. This ensures the RNA-Seq results were considerably reliable for the identification of DEGs under temperature stress, and also the feasibility and sustainability of our further research on these or other DEGs from the transcriptome Data.

Discussion

Comprehensive investigation of gene expression regulation under temperature stress is very important to understand the biochemical and physiological adaptation processes in invasive insect pests[16]. In this study, a comprehensive transcriptome analysis and characterization of the gene expression profiles of *S. invicta* under cold and high temperature stress were evaluated. Through the analysis of DEGs, transcriptome changes in *S. invicta* adult ants were revealed. Using RNA-seq techniques four transcriptomes was *de novo* assembled from the adult stages of RIFA which exposed to four different temperatures (10, 20, 30, and 40°C), and 19,154 unigenes (19.33%) were successfully annotated from at least one public databases (UniProt) (Table 1). The results are in line with other transcriptome projects using Illumina technology[25–27]. 56.80% of the unigene sequences were most similar to gene sequences from *Solenopsis invicta* and more than 70% similarity with ant genus were observed. In this study, DEGs from adult RIFA which were exposed to different treatment temperatures (10, 20, and 40°C) were compared to that of a 30°C control group. The majority of DEGs were observed at T10, follow by T40, both of which expressed a greater DEG distribution than T20 (Fig. 3A). As mentioned earlier, this is consistent with proteomics data from *Locusta migratoria* under high and low temperature stress[28]. To identify specific genes associated with response to temperature, the number of unigenes with $\log_2 FC \geq 10$ was clarified by Venn diagram, and KEGG analysis was conducted to determine the probability of function in pathway enrichment. KEGG analysis revealed that from 203 specific cold-regulated DEGs (associated with T10), 41 DEGs were enriched in the KEGG pathways, and most of them were classified to following pathways: ‘Metabolic pathway’, ‘Carbon metabolism’, ‘Citrate cycle (TCA)’, ‘RNA transport’, and ‘Lysosome’. Interestingly in T20 and T40, ‘Metabolic pathway’ included more DEGs than other pathways (Fig. 5). ‘Purine metabolism’, ‘Spliceosome’, ‘Lysosome’, ‘RNA degradation’, ‘Glycolysis/Gluconeogenesis’, ‘Pyruvate metabolism’, ‘Phagosome’, ‘Sphingolipid metabolism’, ‘RNA transport’, ‘Glycerolipid metabolism’, ‘Carbon metabolism’, ‘ECM-receptor interaction’ are pathways that demonstrate similar results as those of

investigations on transcriptome responses to cold stress in the carpenter moth, *Eogystia hippophaecolus*[25], and the a chrysomelid beetle, *Galeruca daurica*[16]. Transcriptome analysis revealed that the expression of 'Glycolysis' and 'TCA cycle pathways' are up-regulated in a similar manner as the braconid wasp, *Aphidius colemani*, when exposed to low temperatures[29]. When RIFA was exposed to the highest temperature (40°C), 'Tyrosine metabolism', 'Phenylalanine metabolism', 'Cysteine and Methionine metabolism', 'Spliceosome', 'Protein processing in endoplasmic reticulum', and 'Metabolic pathway', were found to be enriched pathways that are similarly enriched in three species' of rice plant hopper when exposed to 37°C[22]. 'Fatty acid synthase' and 'Fatty acid metabolism', are two of the main pathways of RIFA when exposed to high temperature. Expression of fatty acids as hydrophobic agents allows insects to avoid water loss in warmer regions of the globe[30]. At high temperatures, *Gomphocerus sibiricus* are known to increase their levels of oleic acid, linoleic acid, linolenic acid and glycerin, and phenomenon can suppress mortality due to excessive evaporation of body moisture[31]. In our study on RIFA, 'Amino acid metabolism' was clearly up-regulated during high-temperature stress. It is suggested that amino acid metabolism provides heat resistance in RIFA similar to that of results that have been reported for *Locusta migratoria*[28] when exposed at 40°C. Due to the synthesis of immune proteins and defense enzymes, insects seek out and consume numerous free amino acids when coping with stress conditions such as high temperature, low temperature and fungal invasion[32]. The synthesis and metabolism of amino acids are necessary to produce a significant number of amino acids, which make available the raw materials necessary for the synthesis of heat-resistant proteins[28].

In this study, two cuticular protein unigenes were identified from 203 co-regulated DEGs under cold temperature stress (Table S9). Cuticular protein gene expression has been observed in studying other insects studies such as beetles, moths, planthoppers, and stick insects when exposed to cold temperature stress[16,22,25,33]. Although the physiological role of cuticular proteins in insect cold hardiness has not yet been identified, it seems insect cuticle may play an important role in insects when coping with low temperature[16,19,22,34,35].

According GO analysis (Fig. 6A), 'Antioxidant activity' was enriched at low temperatures. Suggesting that is might contribute to RIFA ability to resist oxidative stress damage at low temperature[16], or their potential for cell preservation via antioxidant defense when in challenged by environmental complexity[36]. In addition, one cytochrome P450 was identified that up-regulated exclusively under low temperature (Table S9). Meanwhile NADH dehydrogenase subunit 1 was up-regulated at all treated temperature (Fig. 5C, Table S8). These two proteins are main enzymes at antioxidant activity pathway[25]. In comparative analysis of the transcriptional responses to low and high temperatures in three rice planthopper species, some cytochrome P450 genes were up-regulated under both low and high temperatures, which suggests cold and heat stress increase oxidative stress in the insect body[22].

Heat shock proteins (HSPs) are another important protein that insects use as critical physiological products when under abiotic stress conditions[28]. From earlier studies it was believed that *Hsp* is associated with biological cold and heat resistance[37,38]. HSPs are molecular chaperones, which play

important physiological roles including: correct folding of proteins, prevention of protein denaturation, and degradation of misfolded or condensed proteins and maintenance of correct protein conformation[39,40]. In this study, we identified one specific *Hsp70* that up-regulated about a 12-fold change when RIFA was exposure at 10°C (Table S9). Two pathways, including 'Protein export' and 'Protein processing in endoplasmic reticulum', were enriched under *Hsp70* gene expression at low temperatures (Fig. 5A, Table S12). Interestingly, heat shock protein 83 was found only in the 40°C treatment group that was up-regulated about 11 times (Table S11). Many studies have confirmed that the expression of *Hsp* genes can be up-regulated by cold and heat stimulus[39,41]. To assist the resistance to temperature stress, the *Hsp60* gene expression in *Stegobium paniceum* significantly increases under high- and low-temperature stress[42]. Three *Hsp90* and four *Hsp70* were up-regulated by cold stress and were differentially expressed at desert beetle, *Microdera punctipennis*[21]. The differences in *Hsp*, insect species, sex of organism, and intensity of temperature are important factors related to *Hsp* expression level in insects[16,43].

In conclusion, we compared the transcriptomes of *S. invicta* under high- and low-temperature stresses using RNA-Seq technology based on the high-throughput sequencing. Comparative transcriptome analysis identified many genes, and a large number of changes were discovered in the metabolic pathways through the GO and KEGG enrichment analysis. Our data will facilitate further molecular investigations and genomic research. Many novel relationships between high- and low-temperature and significantly up-regulated genes were identified in this study (Table S7-11). These newly found genes may be important to RIFA overwintering and adaptation potential in new environments as well as quarantine area.

Materials And Methods

Insect rearing, exposure temperatures and sample preparation

S. invicta colonies were collected in Somerville, TX, US (30°31'13" N, 96°25'33"W) and imported to Korea according to the Plant Protection Act. All insects were reared at Plant Quarantine Technology Center laboratories (Animal and Plant Quarantine Agency, Gimcheon, Korea). To prevent insect escape, all laboratories are equipped with automatic wind curtain and sticky floor mats at entrance. Ant colonies were maintained at 25 ± 1℃. Plastic trays (25 (H) · 30 · 35 cm^3) containing each test colony were placed in larger holding trays (35 (H) · 45 · 65 cm^3). Talcum powder was dusted on the top 10 cm interior edge of the trays and along the bottom of the larger holding trays so that any escaping ants escaping would be trapped inside the holding trays. Ants were fed a 20% sucrose solution, mealworms (*Tenebrio molitor* larvae), and an artificial diet described by Dussutour and Simpson (2008)[44]. Water was provided *ad libitum*. Colonies contained dealated mated queens, alate queens, males, brood (eggs, larvae, and pupae) and workers. To perform transcriptomic analysis, medium sized workers were incubated at 10, 20, 30, and 40℃ for 24 h. Untreated control ants were incubated at 30℃. Each treatment was replicated three times.

After the temperature treatment, ten ants from each group were immediately frozen in liquid nitrogen and stored at -80℃ for subsequent experiments.

RNA extraction and RT-qPCR

RNAs samples were extracted from the whole body of *S. invicta* adults using Trizol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. After RNA extraction, it was resuspended in nuclease-free water and quantified using a spectrophotometer (NanoDrop, Thermo Scientific, Wilmington, DE, USA). cDNA was then synthesized from RNA (1 μg) using RT PreMix (Intron Biotechnology, Seoul, Korea) containing *oligo dT* primer according to the manufacturer's instruction. All quantitative PCRs (qPCRs) in this study were determined using a real time PCR machine (CFX Connect Real-Time PCR Detection System, Bio-Rad, Hercules, CA, USA) and iQ SYBR Green Supermix (Bio-Rad, Hercules, CA, USA) according to the guidelines provided by the manufacturer. The reaction mixture (20 μL) contained 10 μL of iQ SYBR Green Supermix, 1 μL of cDNA template (100 ng), 1 μL each of forward and reverse primers (Table S13), and 7 μL nuclease free water. RT-qPCR cycling began with 95°C heat treatment for 10 min followed by 40 cycles of denaturation at 94°C for 30 s, annealing at 52°C for 30 s, and extension at 72°C for 20 s. The expression level of *Ef1_β* as a reference gene was used to normalize target gene expression levels[45] under different treatments. PCR products were assessed by melting curve analysis. Quantitative analysis was performed using comparative CT ($2^{-\Delta\Delta CT}$) method[46].

Illumina sequencing

To obtain short-read RNA sequences, Illumina sequencing was performed at Macrogen (Seoul, Korea). Each library was constructed from 1 μg total RNA from the whole body of 5 individuals (not pooled) of *S. invicta* adults per treatment using the TruSeq Stranded mRNA LT Sample Prep Kit (Illumina, San Diego, USA) and sequenced using HiSeq 4000 System (Illumina, San Diego, USA) with 101 bp pair end read (Table S1).

De Novo **Assembly**

Illumina short reads were quality-filtered and adapter-trimmed using Trimmomatic v0.38 (http://www.usadellab.org/cms/?page=trimmomatic). FastQC v0.11.7 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to check data quality before and after trimming. After the removal of low-quality reads, an Illumina-based *de novo* transcriptome assembly was performed using Trinity version trinity rnaseq r20140717, bowtie 1.1.2[47]. Trimmed reads for every sample were merged into one file to construct a combined reference. The *de novo* assembly of merged data was carried out using Trinity with default parameters and assembled into transcript contigs[48]. Total number of genes, transcripts, GC content max/min/median/average contig length and total assembled bases were summarized. Trinity groups transcripts into clusters based on shared sequence content. For assembled genes, the longest contigs of the assembled contigs are filtered and clustered into the non-redundant transcripts using CD-HIT version 4.6 (http://weizhongli-lab.org/cd-hit)[49]. These transcripts were defined as ‘unigenes’ which are used for predicting the ORFs (Open Reading Frames), annotating

against several known sequence databases, and analyzing differentially expressed genes (DEGs). ORF prediction for unigenes was performed using TransDecoder version 3.0.1 (https://github.com/TransDecoder/TransDecoder/wiki)[50] to identify candidate coding regions within transcript sequences. After extracting ORFs that were at least 100 amino acids long, TransDecoder predicted the likely coding regions. Trimmed reads for each sample were aligned to the assembled reference using Bowtie program. For the differentially expressed gene analysis, the abundances of unigenes across samples were estimated into read count as an expression measure by RSEM algorithm (RSEM version v1.2.29, bowtie 1.1.2, http://deweylab.github.io/RSEM/, (Li and Dewey 2011)[51]).

Gene Functional Annotation

For functional annotation, unigenes were searched against Kyoto Encyclopedia of Genes and Genomes (KEGG) v20190104 (http://www.genome.jp/kegg/ko.html)[52], NCBI Nucleotide (NT) v20180116 (https://www.ncbi.nlm.nih.gov/nucleotide/)[16], Pfam v20160316 (https://pfam.xfam.org/)[53], Gene ontology (GO) v20180319 (http://www.geneontology.org/)[54], NCBI non-redundant Protein (NR) v20180503 (https://www.ncbi.nlm.nih.gov/protein/)[55], UniProt v20180116 (http://www.uniprot.org/)[56] and EggNOG (http://eggnogdb.embl.de/)[57] using BLASTN of NCBI BLAST and BLASTX of DIAMOND version 0.9.21 (https://github.com/bbuchfink/diamond) with an *E*-value default cutoff of 10^{-5}.

Differential Gene Expression Analysis

Quality check was conducted for all samples, so that if more than one read count value was zero, it was not included in the analysis. In order to reduce systematic bias, we estimated the size factors from the count data and applied Relative Log Expression (RLE) normalization with $DESeq_2$ R library. Using each sample's normalized value, the high expression similarities were grouped together by Hierarchical Clustering Analysis and graphically shown in a 2D plot to show the variability of the total data using Multidimensional Scaling Analysis. Significant unigene results were analyzed as Up and Down-regulated count by $\log_2 FC \geq 5$, ≤ -5 and ≥ 10, distribution of expression level between two groups was plotted as Volcano plot and simple bar plots.

Quantitative RT-PCR validation

The twenty genes in response to cold treatment (T10) were chosen for validation using qRT-PCR. To do that, *S. invicta* adults were incubated at 10° and 30°C for 24 hr in two separate groups that included 10 ants. RNA extraction and cDNA synthesis were performed according to **'RNA extraction and RT-qPCR'** section. Specific primers were designed using Primer Quest tool (www.idtdna.com) (Table S13). Expression level of *Ef1_β* was used as a reference gene and to normalize target gene expression levels under different treatments[45]. PCR products were assessed by melting curve analysis. Quantitative analysis was performed using comparative CT ($2^{-\Delta\Delta CT}$) method[46]. Finally, the data were compared according ratio of FPKM and ratio of mRNA expression levels for all selected genes.

Declarations

Competing interests

The authors declare no competing interests.

Author Contributions

M.V. and Y.P. conceived the idea, designed the experiments; M.V. and Y.P. performed the experiments; M.V. and Y.P. analyzed the data; M.V., R.T.P., and Y.P. co-wrote the manuscript; M.V., R.T.P., D.C., and Y.P. discussed the results and commented on the manuscript.

Acknowledgements

This research was supported by a fund (PQ20180A006) by Research of Animal and Plant Quarantine Agency, Korea.

References

1. Ascunce, M. S. *et al.* Global Invasion History of the Fire Ant *Solenopsis invicta*. *Science.* **331**, 1066-1068 (2011).
2. IUCN. View 100 of the World's Worst Invasive Alien Species.
3. Wang, L., Xu, Y.J., Zeng, L. & Lu, Y.Y. Impact of the red imported fire ant *Solenopsis invicta* Buren on biodiversity in South China: A review. *J. Integr. Agric.* **18**, 788-796 (2019).
4. GISD. (Global Invasive Species Database). http://www.iucngisd.org/gisd/speciesname/Solenopsis+invicta# (2018).
5. Vinson, S. B. Impact of the invasion of the imported fire ant. *Insect Sci.* **20**, 439-455 (2013).
6. Wang, L., Chen, K.W. & Lu, Y.Y. Long-distance spreading speed and trend predication of fall armyworm, *Spodoptera frugiperda*, in China. *J. Environ. Entomol.* **41**, 683-694 (2019).
7. Li, H. *et al.* Comparative Transcriptome Analysis of the Heat Stress Response in *Monochamus alternatus* Hope (Coleoptera: Cerambycidae). *Front. Physiol.* **10**, 1568 (2020).
8. Bale, J. S. Classes of insect cold hardiness. (1993).
9. Garrad, R., Booth, D. T. & Furlong, M. J. The effect of rearing temperature on development, body size, energetics and fecundity of the diamondback moth. *Bull. Entomol. Res.* **106**, 175 (2016).
10. Bale, J. S. Insects and low temperatures: from molecular biology to distributions and abundance. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **357**, 849-862 (2002).
11. Lee, R. E. Principles of insect low temperature tolerance. in *Insects at low temperature* 17-46 (Springer, 1991).
12. Turnock, W. J., Lamb, R. J. & Bodnaryk, R. P. Effects of cold stress during pupal diapause on the survival and development of *Mamestra configurata* (Lepidoptera: Noctuidae). *Oecologia* **56**, 185-192

(1983).

13. Hutchinson, L. A. & Bale, J. S. Effects of sublethal cold stress on the aphid *Rhopalosiphum padi*. *J. Appl. Ecol.* 102-108 (1994).
14. Xu, Y.J., Lu, Y.Y., Pan, Z.P., Zeng, L. & Liang, G.W. Heat tolerance of the red imported fire ant, *Solenopsis invicta* (Hymenoptera: Formicidae) in mainland China. *Sociobiology* **54**, 115 (2009).
15. Morrison, J.E.J., Williams, D.F., Oi, D.H, & Potter, K.N. Damage to Dry Crop Seed by Red Imported Fire Ant (Hymenoptera: Formicidae). *J. Econ. Entomol.* **90**, 218-222 (1997).
16. Zhou, X.R., Shan, Y.M., Tan, Y., Zhang, Z.R, & Pang, B.P. Comparative Analysis of Transcriptome Responses to Cold Stress in *Galeruca daurica* (Coleoptera: Chrysomelidae). *J. Insect Sci.* **19**, (2019).
17. Hayward, S. A. L. Application of functional 'Omics' in environmental stress physiology: insights, limitations, and future challenges. *Curr. Opin. insect Sci.* **4**, 35-41 (2014).
18. Zhang, Y. *et al.* Transcriptome responses to heat-and cold-stress in ladybirds (*Cryptolaemus montrouzieri* Mulasnt) analyzed by deep-sequencing. *Biol. Res.* **48**, 66 (2015).
19. Dunning, L. T., Dennis, A. B., Sinclair, B. J., Newcomb, R. D. & Buckley, T. R. Divergent transcriptional responses to low temperature among populations of alpine and lowland species of Newzealand stick insects (Micrarchus). *Mol. Ecol.* **23**, 2712-2726 (2014).
20. Parker, D. J. *et al.* How consistent are the transcriptome changes associated with cold acclimation in two species of the *Drosophila virilis* group? *Heredity (Edinb).* **115**, 13-21 (2015).
21. Tusong, K., Lu, X., Liu, X. & Ma, J. Transcriptomic analysis of the desert beetle *Microdera punctipennis* (Coleoptera: Tenebrionidae) in response to short-term cold stress. *Acta Entomol. Sin.* **59**, 581-591 (2016).
22. Huang, H. *et al.* Comparative analysis of the transcriptional responses to low and high temperatures in three rice planthopper species. *Mol. Ecol.* **26**, 2726-2737 (2017).
23. 23. Wei, D. *et al.* Comparative proteomic analysis of *Bactrocera dorsalis* (Hendel) in response to thermal stress. *Insect Physiol.* **74**, 16-24 (2015).
24. Liu, Y. *et al.* Comparative transcriptome analysis of *Glyphodes pyloalis* Walker (Lepidoptera: Pyralidae) reveals novel insights into heat stress tolerance in insects. *BMC Genomics* **18**, 1-13 (2017).
25. Cui, M., Hu, P., Wang, T., Tao, J. & Zong, S. Differential transcriptome analysis reveals genes related to cold tolerance in seabuckthorn carpenter moth, *Eogystia hippophaecolus*. *PLoS One* **12**, e0187105 (2017).
26. Wu, W. *et al.* De novo transcriptome characterization of the ghost moth, *Thitarodes pui*, and elevation-based differences in the gene expression of its larvae. *Gene* **574**, 95-105 (2015).
27. Zhu, J.-Y., Zhao, N. & Yang, B. Global transcriptome profiling of the pine shoot beetle, Tomicus yunnanensis (Coleoptera: Scolytinae). *PLoS One* **7**, e32291 (2012).
28. Ran, L. *et al.* Comparative transcriptome analysis of *Locusta migratoria tibetensis* Chen (Orthoptera: Oedipodidae) under high-and low-temperature stress. *J. Appl. Entomol.* **144**, 181-190 (2020).

29. Colinet, H., Nguyen, T. T. A., Cloutier, C., Michaud, D. & Hance, T. Proteomic profiling of a parasitic wasp exposed to constant and fluctuating cold exposure. *Insect Biochem. Mol. Biol.* **37**, 1177-1188 (2007).
30. Stanley-Samuelson, D. W., Jurenka, R. A., Cripps, C., Blomquist, G. J. & de Renobales, M. Fatty acids in insects: composition, metabolism, and biological significance. *Arch. Insect Biochem. Physiol.* **9**, 1-33 (1988).
31. Li, J., Li, S., Wang, D. & Ji, R. Changes in the contents of stress resistant substances in *Gomphocerus sibiricus* (Orthoptera: Acrididae) under high temperature stress. *Acta Entomol. Sin.* **57**, 1155-1161 (2014).
32. Ma, X. J., Zhang, Z. H., Wang, Z., Zhang, Y. X. & Chen, J. Effects of brief exposure to high temperature on free amino acids, total sugar and trehalose of female adult *Monolepta hieroglyphica* (Motschulsky). *Plant Prot* **44**, 111-115 (2018).
33. Dunning, L. T. *et al.* Identification of cold-responsive genes in a new zealand alpine stick insect using RNA-Seq. *Comp. Biochem. Physiol. Part D Genomics Proteomics* **8**, 24-31 (2013).
34. Carrasco, M. A. *et al.* Elucidating the biochemical overwintering adaptations of larval *Cucujus clavipes puniceus*, a nonmodel organism, via high throughput proteomics. *J. Proteome Res.* **10**, 4634-4646 (2011).
35. Qin, W., Neal, S. J., Robertson, R. M., Westwood, J. T. & Walker, V. K. Cold hardening and transcriptional change in *Drosophila melanogaster*. *Insect Mol. Biol.* **14**, 607-613 (2005).
36. Zhang, J., Goyer, C. & Pelletier, Y. Environmental stresses induce the expression of putative glycine-rich insect cuticular protein genes in adult *Leptinotarsa decemlineata* (Say). *Insect Mol. Biol.* **17**, 209-216 (2008).
37. Guttman, S. D., Glover, C. V. C., Allis, C. D. & Gorovsky, M. A. Heat shock, deciliation and release from anoxia induce the synthesis of the same set of polypeptides in starved T. pyriformis. *Cell* **22**, 299-307 (1980).
38. Hahn, G. M. & Li, G. C. Thermotolerance and heat shock proteins in mammalian cells. *Radiat. Res.* **92**, 452-457 (1982).
39. King, A. M. & MacRae, T. H. Insect heat shock proteins during stress and diapause. *Annu. Rev. Entomol.* **60**, 59-75 (2015).
40. Gething, M.-J. & Sambrook, J. Protein folding in the cell. *Nature* **355**, 33-45 (1992).
41. Lang, R. P. *et al.* Transcriptome profiling of selectively bred Pacific oyster *Crassostrea gigas* families that differ in tolerance of heat shock. *Mar. Biotechnol.* **11**, 650-668 (2009).
42. Zhu, X. Y., Ding, T. B., Xu, K. K., Yan, Y., Li, C., & Yang, W. J. cDNA cloning of heat shock protein 60(SpHsp60) and its expression under temperature stress in *Stegobium paniceum*. *J. Sichuan Univ.* **54**, 1311-1316 (2017).
43. Li, J., Moghaddam, S. H. H., Du, X., Zhong, B. & Chen, Y. Y. Comparative analysis on the expression of inducible HSPs in the silkworm, *Bombyx mori*. *Mol. Biol. Rep.* **39**, 3915-3923 (2012).

44. Dussutour, A. & Simpson, S. J. Description of a simple synthetic diet for studying nutritional responses in ants. *Insectes Soc.* **55**, 329-333 (2008).
45. Cheng, D., Zhang, Z., He, X. & Liang, G. Validation of reference genes in *Solenopsis invicta* in different developmental stages, castes and tissues. *PLoS One* **8**, e57718 (2013).
46. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**, 402-408 (2001).
47. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
48. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644-652 (2011).
49. Pertea, G. *et al.* TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* **19**, 651-652 (2003).
50. Haas, B. J. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494-1512 (2013).
51. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
52. Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. & Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32**, 277-280 (2004).
53. Finn, R. D., Clements, J. & Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, W29-W37 (2011).
54. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25-29 (2000).
55. Deng, Y. Y. *et al.* Integrated nr database in protein annotation system and its localization. *Comput Eng* **32**, 71-74 (2006).
56. Apweiler, R. *et al.* UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **32**, 115-119 (2004).
57. Powell, S. *et al.* eggNOG v4.0: nested orthology inference across 3686 organisms. *Nucleic Acids Res.* **42**, 231-239 (2013).

Figures

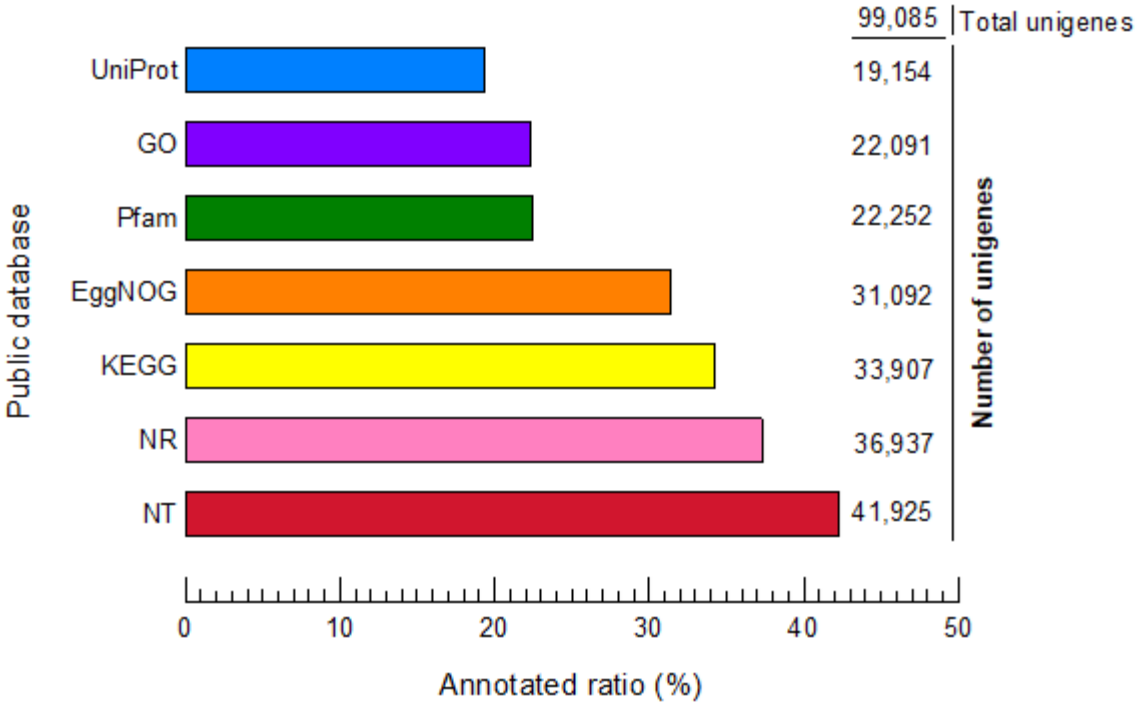

Figure 1

Blasted results for unigenes based on public databases with an E-value default cutoff of E-value 10-5. X axis shows the ratio (%) of annotated unigenes for each database based on total unigenes. Y axis shows the number of unigenes that annotated based on each database.

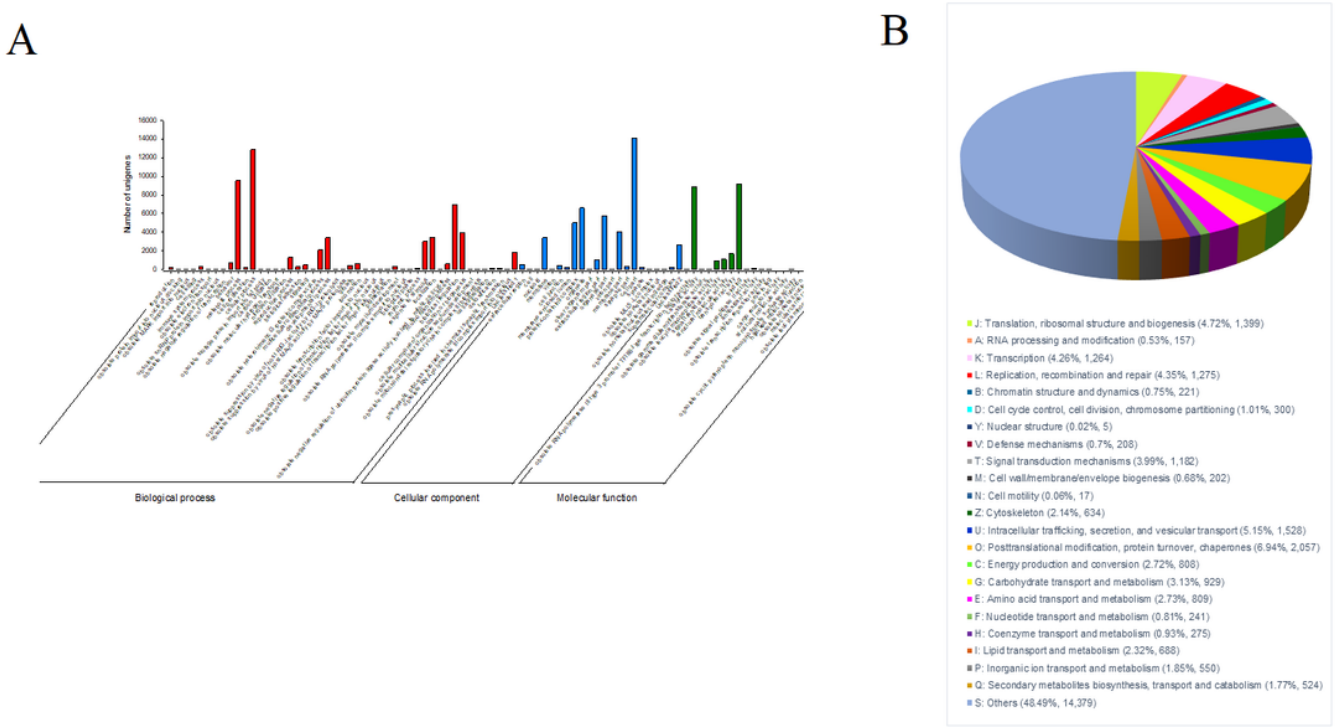

Figure 2

Annotations of the annotated unigenes. (A) The Gene Ontology (GO) classification of all annotated unigenes. 22,091 unigenes were assigned to three main GO categories of ‘Biological process with 46 functional subcategories’, ‘Cellular component with 21 functional subcategories’, and ‘Molecular function with 19 functional subcategories. The Y axis shows the number of genes in each category. Some unigenes was not classified and showed as ‘Unclassified’ in each category. (B) The annotated unigenes are mapped to the annotation of the corresponding orthologous groups in EggNOG (Evolutionary genealogy of genes) database.

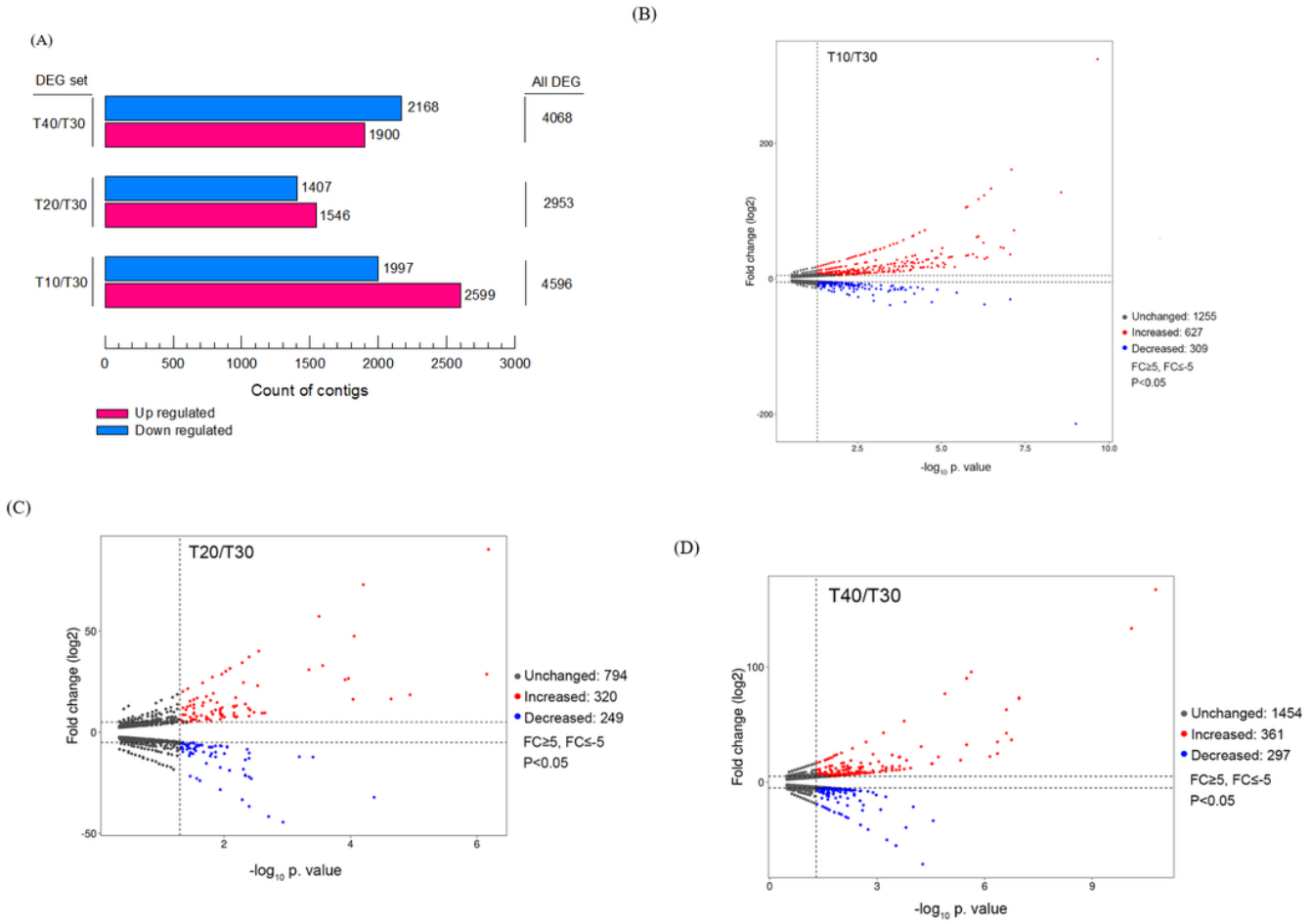

Figure 3

Significant Differentially Expressed Gene (DEG) analysis. (A) The number of all up and down-regulated contigs based on p-value < 0.05 and log2FC ≥ 2 or log2FC ≤ -2 of comparison pairs was plotted. (B-D) Volcano diagram for distribution of the identified DEGs in three different treatment groups in comparison with T30 as control. Red and blue points represent the significant DEGs with p-value < 0.05 and log2FC ≥ 5 or log2FC ≤ -5 and grey ones show those without significant, respectively.

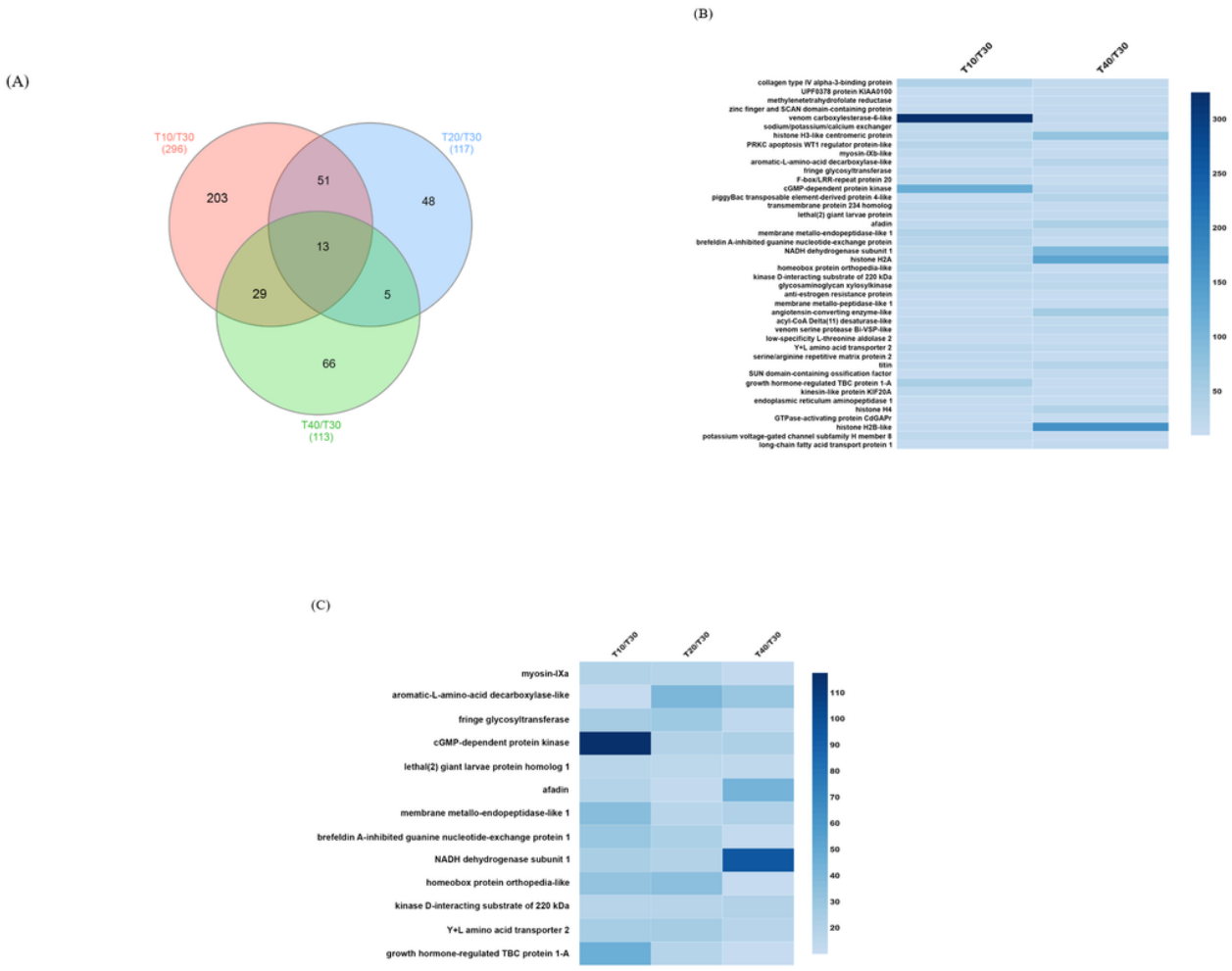

Figure 4

Distribution of same Differentially Expressed Genes (DEGs) between treatment groups. (A) Venn diagram represents for same unigenes (log2FC ≥ 10) between all three groups in comparison with T30 as control. (B) Heatmap shows expression patterns for same DEGs (p-value < 0.05, log2FC ≥ 10) involved in low (10℃) and high (40℃) temperature treatment groups in comparison with T30 as control. (C) Heatmap represents expression of same DEGs (p-value < 0.05, log2FC ≥ 10) involved in all three groups of temperature treatment groups in compare with T30 as control.

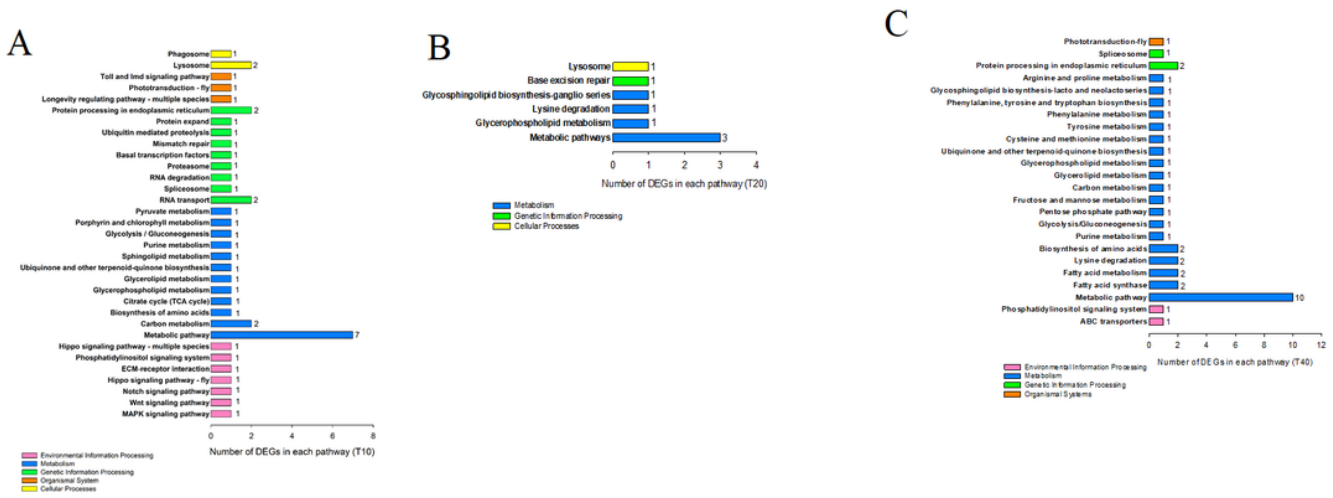

Figure 5

(A-C) Analysis of significant KEGG pathways enrichment in the specific Differentially Expressed Genes (DEGs) (p-value < 0.05, log2FC ≥ 10) co-regulated by temperature stresses including 10, 20, and 40℃, respectively. The Y axis indicates the KEGG pathways, and the X axis shows the number of specific DEGs for each treatment group in each pathway that indicated on the right side of bars.

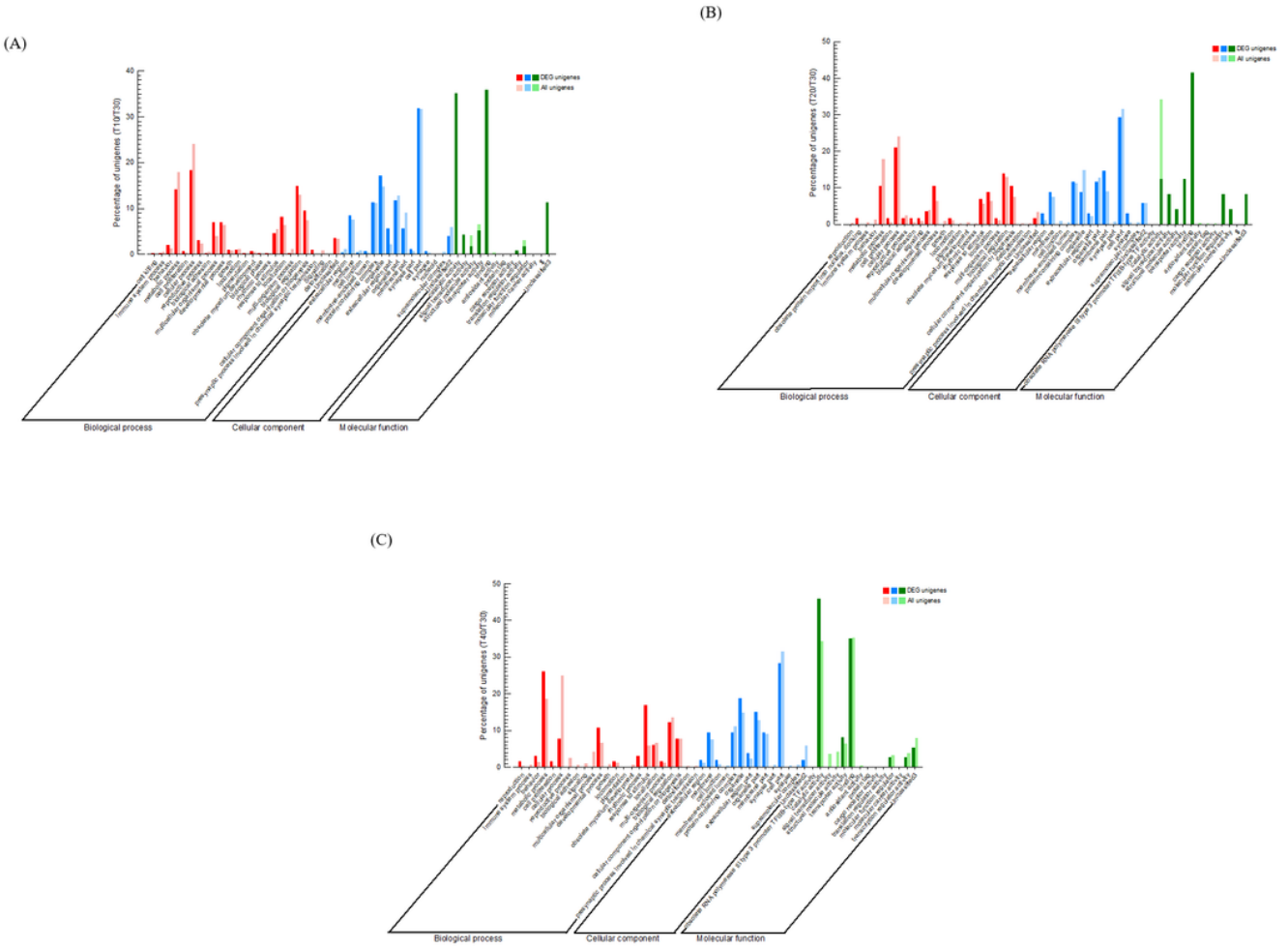

Figure 6

(A-C) Gene Ontology (GO) annotations of the specific Differentially Expressed Genes (DEGs) (p-value < 0.05, log2FC ≥ 10) co-regulated by temperature stresses including 10, 20, and 40°C, respectively, in comparison to all DEGs. Go terms are summarized in three main categories of 'Biological process', 'Cellular component', and 'Molecular function'. The Y axis shows the percentage (%) of unigenes in each category.

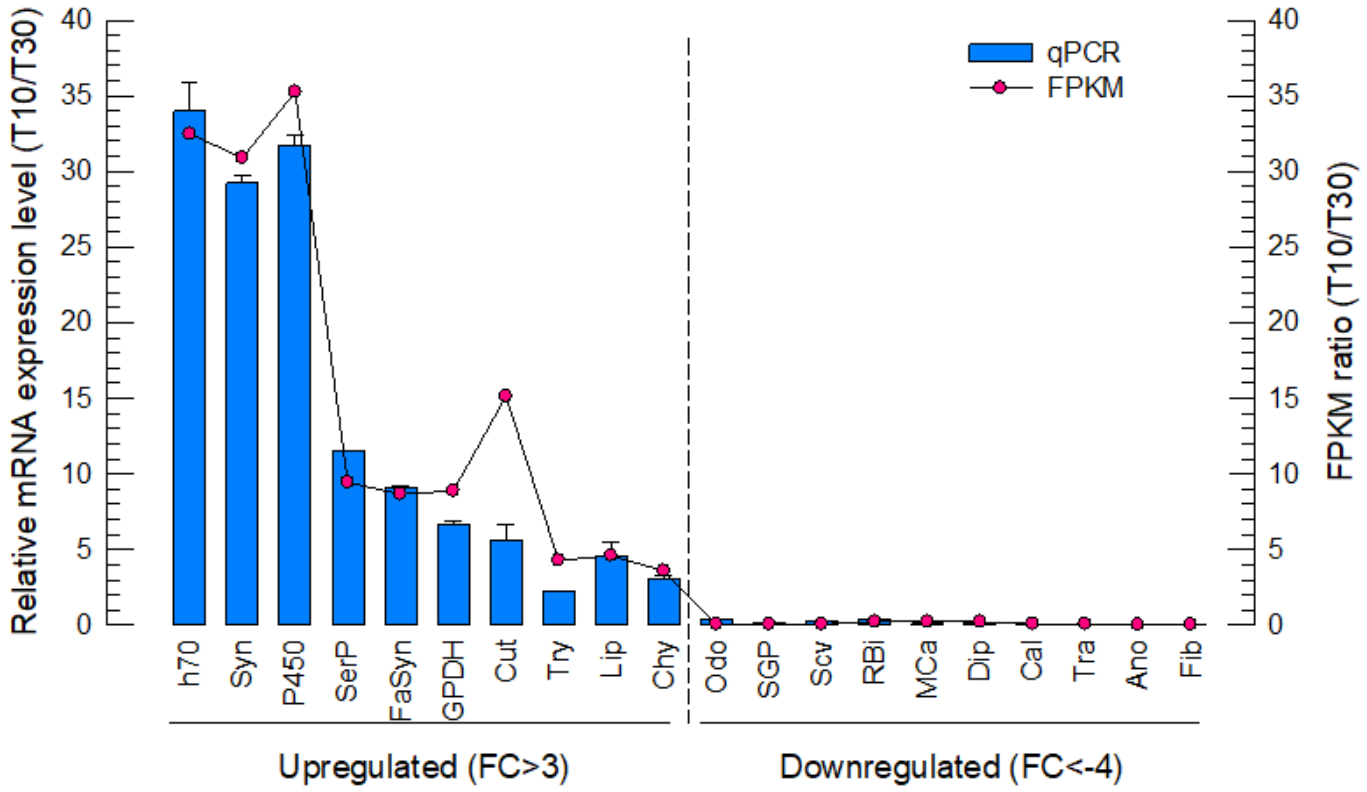

Figure 7

Differentially Expressed Genes (DEGs) validation by qRT-PCR in comparison to corresponding FPKM data detected in RNA-Seq. Relative expression level of each DEG is presented as blue bars and ratio of -log2 FPKM (T10/T30) is plotted as simple line and red scatters. Relative genes expression (T10/T30) is shown as the ratio of them per ef1_β. Each treatment was replicated three times.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformationComparisonofgeneexpressioninRIFA.docx