Insect science

Insect Science (2020) 00, 1–12, DOI 10.1111/1744-7917.12840

ORIGINAL ARTICLE

Characterization of protein–protein interactions between rice viruses and vector insects

Junjie Zhu[1,2], Fatma Elzahraa Eid[3], Lu Tong[1,2], Wan Zhao[1,2], Wei Wang[1,2], Lenwood S. Heath[4], Le Kang[1,2] and Feng Cui[1,2]

[1]*State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing, China;* [2]*CAS Center for Excellence in Biotic Interactions, University of Chinese Academy of Sciences, Beijing, China;* [3]*Broad Institute of MIT and Harvard, Cambridge, MA, United States and* [4]*Department of Computer Science, Virginia Tech, Blacksburg, VA, United States*

Abstract Planthoppers are the most notorious rice pests, because they transmit various rice viruses in a persistent-propagative manner. Protein–protein interactions (PPIs) between virus and vector are crucial for virus transmission by vector insects. However, the number of known PPIs for pairs of rice viruses and planthoppers is restricted by low throughput research methods. In this study, we applied DeNovo, a virus-host sequence-based PPI predictor, to predict potential PPIs at a genome-wide scale between three planthoppers and five rice viruses. PPIs were identified at two different confidence thresholds, referred to as low and high modes. The number of PPIs for the five planthopper-virus pairs ranged from 506 to 1985 in the low mode and from 1254 to 4286 in the high mode. After eliminating the "one-too-many" redundant interacting information, the PPIs with unique planthopper proteins were reduced to 343–724 in the low mode and 758–1671 in the high mode. Homologous analysis showed that 11 sets and 31 sets of homologous planthopper proteins were shared by all planthopper-virus interactions in the two modes, indicating that they are potential conserved vector factors essential for transmission of rice viruses. Ten PPIs between small brown planthopper and rice stripe virus (RSV) were verified using glutathione-S-transferase (GST)/His-pull down or co-immunoprecipitation assay. Five of the ten PPIs were proven positive, and three of the five SBPH proteins were confirmed to interact with RSV. The predicted PPIs provide new clues for further studies of the complicated relationship between rice viruses and their vector insects.

Key words computational prediction; planthopper; plant virus; protein-protein interaction; rice virus; vector insect

Introduction

Approximately 80% of plant viruses rely on insect vectors for transmission (Hohn, 2007). Persistent-propagative plant viruses infect a wide range of food crops and economical crops, representing a serious threat to agriculturally sustainable development (Soosaar *et al*., 2005). During transmission, persistent-propagative plant viruses replicate in the gut epithelial cells of their vector insects and then are released to the hemolymph, spreading to other tissues and finally to the salivary glands, where viruses are secreted into the host plants with the saliva. Viral infection and transmission are involved in a large number of protein-protein interactions (PPIs) between virus and vector (Hogenhout *et al*., 2008; Whitfield *et al*., 2015). Therefore, identifying PPIs between plant viruses and insect vectors will broaden our understanding of the transmission mechanisms and is essential for effective control of plant viral diseases.

Planthoppers are important rice pests, among which small brown planthopper (SBPH, *Laodelphax*

Correspondence: Feng Cui, State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China. Email: cuif@ioz.ac.cn

striatellus), brown planthopper (BPH, *Nilaparvata lugens*), and white-backed planthopper (WBPH, *Sogatella furcifera*) are the most notorious pests, because they transmit various rice viruses in a persistent-propagative manner (Uehara-Ichiki *et al*., 2013). SBPH transmits rice stripe virus (RSV) and rice black-streaked dwarf virus (RBSDV). BPH transmits rice grassy stunt virus (RGSV) and rice ragged stunt virus (RRSV). WBPH only transmits southern rice black-streaked dwarf virus (SRBSDV) (Uehara-Ichiki *et al*., 2013). RSV and RGSV belong to the genus *Tenuivirus*, the genomes of which contain segmented, ambisense single-stranded RNAs. RSV has a genome of ∼17 kb, consisting of four single-stranded RNA segments, and encodes seven proteins: RNA-dependent RNA polymerase (RdRp), capsid protein (CP), and non-structural proteins NS2, NSvc2, NS3, SP, and NSvc4 (Cui *et al*., 2016). RGSV has a genome of ∼25 kb, consisting of six single-stranded RNA segments, and encodes 12 proteins (Toriyama *et al*., 1998). RBSDV, SRBSDV, and RRSV belong to *Reoviridae*, containing segmented double-stranded RNAs in a double-shelled, icosahedral particle structure. The genomes of RBSDV, SRBSDV, and RRSV consist of ten double-stranded RNAs, encoding 13, 13, and 11 proteins, respectively (Hagiwara *et al*., 1986; Zhou *et al*., 2013; Zhou *et al*., 2017). The PPIs between each planthopper and the transmitted rice virus are still under-characterized.

With the high cost and/or low throughput of traditional PPI identification methods, computational PPI prediction frameworks are gaining ground in characterizing potential PPIs, especially with the accumulation of experimentally validated PPIs needed for training these frameworks (Nicod *et al*., 2017). Many computational methods have been developed and widely used to predict PPIs between viruses and hosts based on amino acid sequences, domain architectures, physiochemical properties, or protein structures (Calderwood *et al*., 2007; Nourani *et al*., 2015; Halder *et al*., 2018; Subramani *et al*., 2018). However, all methods were developed to predict PPIs for viruses with 100s of characterized PPIs to utilize for training their machine learning models. DeNovo, a sequence-based virus-human PPI prediction framework, has been the first to enable prediction for any virus with no characterized PPIs. DeNovo achieved an accuracy >80% even when predicting PPIs of viral proteins that have no sequence similarity to the ones used for training (Eid *et al*., 2016). However, no PPI predictor has been used to characterize PPIs in the plant virus-vector system.

In this study, we applied DeNovo to predict potential PPIs between the three planthoppers and five rice viruses with the availability of the genome resources for the three planthoppers (Xue *et al*., 2014; Wang *et al*., 2017a; Zhu *et al*., 2017) and five rice viruses. The interaction proteins shared by all planthopper-virus pairs are summarized. Experimental validations for several PPIs were carried out.

Materials and methods

Sequence information of rice viruses and planthoppers

Genome sequences of the three planthoppers were downloaded from the GigaScience repository, GigaDB (SBPH: http://doi.org/10.5524/100361, BPH: http://doi.org/10.5524/100139, WBPH: http://doi.org/10.5524/100255). Genome sequences of the five rice viruses were downloaded from the National Center for Biotechnology Information under the accession numbers GCF_000851385.1 for RSV, GCF_000852945.1 for RBSDV, GCF_000852965.1 for RRSV, GCF_000847645.1 for RGSV, and GCF_000889455.1 for SRBSDV. Protein sequences of planthoppers and viruses were extracted and used for prediction of PPIs.

Prediction of PPIs

PPIs between planthoppers and the rice viruses were predicted using the DeNovo method as described previously (Eid *et al*., 2016). Briefly, DeNovo utilizes experimentally characterized virus-human PPIs and their protein amino acid sequences to train a machine learning model to predict novel PPIs for uncharacterized viral proteins. We used DeNovo to predict potentially occurring PPIs between rice viruses and planthoppers by feeding the DeNovo predictor with the amino acid sequences of their proteins. The predictor scores how confident it is that each pair of virus-planthopper proteins may interact. We considered pairs of proteins to be potentially interacting when their prediction score exceeds a confidence threshold. We examined the predicted interactions at two thresholds: high and low. In the high-threshold mode (high mode for short), a smaller number of false positive PPIs are labeled as interacting compared to the low-threshold mode (low mode for short), but this comes at the cost of identifying fewer true positive PPIs in the high mode.

Functional analysis of planthopper interaction proteins

Gene Ontology (GO) annotations of planthopper interaction proteins were performed using Blast2GO software with default parameters (Götz *et al*., 2008). Enrichment

analyses of GO terms were performed using TBtools software. The GO terms with *P*-values < 0.05 were considered as significantly enriched (Chen *et al*., 2018). GO classifications were performed and plotted using WEGO 2.0 (Ye *et al*., 2018).

Homologous analysis of planthopper proteins

To identify the common PPIs shared by all planthopper-virus pairs, SBPH proteins interacting with both RSV and RBSDV, BPH proteins interacting with both RRSV and RGSV, and WBPH proteins interacting with SRBSDV were separately extracted. Then, the extracted proteins of BPH and WBPH were blasted against the extracted SBPH proteins. Proteins with E values < 1E-6 were considered homologous to SBPH proteins. Similarly, homologous sequences to those of BPH and WBPH were identified. Finally, homologous proteins of the three planthoppers were manually curated and the combined set of homologous proteins were considered as the common PPIs for rice virus transmission.

Small brown planthopper samples

The viruliferous small brown planthopper strain was derived from a field population collected in Hai'an, Jiangsu Province, China. The planthoppers were reared on 2 cm to 5 cm seedlings of rice (*Wuyujing*) in glass incubators at 25 °C and a light period of 16 : 8 (L : D) (Zhao *et al*., 2016).

RNA extraction and cDNA synthesis

Total RNAs were extracted from five SBPHs using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. After removal of genomic DNA contamination with a TURBO DNA-free kit (Ambion, Austin, TX, USA), 1 μg of RNA was reverse-transcribed to complementary DNA (cDNA) using random primers and an Moloney murine leukemia virus (M-MLV) Reverse Transcription System (Promega, Madison, WI, USA), following the manufacturer's instructions.

Gene cloning and protein expression

Full-length open reading frames (ORFs) of SBPH genes *flotillin2*, *importin α*, *translin*, *vacuolar protein sorting-associated protein 52* (*VPS52*), and *Ran GTPase*, RSV *CP*, *SP*, *NS3*, and *NSvc4* were cloned from the cDNA library of viruliferous planthoppers and sequenced. RSV *RdRp* was divided into five fragments to clone according to the domain architectures due to the long sequence (8760 bp). Then *importin α* and *translin* were constructed in the pGEX-3X vector between the restriction sites *Eco*RI and *Bam*HI to generate the recombinant plasmids with an N-terminal glutathione-S-transferase (GST)-tag. *Flotillin2*, *VPS52*, and *Ran GTPase* were constructed in the pET28a vector between NcoI and EcoRI to generate the recombinant plasmids with a C-terminal Flag-tag. RSV *CP*, *SP*, *NS3*, *NSvc4*, and five fragments of *RdRp* were cloned into the pET28a vector between *Eco*RI and *Xho*I to generate the recombinant plasmids with a N-terminal His-tag. The corresponding primers are listed in Supporting file 1.

The recombinant plasmids were transfected to *Escherichia coli* strain BL21 (DE3) for protein expression. After overnight induction with 0.5 mmol/L isopropyl β-D-thiogalactoside at 16 °C, cells were collected by centrifugation and sonicated for 30 min in ice water. The supernatant was kept for pull down and co-immunoprecipitation assays.

Pull down and co-immunoprecipitation assay

For GST-tag pull down, the GST-tagged recombinant proteins were bound to glutathione (GSH)-coupled particles at 4 °C for 2 h. After washing with PBST buffer (1 × phosphate buffer saline containing 0.1% triton ×100, pH 8.0), the His-tagged recombinant proteins were added and incubated for 2 h at 4 °C. After washing with PBST buffer, proteins were released with the elution buffer (50 mmol/L Tris-HCl containing 10 mmol/L reduced glutathione, pH 8.0). The expression products of pGEX-3X and pET28a vector served as negative controls.

For His-tag pull down, the His-tagged recombinant proteins were bound to Ni Sepharose (GE Healthcare, Buckinghamshire, UK) for 2 h at 4 °C. Then the Flag-tagged recombinant proteins were added and incubated for 2 h at 4 °C. After washing with lysis buffer (20 mmol/L sodium phosphate containing 50 mmol/L imidazole, pH 7.4), proteins were released with elution buffer (20 mmol/L sodium phosphate containing 250 mmol/L imidazole, pH 7.4). The expression product of pET28a vector served as negative control.

For co-immunoprecipitation (Co-IP), 30 μL of Dynabeads Protein G (Thermo Fisher Scientific, Waltham, MA, USA) was incubated with 5 μg of Flag monoclonal antibody (CWBiotech, Beijing, China) at room temperature for 15 min and then incubated with 200 μL of Flag-tagged protein for 30 min; 200 μL of bait proteins

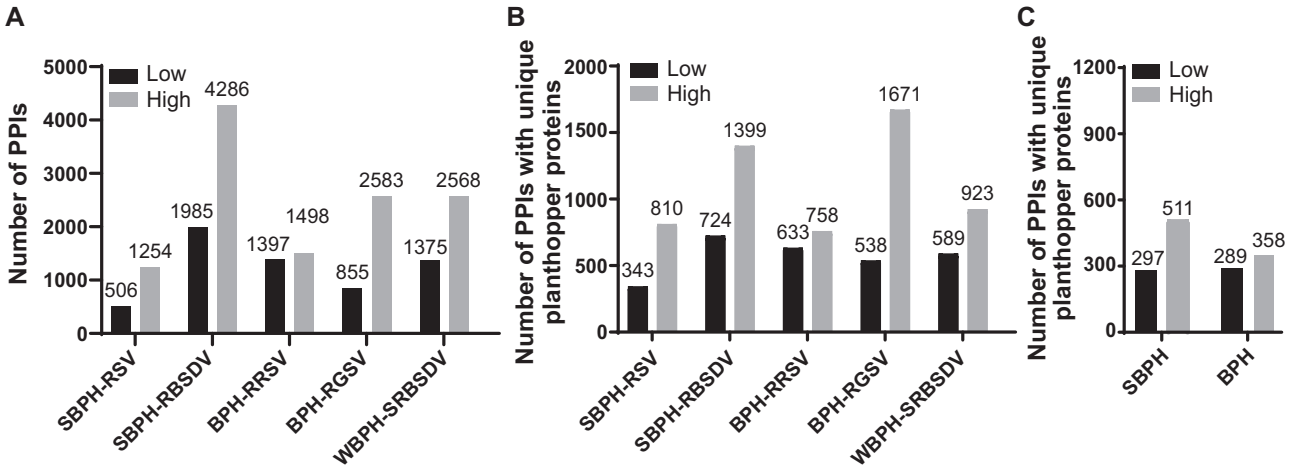

Fig. 1 Statistics of protein–protein interactions (PPIs) between planthoppers and rice viruses predicted with the low or high mode. (A) Number of PPIs. (B) Number of PPIs with unique planthopper proteins. (C) Number of PPIs with unique planthopper proteins shared by SBPH-RSV and SBPH-RBSDV pairs, and BPH-RRSV and BPH-RGSV pairs. SBPH, *Laodelphax striatellus*, BPH, *Nilaparvata lugens*, WBPH, *Sogatella furcifera*.

were added and incubated with the bead-antibody-protein complex for 30 min. The mixture was finally dissociated from the beads with elution buffer (Thermo Fisher Scientific). The mouse immunoglobulin G (Merck Millipore, Billerica, MA, USA) was used as negative control.

The input proteins and pulled-down and co-immunoprecipitated proteins were analyzed with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) followed by western blot using the anti-His monoclonal antibody, anti-GST polyclonal antibody, or anti-Flag monoclonal antibody (CWBiotech).

Results

PPIs between each pair of planthopper and rice virus

According to the genome annotations, 17 736, 27 571, and 21 254 proteins were retrieved from SBPH, BPH, and WBPH genomes, respectively, and used for PPIs prediction with rice virus proteins. We utilized DeNovo, the first PPI predictor to enable virus-host PPI prediction for any viral protein, to predict potential interactions between the proteins of the two organisms. We thresholded the confidence scores retrieved for all protein pairs at two thresholds: low and high, creating two predicted PPI sets. In the low mode, we predicted 506 PPIs between SBPH and RSV, 1985 PPIs between SBPH and RBSDV, 1397 PPIs between BPH and RRSV, 855 PPIs between BPH and RGSV, and 1375 PPIs between WBPH and SRBSDV (Fig. 1A, Supporting file 2). In the high mode, more PPIs were predicted for each pair of planthopper and virus (Fig. 1A, Supporting file 3). These PPIs contain the "one-to-many" situation, in which one planthopper protein interacts with multiple viral proteins. After eliminating the "one-to-many" redundant interacting information, that is, the planthopper protein interacting with multiple proteins of one rice virus was counted only once, the PPIs with unique planthopper proteins in the low mode were 343 for SBPH-RSV, 724 for SBPH-RBSDV, 633 for BPH-RRSV, 538 for BPH-RGSV, and 589 for WBPH-SRBSDV; in the high mode, 810 for SBPH-RSV, 1399 for SBPH-RBSDV, 758 for BPH-RRSV, 1671 for BPH-RGSV, and 923 for WBPH-SRBSDV (Fig. 1B, Supporting files 2 and 3).

GO assignments of molecular function at level 3 and enrichment analysis were used to classify the most represented functions of these unique planthopper proteins in PPIs. In the low mode, structural constituent of ribosome and lyase activity were enriched for PPIs of SBPH-RSV; structural constituent of cytoskeleton, protein binding, DNA-binding transcription factor activity, catalytic activity acting on DNA, and transcription coregulator activity were enriched for PPIs of SBPH-RBSDV; carbohydrate derivative binding, small molecule binding, and drug binding were enriched for PPIs of BPH-RRSV; heterocyclic compound binding and organic cyclic compound binding were enriched for PPIs of BPH-RGSV; metal cluster binding and protein binding were enriched for PPIs of WBPH-SRBSDV (Supporting file 4). The most represented functions in PPIs are different when planthoppers transmit different rice viruses. In the high mode, more enriched GO terms of molecular function

were observed and overlapped for PPIs of planthopper-rice virus (Supporting file 5).

The conserved PPIs shared by all pairs of planthoppers and rice viruses

The common proteins shared by one planthopper for transmitting different rice viruses were examined (Fig. 1C). Except WBPH that only transmits SRBSDV, SBPH can transmit RSV and RBSDV, and BPH transmits RRSV and RGSV. In the low mode, 297 SBPH proteins were involved in the potential interactions with both RSV and RBSDV, and, in the high mode, it was 511 SBPH proteins; 289 BPH proteins were involved in the potential interactions with both RRSV and RGSV in the low mode and 358 BPH proteins in the high mode.

To predict the common PPIs shared by all planthopper-virus pairs, we identified the homologous proteins from 297 SBPH proteins interactive with both RSV and RBSDV, 289 BPH proteins interactive with both RRSV and RGSV, and 589 WBPH proteins interactive with SRBSDV in the low mode. There were 11 sets of planthopper homologous proteins shared by all planthopper-virus interactions. They were annotated as elongation of very long chain fatty acids protein, N-alpha-acetyltransferase 50, RuvB, flotillin2, translin, adenosylhomocysteinase, retinol dehydrogenase 12, ras-like protein 2/3, dioxygenase, serine/threonine-protein kinase WNK1, vacuolar protein sorting-associated protein 37, and an uncharacterized protein (Table 1). Similarly, for the common PPIs from 511 SBPH proteins, 358 BPH proteins, and 923 WBPH proteins predicted with the high mode, there were 31 sets of homologous proteins shared by all planthopper-virus interactions. They were Ran GTPase, RGK family of GTP-binding proteins, trehalose-6-phosphate synthase, rab3 GTPase-activating protein non-catalytic subunit, threonylcarbamoyladenosine tRNA methylthiotransferase, aspartate-tRNA ligase, replication factor C subunit 4, importin alpha, heparan sulfate 2-O-sulfotransferase pipe, fumarylacetoacetase, DNA-directed RNA polymerase I subunit RPA2, Tubulin–tyrosine ligase-like protein 12, 28S ribosomal protein S5, N(G),N(G)-dimethylarginine dimethylaminohydrolase 1, T-complex protein 1 subunit beta, DnaJ-like/homolog subfamily C members, nitric oxide synthase, keratin 9, E3 ubiquitin-protein ligase, PAX-interacting protein 1, vacuolar protein sorting-associated protein 52 (VPS52), neuropeptide GPCR family, glutamine-dependent NAD(+) synthetase, hemicentin, Sp1-like transcription factor family, cyclin-dependent kinase 9/10, fork head box proteins or fork head domain-containing protein crocodile, receptor-type tyrosine-protein phosphatase N2, and elongation of very long chain fatty acids protein (Table 2). Elongation of very long chain fatty acids protein and vacuolar protein sorting-associated protein family were predicted by both of the two modes.

GO assignments at level 3 were used to classify the functions of these homologous planthopper proteins (Fig. 2). For the 11 sets of homologous proteins from the low mode, 16 GO terms were assigned in total. The main terms included intrinsic component of membrane and membrane part for the cellular component; transferase activity, hydrolase activity, and catalytic activity acting on DNA for the molecular function; and cellular metabolic process, organic substance metabolic process, primary metabolic process, nitrogen compound metabolic process, and small molecule metabolic process for the biological process. For the 31 sets of homologous proteins from the high mode, 30 GO terms were assigned in total. More molecular function terms were assigned and only the transferase activity was shared by the two modes. The main cellular component terms and biological process terms were similar to those from the low mode.

Validation of PPIs between planthopper proteins and virus proteins

To verify the predicted PPIs, we tested ten PPIs between SBPH and RSV using GST/His-pull down or Co-IP assay. The ten PPIs were flotillin2-CP/RdRp, translin-RdRp, VPS52-NSvc4, importin α-CP/NS3/SP/NSvc4/RdRp, and Ran GTPase-SP (Supporting files 2 and 3). Flotillin2 (evm.model.Contig20.196) and translin (evm.model.Contig223.14) belonged to the conserved homologous proteins shared by all planthopper-virus interactions predicted with the low mode (Table 1). VPS52 (evm.model.Contig64.113), importin α (evm.model.Contig1616.7), and Ran GTPase (evm.model.Contig1.118) were the conserved homologous proteins predicted with the high mode (Table 2). The SBPH proteins were recombinantly expressed with a GST- or Flag-tag. RSV proteins were recombinantly expressed with a His-tag.

The *in vitro* pull down or Co-IP assay showed that importin α-GST could bind His-CP, His-NS3, and His-NSvc4 (Fig. 3A), but failed to bind the five fragments of RdRp or His-SP (Supporting File 6A, 6B). GST-translin only bound the fragment 2 of RdRp (RdRp2) (Fig. 3B and Supporting file 6C). GST-flotillin2 was observed to bind His-CP (Fig. 3C), not binding with either of

Table 1 The homologous planthopper proteins shared by all planthopper-virus interactions predicted with the low mode.

Set	Planthopper ID[†]	Annotation
1	NLU005905.1	Elongation of very long chain fatty acids protein AAEL008004 [*Acyrthosiphon pisum*]
	Sfur-4.85	Elongation of very long chain fatty acids protein AAEL008004 isoform X3 [*Megachile rotundata*]
	evm.model.Contig143.52	Elongation of very long chain fatty acids protein [*Zootermopsis nevadensis*]
2	NLU002938.1	N-alpha-acetyltransferase 50 isoform X3 [*Halyomorpha halys*]
	Sfur-255.21	N-alpha-acetyltransferase 50 isoform X3 [*Halyomorpha halys*]
	evm.model.Contig152.41	N-alpha-acetyltransferase 50 isoform X4 [*Cimex lectularius*]
3	NLU022088.1	RuvB-like helicase 1 [Bombyx mori]
	Sfur-490.23	RuvB-like 2 [*Eufriesea mexicana*]
	evm.model.Contig193.8	RuvB-like 1 [*Zootermopsis nevadensis*]
4	NLU012648.1	Flotillin2 isoform X2 [*Aedes albopictus*]
	Sfur-229.21	Flotillin2 [*Bemisia tabaci*]
	evm.model.Contig20.196	Flotillin2 [*Bemisia tabaci*]
5	NLU007670.1	Translin isoform X1 [*Sorex araneus*]
	Sfur-1146.1	Translin [*Ictalurus punctatus*]
	Sfur-450.9	Translin [*Ictalurus punctatus*]
	evm.model.Contig223.14	Translin [*Plutella xylostella*]
6	NLU012841.3	Adenosylhomocysteinase-like [*Scleropages formosus*]
	Sfur-621.8	Adenosylhomocysteinase-like [*Bemisia tabaci*]
	evm.model.Contig32.109	Adenosylhomocysteinase-like [*Bemisia tabaci*]
7	NLU004958.1	Retinol dehydrogenase 12 isoform X2 [*Diaphorina citri*]
	NLU005829.1	Retinol dehydrogenase 12-like [*Halyomorpha halys*]
	Sfur-1002.6	Retinol dehydrogenase 12-like [*Halyomorpha halys*]
	evm.model.Contig325.14	Retinol dehydrogenase 12-like [*Halyomorpha halys*]
8	NLU009623.1	Uncharacterized protein LOC100867628 isoform X4 [*Apis florea*]
	Sfur-25.84	Uncharacterized protein LOC100867628 isoform X4 [*Apis florea*]
	evm.model.Contig45.146	Uncharacterized protein LOC100867628 isoform X2 [*Apis florea*]
9	NLU008048.1	Ras-like protein 2 [*Zootermopsis nevadensis*]
	Sfur-20.207	Ras-like protein 3 [*Asbolus verrucosus*]
	evm.model.Contig489.28	Ras-like protein 3 [*Cimex lectularius*]
10	NLU020849.1	Putative dioxygenase Mb0100 [*Halyomorpha halys*]
	Sfur-234.23	Putative dioxygenase Mb0100 [*Halyomorpha halys*]
	Sfur-395.11	Putative dioxygenase Mb0100 [*Halyomorpha halys*]
	evm.model.Contig70.33	Putative dioxygenase Mb0100 [*Halyomorpha halys*]
11	NLU027631.1	Vacuolar protein sorting-associated protein 37B [*Polistes dominula*]
	Sfur-848.4	Vacuolar protein sorting-associated protein 37A [*Zootermopsis nevadensis*]
	evm.model.Contig226.21	Vacuolar protein sorting-associated protein 37A [*Zootermopsis nevadensis*]

[†]IDs beginning with NLU, Sfur and evm stand for proteins from the genome of *Nilaparvata lugens*, *Sogatella furcifera*, and *Laodelphax striatellus*, respectively.

the RdRp fragments (Supporting file 6D). Negative interactions were observed between Flag-VPS52 and His-NSvc4 (Supporting file 6E) and between Flag-Ran GTPase and His-SP (Supporting file 6F). Thus, five of the ten PPIs were proven positive for the pair of SBPH-RSV, and three of the five SBPH proteins were confirmed to interact with RSV.

Discussion

For the first time, we predicted PPIs between rice viruses and their vector insects at a genome-wide scale using the DeNovo method and further provided a list of potential conserved vector proteins essential for transmission of rice viruses. Although these PPIs require further

Table 2 The homologous planthopper proteins shared by all planthopper-virus interactions predicted with the high mode.

Set	Planthopper ID[†]	Annotation
1	NLU004191.1, Sfur-2.162, evm.model.Contig1.118	RanGTPase
2	NLU006971.1, Sfur-308.11, evm.model.Contig109.25	GTP-binding protein REM 1, RAD, GEM
3	NLU015790.1, Sfur-79.5, evm.model.Contig110.14	Trehalose 6-phosphate synthase
4	NLU003954.1, Sfur-177.21, evm.model.Contig115.107	Rab3 GTPase-activating protein non-catalytic subunit
5	NLU012587.1, Sfur-594.2, evm.model.Contig13.334	Threonylcarbamoyladenosine tRNA methylthiotransferase
6	NLU022130.1, Sfur-107.89, evm.model.Contig1409.2	Aspartate–tRNA ligase cytoplasmic
7	NLU026574.1, Sfur-456.7, Sfur-691.8, evm.model.Contig15902.1	Replication factor C subunit 4-like
8	NLU006108.1, Sfur-23.119, evm.model.Contig1616.7	Importin subunit alpha
9	NLU014543.1, Sfur-119.20, evm.model.Contig1617.5	Heparan sulfate 2-O-sulfotransferase pipe, heparan sulfate 2-O-sulfotransferase 1
10	NLU028405.1, Sfur-20.87, evm.model.Contig175.56	Fumarylacetoacetase
11	NLU020781.1, Sfur-77.74, evm.model.Contig18.98	DNA-directed RNA polymerase I subunit RPA2
12	NLU004591.1, Sfur-114.8, evm.model.Contig21.144	Tubulin–tyrosine ligase-like protein 12
13	NLU016820.1, Sfur-9.312, evm.model.Contig21.71	28S ribosomal protein S5 mitochondrial
14	NLU011193.1, Sfur-24.345, Sfur-87.66, evm.model.Contig252.4	Neuropeptide GPCR A21, A33, A20; tachykinin-like peptides receptor 86C
15	NLU008920.1, Sfur-122.53, evm.model.Contig259.10	N(G),N(G)-dimethylarginine dimethylaminohydrolase 1
16	NLU021161.1, Sfur-1.99, evm.model.Contig31860.1	T-complex protein 1 subunit beta
17	NLU005039.1, Sfur-70.73, evm.model.Contig34949.1	DnaJ-like protein subfamily C member 7
18	NLU025650.1, Sfur-9.52, evm.model.Contig385.13	Nitric oxide synthase
19	NLU010038.1, Sfur-323.24, evm.model.Contig389.19	Keratin type I cytoskeletal 9-like
20	NLU010023.1, NLU023516.1, Sfur-31.96, evm.model.Contig41.48.3	Ubiquitin-protein ligase HECTD2, E3C, E3A
21	NLU027477.1, Sfur-49.48, evm.model.Contig480.51.1	PAX-interacting protein 1
22	NLU008758.1, NLU028397.1, Sfur-422.8, evm.model.Contig64.113	Vacuolar protein sorting-associated protein 52 homolog
23	NLU009603.1, Sfur-86.8, evm.model.Contig64.45	Neuropeptide GPCR A27, A48
24	NLU006122.1, Sfur-693.2, evm.model.Contig79.132	Putative glutamine-dependent NAD(+) synthetase
25	NLU007693.1, NLU016547.1, Sfur-119.7, evm.model.Contig262.47	Hemicentin-1, -2
26	NLU011964.1, NLU013169.1, Sfur-9.110, evm.model.Contig14.29	Transcription factor Sp4, Sp9, Sp3
27	NLU023154.2, Sfur-705.2, evm.model.Contig651.2	DnaJ homolog subfamily C member 7, 16
28	NLU013975.1, Sfur-357.21, evm.model.Contig70.24.1	Cyclin-dependent kinase 9, 10
29	NLU018180.1, Sfur-637.13, evm.model.Contig427.28, evm.model.Contig64.116, evm.model.Contig90.11	Forkhead box protein f2, E1, D3-B; fork head domain-containing protein crocodile
30	NLU008807.1, Sfur-824.1, evm.model.Contig636.1	Receptor-type tyrosine-protein phosphatase N2
31	NLU005905.1, Sfur-10.225, Sfur-212.49, Sfur-4.85, Sfur-45.66, evm.model.Contig143.52, evm.model.Contig158.46,evm.model.Contig696.3	Elongation of very long chain fatty acids protein AAEL008004

[†]IDs beginning with NLU, Sfur and evm stand for proteins from the genome of *Nilaparvata lugens*, *Sogatella furcifera*, and *Laodelphax striatellus*, respectively.

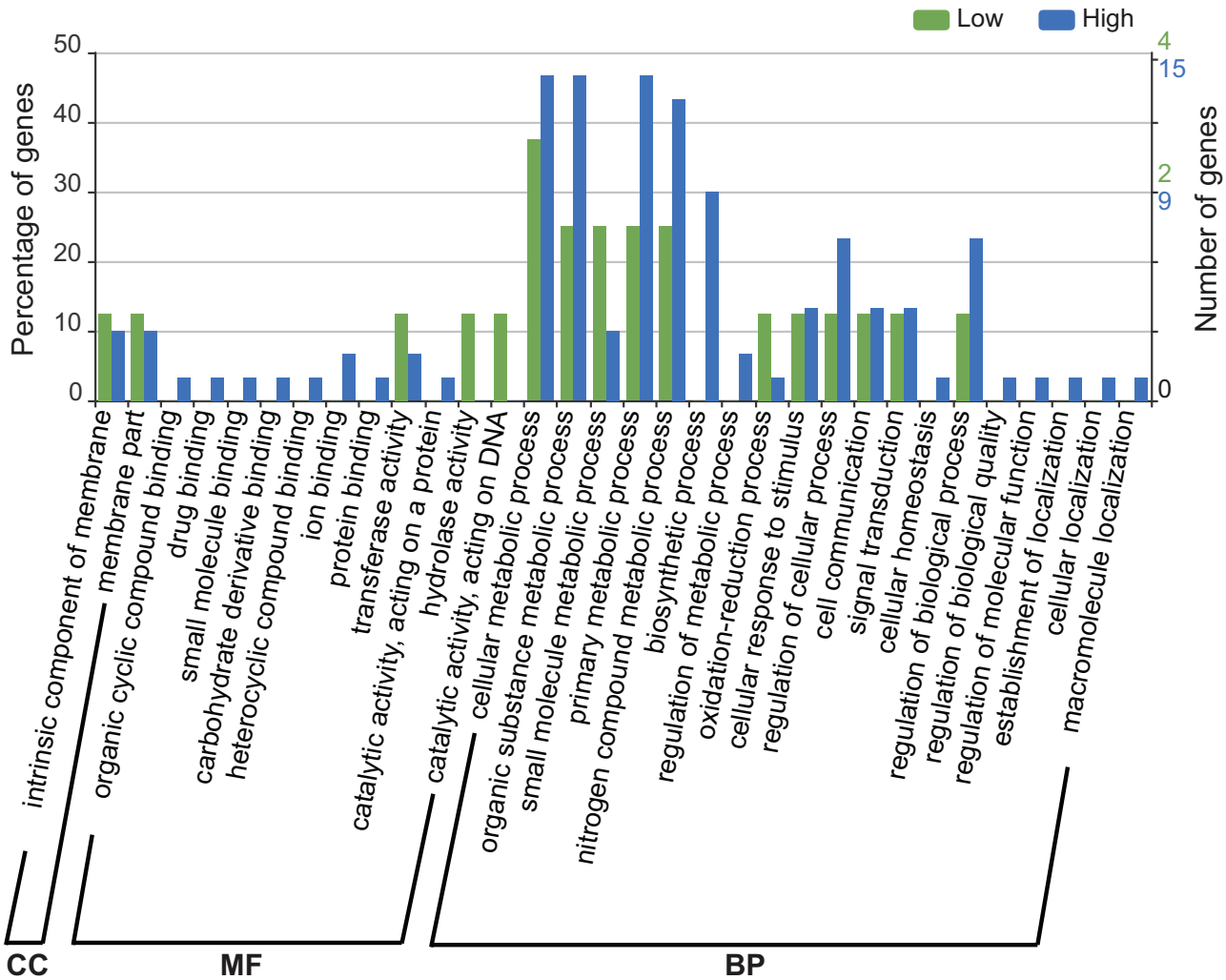

Fig. 2 Gene Ontology (GO) classifications at level 3 for the homologous planthopper proteins shared by all planthopper-virus interactions predicted with the low or high mode. CC, cellular component. MF, molecular function. BP, biological process. In the right vertical axis, green numbers account for the low mode and blue numbers account for the high mode.

verification with experimental data, they provide new clues for downstream studies and reflect the complexity of vector-virus interactions.

The true discovery rate for DeNovo predictions of the PPI between rice viruses and planthoppers is not perfectly 100%. Using the *in vitro* pull down or Co-IP assay, we found that three of the five SBPH proteins and five of the ten PPIs were positive for the pair of SBPH-RSV. However, some SBPH proteins, which were previously proven to interact with RSV's CP, do not appear in the PPI list of DeNovo prediction, such as GPS2, cuticular protein CPR1, vitellogenin (Li *et al*., 2011; Huo *et al*., 2014; Liu *et al*., 2015; Wang *et al*., 2017b). Machine learning predictors are limited by what they are trained on; while they are designed to generalize, they learn generalizable rules present in the data but cannot generalize to new rules with no examples in the training data. DeNovo was trained on interactions of human proteins and human viral proteins. The missed interactions may be due to the sequence dissimilarity between human and insect proteins, the sequence dissimilarity between human virus and rice virus proteins, or the interaction rules governing those characterized interactions are different. Moreover, even for those known interactions sharing sequence similarity with insect or rice virus proteins, DeNovo needs enough number of known interactions for training to make proper predictions. Among the 42 sets of conserved homologous proteins shared by all planthopper–virus interactions,

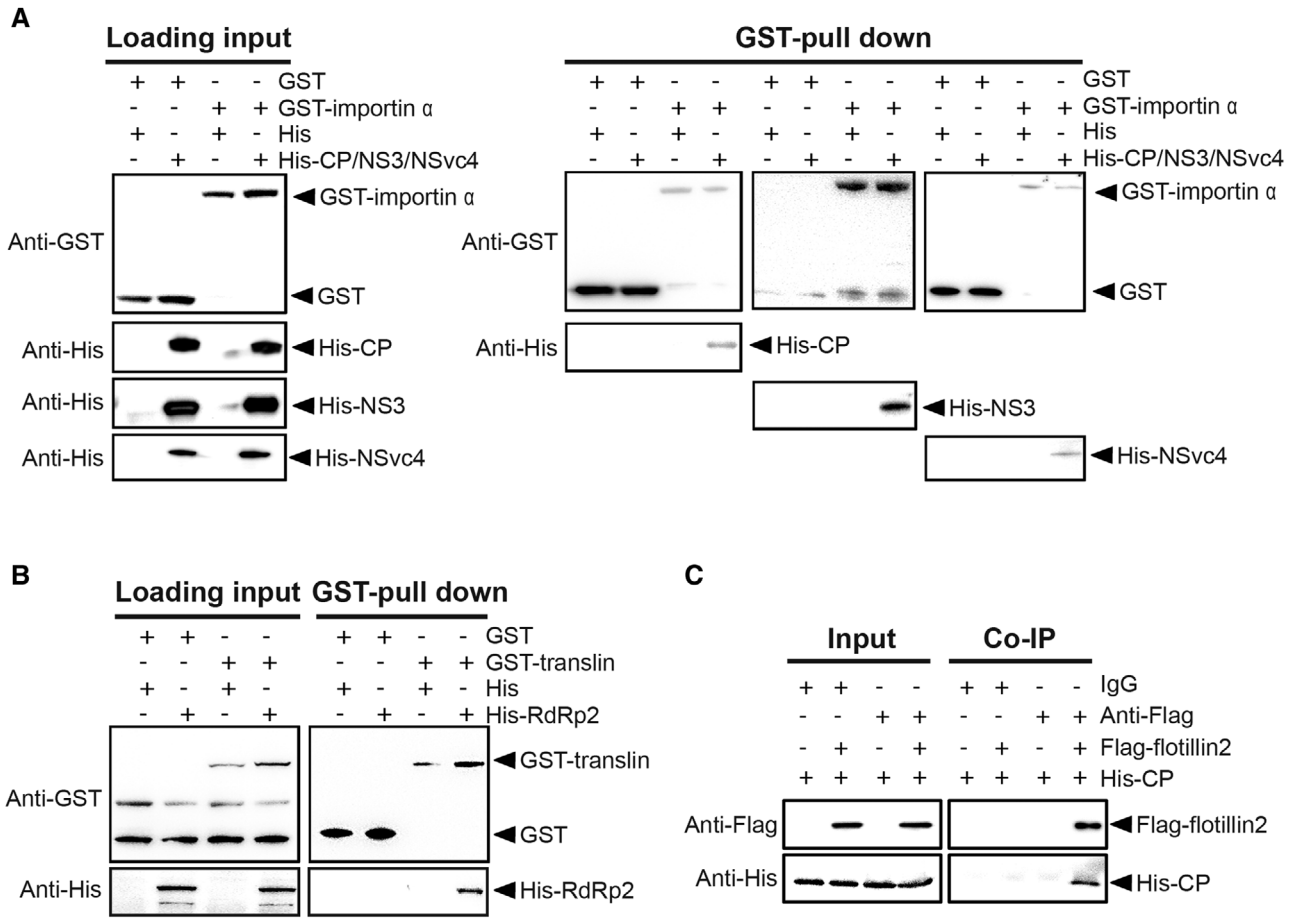

Fig. 3 Glutathione-S-transferase (GST)/His-pull down or co-immunoprecipitation assay on the interactions between small brown planthopper proteins and rice stripe virus (RSV) proteins. (A) GST-pull down showing the interaction between importin α and capsid protein (CP), NS3, or NSvc4. (B) GST-pull down showing the interaction between translin and RdRp. (C) Co-immunoprecipitation (Co-IP) showing the interaction between flotillin2 and CP. The products from pET28a vector and pGEX-3X vector served as negative controls. Mouse immunoglobulin G was used as negative control in Co-IP. The anti-His monoclonal antibody, anti-GST polyclonal antibody, and anti-Flag monoclonal antibody were used in western blot.

88% (37/42) are homologous to human proteins. This indicates that DeNovo emphasizes the prediction of human homologous proteins in PPIs. With the accumulation of known insect–plant virus interaction data, DeNovo can be trained to achieve a higher accuracy when predicting PPIs between viruses and insects.

The conserved homologous planthopper proteins characterized as interacting with rice viruses have extensive functions essential to viral infection or transmission. For example, importin α belongs to the classical importin (karyopherin) α/β nuclear transport pathway (Hiscox, 2007). The interaction between importin α and virus proteins has been widely reported in host cells and found essential for viral infection or pathogenicity. The polymerase subunit PB2 and nucleoprotein of influenza A virus interact with importin α for nuclear localization to benefit viral replication (Gabriel *et al*., 2008; Gabriel *et al*., 2011). Nuclear targeting of potato mop-top virus TGB1 mediated by importin α is vital for the efficient systematic infection of the virus in the entire plant (Lukhovitskaya *et al*., 2015). The interaction of rice viruses with planthopper importin α in this work suggests that importin α may play an essential role in viral replication or transmission in vector insects. In addition, flotillins are membrane organizing proteins and take part in the endocytosis to support the recycling of cell surface proteins (Babuke & Tikkanen, 2007). Recently it was reported that flotillins are necessary for recycling of T cell receptor to the immunological synapse and thus essential for T cell activation (Compeer *et al*., 2018). There have been no reports about the interaction between flotillins and viral proteins. In our work, we predict that flotillin2

of planthoppers can bind multiple proteins of rice viruses, and the interaction of flotillin2 and CP is verified in the SBPH-RSV system. Probably flotillin2 plays a role in the viral intercellular transport in vector insects, which needs further investigation. Another positive protein, translin, is a highly conserved DNA/RNA binding protein. It forms C3PO complex with translin associated factor-X (TRAX). The C3PO complex is involved in RNA silencing process through suppressing synthesis of miRNA (Asada *et al*., 2014) or promoting the siRNA pathway (Liu *et al*., 2009). The interaction between RSV and translin indicates that RSV would manipulate the RNA silencing system of vector insects to attain a balanced infection. Importantly, the three positive proteins have homologous proteins in rice. There are five importin α (XP_015639761.1, XP_015621365.1, XP_015619230.1, XP_015644757.1 and XP_015621115.1 in GenBank), three flotillin (XP_015613363.1, XP_025876759.1, and XP_015613707.1), and one translin (XP_015613233.1) homologous proteins in rice. Probably the viral infection mechanisms involved in these homologous proteins are similar in planthoppers and rice.

In conclusion, our study characterized the potential PPIs between rice viruses and vector insects using a computational virus-host PPI prediction method. These PPIs provide new candidate targets for us to study the complicated relationship between viruses and their vector insects and may serve as potential targets for plant virus prevention in future.

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China (No. 31772162) and the Chinese Academy of Sciences (No. ZDBS-LY-SM027).

Disclosure

The authors declare no competing interests.

References

Asada, K., Canestrari, E., Fu, X., Li, Z., Makowski, E., Wu, Y.C. *et al*. (2014) Rescuing dicer defects via inhibition of an anti-dicing nuclease. *Cell Reports*, 9, 1471–1481.

Babuke, T. and Tikkanen, R. (2007) Dissecting the molecular function of reggie/flotillin proteins. *European Journal of Cell Biology*, 86, 525–532.

Calderwood, M.A., Venkatesan, K., Xing, L., Chase, M.R., Vazquez, A., Holthaus, A.M. *et al*. (2007) Epstein-Barr virus and virus human protein interaction maps. *Proceedings of the National Academy of Sciences USA*, 104, 7606–7611.

Chen, C., Chen, H., He, Y. and Xia, R. (2018) TBtools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. *BioRxiv*, 289660, 1–6.

Compeer, E.B., Kraus, F., Ecker, M., Redpath, G., Amiezer, M., Rother, N. *et al*. (2018) A mobile endocytic network connects clathrin-independent receptor endocytosis to recycling and promotes T cell activation. *Nature Communications*, 9, 1–15.

Cui, F., Zhao, W., Luo, L. and Kang, L. (2016) Rice responses and resistance to planthopper-borne viruses at transcriptomic and proteomic levels. *Current Issues in Molecular Biology*, 19, 43–52.

Eid, F.E., Elhefnawi, M. and Heath, L.S. (2016) DeNovo: virus-host sequence-based protein-protein interaction prediction. *Bioinformatics*, 32, 1144–1150.

Gabriel, G., Herwig, A. and Klenk, H.D. (2008) Interaction of polymerase subunit PB2 and NP with importin alpha1 is a determinant of host range of influenza A virus. *PLoS Pathogens*, 4, e11.

Gabriel, G., Klingel, K., Otte, A., Thiele, S., Hudjetz, B., Arman-Kalcek, G. *et al*. (2011) Differential use of importin-α isoforms governs cell tropism and host adaptation of influenza virus. *Nature Communications*, 2, 1–7.

Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J. *et al*. (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, 36, 3420–3435.

Hagiwara, K., Minobe, Y., Nozu, Y., Hibino, H., Kimura, I. and Omura, T. (1986) Component proteins and structure of rice ragged stunt virus. *Journal of General Virology*, 67, 1711–1715.

Halder, A.K., Dutta, P., Kundu, M., Basu, S. and Nasipuri, M. (2018) Review of computational methods for virus–host protein interaction prediction: a case study on novel Ebola–human interactions. *Briefings in Functional Genomics*, 17, 381–391.

Hiscox, J.A. (2007) RNA viruses: hijacking the dynamic nucleolus. *Nature Reviews Microbiology*, 5, 119–127.

Hogenhout, S.A., Ammar El, D., Whitfield, A.E. and Redinbaugh, M.G. (2008) Insect vector interactions with persistently transmitted viruses. *Annual Review of Phytopathology*, 46, 327–359.

Hohn, T. (2007) Plant virus transmission from the insect point of view. *Proceedings of the National Academy of Sciences USA*, 104, 17905–17906.

Huo, Y., Liu, W., Zhang, F., Chen, X., Li, L., Liu, Q. *et al*. (2014) Transovarial transmission of a plant virus is mediated by vitellogenin of its insect vector. *PLoS Pathogens*, 10, e1003949.

Li, S., Xiong, R., Wang, X. and Zhou, Y. (2011) Five proteins of *Laodelphax striatellus* are potentially involved in the interactions between rice stripe virus and vector. *PLoS ONE*, 6, e26585.

Liu, Y., Ye, X., Jiang, F., Liang, C., Chen, D., Peng, J. *et al*. (2009) C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation. *Science*, 325, 750–753.

Liu, W., Gray, S., Huo, Y., Li, L., Wei, T. and Wang, X. (2015) Proteomic analysis of interaction between a plant virus and its vector insect reveals new functions of Hemipteran cuticular protein. *Molecular & Cellular Proteomics*, 14, 2229–2242.

Lukhovitskaya, N.I., Cowan, G.H., Vetukuri, R.R., Tilsner, J., Torrance, L. and Savenkov, E.I. (2015) Importin-alpha-mediated nucleolar localization of potato mop-top virus TRIPLE GENE BLOCK1 (TGB1) protein facilitates virus systemic movement, whereas TGB1 self-interaction is required for cell-to-cell movement in *Nicotiana benthamiana*. *Plant Physiology*, 167, 738–752.

Nicod, C., Banaei-Esfahani, A. and Collins, B.C. (2017) Elucidation of host–pathogen protein–protein interactions to uncover mechanisms of host cell rewiring. *Current Opinion in Microbiology*, 39, 7–15.

Nourani, E., Khunjush, F. and Durmuş, S. (2015) Computational approaches for prediction of pathogen-host protein-protein interactions. *Frontiers in Microbiology*, 6, 94.

Soosaar, J.L.M., Burch-Smith, T.M. and Dinesh-Kumar, S.P. (2005) Mechanisms of plant resistance to viruses. *Nature Reviews Microbiology*, 3, 789–798.

Subramani, C., Nair, V.P., Anang, S., Mandal, S.D., Pareek, M., Kaushik, N. *et al*. (2018) Host-virus protein interaction network reveals the involvement of multiple host processes in the life cycle of Hepatitis E virus. *mSystems*, 3, e00135-17.

Toriyama, S., Kimishima, T., Takahashi, M., Shimizu, T., Minaka, N. and Akutsu, K. (1998) The complete nucleotide sequence of the rice grassy stunt virus genome and genomic comparisons with viruses of the genus Tenuivirus. *Journal of General Virology*, 1998, 2051–2058.

Uehara-Ichiki, T., Shiba, T., Matsukura, K., Ueno, T., Hirae, M. and Sasaya, T. (2013) Detection and diagnosis of rice-infecting viruses. *Frontiers in Microbiology*, 4, 289.

Wang, L., Tang, N., Gao, X., Chang, Z., Zhang, L., Zhou, G. *et al*. (2017a) Genome sequence of a rice pest, the white-backed planthopper (*Sogatella furcifera*). *Gigascience*, 6, 1–9.

Wang, W., Zhao, W., Li, J., Luo, L., Kang, L. and Cui, F. (2017b) The c-Jun N-terminal kinase pathway of a vector insect is activated by virus capsid protein and promotes viral replication. *eLife*, 6, e26591.

Whitfield, A.E., Falk, B.W. and Rotenberg, D. (2015) Insect vector-mediated transmission of plant viruses. *Virology*, 479, 278–289.

Xue, J., Zhou, X., Zhang, C.X., Yu, L.L., Fan, H.W., Wang, Z. *et al*. (2014) Genomes of the rice pest brown planthopper and its endosymbionts reveal complex complementary contributions for host adaptation. *Genome Biology*, 15, 521.

Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y. *et al*. (2018) WEGO 2.0: a web tool for analyzing and plotting GO annotations, 2018 update. *Nucleic Acids Research*, 46, W71–W75.

Zhao, W., Yang, P., Kang, L. and Cui, F. (2016) Different pathogenicities of Rice stripe virus from the insect vector and from viruliferous plants. *New Phytologist*, 210, 196–207.

Zhou, G., Xu, D., Xu, D. and Zhang, M. (2013) Southern rice black-streaked dwarf virus: a white-backed planthopper-transmitted fijivirus threatening rice production in Asia. *Frontiers in Microbiology*, 4, 270.

Zhou, Y., Zhang, L., Zhang, X., Zu, H., Di, H., Dong, L. *et al*. (2017) Rice black-streaked dwarf virus genome in China: Diversification, phylogeny, and selection. *Plant Disease*, 101, 1588–1596.

Zhu, J., Jiang, F., Wang, X., Yang, P., Bao, Y., Zhao, W. *et al*. (2017) Genome sequence of the small brown planthopper, *Laodelphax striatellus*. *Gigascience*, 6, 1–12.

Manuscript received April 17, 2020
Final version received June 9, 2020
Accepted June 10, 2020

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supporting file 1. Primers used in this study.

Supporting file 2. Protein-protein interactions (PPIs) between planthoppers and rice viruses predicted with the low mode.

Supporting file 3. Protein-protein interactions (PPIs) between planthoppers and rice viruses predicted with the high mode.

Supporting file 4. Enriched Gene Ontology (GO) terms of molecular function at level 3 for the planthopper proteins interactive with rice virus proteins predicted with the low mode.

Supporting file 5. Enriched Gene Ontology (GO) terms of molecular function at level 3 for the planthopper proteins interactive with rice virus proteins predicted with the high mode.

Supporting file 6. Glutathione-S-transferase (GST)/His-pull down or co-immunoprecipitation assay on the interactions between small brown planthopper proteins and rice stripe virus (RSV) proteins.