

Transcriptome Response Comparison Between Vector and Non-Vector Aphids After Feeding on Virus-Infected Wheat Plants

Dandan Li

Northwest Agriculture and Forestry University

Chi Zhang

Northwest Agriculture and Forestry University

Zeqian Tong

Northwest Agriculture and Forestry University

Dan Su

Northwest Agriculture and Forestry University

Gaisheng Zhang

Northwest Agriculture and Forestry University

Shize Zhang

Northwest Agriculture and Forestry University

Huiyan Zhao

Northwest Agriculture and Forestry University

Zuqing Hu (✉ huzuqing@nwsuaf.edu.cn)

<https://orcid.org/0000-0003-2496-339X>

Research article

Keywords: BYDV, WDV, vector, non-vector, transcriptome response, gene clone

Posted Date: July 1st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-37443/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on September 15th, 2020. See the published version at <https://doi.org/10.1186/s12864-020-07057-0>.

Abstract

Background: Plant viruses maintain intricate interactions with their vector and non-vector insects and can impact the fitness of insects. However, the details of their molecular and cellular mechanisms have not been studied well. We compared the transcriptome-level responses in vector and non-vector aphids (*Schizaphis graminum* and *Rhopalosiphum padi*) after feeding on wheat plants with viral infections (*Barley Yellow Dwarf Virus* (BYDV)/ *Wheat dwarf virus* (WDV)). We conducted differentially expressed gene (DEG) annotation analyses and observed DEGs related to immune pathway, growth, development, and reproduction. And we conducted cloning and bioinformatic analyses of the key DEG involved in immune.

Results: For all differentially expressed gene analyses, the numbers of DEGs related to immune, growth, development, reproduction and cuticle were higher in vector aphids than in non-vector aphids. STAT5B (signal transducer and activator of transcription 5B), which is involved in the JAK-STAT pathway, was upregulated in *R. padi* exposed to WDV. The cloning and bioinformatic results indicated that the *RpSTAT5B* sequence contains a 2082 bp ORF encoding 693 amino acids. The protein molecular weight is 79.1 kD and pI is 8.13. Analysis indicated that *RpSTAT5B* is a non-transmembrane protein and a non-secreted protein. Homology and evolutionary analysis indicated that *RpSTAT5B* was closely related to *Rhopalosiphum maidis*.

Conclusions: Unigene expression analysis showed that the total number of differentially expressed genes (DEGs) in the vector aphids was larger than that in the non-vector aphids. Functional enrichment analysis showed that the DEGs related to immunity, growth and reproduction in vector aphids were larger than those in non-vector aphids, and the differentially expressed genes related to immune were up-regulated. This study provides a basis for the evaluation of the response mechanisms of vector/non-vector insects to plant viruses.

Background

Schizaphis graminum and *Rhopalosiphum padi* are serious pests of cereal crops, particularly wheat that may cause harm to plants by feeding on them and by transmitting the *Barley yellow dwarf virus* (BYDV) (Luteoviridae: Luteovirus) in a persistent and non-proliferative manner [1–4]. That causes one of the most economically important viral diseases of cereal plants [5, 6]. In China, there were 4 isolated strains be identified as GPV, GAV, PAV and RMV and the BYDV-GAV is the major isolate [7]. *Wheat dwarf virus* (WDV) (Geminiviridae: Mastrevirus) is another serious virus of wheat in China that causes significant losses and is mainly transmitted by the leafhopper *Psammotettix alienus* (Dahlbom) in a persistent and non-proliferative manner [8, 9]. *S. graminum* is a vector of BYDV-GAV and is a non-vector of WDV, while *R. padi* is a non-vector of both BYDV-GAV and WDV.

The aphids (*S. graminum* and *R. padi*) and viruses (BYDV-GAV and WDV) used in our study are important insect pests and viral pathogens that often occur together on the same host plant in agro-ecosystems.

The plant-arthropod-pathogen interactions are complex, and some studies have shown that plant viruses can produce favorable or unfavorable effects on vector/or non-vector insects [10, 11]. Exploring the effect of viruses on aphids can provide a better understanding of these three-way interactions and improve integrated pest management.

Previous studies showed that plant virus infection of its host plant could result in either beneficial or adverse effects on its vector via direct feeding or other indirect ways [12–13]. Transcriptional induction or repression is an important mechanism for insects to regulate their innate immune responses [14]. There are a few previous reports that have studied the global transcription profiles of insect vectors fed on virus-infected plants. Brault conducted the first large-scale analysis of aphid gene regulation following virus acquisition [15]. Southern rice black-streaked dwarf virus (SRBSDV) infection activated the immune regulatory systems of the white-backed planthopper (WBPH) [16]. Transcriptome analysis of the thrips response to Tomato spotted wilt virus (TSWV) infection showed that the diversity of the innate immune-related transcripts in response to viral infection was most pronounced in the P1 stage [17]. Reproduction, embryo development and growth were associated with upregulated contigs in virus-exposed *Frankliniella fusca* adults [18].

Compared with vectors, there are fewer studies on non-vectors. Although less studied, there is some evidence that plant viruses have an impact on non-vectors. The nymph survival rate was decreased and the longevity of female adults was shortened, while egg hatchability increased in response to non-vector brown planthoppers feeding on SRBSDV-infected plants [19]. In contrast, rice plants infected by rice black-streaked dwarf virus (RBSDV) improved the ecological fitness of the brown planthopper [20]. In addition, TSWV infection reduced the fecundity and longevity of *B. tabaci* and increased the developmental time of *Bemisia tabaci* [21]. Additionally, feeding on plants inoculated with TSWV enhanced the developmental and oviposition rates of spider mites [22].

Pathogenic infections may affect the physiology of the host insect and even cause death [23]. In the process of long-term evolution, insects have formed defense mechanisms to achieve a dynamic balance between viruses and insects. Insects rely on their immune system to resist the invasion of pathogens, which is not conducive to the replication and transmission of the virus. One of the long-term goals of studying virus-insect interactions is to elucidate the molecular mechanism by which viruses affect the innate immune system of insects to avoid damage from the immune defense response. Therefore, understanding the insect's immune system is key to solving this problem. Immune-related signaling pathways that have been studied in insects include the RNAi, Toll, JAK/STAT, phagocytosis, apoptosis, proteolysis and JNK pathways [18, 24, 25].

Because the persistent plant viruses cannot break through the midgut barrier or salivary gland barrier of non-vector insects, they cannot be transmitted by the non-vector insects [26]. We previously have found that plant virus (e. g. BYDV-GAV and WDV) can be detected in non-vector aphids (e. g. *R. padi*) (unpublished data). Previous studies also have shown that plant viruses can improve the fitness of non-vector insects (as well as vectors) [20–23]. We speculate that plant viruses in non-vector aphids can also

affect their biochemical response. At present, the molecular and cytological mechanism of plant viruses to the vector and non-vector insects is not clear, so BYDV-GAV and WDV were used as the test viruses, *R. padi* and *S. graminum* were used as the test insects to compare the transcriptome changes of the vector and non-vector insects to the plant viruses, so as to find out the differences between the plant viruses affecting the gene expression of the vector and non-vector insects. This study will lay the foundation for the molecular mechanism of plant viruses to improve the fitness of non-mediator insects.

Results

Sequence assembly and annotation

To explore the transcriptome response to viral infection by vector and non-vector aphids, we performed RNA-Seq analysis on adult *S. graminum* and *R. padi* exposed to BYDV-infected, WDV-infected and uninfected wheat plants. The *S. graminum* and *R. padi* cDNA libraries were sequenced, which generated 72,092 and 68,996 unigenes, respectively (Table 1). The N50 sizes were 1,481 kb and 1,497 kb, respectively. The Q30 of each transcriptome library was above 85.74% and 85.62% for *S. graminum* and *R. padi*, respectively.

Table 1 Summary statistics for *S. graminum* and *R. padi* transcriptome assembly

Statistics	<i>S. graminum</i>	<i>R. padi</i>
Total number of transcripts	160,522	142,240
Total number of unigenes	72,092	68,996
Unigenes length >1 kb	13,758	12,588
Mean length of unigenes	833.11	805.47
N50 unigene length	1,481	1,497

Nr (non-redundant protein sequence database) homologous species distribution revealed that 71.64% of *S. graminum* sequences matched with *Acyrtosiphon pisum* (Fig. 1 A), and 67.81% of *R. padi* sequences matched with *A. pisum* (Fig. 1 B). For the Gene ontology assignment, 7263 and 8164 unigenes in *S. graminum* and *R. padi* could be annotated. For *S. graminum*, 6200 unigenes were associated with biological process, 3080 with cellular component, and 6344 with molecular function (Fig. S1 A). In *R. padi*, 7104 unigenes were associated with biological process, 3597 with cellular component, and 7173 with molecular function (Fig. S1 B). These sequences have been submitted to the NCBI Sequence Read Archive (SRA), and the accession number is PRJNA490258.

DEGs in S. graminum and R. padi, as the vector/non-vector, in response to feeding on virus-infected wheat plants

For *S. graminum* fed on BYDV-infected wheat, 1525 DEGs were identified, with 693 upregulated and 832 downregulated. There were fewer DEGs in the *S. graminum* fed on WDV-infected wheat (494), and the majority were downregulated (73.28%). There were 589 DEGs in *R. padi* exposed to BYDV-infected wheat, with 247 upregulated and 342 downregulated. A total of 343 DEGs were identified in *R. padi* exposed to WDV-infected wheat, with 75.8% downregulated. The change ratios of most DEGs were between 2^{-5} and 2^5 (Fig. 2). Gene ontology (GO) and KEGG enrichment analyses were conducted for all DEGs.

Distribution and enrichment of DEGs in vector and non-vector aphids into gene ontologies (GOs)

Gene ontology (GO) analysis of all the DEGs in *S. graminum* and *R. padi* was conducted to classify the sequences into the biological process, molecular function and cellular component categories (Fig. S2–Fig. S5). The most enriched category in both vector and non-vector aphids was metabolic process (Fig. S2–Fig. S5). Among all comparisons, the differentially expressed genes contained a high proportion of genes involved in development, metabolism, reproduction and growth, and the number in vector aphids was larger (29 DEGs in Sg-BYDV) than that in non-vector aphids (13, 13, and 3 DEGs in Sg-WDV, Rp-BYDV, and Rp-WDV, respectively) (Fig. 3 and Table S5). In addition, genes associated with the cuticle (structural constituent of the cuticle and structural constituent of the chitin-based cuticle) were also differentially expressed, and there were 16 DEGs in the vector treatment (Sg-BYDV), while there were fewer in the non-vector treatments (with 2 in Sg-WDV, 5 in Rp-BYDV, and 2 in Rp-WDV). In Sg-BYDV, there were 15 DEGs associated with the cuticle, and the majority of them were downregulated (14 downregulated) (Fig. 3 and Table 2). Several genes involved in the cytoskeleton were also differentially expressed, with 6, 1, and 2 related DEGs in Sg-BYDV, Sg-WDV, and Rp-BYDV, respectively (Fig. 3 and Table 3). The results of GO enrichment analysis showed that the number of categories of DEGs enriched in vector aphids was larger than that in non-vector aphids (Fig. S2–S5). The COG enrichment analysis showed a similar pattern in that the number of categories in the vector aphids was greater than that in the non-vector aphids (Fig. S6–S9).

Table 3 DEGs involved in cytoskeleton

	Gene ID	Annotation	log2FC	FDR
Sg-BYDV	Sg53134	Cytoskeleton	1.098	3.88E-08
	Sg55484	Cytoskeleton	-2.175	3.45E-05
	Sg32025	Cytoskeleton	Inf	8.62E-08
	Sg56027	Cytoskeleton	-1.346	0.000376
	Sg49676	Cytoskeleton	-1.527	0.008734
	Sg55690	cytoskeleton	-1.879	5.80E-18
Sg-WDV	Sg55690	Cytoskeleton	-2.033	7.56E-07
Rp-BYDV	Rp29892	Cytoskeleton	-4.719	2.19E-05
	Rp50802	Cytoskeleton	-3.857	0.000136

KEGG pathways analysis

To further understand the metabolic pathways that contained differentially expressed genes, KEGG significant enrichment analysis was performed on the differentially expressed genes between virus-treated vector and non-vector aphids. Among all of the DEGs, there were 287, 74, 91 and 27 DEGs that were assigned to KEGG pathways in Sg-BYDV, Sg-WDV, Rp-BYDV and Rp-WDV, respectively (Fig. S10–S13). The DEGs of the richest and shared pathways were associated with metabolism, including translation and fatty acid and carboxylic acid metabolic pathways. Genes involved in energy production pathways may be important for immune defense responses [27]. In the KEGG pathway analyses, we also identified DEGs involved in immune pathways.

Immune-related DEGs of vector and non-vector aphids feeding on virus-infected wheat

Feeding on virus-infected plants resulted in DEGs related to the innate immunity pathway being upregulated. The DEGs related to immunity included several members of the pathway of the lysosomes, the JAK-STAT signaling pathway, antigen processing and presentation, ubiquitin mediated proteolysis, and the peroxisome (Table 4). For the DEG analyses, we identified 12 DEGs involved in immune pathways in the vector aphids (Sg-BYDV), the majority of which were upregulated (11 DEGs). For the non-vector aphids, there were fewer DEGs related to immunity, and only one DEG involved in immunity was identified in each of the Sg-WDV, Rp-BYDV, and Rp-WDV conditions (Table 4).

RT-qPCR

RT-qPCR was used to confirm the expression of 13 selected DEGs of *S. graminum* and *R. padi*. These genes were selected because we were interested in their function, since they were involved in immunity, development, growth and reproduction. The 13 genes included eight from the vector aphids (Sg35675, Sg36590, Sg40066, Sg43786, Sg48920, Sg51592, Sg53134, and Sg56027) and five from the non-vector aphids (Sg30610, Sg35800, Rp22910, Rp22931, and Rp46740). The results of RT-qPCR showed that all 13 selected genes showed the same expression patterns (upregulated or downregulated) as the RNA-Seq analyses (Fig. 4).

Cloning and characterization of STAT5B in R. padi

For the immunity related genes analysis, STAT5B (signal transducer and activator of transcription 5B) was the only related gene that was upregulated in *R. padi* fed on WDV infected wheat. STAT5B is a key gene involved in the JAK-STAT pathway, which is important for the innate immune response of insects. Based on the sequence of the transcriptome, we cloned STAT5B from *R. padi* (designated as *RpSTAT5B*). The cloned product was validated using RT-PCR and sequencing. The sequencing results of the *RpSTAT5B* individual isolated colonies confirmed a single transcript identical to that found in the transcriptome sequencing. We have submitted the sequence of STAT5B to the NCBI GenBank database (accession number MK931299). The cloned product of *RpSTAT5B* is 2364 bp, and it has a complete open reading frame of 2082 bp encoding 693 amino acids, with a molecular weight of 83.2 kDa and a theoretical isoelectric point of 8.31.

Multiple protein sequence alignment of *RpSTAT5B* with six STAT5B proteins from other insects was conducted. Structural analysis showed that the *RpSTAT5B* protein has the typical structural features of the STAT5B family (Fig. 5). Phylogenetic analysis indicated that the amino acid sequence of the *RpSTAT5B* protein showed high identity to that of the corresponding proteins from other aphids of Homoptera, particularly *Rhopalosiphum maidis* (Fig. 6). Both the RT-qPCR and transcriptome results showed that *RpSTAT5B* was upregulated in *R. padi* fed on WDV infected wheat (Fig. 4), suggesting that *RpSTAT5B* is critical to the response to feeding on infected wheat plants. However, the mechanism of *RpSTAT5B* in aphid immune defense remains to be further studied.

Discussion

Our study investigated the effects of plant viruses on vector and non-vector insects at the transcriptional level. The number of DEGs was greater in the vector aphids (Sg-BYDV with 1525) than in the non-vector aphids (Sg-WDV, Rp-BYDV, and Rp-WDV with 494, 589 and 343, respectively). For all DEGs, the analysis showed that many of the DEGs were involved in development, growth and reproduction. The numbers of genes related to these processes were higher in the vectors (Sg-BYDV) than in the non-vectors (Sg-WDV, Rp-BYDV, and Rp-WDV). For the GO analyses of all DEGs, the genes associated with the cuticles were mostly downregulated. Particularly in the vector aphids (Sg-BYDV), there were 16 genes associated with cuticles that were differentially expressed, among which 15 genes were downregulated. Moreover, fewer

related DEGs were identified in the non-vectors, i.e., two DEGs in Sg-WDV, five DEGs in Rp-BYDV, and two DEGs in Rp-WDV. The cuticle is of great importance for the survival of insects, since it is the articulated exoskeleton that protects the body against the invasion of pathogens [28, 29].

Insects rely on their immune system to fight against pathogens [30]. After feeding on virus-infected wheat plants, the DEGs related to immunity in *S. graminum* and *R. padi* were upregulated, including the JAK-STAT signaling pathway, the lysosome, antigen processing and presentation, ubiquitin mediated proteolysis and the peroxisome. As shown in our results, the number of DEGs related to immunity in *S. graminum* exposed to BYDV was higher than that in *S. graminum* exposed to WDV and in *R. padi* exposed to BYDV and WDV. These results suggest that feeding on virus-infected plants has a greater effect on the immune system of vector insects than that of non-vector insects.

The DEGs involved in the cytoskeleton were also differentially expressed, which may be related to the immune response [31]. There have been previous studies showing that viruses can interact with and reorganize host cytoskeleton components for intercellular trafficking and infection processes [32–35]. In addition, the cytoskeleton is also commonly involved in the intracellular transport of viruses [36–39].

There were 3 DEGs upregulated in Sg-BYDV that were related to ubiquitin mediated proteolysis, which is important for insect defenses against pathogens [40]. In this pathway, we identified that UBE2D (ubiquitin-conjugating enzyme E2 D), BIRC2_3 (baculoviral IAP repeat-containing protein 2/3) and PIAS1 (E3 SUMO-protein ligase PIAS1) were upregulated. The E2 enzyme is the ubiquitin carrier protein or ubiquitin-conjugating enzyme that can transfer ubiquitin from E1 to the substrate [41]. Inhibitors of apoptosis proteins (IAPs), also known as baculovirus IAP repeat (BIR)-containing proteins (BIRCs), generally display anti-apoptotic properties when overexpressed [42]. PIAS1 has been recognized as a small ubiquitin-like modifier (SUMO) ligase [43].

The JAK-STAT signaling pathway is important for innate immunity and antiviral responses in insects [44–47]. This pathway consists of the ligand unpaired (UPD), the receptor Domeless, JAK, STAT, protein inhibitor of activated STAT (PIAS), suppressor of cytokine signaling (SOCS), and signal transducing adaptor molecule (STAM) [40, 48, 49]. We found that PIAS1 (Sg55104) was up-regulated in the Sg-BYDV aphids; this gene encodes E3 SUMO-protein ligase PIAS1. A previous study has shown that PIAS1 is a specific inhibitor of Stat1-mediated gene activation [50].

STAT5B was the only related gene that was differentially expressed in *R. padi* fed on WDV infected wheat. *RpSTAT5B* possesses the SH2 (Src-homology Domain), STAT5_CCD (Coiled-coiled Domain), and the STAT_DBD (STAT_bind) conserved domains, which are the characteristic conserved domains of STAT5 (Fig. 5) [51-55]. For prediction of its three-dimensional structure, the SWISS-MODEL template library (SMTL version 2019-06-27, PDB release 2019-06-21) was searched with BLAST and HHblits for evolutionary related structures matching *RpSTAT5B*. The predicted three-dimensional structure of *RpSTAT5B* is shown in Fig. 7, and its identity with the template (SMTL ID: 1y1u.1.A) is 45.47%. Overall, the identification of *RpSTAT5B*, and its increased expression in *R. padi* fed on WDV infected wheat,

suggests it may fulfill a specific physiological role. Future studies are needed to understand the function of *RpSTAT5B* in *R. padi* fed on WDV infected wheat.

Insect pests and viral diseases are important factors affecting wheat yield. In this study, we analyzed the response of vector/non-vector aphids to wheat virus at the transcriptional level. Among all DEGs analyzed, we identified several genes associated with immunity, growth, development, reproduction and the cuticle of aphids. The results revealed that the number of DEGs and the functions related to the DEGs were both larger in vector aphids than that in non-vector aphids. This may imply that wheat viruses have a larger impact on vector than non-vector aphids. Exposure to the virus activated the immunity response of the aphids, particularly vector aphids. These results will provide a reference for investigation of new methods to improve the efficiency of pest control, including prokaryotic expression and gene silencing.

Methods

Insect rearing and plant infection by the viruses

A single parthenogenetic female adult from each of the species *S. graminum* and *R. padi* was collected from Yangpingguan (34.85°N, 105.63°E), Shaanxi Province, and they were maintained on wheat seedlings in growth chambers kept at $20 \pm 1^\circ\text{C}$, $65\% \pm 5\%$ relative humidity and a photoperiod of 16 : 8 h (L : D). BYDV (isolate BYDV-GAV) and WDV were collected from Yangling (34.28°N, 108.22°E) and Hancheng (35.47°N, 110.45°E), Shaanxi Province, China, from winter wheat plants. The viruses were identified using RT-PCR and agarose gel electrophoresis using the methods briefly described below.

We isolated total RNA from the wheat using TRIzol reagent (TaKaRa, Japan) following the manufacturers' instructions to test whether wheat was infected by BYDV. The cDNA was synthesized using PrimeScript RT Master Mix (TaKaRa, Japan). We isolated total gDNA from the wheat using DNase Plant Kit (TIANGEN, China) following the manufacturers' instructions to test whether wheat was infected by WDV. Then PCR was performed; the primers used for the virus identification are shown in Table S1. The PCRs consisted of the following: 6.5 μl of Reaction Mix (TIANGEN, China), 1 μl of each forward and reverse primer, 1 μl of cDNA/gDNA and 4.5 μl of ddH₂O. The cycling conditions were 94°C for 3 min; 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min; and finally, 72°C for 5 min. Then, agarose gel electrophoresis was performed, and we sequenced the PCR products to confirm that the wheat was infected by BYDV or WDV. The detailed methods of the BYDV-infected and WDV-infected plant treatments, as transmitted by the leafhopper and *S. graminum*, respectively, have been previously described as follows [56]. Wheat plants at the second leaf stage were first covered with transparent, plastic, tube-shaped cages (30 cm in height, 13.5 cm in diameter, and with a mesh screen cover on the top), and then 5 late instar nymphs of the vector were introduced into the cage and allowed to feed on the leaves. All the nymphs were removed after 5 days, and the treated plants were maintained under the growth chamber described as above.

Feeding assays and RNA isolation

The young first instar nymphs of *S. graminum* and *R. padi* were fed on BYDV-infected, WDV-infected, and healthy wheat plants to achieve six experimental treatments: the vector treatment was defined as *S. graminum* fed on BYDV-infected wheat (Sg-BYDV); the non-vector treatment was defined as *S. graminum* fed on WDV-infected wheat and *R. padi* fed on BYDV-GAV-infected, WDV-infected wheat (Sg-WDV, Rp-BYDV, Rp-WDV); and the control treatments were *S. graminum* and *R. padi* fed on healthy wheat seedlings (Sg-ck, Rp-ck).

Three biological replications of each treatment were conducted and each replicate contained 20 individual aphids. For each treatment-replicate, they were maintained in a growth chamber under the same environmental conditions as described above. We collected 20 adult aphids into a 1.5 ml tube, and they were immediately flash frozen in liquid nitrogen and stored at -80 °C. Total RNA was extracted from each sample using TRIzol reagent (TaKaRa, Japan) following the manufacturers' instructions. RNA degradation and contamination were monitored on 1% agarose gels. RNA purity was checked using the NanoPhotometer® spectrophotometer (IMPLEN, CA, USA). RNA concentration was measured using a Qubit® RNA Assay Kit in a Qubit®2.0 Fluorometer (Life Technologies, CA, USA). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

Transcriptome sequencing and analysis

A total amount of 3 µg of RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following the manufacturer's recommendations. Briefly, mRNA was purified from the total RNA using poly-T oligo-attached magnetic beads. Fragmentation was carried out, and the first and second strands were synthesized. The purified and adaptor-ligated cDNA was subjected to PCR amplification. Finally, the PCR products were purified (AMPure XP system), and the library quality was assessed on the Agilent Bioanalyzer 2100 system.

The clustering of the index-coded samples was performed on a cBot Cluster Generation System using a TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina HiSeq 2000 platform and paired-end reads were generated. Clean data were obtained by removing reads containing adapters, reads containing poly-N and low-quality reads. At the same time, the Q20, Q30, GC-content and sequence duplication level of the clean data were calculated. All of the downstream analyses were based on clean data with high quality. Transcriptome assembly was accomplished using Trinity (r20131110) with min_kmer_cov set to 2 by default and all other parameters set to default [57].

Differential expression analysis of the two groups was performed using the DESeq R package (1.10.1). DESeq provides statistical routines for determining differential expression of digital gene expression data

using a model based on the negative binomial distribution. The resulting P values were adjusted using the Benjamini and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted P-value <0.05 found by DESeq were considered differentially expressed.

Gene ontology (GO) enrichment analysis of the differentially expressed genes (DEGs) was implemented by the topGO R packages based on the Kolmogorov–Smirnov test. We used KOBAS software to test the statistical enrichment of the differentially expressed genes in the KEGG pathways. Genes with similar expression patterns are usually functionally related. Cluster 3.0 software was used to analyze differentially expressed genes by hierarchical clustering with Euclidean distance as the distance matrix. The clustering results were displayed by Java Treeview.

RT-qPCR validation

To validate differential expression in response to the viruses, 13 transcript sequences were selected in *S. graminum* and *R. padi* and then compared with the expression levels between the virus-exposed and non-virus-exposed aphids. Actin was used as an internal reference gene to normalize the expression level. The primers were designed using Primer Premier 5.0 software and are shown in Table S2. The total RNA extraction and cDNA synthesis methods were performed as described above. Total RNA was isolated from three biological replicates, and each targeted gene included three technical replicates.

The reactions of real-time RT-qPCR were performed in 25 µl volumes using TB Green Premix Ex Taq (TaKaRa, Japan) with 12.5 µl of TB Green Premix Ex Taq II, 1 µl of each forward and reverse primer (10 µM), 0.5 µl of ROX Reference Dye II, 2 µl of cDNA and 8 µl of nuclease free water. The cycling parameters were as follows: 95°C for 30 s and 40 cycles of 95°C for 5 s and 60°C for 30 s. A standard curve experiment was set up by performing a dilution series with 5⁻¹ dilution times and 5 dilution points to evaluate the PCR efficiency (E). The Q-PCR was performed on a QuantStudio 3 Real-Time PCR System (Applied Biosystems), and QuantStudio™ Design and Analysis Software was used to analyze the PCR assays. The relative quantitative data analysis was performed using the 2^{-ΔΔct} method.

Cloning and characterization of the STAT5B

STAT5B (signal transducer and activator of transcription 5B) was the only relevant gene that was differentially expressed in *R. padi* fed on WDV infected wheat. To understand the mechanism of STAT5B in the insect immune defense response, we designed primers using Primer Premier 5.0 to clone STAT5B based on the sequence found in the transcriptome (Table S3). PCR analyses were conducted on a PCR Amplifier (purchased from Eppendorf, Yangling, China) using Tks Gflex™ DNA Polymerase (purchased from TaKaRa, Yangling, China). Cloned products were validated by RT-PCR and sequencing. Multiple sequence alignments were performed using Clustal W and DNAMAN. The phylogenetic tree was generated in Mega 6 using the Neighbor-Joining method. All sequences used in the analysis are listed in

Table S4. The protein structure model was predicted using WISS-MODEL (<https://swissmodel.expasy.org/>) and visualized in Pymol (v1.3r1).

Abbreviations

BYDV: barley yellow dwarf virus; WDV: Wheat dwarf virus; *S. graminum*: *Schizaphis graminum*; *R. padi*: *Rhopalosiphum padi*; DEGs: differentially expressed genes; Sg-BYDV: *Schizaphis graminum* feeding on wheat infected by BYDV; Sg-WDV: *Schizaphis graminum* feeding on wheat infected by WDV; Rp-BYDV: *Rhopalosiphum padi* feeding on wheat infected by BYDV; Rp-WDV: *Rhopalosiphum padi* feeding on wheat infected by WDV; STAT5B: signal transducer and activator of transcription 5B; TSWV: Tomato spotted wilt virus; SRBSDV: Southern rice black-streaked dwarf virus; WBPH: white backed planthopper; RBSDV: rice black streaked dwarf virus; RT-qPCR: quantitative real-time PCR; GO: gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes.

Declaration

Acknowledgments

The authors are grateful to Dr. X. Ni, USDA-ARS Tifton, Georgia, USA and Dr. Robert L. Stoltz, University of Idaho, for their critical comments on earlier versions of this manuscript and the anonymous referees for their critical comments on this manuscript.

Authors' contributions

ZQH, HYZ and DDL designed the experiment. DDL, ZQT, CZ, and DS collected samples. DDL and ZQH performed the RNAseq and analyzed the data. The manuscripts were written and revised by ZQH, DDL, GSZ, SZZ and HYZ. All authors have read and approved the final manuscript.

Funding

The Project Supported by the National Key Research and Development Program of China (2018YFD0200402), the National Natural Science Foundation of China (Program No. 31401735), and Natural Science Basic Research Plan in Shaanxi Province of China (Program No. 2017JM3030).

Availability of data and materials

Transcriptome data are available at NCBI with raw sequencing reads deposited in the Sequence Read Archive (accession # PRJNA490258). We have submitted the sequence of *RpSTAT5B* to NCBI GeneBank database and the accession number is MK931299.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹State Key Laboratory of Crop Stress Biology in Arid Areas, College of Plant Protection, Northwest A&F University, Yangling, China

²Ministry of Education, Key Laboratory of Crop Heterosis of Shaanxi Province, National Yangling Agricultural Biotechnology and Breeding Center, Yangling Branch of State Wheat Improvement Centre/Wheat Breeding Engineering Research Center, Northwest A&F University, Yangling, China

References

1. Al-Mousawi AH, Richardson PE, Burton Ultrastructural studies of greenbug (Hemiptera: Aphididae) feeding damage to susceptible and resistant wheat cultivars. *Annals of the Entomological Society of America*. 1983;76:964–971.
2. Song S, Tao Y, Zhang H, Wu Y. *Psathyrostachys huashanica*, a potential resource for resistance to barley yellow dwarf virus-gav. *European Journal of Plant Pathology*. 2013;137:217–221.
3. Zhang QF, Guan WN, Ren ZY, Zhu XS. Transmission of barley yellow dwarf virus strains from northwestern China by four aphid species. *Plant Disease*. 1983;67:895–899.
4. Nuessly GS, Nagata RT, Burd JD, Hentz MG, Carroll AS, Halbert SE. Biology and biotype determination of greenbug, *Schizaphis graminum* (hemiptera: aphididae), on seashore paspalum turfgrass (*paspalum vaginatum*). *Environmental Entomology*. 2008;37:586–591.
5. Hu ZQ, Zhao HY, Thieme T. Comparison of the potential rate of population increase of brown and green color morphs of *Sitobionavenae* (Homoptera: Aphididae) on barley infected and uninfected with Barley yellow dwarf virus. *Insect Science*. 2014;21:326–333.
6. Song S, Tao Y, Zhao L, Liu P, Wu Y. Isolation and characterization of the complete genome of Barley yellow dwarf virus-GAV from wheat in Northwestern China. *Cereal Research Communications*. 2014;1:1–9.

7. Zhou G, Cheng Z, Qian Y, Zhang X, Rochow Serological identification of luteoviruses of small grains in China. *Plant Disease*. 1984;68:710–713.
8. Liu Y, Wang B, Gyula V, Mónica CK, Wu BL, Wu YH, Wang Genomic analysis of the natural population of Wheat dwarf virus in wheat from China and Hungary. *Journal of integrative agriculture*. 2012;11:2020–2027.
9. Xie J, Wang X, Liu Y, Peng Y, Zhou G. First report of the occurrence of wheat dwarf virus in wheat in china. *Plant Disease An International Journal of Applied Plant Pathology*. 2007;91:111–111.
10. Fereres A, Moreno Behavioural aspects influencing plant virus transmission by homopteran insects. *Virus Research*. 2009;141:158–168.
11. Hohn T. Plant virus transmission from the insect point of view. *Proceedings of the National Academy of Sciences of the United States of America*. 2007;104:17905–17906.
12. Colvin J, Omongo CA, Govindappa MR, Stevenson PC, Maruthi MN, Gibson G. Seal SE, Muniyappa Host-plant viral infection effects on arthropod-vector population growth, development and behaviour: management and epidemiological implications. *Advances in Virus Research*. 2006;67:419–452.
13. Stout MJ, Thaler JS, Thomma BPHJ. Plant-mediated interactions between pathogenic microorganisms and herbivorous arthropods. *Annual Review of Entomology*. 2006;51:663–689.
14. Baton LA, Garver L, Xi Z, Dimopoulos G. Functional genomics studies on the innate immunity of disease vectors. *Insect Science*. 2008;15:15–27.
15. Brault V, Tanguy S, Reinbold C, Le Trionnaire G, Arneodo J, Jaubert-Possamai S, Guernec G, Tagu D. Transcriptomic analysis of intestinal genes following acquisition of pea enation mosaic virus by the pea aphid *Acyrtosiphon pisum*. *Journal of General Virology*. 2010;91:802–808.
16. Xu Y, Zhou W, Zhou Y, Wu J, Zhou X. Transcriptome and comparative gene expression analysis of *Sogatella furcifera* (Horváth) in response to Southern rice black-streaked dwarf virus. *Plos One*. 2012;7:e36238.
17. Schneweis DJ, Whitfield AE, Rotenberg D. Thrips developmental stage-specific transcriptome response to tomato spotted wilt virus during the virus infection cycle in *Frankliniella occidentalis*, the primary vector. *Virology*. 2017;500:226–237.
18. Shrestha A, Champagne DE, Culbreath AK, Rotenberg D, Srinivasan R. Transcriptome changes associated with tomato spotted wilt virus infection in various life stages of its thrips vector, *Frankliniella fusca* (Hinds). *Journal of General Virology*. 2017;98:2156–2170.
19. Xu HX, He XC, Zheng XS, Yang YJ, Zhang JF, Lu ZX. Effects of SRBSDV-infected rice plants on the fitness of vector and non-vector rice planthoppers. *Journal of Asia-Pacific Entomology*. 2016;19:707–710.
20. Xu H, He X, Zheng X, Yang Y, Lu Z. Influence of rice black streaked dwarf virus on the ecological fitness of non-vector planthopper *Nilaparvata lugens* (Hemiptera: Delphacidae). *Insect Science*. 2014;21:507–514.
21. Pan H, Chen G, Li F, Wu Q, Wang S, Xie W, Liu B, Xu B, Zhang Y. Tomato spotted wilt virus infection reduces the fitness of a nonvector herbivore on pepper. *Journal of Economic Entomology*.

- 2013;106:924–928.
22. Belliure B, Sabelis MW, Janssen A. Vector and virus induce plant responses that benefit a non-vector herbivore. *Basic and Applied Ecology*. 2010;11:162–169.
 23. Hoffmann JA. Innate immunity of insects. *Current Opinion in Immunology*. 1995;7:4–10.
 24. Badillovargas IE; Rotenberg D. Schneeweis DJ, Hiromasa Y, Tomich JM, Whitfield AE. Proteomic analysis of *Frankliniella occidentalis* and differentially expressed proteins in response to tomato spotted wilt virus infection. *Journal of Virology*. 2012;86:8793.
 25. Medeiros RB, Resende RDO, De Avila AC. The plant virus tomato spotted wilt tospovirus activates the immune system of its main insect vector, *Frankliniella occidentalis*. *Journal of Virology*. 2004;78:4976–4982.
 26. Filho FMA, Stavisky J, Reitz SR, Deom CM, Sherwood JL. Midgut infection by tomato spotted wilt virus and vector incompetence of *Frankliniella tritici*. *Journal of Applied Entomology*. 2005;129:548-550.
 27. Schmid-Hempel P. Evolutionary ecology of insect immune defenses. *Annual Review of Entomology*. 2005;50:529–551.
 28. Tang L, Liang J, Zhan Z, Xiang Z, He N. Identification of the chitin-binding proteins from the larval proteins of silkworm, *Bombyx mori*. *Insect Biochemistry and Molecular Biology*. 2010;40:228–234.
 29. Gallot A, Rispe C, Leterme N, Gauthier JP, Jaubert-Possamai S, Tagu D. Cuticular proteins and seasonal photoperiodism in aphids. *Insect Biochemistry and Molecular Biology*. 2010;40:235e240.
 30. Wang L, Tang N, Gao X, Guo D, Chang Z, Fu Y, Akinyemi IA, Wu. Understanding the immune system architecture and transcriptome responses to southern rice black-streaked dwarf virus in *Sogatella furcifera*. *Scientific Reports*. 2016;6:36254.
 31. Steinert S, Levashina EA. Intracellular immune responses of dipteran insects. *Immunological reviews*. 2011;240:129–140.
 32. Radtke K, Döhner K, Sodeik B. Viral interactions with the cytoskeleton: a hitchhiker's guide to the cell, *Cell Microbiol*. 2006;8:387–400.
 33. Simpson-Holley M, Ellis D, Fisher D, Elton D, Mccauley J, Digard P. A functional link between the actin cytoskeleton and lipid rafts during budding of filamentous influenza virions. *Virology*. 2002;301:212–225.
 34. Taylor MP, Koyuncu OO, Enquist LW. Subversion of the actin cytoskeleton during viral infection. *Nature Reviews Microbiology*. 2011;9:427–439.
 35. Shrestha A, Champagne DE, Culbreath AK, Rotenberg D, Srinivasan R. Transcriptome changes associated with Tomato spotted wilt virus infection in various life stages of its thrips vector, *Frankliniella fusca* (Hinds). *Journal of General Virology*. 2017;98:2156–2170.
 36. Laporte C, Vetter G, Loudes AM, Robinson DG, Ritzenthaler C. Involvement of the secretory pathway and the cytoskeleton in intracellular targeting and tubule assembly of Grapevine fanleaf virus movement protein in tobacco BY-2 cells. *Plant Cell*. 2003;15:2058–2075.

37. Harries PA, Schoelz JE, Nelson RS. Intracellular transport of viruses and their components: utilizing the cytoskeleton and membrane highways. *Molecular Plant-Microbe Interactions*. 2010;23:1381–1393.
38. Schoelz JE, Harries PA, Nelson RS. Intracellular transport of plant viruses: Finding the door out of the cell. *Molecular Plant*. 2011;4:813–831.
39. Cassone BJ, Wijeratne S, Michel AP, Stewart LR, Chen Y, Yan P, Redinbaugh MG. Virus-independent and common transcriptome responses of leafhopper vectors feeding on maize infected with semi-persistently and persistent propagatively transmitted viruses. *BMC Genomics*. 2014;15:133.
40. Lemaitre B, Hoffmann J. The host defense of *Drosophila melanogaster*. *Annual Review of Immunology*. 2007;25:697–743.
41. Gonen H, Bercovich B, Orian A, Carrano A, Takizawa C, Yamanaka K, Pagano M, Iwai K, Ciechanover A. Identification of the Ubiquitin Carrier Proteins, E2s, Involved in Signal-induced Conjugation and Subsequent Degradation of I κ B α . *Journal of biological chemistry*. 1999;274:14823–14830.
42. Dubrez-Daloz L, Dupoux A, Cartie J. IAPs: more than just inhibitors of apoptosis proteins. *Cell Cycle*. 2008;7:1036–1046.
43. Kadaré G, Toutant M, Formstecher E, Corvol JC, Carnaud M, Bouterin MC, Girault JA. PIAS1-mediated sumoylation of focal adhesion kinase activates its autophosphorylation. *The journal of biological chemistry*. 2003;278:47434–47440.
44. Dostert C, Jouanguy E, Irving P, Troxler L, Galiana-Arnoux D, Hetru C, Hoffmann JA, Imler The Jak-STAT signaling pathway is required but not sufficient for the antiviral response of *drosophila*. *Nature Immunology*. 2005;6:946–953.
45. Hedges LM, Johnson KN. Induction of host defence responses by *Drosophila C* virus. *Journal of General Virology*. 2008;89:1497–1501.
46. Kingsolver MB, Huang Z, Hardy RW. Insect antiviral innate immunity: pathways, effectors, and connections. *Journal of Molecular Biology*. 2013;425:4921–4936.
47. Zhang Z, Zhang P, Li W, Zhang J, Huang F, Yang J, Bei Y, Lu Y. De novo transcriptome sequencing in *Frankliniella occidentalis* to identify genes involved in plant virus transmission and insecticide resistance. *Genomics*. 2013;101:296–305.
48. Hou S, Zheng Z, Chen X, Perrimon N. The JAK/STAT pathway review in model organisms: emerging roles in cell movement. *Developmental Cell*. 2002;3:765–778.
49. Kaneko T, Yano T, Aggarwal K, Lim JH, Ueda K, Oshima Y, Peach C, Erturk-Hasdemir D, Goldman WE, Oh B, Kurata S Silverman PGRP-LC and PGRP-LE have essential yet distinct functions in the *drosophila* immune response to monomeric DAP-type peptidoglycan. *Nature Immunology*. 2006;7:715–723.
50. Liu B, Liao J, Rao X, Kushner SA, Chung CD, Chang DD, Shuai Inhibition of Stat1-mediated gene activation by PIAS1. *Proceedings of the National Academy of Sciences*. 1998;95:10626–10631.
51. Boucheron CA. Single Amino Acid in the DNA Binding Regions of STAT5A and STAT5B Confers Distinct DNA Binding Specificities. *Journal of Biological Chemistry*. 1998;273:33936–33941.

52. Zhang T, Kee WH, Seow KT, Fung W, Cao X The coiled-coil domain of STAT3 is essential for its SH2 domain-mediated receptor binding and subsequent activation induced by epidermal growth factor and interleukin-6, *Molecular and Cellular Biol* 2000;20:7132–7139.
53. Wakao H, Gouilleux F, Groner B. Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. *The EMBO Journal*. 1995;14:854–855.
54. Shuai K, Horvath CM, Huang LHT, Qureshi SA., Cowburn D, Darnell JE. Interferon activation of the transcription factor stat91 involves dimerization through sh2-phosphotyrosyl peptide interactions. 1994;76:821–828.
55. Kisseleva T, Bhattacharya S, Braunstein J, Schindler CW. Signaling through the jak/stat pathway, recent advances and future challenges. *Gene (Amsterdam)*. 2002;285:0–24.
56. Hu Z, Zhao H, Thieme T. Modification of non-vector aphid feeding behavior on virus-infected host plant. *Journal of Insect Science*. 2013;13:1–11.
57. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Nir Friedman, Regev Full length transcriptome assembly from RNA Seq data without a reference genome. *Nature Biotechnology* Italic. 2011;29:644–652.

Tables

Table 2 DEGs involved in the cuticle in vector (Sg-BYDV) and non-vector (Sg-WDV, Rp-BYDV, Rp-WDV and Rp-WDV) aphids

	Gene ID	Annotation	log2FC	FDR
Sg-BYDV	Sg39301	structural constituent of cuticle	-1.646	8.07E-05
	Sg27246	structural constituent of cuticle	-1.028	5.62E-07
	Sg19143	structural constituent of cuticle	-1.872	6.79E-08
	Sg56253	structural constituent of cuticle	-2.454	7.80E-14
	Sg17654	structural constituent of cuticle	2.147	1.02E-06
	Sg27360	structural constituent of cuticle	-1.646	1.08E-15
	Sg56455	structural constituent of cuticle	-1.722	0.001659
	Sg56015	structural constituent of cuticle	-1.178	0.000655
	Sg35648	structural constituent of cuticle	-1.328	1.59E-08
	Sg53861	structural constituent of cuticle	-1.740	0.00481
	Sg47817	structural constituent of cuticle	-2.061	0.002456
	Sg36013	structural constituent of cuticle	-1.347	0.001048
	Sg54274	structural constituent of cuticle	-1.190	7.25E-12
	Sg55542	structural constituent of cuticle	-1.864	0.000179
	Sg30508	structural constituent of chitin-based cuticle	-1.155	0.001031
Sg-WDV	Sg17654	structural constituent of cuticle	2.458	2.12E-15
	Sg54274	structural constituent of cuticle	-1.892	0.00097
Rp-BYDV	Rp23024	structural constituent of cuticle	2.045	0.000696
	Rp23495	structural constituent of cuticle	-2.608	0.000335
	Rp37328	structural constituent of cuticle	-2.086	9.10E-05
	Rp18153	structural constituent of cuticle	4.732	8.00E-13
	Rp22931	structural constituent of chitin-based cuticle	3.899	1.19E-12
Rp-WDV	Rp48839	structural constituent of cuticle	2.603	0.001872
	Rp18153	structural constituent of cuticle	3.776	0.001354

Annotation: determined by BLAST

Table 4 Immune-related DEGs expressed in virus-infected *S. graminum* and *R. padi*

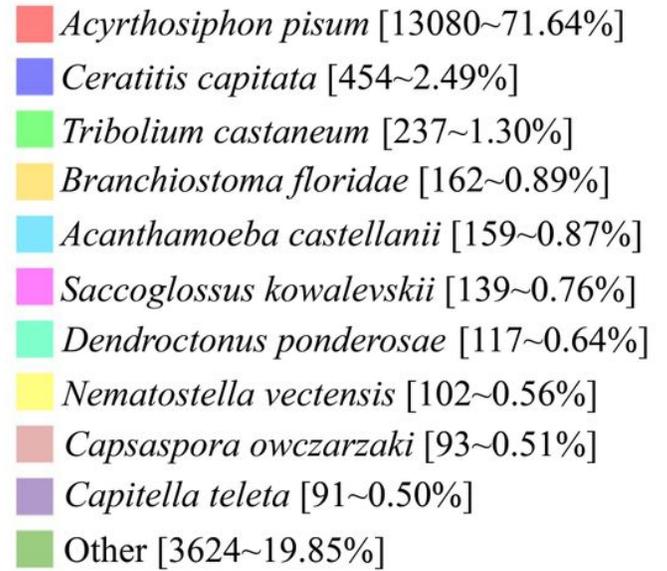
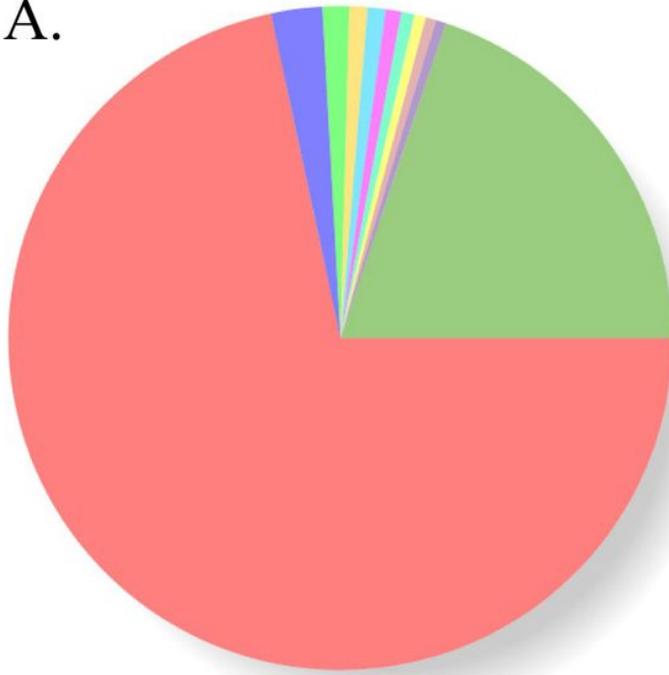
Treatment	Gene ID	Gene Name	Pathways	log2FC	FDR
Sg-BYDV	Sg55104	PIAS1	JAK-STAT signaling pathway	1.305	0.00022237
	Sg50416	UBE2D	Ubiquitin mediated proteolysis	5.042	0.001199371
	Sg49574	BIRC2_3	Ubiquitin mediated proteolysis	1.164	4.42E-09
	Sg55104	PIAS1	Ubiquitin mediated proteolysis	1.304	0.00022237
	Sg42560	HAO	Peroxisome	1.287	6.39E-08
	Sg43235	cathepsin	Lysosome	3.728	4.94E-06
	Sg45052	cathepsin	Lysosome	4.559	0.00263105
	Sg47282	cathepsin	Lysosome	2.494	2.06E-05
	Sg56271	cathepsin	Lysosome	-2.725	1.22E-05
	Sg53565	Legumain	Lysosome	2.621	4.84E-08
	Sg40066	LAMAN	Lysosome	1.223	2.49E-07
	Sg55760	NPC2	Lysosome	1.436	0.009359831
Sg-WDV	Sg40066	LAMAN	Lysosome	1.312	2.49E-07
Rp-BYDV	Rp46740	LAMAN	Lysosome	1.514	0.00928585
Rp-WDV	Rp49092	STAT5B	JAK-STAT signaling pathway	2.138	0.00295669

FC (fold change): calculated by RPKM.

FDR: false discovery rate

Figures

A.



B.

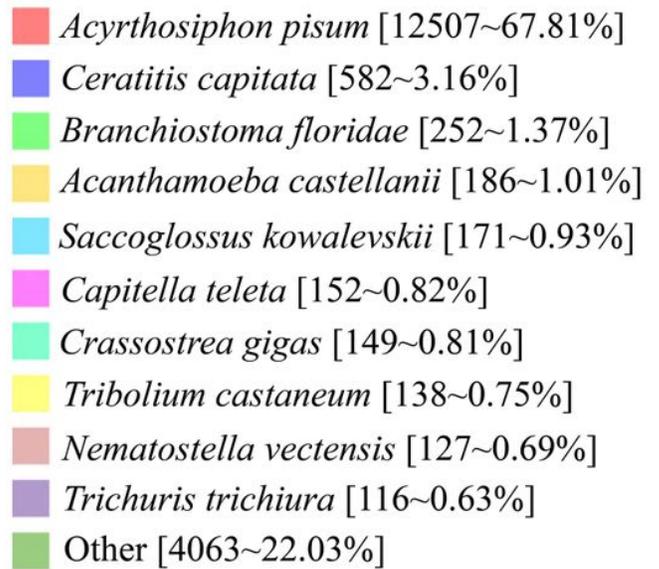
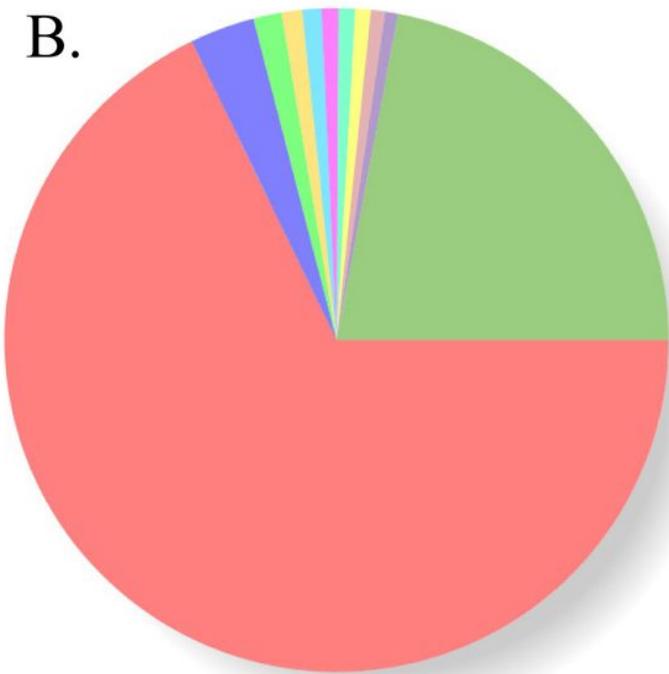


Figure 1

Nr homologous species distribution of *S. graminum* (A) and *R. padi* (B).

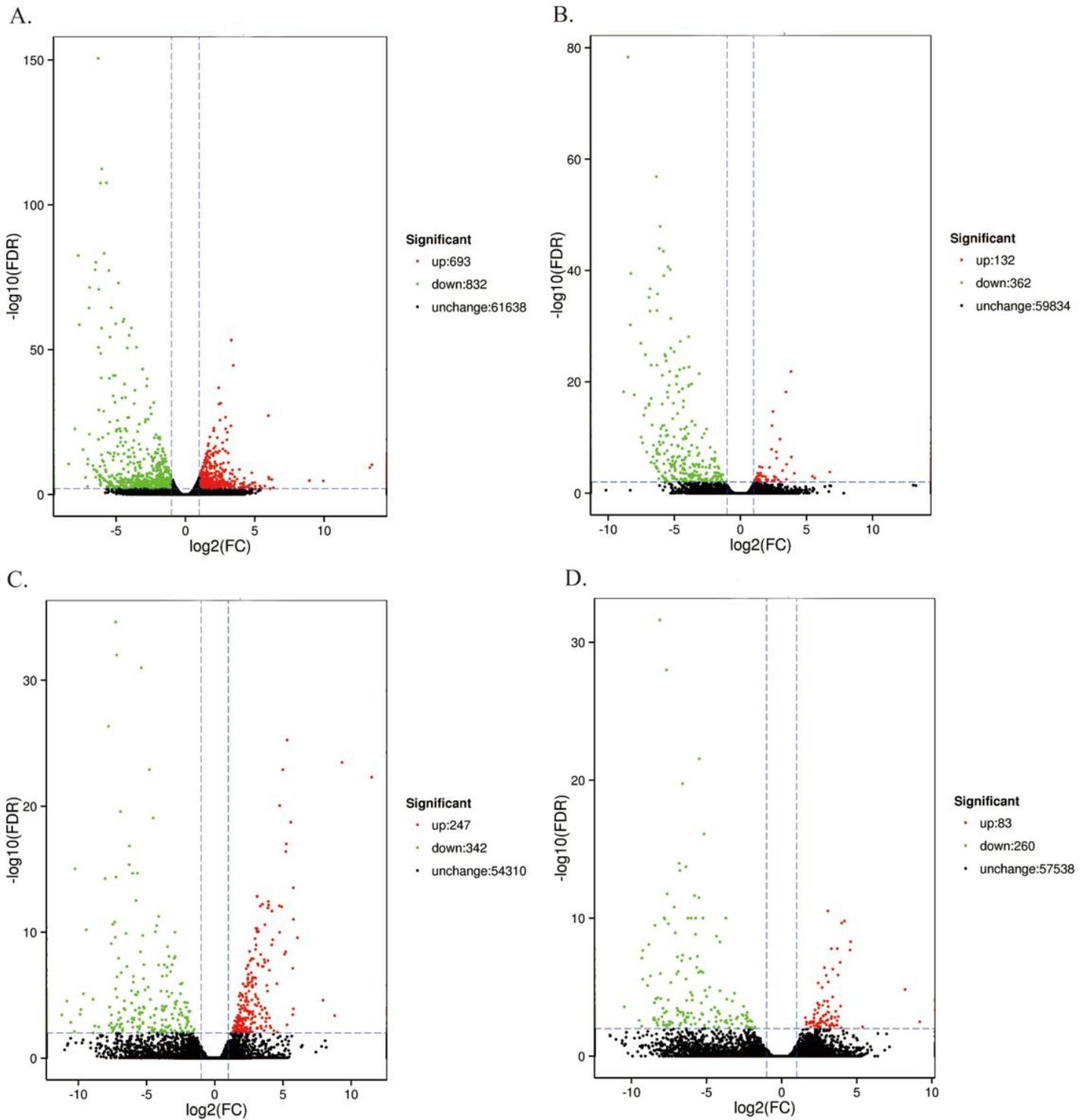


Figure 2

Volcano Plot of DEGs in *S. graminum* fed on wheat infected with BYDV (A), *S. graminum* fed on wheat infected with WDV (B), *R. padi* fed on wheat infected with BYDV (C) and *S. graminum* fed on wheat infected with BYDV (D).

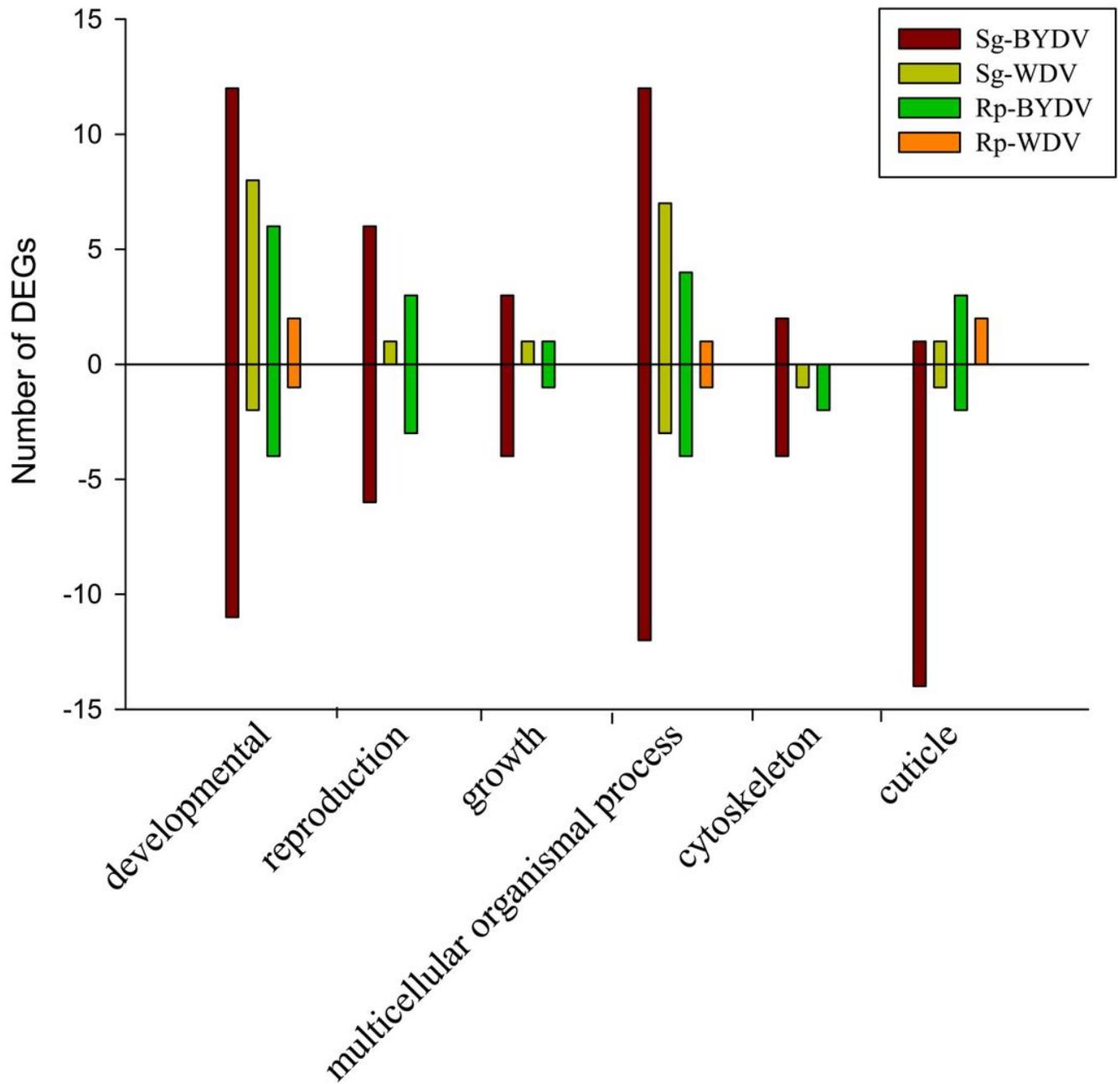


Figure 3

Number of differentially expressed genes (DEGs) involved in the development, reproduction, growth, multicellular organismal process, cytoskeleton and cuticle of vector (Sg-BYDV) and non-vector (Sg-WDV, Rp-BYDV, and Rp-WDV) aphids.

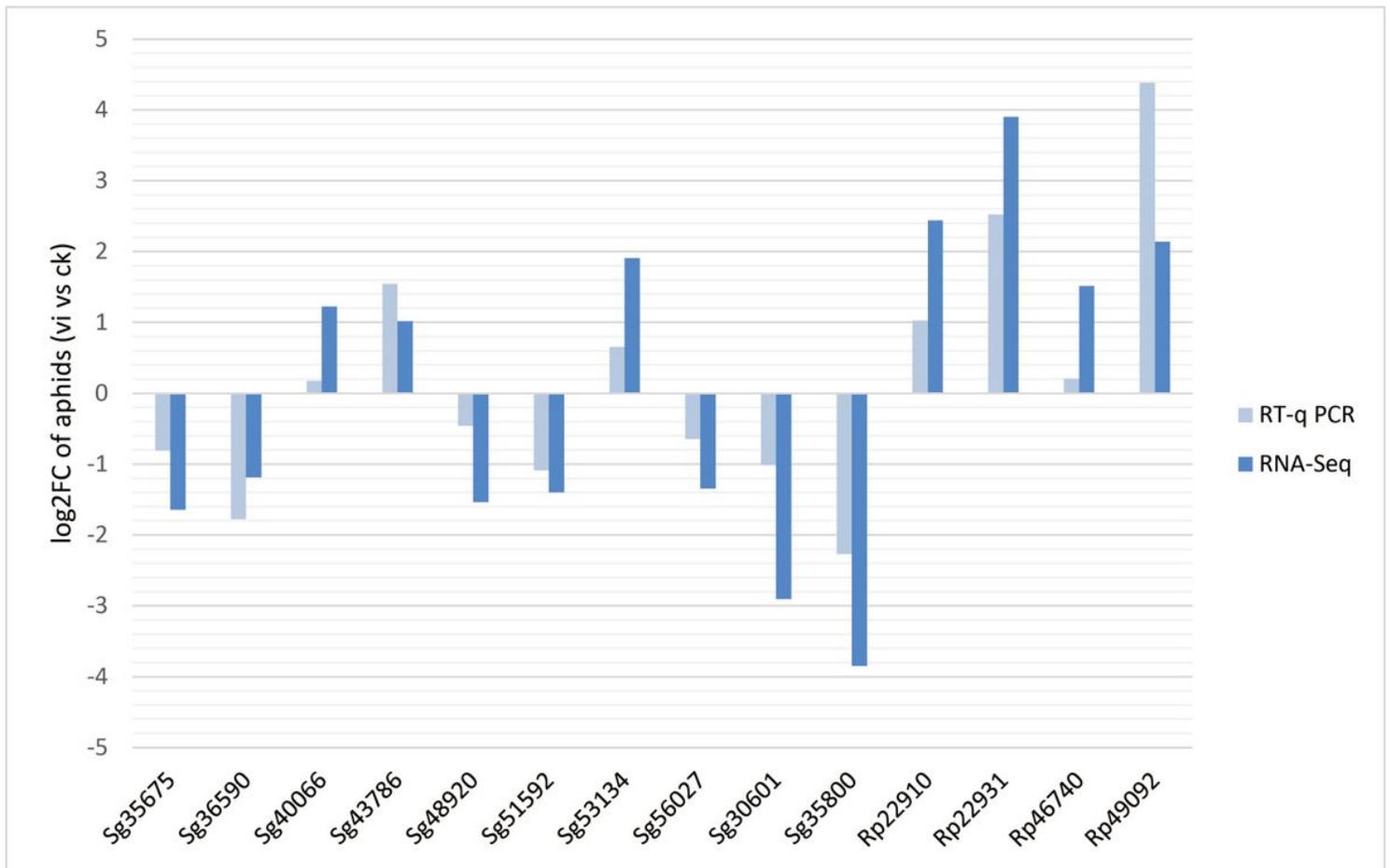


Figure 4

Quantitative reverse transcription-polymerase chain reaction (RT-qPCR) validation of selected genes differentially regulated in *S. graminum* and *R. padi* fed on virus-infected wheat (vi) compared with a virus-free control (ck) *S. avenae*.

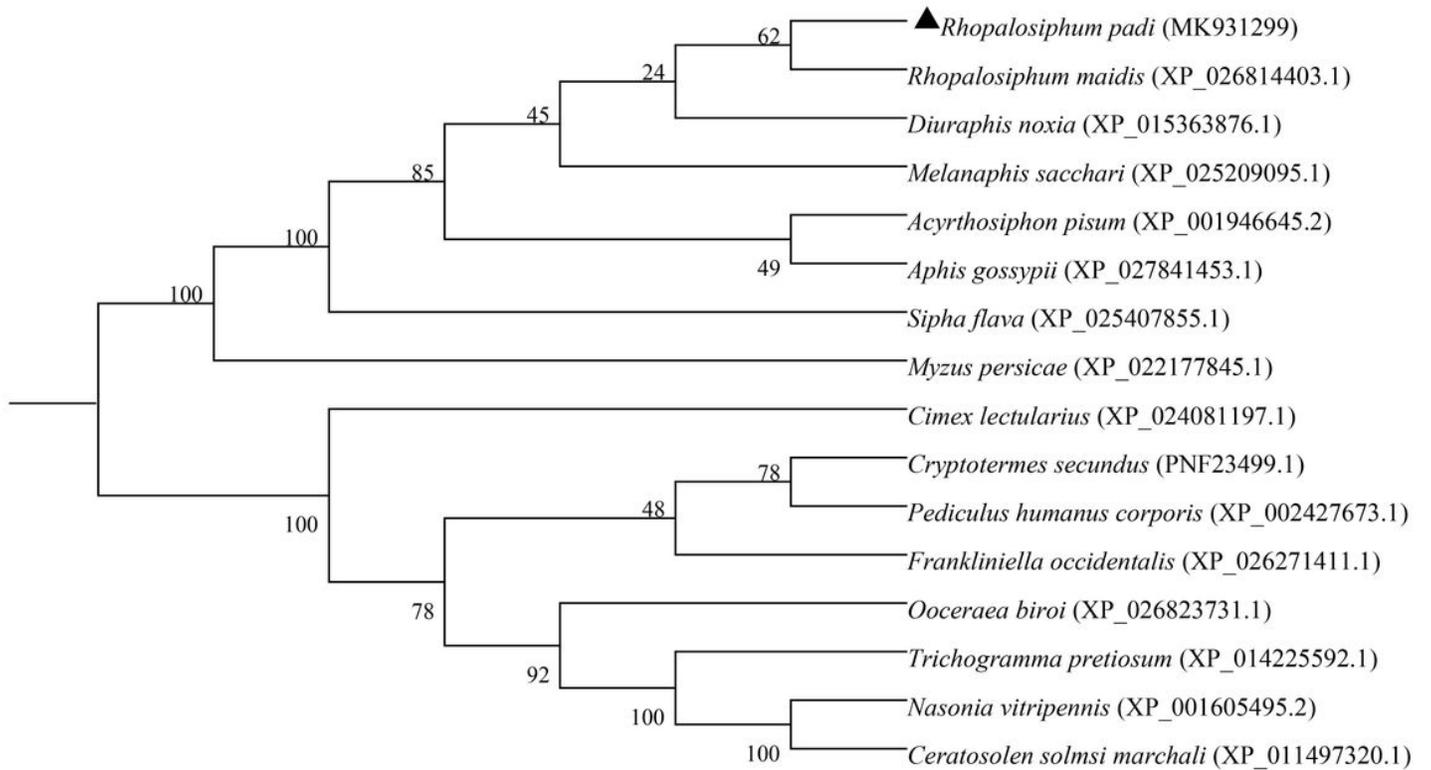


Figure 6

Phylogenetic relationship of RpSTAT5B with other insect STAT5Bs. Sequences were aligned using Clustal Omega and the neighbor-joining tree was constructed using MEGA6.0. Sequences used in the analysis are listed in Table S4.

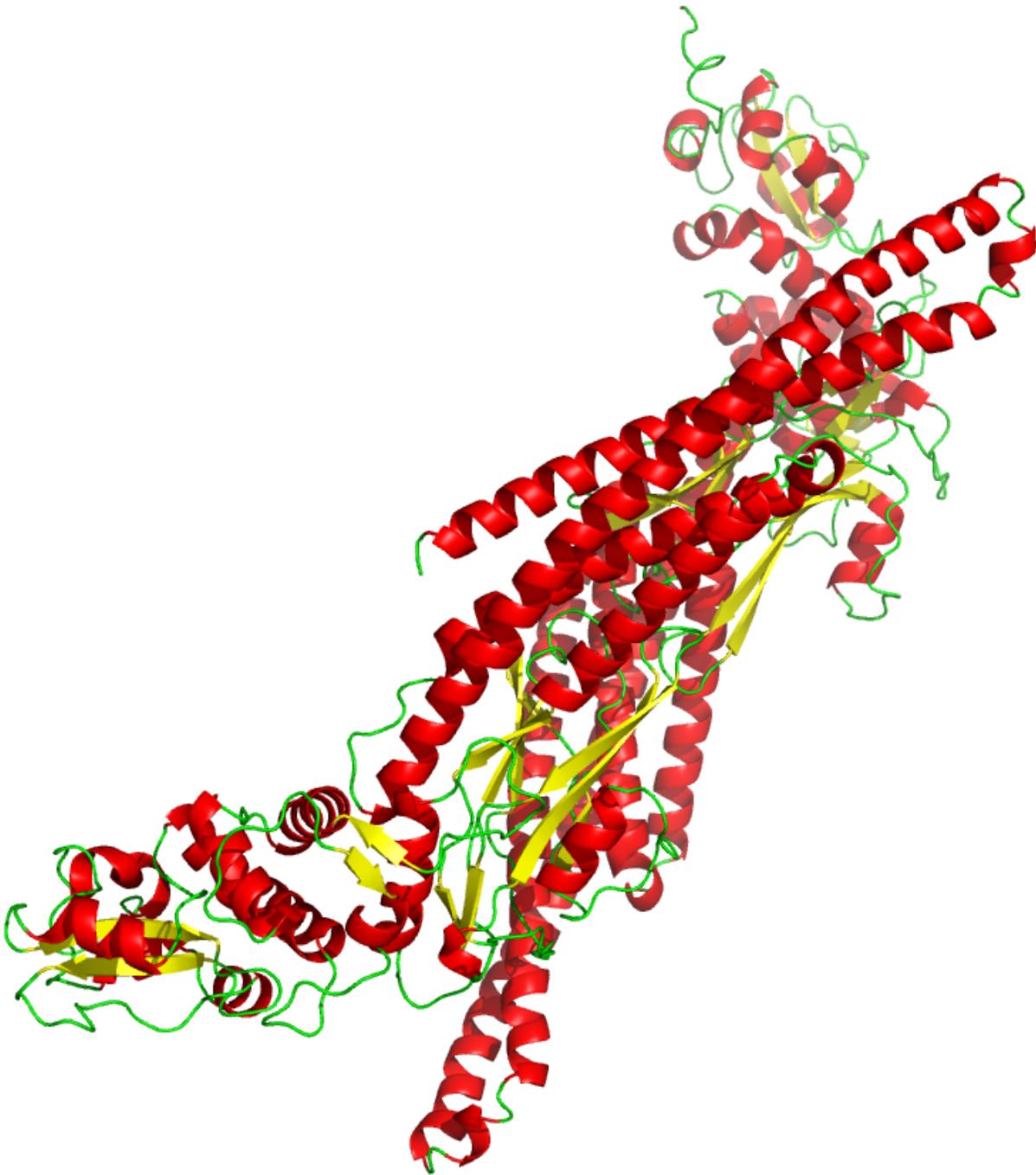


Figure 7

Predicted protein structure of RpSTAT5B.

Supplementary Files

This is a list of supplementary files associated with this preprint. [Click to download.](#)

- [Supplementarytables.docx](#)
- [Supplementaryfigure.pdf](#)