New Phytologist

Research

An aspartic protease 47 causes quantitative recessive resistance to rice black-streaked dwarf virus disease and southern rice black-streaked dwarf virus disease

Zhaoyun Wang[1]*, Lian Zhou[2,3]*, Ying Lan[1], Xuejuan Li[1], Jian Wang[2,3], Jingfang Dong[2,3], Wei Guo[1,4], Dedao Jing[5], Qing Liu[2,3], Shaohong Zhang[2,3], Zhiyang Liu[1], Wenjuan Shi[1], Wu Yang[2,3], Tifeng Yang[2,3], Feng Sun[1], Linlin Du[1], Hua Fu[2,3], Yamei Ma[2,3], Yudong Shao[1], Luo Chen[2,3], Jitong Li[1], Shuo Li[1], Yongjian Fan[1], Yunyue Wang[4], Hei Leung[6], Bin Liu[2,3], Yijun Zhou[1], Junliang Zhao[2,3] and Tong Zhou[1,7]

[1]Key Laboratory of Food Quality and Safety, Institute of Plant Protection, Jiangsu Academy of Agricultural Sciences, Nanjing, 210014, Jiangsu Province, China; [2]Rice Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, Guangdong Province, China; [3]Guangdong Key Laboratory of New Technology in Rice Breeding, Guangzhou 510640, Guangdong Province, China; [4]Key Laboratory of Agricultural Biodiversity and Disease Control of Ministry of Education, College of Plant Protection, Yunnan Agricultural University, Kunming, 650201, Yunnan Province, China; [5]Zhenjiang Institute of Agricultural Sciences of the Ning-Zhen Hilly District, Jurong, 212400, Jiangsu Province, China; [6]International Rice Research Institute, Metro Manila 1301, Philippines; [7]International Rice Research Institute and Jiangsu Academy of Agricultural Sciences Joint Laboratory, Nanjing 210014, Jiangsu Province, China

Authors for correspondence:
Bin Liu
Email: liubin@gdaas.cn

Tong Zhou
Email: zhoutong@jaas.ac.cn

Junliang Zhao
Email: zhao_junliang@gdaas.cn

Yijun Zhou
Email: yjzhou@jaas.ac.cn

Received: *24 September 2021*
Accepted: *23 December 2021*

New Phytologist (2022)
doi: 10.1111/nph.17961

Key words: aspartic proteinase, quantitative trait loci (QTLs), resistance variety, *rice black-streaked dwarf virus* (RBSDV), *southern rice black-streaked dwarf virus* (SRBSDV).

Summary

- Rice black-streaked dwarf virus disease (RBSDVD) and southern rice black-streaked dwarf virus disease (SRBSDVD) are the most destructive viral diseases in rice. Progress is limited in breeding due to lack of resistance resource and inadequate knowledge on the underlying functional gene.
- Using genome-wide association study (GWAS), linkage disequilibrium (LD) decay analyses, RNA-sequencing, and genome editing, we identified a highly RBSDVD-resistant variety and its first functional gene.
- A highly RBSDVD-resistant variety W44 was identified through extensive evaluation of a diverse international rice panel. Seventeen quantitative trait loci (QTLs) were identified among which *qRBSDV6-1* had the largest phenotypic effect. It was finely mapped to a 0.8–1.2 Mb region on chromosome 6, with 62 annotated genes. Analysis of the candidate genes underlying *qRBSDV6-1* showed high expression of *aspartic proteinase 47* (*OsAP47*) in a susceptible variety, W122, and a low resistance variety, W44. *OsAP47* overexpressing lines exhibited significantly reduced resistance, while the knockout mutants exhibited significantly reduced SRBSDVD and RBSDVD severity. Furthermore, the resistant allele Hap1 of *OsAP47* is almost exclusive to *Indica*, but rare in *Japonica*.
- Results suggest that *OsAP47* knockout by editing is effective for improving RBSDVD and SRBSDVD resistance. This study provides genetic information for breeding resistant cultivars.

Introduction

Rice black-streaked dwarf virus (RBSDV) and *southern rice black-streaked dwarf virus* (SRBSDV), which belong to the genus *Fijivirus* (family *Reoviridae*), are transmitted by small brown planthoppers (SBPH, *Laodelphax striatellus* Fallén) and white-backed planthoppers (WBPH, *Sogatella furcifera* Horváth), respectively. After rice is infected, these viruses can cause serious stunted growth, failure in heading, poor seed setting rate, and even seedling death or total loss of grain yield under severe conditions. Rice black-streaked dwarf virus disease (RBSDVD) rapidly increased the area of damage by 13 times to 267 000 ha from 2007 to 2008 in Jiangsu Province, China (Zhou *et al.*, 2011). Meanwhile, southern rice black-streaked dwarf virus disease (SRBSDVD) had spread to 300 000 ha, an increase of 100 000 times in 8 yr since it was first discovered in China in 2001, and infected an area of 1601 600 ha only in China and Vietnam in 2010 (Hoang *et al.*, 2011; Wang *et al.*, 2014). SRBSDVD and

*These authors contributed equally to this work.

rice blast are the only two rice diseases that are listed in Category A of Crop Diseases and Insect Pests by the Ministry of Agriculture and Rural Affairs of China in 2020 (http://www.gov.cn/zhengce/zhengceku/2020-09/17/content_5544165.htm) because of the intensive harm inflicted by them and the difficulty in disease control. More importantly, some studies have reported that, aside from rice, RBSDV can also infect maize, wheat, barley, and other cereal crops in Asia, Europe, and South America, making it the most destructive crop virus worldwide (Lenardon *et al.*, 1998; Dovas *et al.*, 2004; Achon *et al.*, 2015; Zhou *et al.*, 2015).

Development of resistant cultivars (cvs) is considered to be an effective, economical, and environmentally sound approach for the management of viral diseases (Shikata & Kitagawa, 1977; Li *et al.*, 2013; Xu & Zhou, 2015; Sun *et al.*, 2017). Several efforts have been made in the search for RBSDVD-resistant resources. However, only a few medium-resistant varieties have been identified, such as Tetep and 9194 (Zhou *et al.*, 2015). The rice variety 9194 displayed high resistance to RBSDVD in field trials; but, it exhibited only medium resistance in the artificial inoculation test, with a disease incidence (DI) of $20.80 \pm 3.22\%$, and was resistant to SBPH (Sun *et al.*, 2017). Owing to the lack of high and stable resistance resources, so far, no breakthrough has been made in the breeding of RBSDVD- and SRBSDVD-resistant rice varieties.

The resistance of rice to RBSDVD and SRBSDVD is complex as it is controlled by multiple genes (Li *et al.*, 2013; Feng *et al.*, 2019; Soares *et al.*, 2019; Xiao *et al.*, 2019), and thus, poses significant difficulties in breeding resistant rice. Efficient disease resistance breeding cannot be achieved without understanding the genetic basis of resistance. Although genetic studies on RBSDVD and SRBSDVD resistance in rice have been conducted, these studies are inadequate compared to those on other major rice diseases such as blast and bacterial blight. Six quantitative trait loci (QTLs) for RBSDVD resistance were identified and mapped on chromosomes 6, 7, and 9 using a recombinant inbred line (RIL) population derived from Zhenshan 97 and Minghui 63. Among them, the *qRBSDV-6*^{MH} has been fine-mapped to a 627.6-kb region (Li *et al.*, 2013). Zheng *et al.* (2012) identified 11 QTLs for RBSDVD resistance, which were distributed on chromosomes 1, 2, 3, 4, 6, 8, 9, and 11, through field evaluation of RBSDVD resistance of the reciprocal introgression lines derived from the cross between Lemont and TeQing. In another study conducted by Sun *et al.* (2017), three QTLs for RBSDVD resistance, *qRBSDV6*, *qRBSDV9*, and *qRBSDV11*, were identified by using an $F_{2:3}$ population derived from a highly RBSDVD-resistant *Indica* variety 9194 (Sun *et al.*, 2017). In addition to bi-parental QTL analysis, genome-wide association study (GWAS) on RBSDVD resistance has been conducted in the field using 1070 diverse rice accessions selected from 3010 genome-sequenced varieties and the rice diversity panel 1 (RDP1). And several loci associated with RBSDVD resistance were identified and mapped (Feng *et al.*, 2019; Xiao *et al.*, 2019). However, only three publications on QTL mapping associated with SRBSDVD resistance in rice are available (Nong *et al.*, 2019a,b; Wei *et al.*, 2019). In a previous study, a gene *RabGDIα* conferring resistance to maize rough dwarf disease (MRDD), which was also caused by RBSDV, was cloned in maize (Q. C. Liu *et al.*, 2020; Q. Liu *et al.*, 2020). However, no gene has been cloned and functionally confirmed for its resistance to RBSDVD or SRBSDVD in rice.

To screen for reliable and highly resistant resources and identify the genes for RBSDVD resistance, a diverse rice panel consisting of 509 accessions from 59 countries, which have been genotyped with 700 K single nucleotide polymorphisms (SNPs), was used for this study. Both field trial and artificial inoculation test were conducted to evaluate RBSDVD resistance of the diverse rice panel. The highly resistant accession was subjected to antibiosis and virus accumulation tests to exclude the insect resistance and preference for plant. GWAS was conducted to identify and map QTLs for RBSDVD resistance. Based on linkage disequilibrium (LD) decay analysis, the region of QTL was narrowed down. A candidate gene was identified by differential expression analysis through RNA-sequencing (RNA-Seq) and its function was confirmed by transgenic experiments. We also conducted haplotype and comparative genomic analysis of the functional gene *OsAP47*. Through these studies, we have identified the highly RBSDVD resistant rice variety W44 and the first functional gene *OsAP47*. Our results suggest that *OsAP47* can negatively regulate RBSDVD and SRBSDVD resistance in rice, and it or its homologs might also function in other crops hosts of RBSDV and SRBSDV. These new findings pave the way for systematic studies on the molecular mechanisms of virus resistance.

Materials and Methods

Plant materials and growth conditions

A total of 509 rice accessions collected from 59 countries or regions from the rice diversity panel 2 (RDP2) were selected to evaluate RBSDVD resistance, including *Indica*, *Japonica*, *Aus*, and *Aromatic* accessions.

The transformation of the transgenic line was performed using *Agrobacterium tumefaciens-mediated* transformation and hygromycin selection. All the rice lines were grown in a glasshouse at 25°C and 70% relative humidity with a 12 h : 12 h, light : dark cycle. *Nicotiana benthamiana* plants were planted at 25°C with a 12 h : 12 h, light : dark cycle and used for *Agrobacterium* infiltration experiments.

Virus source and inoculation

Rice plants showing typical symptoms of rice dark-green dwarf disease were collected from a field in Jiangsu Province and tested through PCR using RBSDV/SRBSDV-specific primers. The RBSDV/SRBSDV-positive plants were grown in a glasshouse, which was used for the artificial inoculation and identification.

SBPH/WBPH nymphs were from a collection maintained in the laboratory and reared on the RBSDV/SRBSDV-positive rice plants for 7 d. The nymphs were then transferred onto healthy rice cv. Wuyujing No. 3 seedlings for 8 d to allow the virus to pass through its circulation period in SBPH/WBPH. The percentage of RBSDV/SRBSDV viruliferous SBPH/WBPH was

determined using dot-enzyme-linked immunosorbent assay (dot-ELISA).

Evaluation of disease resistance

In 2015 and 2016, the natural infection was carried out in Kaifeng, Henan Province, China due to RBSDVD prevalence in this region. To ensure an adequate viruliferous insect source, a rice seedling bed was encircled with a wheat field where the wheat green dwarf disease (also caused by RBSDV) had broken out. Two weeks before wheat harvest, 509 rice varieties were sown. About one month after sowing seeds, rice seedlings were transplanted into a field at Jiangsu Province. The cultivation was the same as that in normal field management except that pesticides and antivirals were not applied. One month after transplanting, the incidence of RBSDVD was recorded. Compared with healthy rice plants, the disease affected plants showed extremely conspicuous dwarfism, stiffness, and dark green leaves. Three replicates were conducted for each accession. The average of replicates was treated as the phenotypic value. Resistance against RBSDVD was evaluated based on the DI (the number of RBSDV-infected plants/the total number of plants counted × 100%). A survey of the DI was conducted during the first epidemic period of RBSDV, and the second survey was conducted 7 d later.

In the artificial inoculation test, 197 *Indica* rice accessions were evaluated to determine their resistance against RBSDVD. Thirty-five seedlings of each variety were sown in 500-ml beakers in the growth room of Jiangsu Academy of Agricultural Sciences with a humidity of 35–45% and temperature of 25–30°C. At approximately the 1.5-leaf stage, 30 healthy seedlings were inoculated for 72 h by inoculating three RBSDV-carrying SBPH per seedling. SBPH individuals were manually removed after 72 h to ensure uniformity in the inoculation intensity. The seedlings were transplanted into an experimental area consisting of cement pools at Jiangsu Academy of Agricultural Sciences, which were managed under our standard practices without pesticide or antiviral spraying during the rice-growing period. The survey and data processing methods were the same as those used for the field test. According to the resistance evaluation criteria, the disease incidence ranged from 0% to 100%; 0% is immune variety, 0.01–5.0% is resistant variety, 5.01–15% is moderate resistant variety, 15.01–30% is moderate susceptible variety, 30.01–50% is susceptible variety, more than 50% is highly susceptible variety (Zhou *et al.*, 2016).

Evaluation of rice resistance to SBPH infestation

To determine the resistance mechanism of W43, W44, W484, and W9, antibiosis and nonpreference tests were performed to evaluate rice resistance to SBPH infestation as follows: (1) Antibiosis test: seedlings were infested with first to second instar SBPH nymphs (10 nymphs per seedling). Five days later, the survival rate of SBPH on each variety was recorded, respectively. (2) Nonpreference test: seedlings were transferred into a beaker, infested with second to third instar SBPH nymphs (10 nymphs per seedling), and then covered with nylon net. Two days later, the average number of SBPH on each line was regarded as the value of feeding preferences. Details of the procedures have been described previously (Guo *et al.*, 2018).

Plasmid construction

For rice stable transformation, full length *OsAP47* sequence of Nipponbare (NPB) was PCR amplified and cloned individually into pEXT06f vector to produce overexpressing transgenic rice. For gene editing, we used the clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9 (CRISPR/Cas9) technology. Briefly, the specific *OsAP47* targeting sequences were amplified and inserted into pYLgRNA-OsU6, as previously described (Ma *et al.*, 2015). The two target sequences of the *OsAP47* guide RNA are GTTGAGCCGACCGTACGGGC and GCGCCGACTACATACAGTGG. All constructs were confirmed by DNA-sequencing, introduced into *Agrobacterium tumefaciens* strain EHA105 and then transformed into W44 and NPB by *Agrobacterium*-mediated transformation methods as described in a previous study (Hiei *et al.*, 1994; Z. Wang *et al.*, 2018). All the primers used in this study are listed in Supporting Information Dataset S1.

To verify the function gene of *qRBSDV6-1*, specific targeting sequences of the five *OsAP* genes, *LOC_Os06g02730*, *LOC_Os06g02780*, *LOC_Os06g02900*, *LOC_Os06g03080*, and *LOC_Os06g03120* were designed on the CRISPR-GE website (http://skl.scau.edu.cn/), namely APs-U6aT1, APs-U6aT2, APs-U6bT3, APs-U6bT4, and APs-U6aT5, respectively. The information for the guide RNA sequence is listed in Dataset S1. The vector construction was carried out according to a previous description (Ma *et al.*, 2015). The plasmids of positive clones were extracted and sequenced with the primers SP-L1, SP-R1, APs-W1F, and APs-W1R. The correct CRISPR/Cas9 vector was transferred into NPB by *Agrobacterium*-mediated genetic transformation. The positive transgenic plants were detected by PCR and sequencing with specific primers, which were *LOC_Os06g02730*-T1-F/R for APs-U6aT1, *LOC_Os06g02780*-T2-F/R for APs-U6aT2, *LOC_Os06g02900*-T3-F/R for APs-U6bT3, *LOC_Os06g03080*-T4-F/R for APs-U6bT4, and *LOC_Os06g03120*-T5-F/R for APs-U6aT5. Decoding was done through CRISPR-GE (http://skl.scau.edu.cn/) or manually. All the primers used in this study are listed in Dataset S1.

Genome-wide association analysis

Genome association and prediction integrated tool (GAPIT) v.2 was used for GWAS analysis (Tang *et al.*, 2016). SNPs were selected for GWAS analysis from the 700-Kb assay of a previous study based on the criteria of missing data < 15% and minor allele frequency of > 0.05 (McCouch *et al.*, 2016). GWAS was conducted using the mixed linear model with kinship matrix, and principal component analysis in GAPIT. Manhattan plots were produced by using R package CMPLOT (Yin *et al.*, 2021). Rice genome sequence version of MSU v.7.0 was used as reference for analysis (Kawahara *et al.*, 2013).

RNA isolation and quantitative real-time PCR analysis

Total RNA was extracted from the collected leaf samples using TRIzol reagent. The resulting RNA samples were treated with DNase prior to reverse transcription. Quantitative real-time PCR analyses were performed by using ChamQ SYBR qPCR Master Mix (Vazyme Biotech Co. Ltd, Nanjing, China) as instructed by the manufacturer. Relative expression levels of RBSDV and host defense genes were calculated using $2^{-\Delta\Delta CT}$ method as previously described (Livak & Schmittgen, 2001). Expression of rice *18s-rRNA* gene was used as an internal control. Three biological replicates with three technical replicates each were used for each treatment. Primers used in this experiment are listed in Dataset S1.

Protein extraction and Western blot assay

The harvested plant tissues were ground individually in liquid nitrogen and then homogenized in a protein extraction buffer (Sigma-Aldrich, St Louis, MO, USA) supplemented with a protease inhibitor cocktail (Roche, Basel, Switzerland). After 15 min centrifugation at 18 000 rcf at 4°C, the supernatant was collected from each sample and boiled for 8 min. Protein in each sample was separated in sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) gels through electrophoresis and then transferred onto polyvinylidene fluoride (PVDF) membranes. Antibody against FLAG was purchased from Sigma-Aldrich (Wang *et al.*, 2021).

Transcriptome analysis

At the 1.5-leaf stage, 35 healthy seedlings of susceptible W122 and resistant W44 were inoculated for 48 h with two RBSDV-carrying SBPH per seedling and inoculated with two RBSDV-free SBPH as mock control. After 48 h inoculation, 10 seedlings of W122 and W44 were collected at 0, 8, and 19 days post-inoculation (dpi), and the samples were ground individually in liquid nitrogen. We took 48 h after inoculation with RBSDV-carrying SBPH as the starting time point and designated as '0 dpi'. The other time points were counted from this onwards. Total RNA was extracted from the collected samples using TRIzol reagent for RNA-Seq.

Data analysis was conducted using the HISAT2-STRINGTIE-DESEQ2 pipeline (Pertea *et al.*, 2016). Briefly, reads of each sample were mapped to MSU7.0 reference sequence by HISAT2 (Kim *et al.*, 2015), and then analyzed by STRINGTIE (Pertea *et al.*, 2015). Raw counts of each sample exported from Stringtie were imported and normalized by DESEQ2 (Love *et al.*, 2014). Differential expression genes were further analyzed by DESEQ2. Cluster analysis of transcriptome data was performed using TBTOOLS. Bubble Map was plotted by http://www.bioinformatics.com.cn, an online platform for data analysis and visualization (Wang *et al.*, 2015).

Comparative mapping and ortholog gene analysis

Comparative genomic analysis was conducted with SYNMAP in COGE (http://genomevolution.org/CoGe/). A syntenic map was generated based on the data sets, including rice (NPB, id3), maize (B73, id333), and sorghum (Tx623, id331), in COGE. The analysis of the OsAP47 protein sequence was performed using the basic local alignment search tool (BLAST) in Gramene (http://www.gramene.org/), and significance thresholds were set at an *E*-value of 0.001 for orthologs in rice, maize, and sorghum. Protein sequences of OsAP47 orthologs from rice, maize and sorghum were sourced from the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) and Gramene databases. Sequences of aspartic proteinase (AP) genes from *Arabidopsis* and rice were downloaded from the NCBI based on results from previous studies. All protein sequences of APs that were identified were aligned by CLUSTALW in MEGAX (Kumar *et al.*, 2018). The phylogenetic tree was derived using IQ-TREE 2 based on the alignment results from MEGAX, and using the statistical method of maximum likelihood (Minh *et al.*, 2020). The phylogenetic tree was drawn using FIGTREE (http://tree.bio.ed.ac.uk/software/figtree/).

DNA diversity analysis

DNA diversity was analyzed using the 400-kb sequences of the studied rice in a global rice population (3k-RG panel) (W. S. Wang *et al.*, 2018). SNPs of the target interval were downloaded from the International River Interface Cooperative, and π analysis was performed by using VCFTOOLS (Danecek *et al.*, 2011).

Results

W44 showed high and stable resistance to RBSDVD by conferring resistance to the virus itself and not the insect

The field evaluation of 509 rice accessions revealed a large variation in the DI from 2.6% to 100% (Dataset S2). Among the 25 most resistant accessions (DI < 15%), 21 accessions were that of *Indica* (Dataset S2). The comparison of RBSDVD resistance among subpopulations also suggested that *Indica* is the most resistant subpopulation, while *Japonica* is the most susceptible subpopulation (Supporting Information Fig. S1). We further evaluated RBSDVD resistance of the rice accessions selected from *Indica* and *Aus* subpopulations (222 accessions) including most of the resistant accessions in the field trial using artificial inoculation (Dataset S3). Combining the results from the field and artificial inoculation tests, six accessions were identified to be RBSDVD resistant accessions (DI < 15%) in both the field and artificial inoculation tests (Dataset S4). Location analysis of resistant accessions demonstrated that eight out of the 25 resistant accessions used in the field trial, and 11 out of the 18 resistant accessions used in the artificial inoculation originate from South Asia (India, Bangladesh, and Sri Lanka) (Datasets S2, S3).

To determine that the resistant varieties are virus resistant rather than insect resistant, we evaluated the SBPH resistance of the accessions that had an average DI < 10% in the field and artificial inoculation tests of W44, W43, and W484. Our insect preference tests showed that there was no significant difference in the number of SBPH per plant between W44 and the SBPH

susceptible cv. NPB. In contrast, W43 and W484 had weak preferences as manifested by the significant difference in the number of SBPH per plant between the two accessions and NPB (Fig. S2a). Furthermore, the antibiosis test showed that the average survival rates of SBPH on W43 and W484 were significantly lower than that on NPB while there was no significant difference in the average survival rates of W44 and NPB (Fig. S2b), suggesting that W44 does not exhibit antibiosis. Therefore, the RBSDVD resistance conferred by W44 is independent of SBPH resistance and is mainly due to its resistance to RBSDV.

Compared to the moderately resistant Tetep and susceptible Huaidao No. 5 (H5) or NPB, W44 showed weaker stunting symptom and had a much lower DI in the repeated experiments in both the field and artificial inoculation tests (Fig. 1a–c). In addition, transcriptional analysis revealed that the accumulation of RBSDV *S10* in the infected W44 was significantly lower than that in the susceptible variety of NPB at 8 and 19 dpi (Fig. 1d). We also detected RBSDV protein with a specific antibody against the P10 epitope, but the expression of P10 at 8 dpi could not be detected in W44, until 19 dpi (Fig. 1e). However, the accumulation of RBSDV P10 increased from 8 dpi to 19 dpi in NPB. These results showed that RBSDV accumulation can be dramatically repressed at messenger RNA (mRNA) and protein levels in W44. These results suggest that W44 is a reliable and highly resistant variety for RBSDV.

Identification of QTLs for RBSDVD resistance using GWAS

To identify the genes associated with RBSDVD resistance in rice, GWAS of RBSDVD resistance in 509 rice accessions was conducted. In total, 17 QTLs for RBSDVD resistance were identified by using the whole population and different subpopulations (Fig. 2a; Table 1). The *qRBSDV1* and *qRBSDV2-2* identified in the whole population overlap with the *qRBSDVia1-2* identified in the *Indica* plus *Aus* subpopulation and the *qRBSDVi2-2* identified in the *Indica* subpopulation, respectively. The *qRBSDVi1* identified in the *Indica* subpopulation overlaps with *qRBSDVia1-1* identified in the *Indica* plus *Aus* subpopulation. Particularly, the QTL, *qRBSDV6-1* identified in the whole population was also detected in the *Indica* subpopulation (*qRBSDVi6*) and *Indica* plus *Aus* subpopulation (*qRBSDVia6-1*).

To confirm the GWAS results in the field trial, we further conducted GWAS using the disease phenotypes collected from artificial inoculation tests. Considering that most of the resistant accessions were from *Indica* and *Aus* subpopulations in the field trials and the population structure, 222 *Inidca* and *Aus* accessions were selected for further artificial inoculation tests and GWAS was conducted with the disease phenotypes. Our results showed that three QTLs for RBSDVD resistance were identified and distributed on chromosomes 3 and 6 (Fig. 2a; Table 1). Notably, the QTL on chromosome 6, *qRBSDVgi6*, identified in the artificial inoculation test overlaps with *qRBSDV6-1* identified in the field trial. Thus, *qRBSDV6-1* was detected in both the field and artificial inoculation tests by using different subpopulations. Interestingly, this QTL also overlaps with the QTL for RBSDVD resistance in previous studies (Zheng *et al.*, 2012; Zhang *et al.*, 2016; Feng *et al.*, 2019, 2020; Xiao *et al.*, 2019), suggesting that it is stable and is a promising target for functional characterization and molecular breeding.

Functional gene exploring in the *qRBSDV6-1* loci

To identify the functional gene underlying *qRBSDV6-1*, we first analyzed the LD block. Based on the LD decay analysis, we delimited *qRBSDV6-1* to a 400-kb region from 0.8 to 1.2 Mb on chromosome 6, in which 62 genes were annotated according to the MSU Rice Genome Annotation Project (RGAP) on rice IRGSP-1.0 genome Release 7.0 (http://rice.plantbiology.msu.edu/) (Fig. 2b; Dataset S5) (Kawahara *et al.*, 2013). Then, we analyzed the haplotypes within the delimited *qRBSDV6-1* region. The results showed that there were four haplotypes in the 222 rice accessions and Hap1 displayed the lowest DI among them (Fig. 2c). Haplotype analysis also indicated that W44 harbored the resistant haplotype Hap1. Therefore, the accession no. W44 was used for the functional characterization of *qRBSDV6-1* thereafter.

To narrow down the list of candidate genes, we performed genome-wide gene expression profiling, of the resistant accession no. W44 and the highly susceptible accession no. W122 after RBSDV inoculation by RNA-Seq. The differential expression of the genes within *qRBSDV6-1* region using RNA-Seq data were analyzed, and the results show that 45 genes can be transcribed in different samples, of which 25 genes encode an expressed protein (Dataset S6). In addition to expressed proteins, our transcriptome data showed that *LOC_Os06g02780* and *LOC_Os06g02900* exhibited differential expression patterns between W44 and W122, with very low expression levels in W44 but high expression levels in W122 (Fig. 3a; Dataset S6). The complementary DNA of *LOC_Os06g02780* and *LOC_Os06g02900* were cloned and sequenced. The sequence of *LOC_Os06g02900* in W44 is consistent with that in W122 and NPB. Sequence comparisons showed that the amino acid sequence of *LOC_Os06g02780* is the same in W122 and NPB, but the sequence in W44 is highly different from its ortholog in W122 and NPB (Fig. S3). According to the gene annotations, *LOC_Os06g02780* is a type of AP, which was designated as *OsAP47* (Dataset S6). Plant typical APs contain two motifs related to enzyme activity, namely the N-terminal DTG and the C-terminal DSG (Xia *et al.*, 2020). The two enzyme activity domains of the OsAP47 are all mutant in W122, NPB, and W44, except that the catalytic residue D is conserved (Fig. S3). In W122 and NPB, two domains were mutated to DTA and DSR. Whereas, these two domains were mutated to DTA and DTR in W44. Gene ontology (GO) analysis of the annotated genes in *qRBSDV6-1* region of W44 and W122 was performed. GO analysis revealed that aspartic-type endopeptidase activity genes were highly enriched (Fig. 3b). Furthermore, the real-time PCR assay also showed that the expression level of *OsAP47* in susceptible variety W122 was significantly higher than that in W44 (Fig. 3c). Thus, *OsAP47* was considered the candidate gene underlying *qRBSDV6-1* to advance to the transgenic functional confirmation.

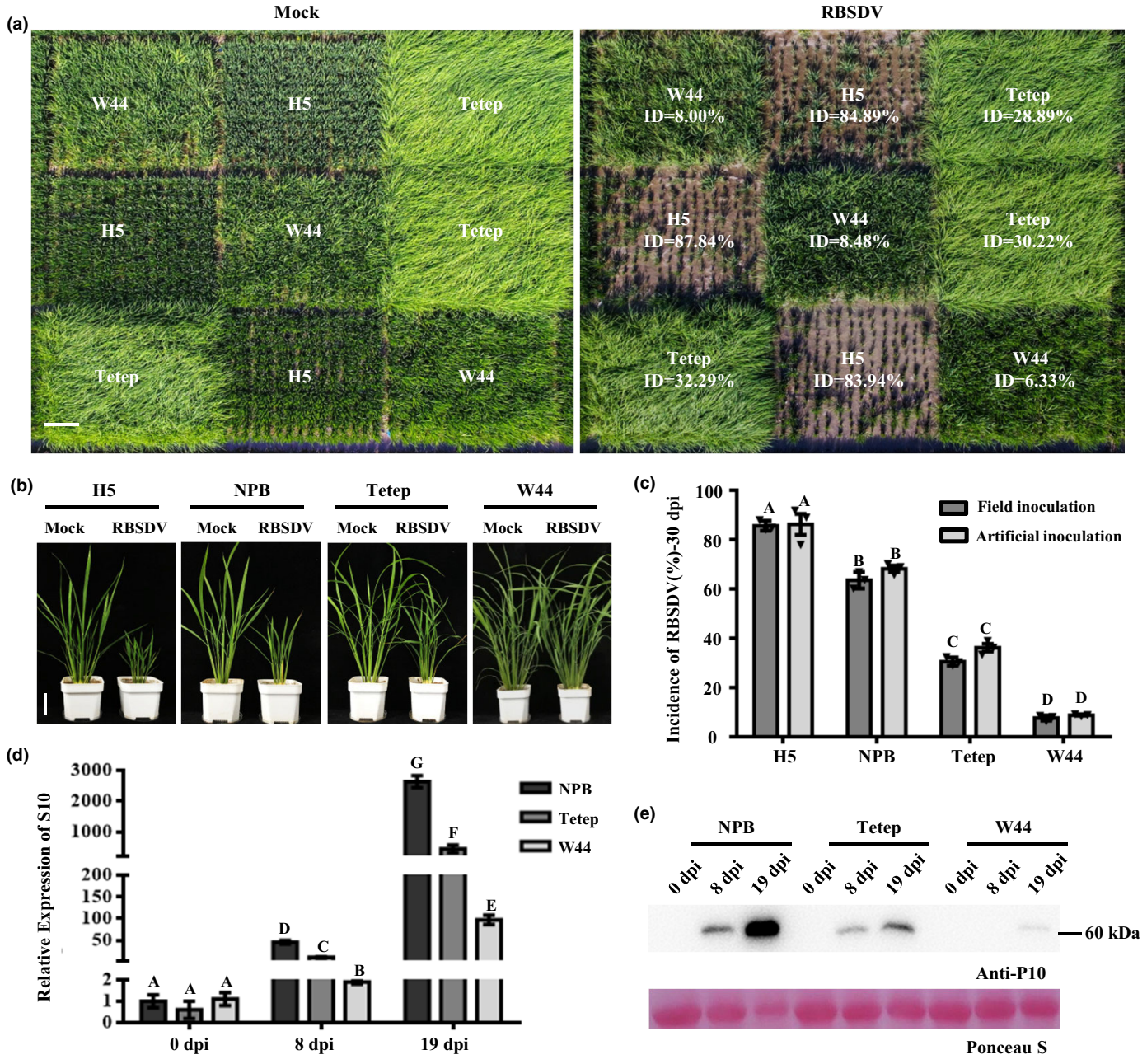
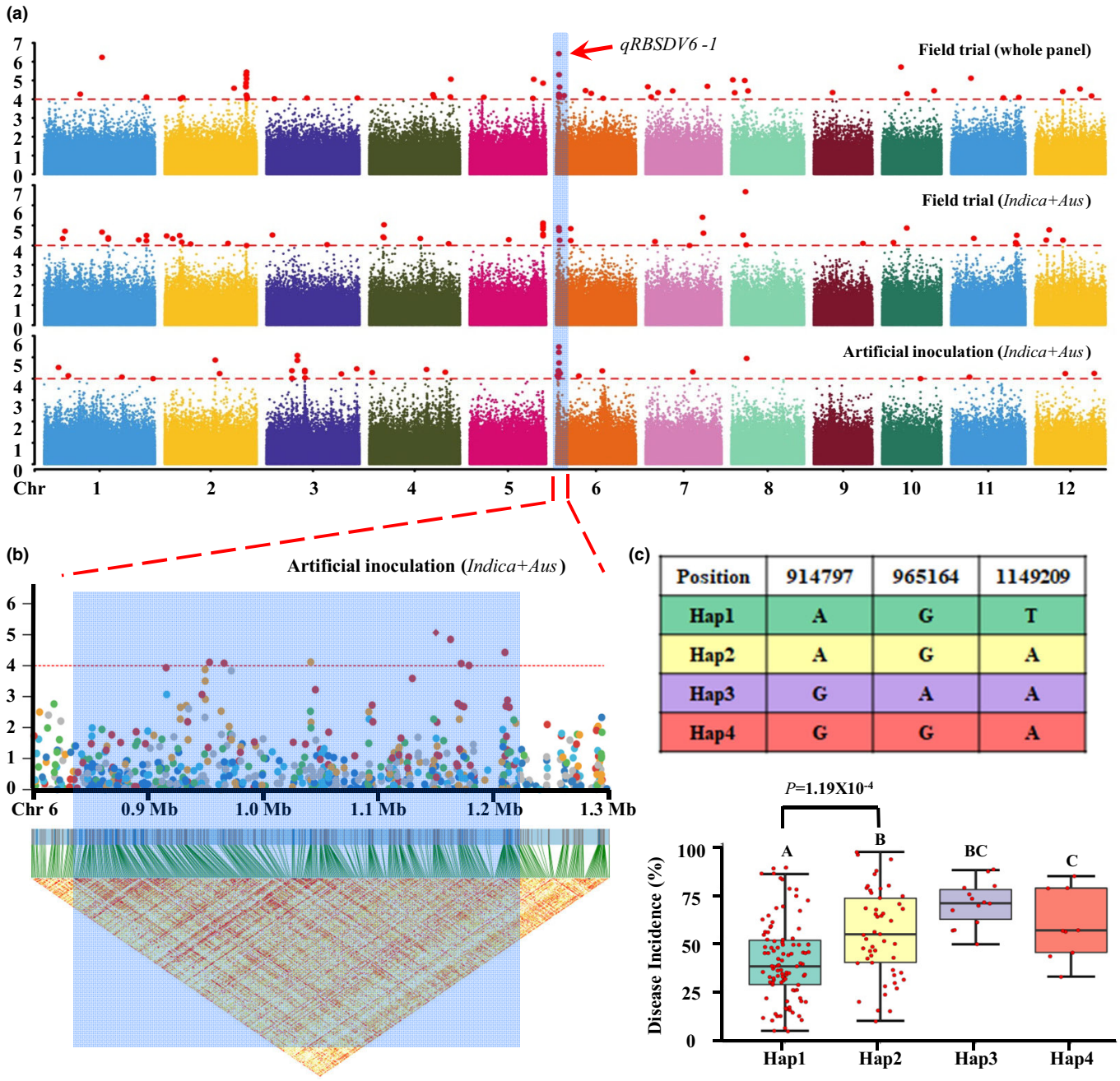


Fig. 1 W44 is a stable resistant accession for *rice black-streaked dwarf virus* (RBSDV). (a) Phenotype and disease incidence (DI) of Huaidao No. 5 (H5), Tetep, and W44 in the natural infection. Bar, 100 cm. (b) Representative Mock- (RBSDV-free SBPH) or RBSDV-inoculated susceptible H5 and Nipponbare (NPB), moderate resistant Tetep, and resistant variety W44 plants were photographed at 30 d post inoculation (dpi). Bar, 4 cm. (c) DI of H5, NPB, Tetep, and W44 in the field and artificial inoculation at 30 dpi. Error bars are the standard deviations (SDs) of different treatments with three biological replicates. Statistical significance between field and aritificial inoculation test was calculated using the Student's *t*-test. Different capital letters, *P*-value < 0.01. (d) Accumulation of RBSDV *S10* RNA transcripts in different assayed plant samples was analyzed through real-time PCR at 0, 8, and 19 dpi. The expression of rice *18s-rRNA* was used as an internal control. Error bars are the SDs of different treatments with three biological replicates. Statistical significance was calculated using the Student's *t*-test. Different capital letters, *P*-value < 0.01. (e) Expression level of RBSDV P10 in different samples was measured through Western blot assay at 0, 8, and 19 dpi. Ponceau-stained Rubisco large protein (RubL) was used to show sample loadings. Similar results were obtained in three independent experiments.

Functional analysis of *OsAP47* involved in RBSDVD and SRBSDVD resistance

To validate the involvement and function of *OsAP47* in RBSDVD resistance, we used the CRISPR/Cas9 technology to knock out the endogenous *OsAP47* gene in the NPB background. We selected 20-nucleotide sequences in the N terminal of the *OsAP47* gene as target sites for Cas9 cleavage, generated multiple putative transgenic lines, and sequenced the target regions after PCR amplification. We found two homozygous



Position	914797	965164	1149209
Hap1	A	G	T
Hap2	A	G	A
Hap3	G	A	A
Hap4	G	G	A

Fig. 2 Genome-wide association study (GWAS) and haplotype analysis of rice black-streaked dwarf virus disease (RBSDVD) resistance. (a) Manhattan plots for GWAS results of RBSDV resistance. Field trial (whole panel): GWAS results using phenotype of whole panel in field trial. Field trial (*Indica*+*Aus*): GWAS results using phenotype of *Indica* and *Aus* subpopulations in field trial. Artificial inoculation (*Indica*+*Aus*): GWAS results using phenotype of *Indica* and *Aus* subpopulations in artificial inoculation. (b) Linkage disequilibrium (LD) heatmaps show the regions surrounding the strong peaks of *qRBSDV6-1* from GWAS results using phenotypes of *Indica* and *Aus* subpopulations in artificial inoculation. (c) Haplotype analysis of *qRBSDV6-1*. Four haplotypes (Hap1, Hap2, Hap3, and Hap4) were identified in *Indica* and *Aus* subpopulation. The box plots show the RBSDV incidence rate of the accession harboring each of the four haplotypes. The center line shows the median, the box shows the interquartile range and the whiskers show no more than 1.5 times the distance between the 25th and 75th percentile. Statistical significance was conducted using the Student's *t*-test. The *y*-axis represents the disease incidence.

mutant lines, namely, *Crispr-OsAP*(2-3) and *Crispr-OsAP*(4-4), that, theoretically, would result in truncated OsAP47 protein and earlier translation stop, respectively (Fig. S4). The main agronomic traits of the homozygous mutant lines, such as grain length, width, thousand grain weight, plant height, and seed setting rate were measured, and no significant differences were observed between the gene editing lines and control lines (Fig. S5). After artificial inoculation of RBSDV for 30 d, the DI of *Crispr-OsAP*(2-3) and *Crispr-OsAP*(4-4) was 30.22% and 33.75%, respectively, while that of NPB was 74.99% (Fig. 4a,b). Our results also show that the accumulation of RBSDV in *Crispr-OsAP* lines decreased remarkably compared to the

Table 1 Quantitative trait loci (QTL) identified for *rice black-streaked dwarf virus* (RBSDV) disease resistance by genome-wide association study (GWAS) using field trial and artificial inoculation phenotypes.

QTL	Significant SNP marker	Chromosome	Position	*P*-value
Field trial (whole panel)				
qRBSDV1	SNP-1.39916258	1	39 917 302	7.62E-05
qRBSDV2-1	SNP-2.7038785	2	7038 786	8.13E-05
qRBSDV2-2	SNP-2.31794353	2	31 800 223	1.38E-05
qRBSDV4	SNP-4.31499590	4	31 684 698	7.49E-05
qRBSDV6-1	SNP-6.1148209	6	1149 209	3.78E-07
qRBSDV6-2	SNP-6.1423236	6	1424 236	2.25E-05
Field trial (*Indica*)				
qRBSDVi1	SNP-1.24984156.	1	24 985 201	3.01E-05
qRBSDVi11	SNP-11.24798031	11	25 264 192	4.95E-05
qRBSDVi2-1	SNP-2.5890232	2	5890 233	2.36E-06
qRBSDVi2-2	SNP-2.31984579	2	31 990 449	2.39E-05
qRBSDVi6	SNP-6.1148209	6	1149 209	5.43E-05
qRBSDVi8	SNP-8.18149010	8	18 151 724	6.42E-05
Field trial (*Indica*+*Aus*)				
qRBSDVia11	SNP-11.25153394	11	25 619 556	3.08E-05
qRBSDVia1-1	SNP-1.25019448.	1	25 020 493	4.12E-05
qRBSDVia1-2	SNP-1.39916258	1	39 917 302	5.78E-05
qRBSDVia4	SNP-4.5698186	4	5702 748	4.11E-05
qRBSDVia5	SNP-5.28580792	5	28 643 438	1.17E-05
qRBSDVia6-1	SNP-6.1148209	6	1149 209	1.29E-05
qRBSDVia6-2	SNP-6.5763300	6	5764 300	1.46E-05
qRBSDVia8	SNP-8.5550018	8	5551 016	2.13E-07
Field trial (*Japonica*)				
qRBSDVj1	SNP-1.38943742	1	38 944 786	2.05E-05
qRBSDVj4	SNP-4.21897884	4	22 051 310	3.21E-05
Artificial inoculation (*Indica*+*Aus*)				
qRBSDVi3-1	SNP-3.10018321	3	10 019 386	4.24E-05
qRBSDVi3-2	SNP-3.15066221	3	15 067 575	4.06E-05
qRBSDVi6	SNP-6.1148209	6	1149 209	3.24E-06

Single nucleotide polymorphism (SNP) marker, the most significant SNP in the identified locus; Chromosome, the chromosome of the QTL locus; Position, the physical position of the most significant SNP based on Reference-IRGPS-1.0.

wild-type plants (Fig. 4c,d). Actually, there were five different genes encoding aspartic protease underlying *qRBSDV6-1* region, including *LOC_Os06g02730*, *LOC_Os06g02780*, *LOC_Os06g02900*, *LOC_Os06g03080*, and *LOC_Os06g03120* (Dataset S5; Fig. 3). In the latest studies, *LOC_Os06g03120* was also considered to be a candidate gene for rice resistance to RBSDVD (Xiao *et al.*, 2019). To further clarify the contribution of other *OsAP* genes in rice resistance to RBSDVD, we generated quintuple mutants of these *OsAP* genes using CRISPR/Cas9 technology (Fig. S6a,b). We analyzed the role of the five *OsAP* genes in rice response to RBSDV infection by using two stable transgenic homozygous lines, *Crispr-OsAP*(14) and *Crispr-OsAP*(23). These results showed that the rice resistance to RBSDV infection in the quintuple mutants of *OsAP* was similar to *Crispr-OsAP47* single mutant plant, as evident by the disease incidence and disease symptoms (Fig. S6c,d). Thus, we conclude that the other four *OsAP* genes may not be involved in RBSDVD resistance.

To further confirm the function of *OsAP47* in RBSDVD resistance, we introduced a fragment containing *OsAP47*-NPB coding sequence driven by the CaMV 35S promoter into W44 background (referred to as *35S:OsAP*). We validated the transgenic lines at transcriptional and translational level by reverse transcription PCR and Western blot (Fig. S7). After inoculation for 30 d, the DI of *35S:OsAP*(#1) and *35S:OsAP*(#2) plants in W44 background were 30.06% and 27.65%, respectively, while that of W44 was 4.56% (Fig. 4e,f). Moreover, overexpressing *OsAP47* enhanced the accumulation of RBSDV at mRNA and protein levels (Fig. 4g,h). These results indicate that *OsAP47* negatively regulates rice resistance to RBSDVD. To our surprise, *OsAP47* also plays an important role in rice resistance to SRBSDVD. Our results indicate that the *OsAP47*-knockout plants exhibited enhanced resistance to SRBSDVD, while the *OsAP47*-overexpression plants were susceptible to SRBSDVD compared to the control plants (Fig. 5a–h). Thus, our results suggest that *OsAP47* negatively regulates rice resistance to RBSDVD and SRBSDVD.

These encouraging results prompted us to investigate if the *OsAP47* also functions in other crops that are hosts of RBSDV and SRBSDV. For this purpose, we conducted a comparative genomic analysis of *qRBSDV6-1* between rice and maize or sorghum. Syntenic blocks corresponding to the *qRBSDV6-1* interval were identified on chromosome 6 in maize and on chromosome 10 in sorghum, and the orthologs of OsAP47 were also identified in both regions (Fig. S8a). Notably, two previous studies had identified a similar locus (*rmrdd6* and *qZZ-MRDD6*) in the maize syntenic region conferring resistance to MRDD, which is also caused by RBSDV in maize (Li *et al.*, 2018; Wang *et al.*, 2019). The maize ortholog of OsAPs (*Zm00001d035981*) is inside this locus. We then further conducted phylogenetic analysis of genes of the *APs* identified from maize and sorghum genomes as well as those from rice and *Arabidopsis* genomes that have been previously identified (Faro & Gal, 2005; Chen *et al.*, 2009). Phylogenetic analysis shows three distinct groups (Groups A, B, and C) of genes of *APs* in plants. OsAPs are in Group C, and especially, the sub-group C1, which is the same as in a previous study (Chen *et al.*, 2009). Two orthologs from maize and sorghum (Zm00001d035981 and SORBI_3010G014600) were similar to that of OsAP47 in the phylogenetic tree (Fig. S8b), and they are also in the same cluster as that of AtAED1 in *Arabidopsis* and CND41 in *N. benthamiana*, which have been proven to function in systemic acquired resistance and programmed cell death in plants (Nakano *et al.*, 2003; Breitenbach *et al.*, 2014).

Nucleotide diversity (π) around the *qRBSDV6-1* region and the location of OsAP47 were analyzed in four subpopulations of rice (*Indica*, *Japonica*, *Aus*, and *Aromatic*) using data from the 3000 Rice Genomes (3k-RG) Project (W. S. Wang *et al.*, 2018). A significantly higher π was seen in *Indica* and *Aus* than in *Japonica* and *Aromatic* (Fig. 6a) indicating that the diversity of *Japonica* in *qRBSDV6-1* is very low. To identify the haplotypes of *OsAP47* that are associated with RBSDV disease incidence and their distribution in *Indica* and *Japonica*, we analyzed the haplotypes of *OsAP47* using our GWAS population. Our results reveal that there were two haplotypes, Hap1 and Hap2 that were associated with RBSDV disease incidence (Fig. 6b). Overall, the average RBSDV disease incidence of the accessions harboring Hap1 was

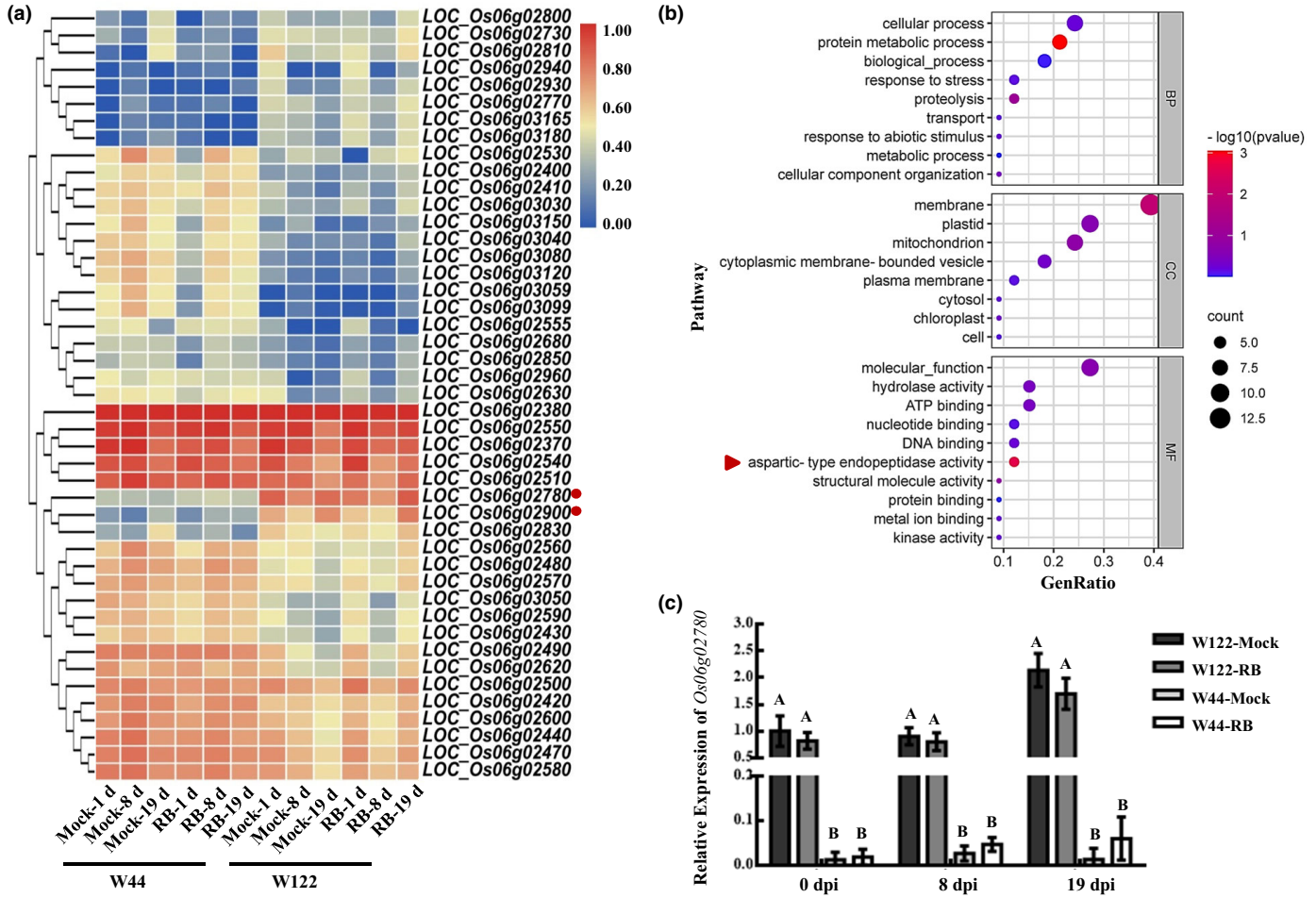


Fig. 3 Screening of candidate gene of *qRBSDV6-1*. (a) Expression profile of the annotated genes in *qRBSDV6-1* region based on RNA-sequencing of W44 (resistant) and W122 (sensitive) at different stages upon Mock or *rice black-streaked dwarf virus* (RBSDV) inoculation. (b) Gene ontology (GO) analysis of the annotated genes in *qRBSDV6-1* region of W44 and W122. BP, biological process; CC, cellular component; MF, molecular function (see also Supporting Information Dataset S6). (c) Real-time PCR verification of the expression of rice *aspartic proteinase 47* (*OsAP47*) between W44 and W122. *18s-rRNA* was taken as the endogenous reference gene. Error bars are the standard deviations (SDs) of different treatments with three biological replicates. Statistical significance was calculated using the Student's *t*-test. Similar results were obtained in three independent experiments. Different capital letters, *P*-value < 0.01.

0.43, which is significantly lower than the accessions harboring Hap2 (0.55) (Fig. 6c). This result suggests that Hap1 is the resistant haplotype. Subpopulation analysis of the two haplotypes showed that among the accessions with Hap1, there were 69 *Indica* accessions and two *Japonica* accessions. However, in Hap2 group, there were 103 *Indica* accessions and 103 *Japonica* accessions (Fig. 6d). These results indicate that Hap1 is almost exclusive in *Indica*, but rare in *Japonica*.

To confirm these results, we further analyzed the two haplotypes in the International Rice Research Institute (IRRI) 3 K dataset containing five *Indica* sub-groups and four *Japonica* subgroups (W. S. Wang *et al.*, 2018). The frequency of these two haplotypes in *Indica* and *Japonica* were calculated according to the 3 K SNP data. For *Indica*, the number of samples harboring Hap1 was 388 out of 1277 (30.38%). For *Japonica*, the number of samples harboring Hap1 was three out of 804 (0.37%) (Fig. 6e). This result is consistent with the earlier results derived from the GWAS population, which signifies that the elite resistant allele of *OsAP47* may mostly exist in *Indica*. We further investigated the worldwide distribution of the two haplotypes (Fig. S9). The results show that the resistant haplotype (Hap1) was present in a considerably high proportion of accessions from mid-east, south Asia and west Africa. In contrast, the presence of Hap1 in accessions from Europe, north America, and east Asia was considerably lower. Accessions from Iran and Pakistan displayed the highest proportion of Hap1.

Discussion

Acquiring reliable and highly resistant germplasms is the prerequisite and basis for the effective breeding of resistant rice. Although considerable efforts have been made toward screening of resistant germplasms in the recent decades, only a few reliable resistant rice accessions have been identified. Specifically, progress

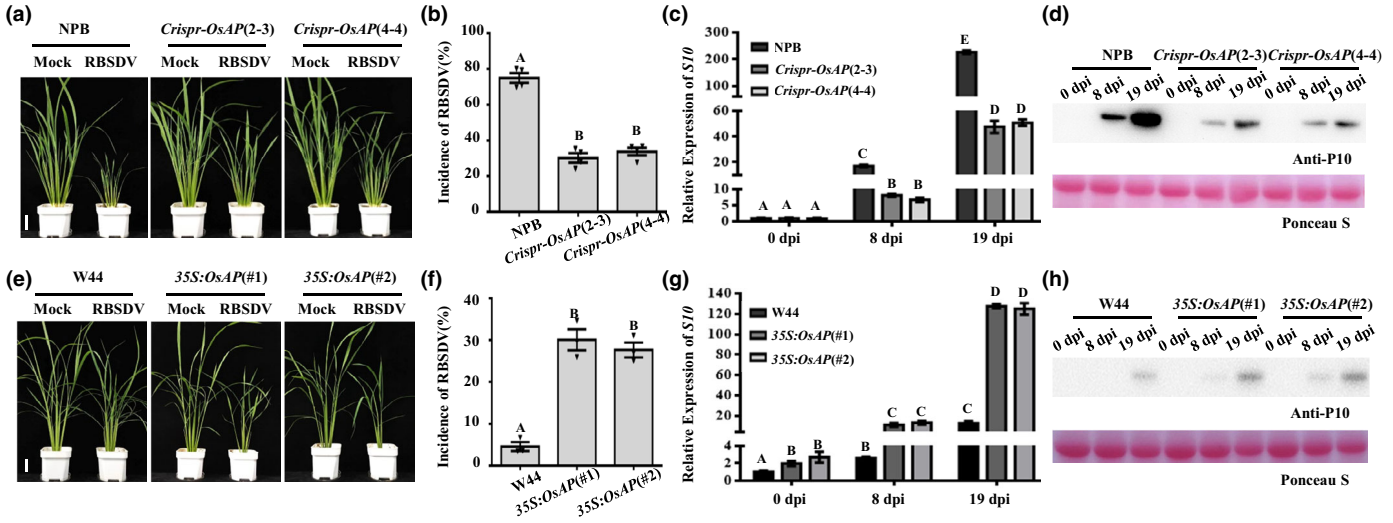


Fig. 4 Knockout of *aspartic proteinase 47* (*OsAP47*) in Nipponbare (NPB) enhances rice resistance to rice black-streaked dwarf virus disease (RBSDVD). (a, b) RBSDV disease incidence (DI), and phenotype of NPB and *Crispr-OsAP* gene editing lines was determined at 30 days post inoculation (dpi). Bar, 4 cm. (e, f) RBSDV DI and phenotype of W44 and OsAP47 overexpressing transgenic lines in W44 background were determined at 30 dpi. Bar, 4 cm. (c, d, g, h) Accumulation of RBSDV *S10* RNA transcript and protein level in different assayed plant samples were analyzed through real-time PCR and Western blot at 0, 8, and 19 dpi. *18s-rRNA* was taken as the endogenous reference gene. Error bars are the standard deviations (SDs) of different treatments with three biological replicates. Statistical significance was calculated using the Student's *t*-test. Ponceau-stained Rubisco large protein (RubL) was used to show sample loadings. Similar results were obtained in three independent experiments. Different capital letters, *P*-value < 0.01.

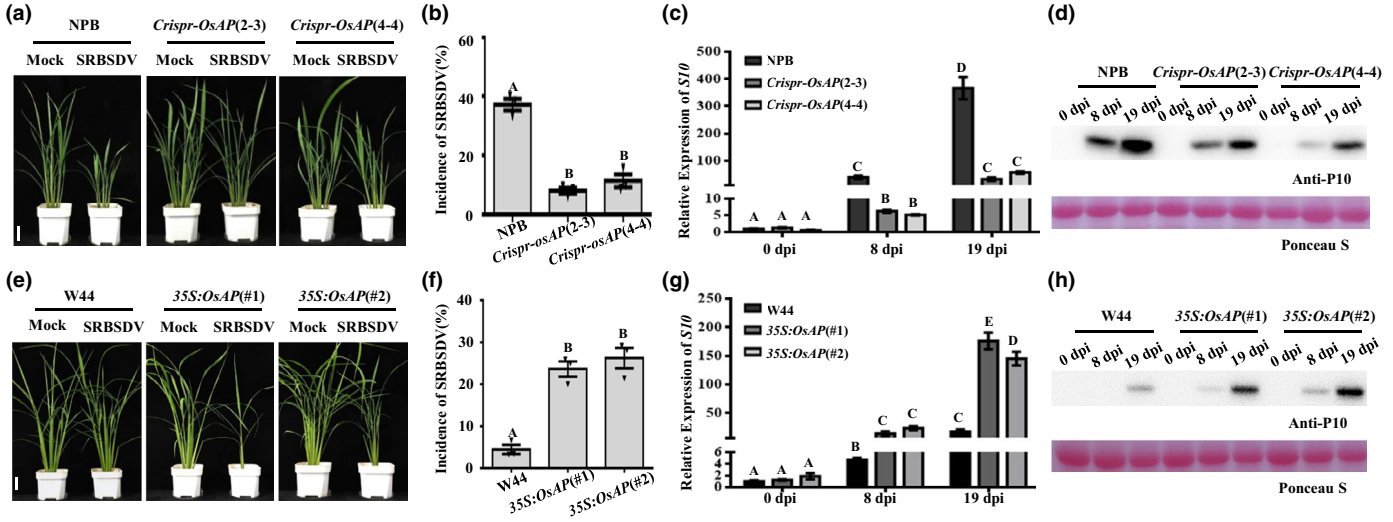


Fig. 5 Knockout of *aspartic proteinase 47* (*OsAP47*) in Nipponbare (NPB) enhances rice resistance to southern rice black-streaked dwarf virus disease (SRBSDVD). (a, b) SRBSDV disease incidence (DI), and phenotype of NPB and *Crispr-OsAP* gene editing lines were determined at 30 days post inoculation (dpi). Bar, 4 cm. (e, f) SRBSDV DI, and phenotype of W44 and OsAP47 overexpressing transgenic lines in W44 background were determined at 30 dpi. Bar, 4 cm. (c, d, g, h) Accumulations of SRBSDV *S10* RNA transcript and protein level in different assayed plant samples were analyzed through real-time PCR and Western blot at 0, 8, and 19 dpi. *18s-rRNA* was taken as the endogenous reference gene. Error bars are the standard deviations (SDs) of different treatments with three biological replicates. Statistical significance was calculated using the Student's *t*-test. Ponceau-stained Rubisco large protein (RubL) was used to show sample loadings. Similar results were obtained in three independent experiments. Different capital letters, *P*-value < 0.01.

in rice breeding for resistance to RBSDVD has not been significant. In the present study, there were several accessions showed resistance to RBSDVD, especially those original from southeast Asia (Dataset S2, S3). These accessions were resistant to another virus also transmitted by SBPH, *rice stripe virus* (RSV), and share a common place of origin. Also, a germplasm belonging to *Indica*, W44 displayed the strongest resistance to RBSDVD with an average DI of 8.75%, which is considerably lower than that of the recognized resistant variety, Tetep that had an average DI of 36.11% (Fig. 1). To avoid the possibility that the resistance of W44 to RBSDVD is from insect resistance, antibiosis and preference tests were conducted, and the results showed no significant difference between W44 and NPB (Fig. S2). Furthermore, RBSDV accumulation can be significantly repressed at mRNA and protein levels in W44 (Fig. 1). All these results suggest that the exhibited RBSDVD resistance in W44 is genuine RBSDV

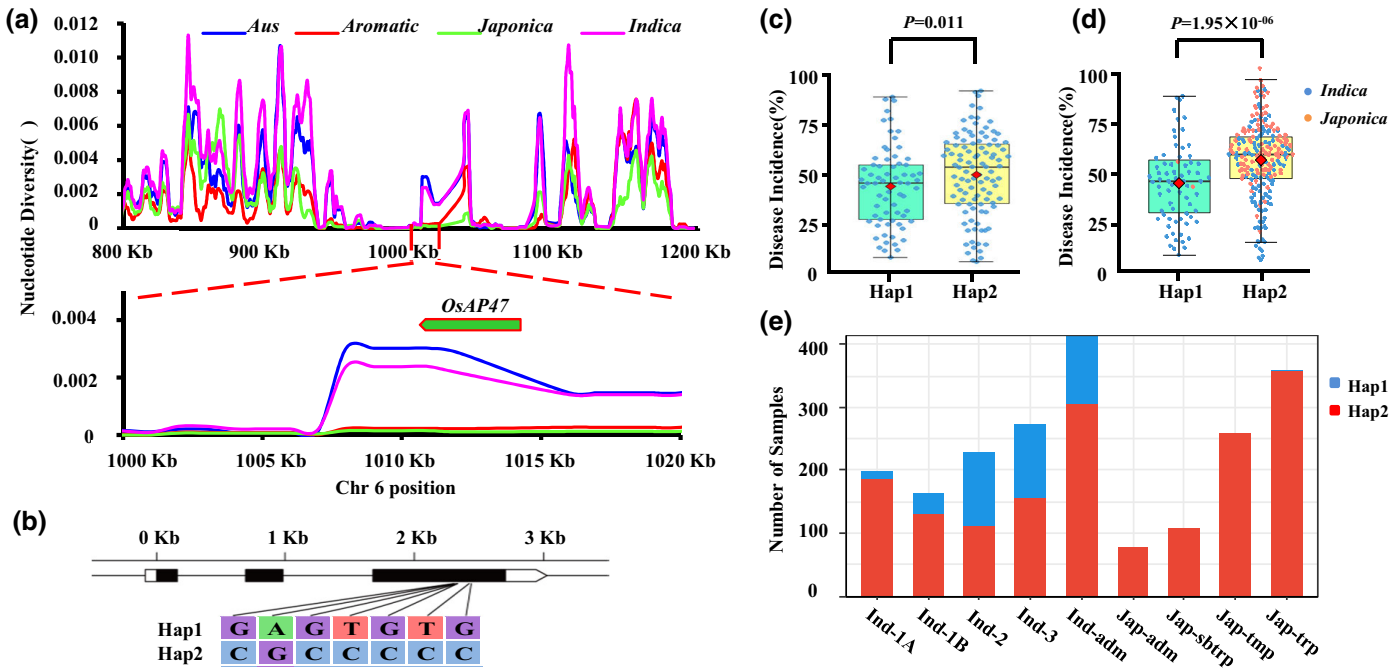


Fig. 6 Haplotype analysis of aspartic proteinase 47 (OsAP47) in rice natural population. (a) Nucleotide diversity (π) of *qRBSDV6-1* region in rice subpopulations. Single nucleotide polymorphism (SNP) from the 3k Rice Genome (RG) project was used to calculate the π in the *qRBSDV6-1* region, and *OsAP47* was placed in its relative position in the chromosome. (b) Sequence comparisons of two haplotypes of *OsAP47* gene region. (c) Box plot of *rice black-streaked dwarf virus* (RBSDV) disease incidence (DI) based on *OsAP47* haplotypes in *Indica* subpopulation in our GWAS population. (d) Box plot of RBSDV DI based on haplotypes *OsAP47* in *Indica* and *Japonica* subpopulation from our genome-wide association study (GWAS) population. Blue dots represent *Indica* accessions. Pink dots represent *Japonica* accessions. Red dot represents the mean of the phenotypes of Hap1 or Hap2. The *P*-values were determined by Student's *t*-test. (e) Haplotype distribution of *OsAP47* in different *Indica* and *Japonica* sub-groups of rice 3 K panel. The blue bars represent the number of accessions with Hap1. The red bars represent the number of accessions with Hap2.

resistance, and not insect resistance. In recent decades, most of the RBSDVD-resistant rice accessions have been identified in field trials, and no evidence has been gathered on whether the disease resistance of these accessions was caused by the plant's resistance to RBSDV or the insect's preference for the plant. Therefore, our current study provides the first confirmed and highly resistant resource for RBSDV in rice.

Rice resistance to RBSDVD is a complex trait that is controlled by not only the genes associated with resistance to the virus but also those to the insect, making RBSDVD-resistant rice breeding much more difficult than other diseases. A complete understanding of the genetic basis of RBSDVD-resistant is essential for resistant rice breeding. Although a few QTLs for RBSDVD resistance have been identified in rice, none of them have been cloned and functionally confirmed. In this study, 17 QTLs for RBSDVD resistance were identified through GWAS in a diverse international rice panel. Notably, among all the QTLs identified, the QTL on chromosome 6, *qRBSDV6-1*, was detected through consideration of different subpopulations (Table 1). Additionally, this QTL overlapped with previously identified QTLs for RBSDVD resistance using different rice germplasms in other studies (Zheng *et al.*, 2012; Zhang *et al.*, 2016; Feng *et al.*, 2019, 2020; Xiao *et al.*, 2019), implying that it can be stably expressed in different genetic backgrounds and environmental conditions, and therefore has significant potential in rice breeding.

The lack of understanding of the genetic basis of functional genes underlying RBSDVD resistance has hindered molecular breeding strategies for disease control. By combining transcriptome analysis of the differentially expressed genes between W44 and W122, *OsAP47* was selected as the candidate gene of *qRBSDV6-1* for further analysis. Compared to their wild-type plants, *OsAP47* knockout mutants exhibited enhanced RBSDVD resistance as manifested by significantly lower DI and accumulation of RBSDV, while the gene overexpressing plants exhibited enhanced RBSDVD susceptibility. To our surprise, the knockout mutants also exhibited enhanced SRBSDVD resistance, while the gene overexpressing plants exhibited enhanced SRBSDVD susceptibility compared to their wild-type plants. From these results, we conclude that *OsAP47* is the right functional gene of *qRBSDV6-1* and it negatively regulates RBSDVD and SRBSDVD resistance in rice. To our knowledge, this is the first cloned and functionally confirmed resistance related gene for RBSDVD and SRBSDVD resistance in rice. *OsAP47* gene editing lines can not only improve plant resistance to RBSDVD and SRBSDVD but will also not affect the main agronomic traits of rice. The gene against RBSDVD and SRBSDVD identified in this study provides innovative information for molecular breeding of resistant rice cultivars.

APs are one of the four mechanistic classes of proteolytic enzymes, and comprise a large family in plants (Dunn, 2002). A total of 51 and 96 APs have been identified in *Arabidopsis* and

rice genomes, respectively (Faro & Gal, 2005; Chen *et al.*, 2009). Based on their sequence features, plant APs are divided into three categories, which are typical APs, nucellin-like APs and atypical APs (Soares *et al.*, 2019). A number of functions of atypical APs have been uncovered, including response to biotic or abiotic stresses (Soares *et al.*, 2019; Xia *et al.*, 2020). Specifically, OsAP47 is designated as an atypical AP and is close to AtAED1 from *Arabidopsis*. AtAED1 had been identified to act downstream of SA to suppress systemic immunity. It was suggested that AtAED1 may degrade apoplastic proteins, including PRs, that accumulated during this systemic response as a feedback mechanism (Breitenbach *et al.*, 2014). The close relation between OsAP47 and AtAED1 may shed light on the molecular function of OsAP47 in the digestion of PRs. Currently, there is no definitive evidence that apoplastic PRs play any essential roles in induced resistance to viruses, while there is abundant evidence that they have a variety of functions in limiting the spread of cellular pathogens (Carr *et al.*, 2019; Murphy *et al.*, 2020). Therefore, it is not possible to speculate that OsAP47, similar to AtAED1, regulates RBSDVD resistance by degrading PR proteins. The cell-to-cell movement of plant virus is an essential step in viral infections. This process is facilitated by specific virus-encoded movement proteins (MPs), which manipulate the cell-wall channels between neighboring cells known as plasmodesmata (Pd). Plant viruses can assemble a hollow tubule extending between cells to allow virus movement. Recently, it was reported that *citrus psorosis virus* (CPsV) induces the formation of tubular structures at Pd during infection in citrus. A specific aspartic protease motif affects tubule formation of Pd and alters the subcellular localization of the MP to support the movement of CPsV (Luna *et al.*, 2018). Perhaps, like the CPsV, the host's susceptibility factor OsAP47 can be utilized by certain RBSDV proteins in favor of the infection and movement of virus.

In a previous study, a helitron-induced *RabGDIα* variant was identified as causing quantitative recessive resistance to MRDD, which is also caused by RBSDV (Q. C. Liu *et al.*, 2020; Q. Liu *et al.*, 2020). However, no homologous gene for *RabGDIα* is found in rice. The role of APs in disease resistance in plants and the fact that OsAP47 functions in RBSDVD and SRBSDVD in rice prompted us to investigate if OsAP47 also functions in other viral diseases in different crops, particularly those that can be infected by RBSDV or SRBSDV. Our comparative genomic analysis showed that two previously identified loci (*rmrdd6* and *qZZ-MRDD6*) for MRDD resistance, which is also caused by RBSDV, overlaps with the maize syntenic region corresponding to the *qRBSDV6-1* interval (Li *et al.*, 2018; Wang *et al.*, 2019) (Fig. S8a). In addition, our phylogenetic analysis of AP genes in *Arabidopsis*, rice, maize, and sorghum revealed that OsAP47 together with its orthologs in maize and sorghum (*Zm00001d035981* and *SORBI_3010G014600*) are very close to AtAED1 from *Arabidopsis* and CND41 from tobacco (Fig. S8b). Based on the role of APs in disease resistance in plants and the results from our comparative genomic analysis, it can be inferred that OsAP47 might not only function in RBSDVD and SRBSDVD in rice, but also in other viral diseases in crops, in particular, those that can be infected by RBSDV or SRBSDV. Transgenic experiments in different crops may address this issue.

In this study, we have identified six RBSDVD-resistant rice accessions including the first confirmed and highly RBSDVD resistant rice accession no. W44, thus, breaking through the bottlenecks of RBSDVD resistant rice breeding. We have detected 17 QTLs for RBSDVD resistance, particularly *qRBSDV6-1* and its resistant haplotype that have a major effect and are stably expressed, thereby improving our understanding of the genetic basis of RBSDVD resistance. Furthermore, we have achieved the first clone and functional confirmation of the resistance related gene for RBSDVD and SRBSDVD in rice. W44 can serve as a good resistant donor parent in rice resistance breeding for RBSDVD. The QTL *qRBSDV6-1* provides a promising target for rice molecular breeding. With the identification of the resistant haplotype of *qRBSDV6-1* (Hap1), haplotype-based marker assisted screening (MAS) for RBSDVD resistance can be performed in rice breeding to confer resistance to both RBSDVD and SRBSDVD. *OsAP47* has two haplotypes associated with RBSDVD incidence, Hap1 and Hap2 (Fig. 6). Hap1 is almost exclusive in *Indica*, and rare in *Japonica*. It provides a promising target for more specific improvement of RBSDVD and SRBSDVD resistance in rice through gene editing. Particularly, W44 and most of the resistant accessions identified in our study belong to *Indica* (Dataset S2) and they are difficult to be used as donor parents for the improvement of RBSDVD resistance in *Japonica* due to the problem of fertility in crossing between the two sub-species. Probably, knockout of *OsAP47* by editing is the easiest and most effective method to improve RBSDVD resistance in *Japonica* rice. Furthermore, in terms of the role of APs in disease resistance and the results from our comparative genomic analysis, it can be inferred that *OsAP47* might also function against RBSDVD or SRBSDVD in other crops. No other resistance related gene for RBSDVD has been cloned and functionally confirmed except for *OsAP47* in rice. Therefore, our finding is significant as RBSDV infects many important crops such as rice, maize, wheat, barley, and sorghum. Once the functions of *OsAP47* in resistance in these crops are confirmed, molecular breeding for RBSDVD and SRBSDVD resistant crops can significantly contribute toward viral disease control and facilitate food security in the world.

Accession numbers

Sequence data from this study can be found in the Rice Genome Database (http://rice.plantbiology.msu.edu), GenBank (https://www.ncbi.nlm.nih.gov/genbank) under the accession nos. *OsAP47* (*LOC_Os06g02780*), RBSDV *S10* (*AF227205*), SRBSDV *S10* (*JQ927009*).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (31761143012), the National Key R&D Program of China (2018YFD0200302), and the Innovation Team Project of Guangdong Modern Agricultural Industrial System

(2020KJ106 and 2021KJ106). The authors thank associate Prof. Dongdong Niu (Nanjing Agricultural University, Nanjing) for stimulating discussion and Dr Yeqiang Xia (Nanjing Agricultural University, Nanjing) for critical technique assistance. The authors declare that they have no competing interests.

Author contributions

JZ, BL, YZ and TZ conceived the project and designed the experiments; ZW and LZ carried out the experiments with assistance from YL, XL, JW, JD, WG, QL, SZ, ZL, WS, FS, LD, YS, JL and SL; ZW analysed the results with assistance from WY, TY, HF, YM, LC, YF, YW and HL; ZW, JZ, LZ, YL and DJ prepared the plant transgenic materials; ZW and LZ wrote the manuscript. All authors reviewed and approved the final article.

ORCID

Shuo Li https://orcid.org/0000-0002-0811-5166
Bin Liu https://orcid.org/0000-0002-3736-7522
Zhaoyun Wang https://orcid.org/0000-0001-5884-7982
Junliang Zhao https://orcid.org/0000-0003-3194-3290
Tong Zhou https://orcid.org/0000-0001-8706-3001

Data availability

All data supporting the conclusions of this article are provided within the article (and the additional Supporting Information).

References

Achon MA, Serrano L, Sabate J, Porta C. 2015. Understanding the epidemiological factors that intensify the incidence of maize rough dwarf disease in Spain. *Annals of Applied Biology* **166**: 311–320.

Breitenbach HH, Wenig M, Wittek F, Jorda L, Maldonado-Alconada AM, Sarioglu H, Colby T, Knappe C, Bichlmeier M, Pabst E *et al.* 2014. Contrasting roles of the apoplastic aspartyl protease apoplastic, enhanced disease susceptibility-dependent and legume lectin-like protein 1 in *Arabidopsis* systemic acquired resistance. *Plant Physiology* **165**: 791–809.

Carr JP, Murphy AM, Tungadi T, Yoon JY. 2019. Plant defense signals: players and pawns in plant-virus-vector interactions. *Plant Science* **279**: 87–95.

Chen J, Ouyang Y, Wang L, Xie W, Zhang Q. 2009. *Aspartic proteases* gene family in rice: gene structure and expression, predicted protein features and phylogenetic relation. *Gene* **442**: 108–118.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST *et al.* 2011. The variant call format and VCFtools. *Bioinformatics* **27**: 2156–2158.

Dovas CI, Eythymiou K, Katis NI. 2004. First report of *Maize rough dwarf virus* (MRDV) on maize crops in Greece. *Plant Pathology* **53**: 238.

Dunn BM. 2002. Structure and mechanism of the pepsin-like family of aspartic peptidases. *Chemical Reviews* **102**: 4431–4458.

Faro C, Gal S. 2005. Aspartic proteinase content of the *Arabidopsis* genome. *Current Protein Peptide Science* **6**: 493–500.

Feng Z, Kang H, Li M, Zou L, Wang X, Zhao J, Wei L, Zhou N, Li Q, Lan Y *et al.* 2019. Identification of new rice cultivars and resistance loci against rice black-streaked dwarf virus disease through genome-wide association study. *Rice* **12**: 49.

Feng Z, Yuan M, Zou J, Wu L-B, Wei L, Chen T, Zhou N, Xue W, Zhang Y, Chen Z *et al.* 2020. Development of marker-free rice with stable and high resistance to rice black-streaked dwarf virus disease through RNA interference. *Plant Biotechnology Journal* **19**: 212–214.

Guo J, Xu C, Wu DI, Zhao Y, Qiu Y, Wang X, Ouyang Y, Cai B, Liu X, Jing S *et al.* 2018. Bph6 encodes an exocyst-localized protein and confers broad resistance to planthoppers in rice. *Nature Genetics* **50**: 297–306.

Hiei Y, Ohta S, Komari T, Kumashiro T. 1994. Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *The Plant Journal* **6**: 271–282.

Hoang AT, Zhang HM, Yang J, Chen JP, Hebrard E, Zhou GH, Vien VN, Cheng JA. 2011. Identification, characterization, and distribution of *Southern rice black-streaked dwarf virus* in Vietnam. *Plant Disease* **95**, 1063–1069.

Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, Schwartz DC, Tanaka T, Wu J, Zhou S *et al.* 2013. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice* **6**: 4.

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* **12**: 357–360.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**: 1547–1549.

Lenardon SL, March GJ, Nome SF, Ornaghi JA. 1998. Recent outbreak of "*Mal de Rio Cuarto*" *virus* on corn in Argentina. *Plant Disease* **82**: 448.

Li A, Pan C, Wu L, Dai Z, Zuo S, Xiao N, Yu L, Li Y, Zhang X, Xue W *et al.* 2013. Identification and fine mapping of qRBSDV-6 (MH), a major QTL for resistance to rice black-streaked dwarf virus disease. *Molecular Breeding* **32**: 1–13.

Li R, Song W, Wang B, Wang J, Zhang D, Zhang Q, Li X, Wei J, Gao Z. 2018. Identification of a locus conferring dominant resistance to maize rough dwarf disease in maize. *Scientific Reports* **8**: 1–10.

Liu QC, Deng SN, Liu BS, Tao YF, Ai HY, Liu JJ, Zhang YZ, Zhao Y, Xu ML. 2020. A helitron-induced *RabGDI alpha* variant causes quantitative recessive resistance to maize rough dwarf disease. *Nature Communications* **11**: 495.

Liu Q, Lan G, Zhu Y, Chen K, Shen C, Zhao X, Zhang F, Xu J, Li Z. 2020. Genome-wide association study on resistance to rice black-streaked dwarf disease caused by *rice black-streaked dwarf virus. Plant Disease* **105**: 607–615.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**: 402–408.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550.

Luna GR, Pena EJ, Borniego MB, Heinlein M, Garcia ML. 2018. *Citrus psorosis virus* movement protein contains an aspartic protease required for autocleavage and the formation of tubule-like structures at plasmodesmata. *Journal of Virology* **92**: e00355–e00418.

Ma X, Zhang Q, Zhu Q, Liu W, Chen Y, Qiu R, Wang B, Yang Z, Li H, Lin Y *et al.* 2015. A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Molecular Plant* **8**: 1274–1284.

McCouch SR, Wright MH, Tung C-W, Maron LG, McNally KL, Fitzgerald M, Singh N, DeClerck G, Agosto-Perez F, Korniliev P *et al.* 2016. Open access resources for genome-wide association mapping in rice. *Nature Communications* **7**: 10532.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology Evolution* **37**: 1530–1534.

Murphy AM, Zhou T, Carr JP. 2020. An update on salicylic acid biosynthesis, its induction and potential exploitation by plant viruses. *Current Opinion in Virology* **42**: 8–17.

Nakano T, Nagata N, Kimura T, Sekimoto M, Kawaide H, Murakami S, Kaneko Y, Matsushima H, Kamiya Y, Sato F *et al.* 2003. CND41, a chloroplast nucleoid protein that regulates plastid development, causes reduced gibberellin content and dwarfism in tobacco. *Physiol Plantarum* **117**: 130–136.

Nong B, Qin B, Xia X, Yang X, Zhang Z, Zeng, Y, Deng G, Cai J, Li Z, Liu P *et al.* 2019a. Genetic analysis and fine mapping of a major QTL for the resistance to southern rice black-streaked dwarf disease. *China Journal Rice Science* **33**: 135–143.

Nong B, Qin B, Xia X, Yang X, Zhang Z, Zeng, Y, Liu, C, Cai J, Xie H, Cui L *et al.* 2019b. Genome-wide association study of seedling resistance of *southern rice black-streaked dwarf virus. Molecular Plant Breeding* **17**: 1069–1079.

Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols* **11**: 1650–1667.

Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology* **33**: 290–295.

Shikata E, Kitagawa Y. 1977. *Rice black-streaked dwarf virus* – its properties, morphology and intracellular-localization. *Virology* **77**: 826–842.

Soares A, Carlton SMR, Simoes I. 2019. Atypical and nucellin-like aspartic proteases: emerging players in plant developmental processes and stress responses. *Journal of Experimental Botany* **70**: 2059–2076.

Sun Z, Liu Y, Xiao S, Hu J, Pan G, He J, Xu T, Huang J, Qiu Z, Fan D *et al.* 2017. Identification of quantitative trait loci for resistance to rice black-streaked dwarf virus disease and small brown planthopper in rice. *Molecular Breeding* **37**: 72.

Tang Y, Liu X, Wang J, Li M, Wang Q, Tian F, Su Z, Pan Y, Liu DI, Lipka AE *et al.* 2016. GAPIT v.2: an enhanced integrated tool for genomic association and prediction. *Plant Genome* **9**. doi: 10.3835/plantgenome2015.11.0120.

Wang H, Xu D, Pu L, Zhou G. 2014. *Southern rice black-streaked dwarf virus* alters insect vectors' host orientation preferences to enhance spread and increase *rice ragged stunt virus* co-infection. *Phytopathology* **104**: 196–201.

Wang J, Qi M, Liu J, Zhang Y. 2015. CARMO: a comprehensive annotation platform for functional exploration of rice multi-omics data. *The Plant Journal* **83**: 359–374.

Wang WS, Mauleon R, Hu ZQ, Chebotarov D, Tai SS, Wu ZC, Li M, Zheng TQ, Fuentes RR, Zhang F *et al.* 2018. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* **557**: 43–49.

Wang X, Yang Q, Dai Z, Wang Y, Zhang Y, Li B, Zhao W, Hao J. 2019. Identification of QTLs for resistance to maize rough dwarf disease using two connected RIL populations in maize. *PLoS ONE* **14**: e0226700.

Wang Z, Chen D, Sun F, Guo W, Wang W, Li X, Lan Y, Du L, Li S, Fan Y *et al.* 2021. *ARGONAUTE 2* increases rice susceptibility to *rice black-streaked dwarf virus* infection by epigenetically regulating *HEXOKINASE 1* expression. *Molecular Plant Pathology* **22**: 1029–1040.

Wang Z, Xia Y, Lin S, Wang Y, Guo B, Song X, Ding S, Zheng L, Feng R, Chen S *et al.* 2018. Osa-miR164a targets *OsNAC60* and negatively regulates rice immunity against the blast fungus *Magnaporthe oryzae*. *The Plant Journal* **95**: 584–598.

Wei Y, Li X, Chen Y, Liu K, Guo S. 2019. QTL analysis and utilization of resistance of infiltration lines of *Oryza sativa* to *southern rice black-streaked dwarf virus*. *Guizhou Agricultural Sciences* **47**: 1–5.

Xia Y, Ma Z, Qiu M, Guo B, Zhang Q, Jiang H, Zhang B, Lin Y, Xuan M, Sun L *et al.* 2020. N-glycosylation shields *Phytophthora sojae* apoplastic effector *PsXEG1* from a specific host aspartic protease. *Proceedings of the National Academy of Sciences, USA* **117**: 27685–27693.

Xiao S, Wang B, Liu Y, Miao T, Zhang H, Wen P, He J, Huang J, Liu D, Qiu Z *et al.* 2019. Genome-wide association study and linkage analysis on resistance to rice black-streaked dwarf virus disease. *Molecular Breeding* **39**: 73.

Xu Q, Zhou Y. 2015. Whole-genome expression analysis of *rice black-streaked dwarf virus* in different plant hosts and small brown planthopper. *Phytopathology* **105**: 152.

Yin L, Zhang H, Tang Z, Xu J, Yin D, Zhang Z, Yuan X, Zhu M, Zhao S, Li X *et al.* 2021. rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics, Proteomics & Bioinformatics*. doi: 10.1016/j.gpb.2020.10.007.

Zhang H, Ge Y, Wang M, Liu J, Si H, Zhang L, Liang G, Gu M, Tang S. 2016. Mapping QTLs conferring resistance to rice black-streaked dwarf disease in rice (*Oryza sativa* L). *Euphytica* **212**: 323–330.

Zheng T, Yang J, Zhong W, Zhai H, Zhu L, Fan F, Ali AJ, Yang J, Wang J, Zhu J *et al.* 2012. Novel loci for field resistance to *black-streaked dwarf* and *stripe viruses* identified in a set of reciprocal introgression lines of rice (*Oryza sativa* L.). *Molecular Breeding* **29**: 925–938.

Zhou T, Du L, Wang L, Wang Y, Gao C, Lan Y, Sun F, Fan Y, Wang G, Zhou Y. 2015. Genetic analysis and molecular mapping of QTLs for resistance to rice black-streaked dwarf disease in rice. *Scientific Report* **5**: 10509.

Zhou T, Nelson SC, Hu J, Wang L, Fan Y, Cheng Z, Zhou Y. 2011. Inheritance and mechanism of resistance to rice stripe disease in cv. Zhendao 88, a Chinese rice cultivar. *Journal of Phytopathology* **159**: 159–164.

Zhou T, Zeng B, Lan Y, Gu T, Hu X, Xu M, Du L. 2016. *Technical regulation for identification and evaluation of rice black-streaked dwarf virus disease resistance in rice variety test*. NY/T 2955-2016. Agricultural Industry Standards of the People's Republic of China.

Supporting Information

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

Dataset S1 Primers used in this study.

Dataset S2 *Rice black-streaked dwarf virus* (RBSDV) disease incidence of 509 rice accessions in field trial.

Dataset S3 *Rice black-streaked dwarf virus* (RBSDV) disease incidence of *Indica* and *Aus* accessions in artificial inoculation.

Dataset S4 Accessions showing disease incidence < 15% in the field and artificial inoculation tests.

Dataset S5 Annotated genes in *qRBSDV6-1* region.

Dataset S6 Normalized counts of RNA-sequencing of *qRBSDV6-1* candidate genes with and without *rice black-streaked dwarf virus* (RBSDV) inoculation.

Fig. S1 Box plots of *rice black-streaked dwarf virus* (RBSDV) disease incidence of four major rice subpopulations in the field.

Fig. S2 Analysis of rice varieties against insect pest.

Fig. S3 Sequence comparison of rice aspartic proteinase 47 (OsAP47) in W122, NPB, and W44.

Fig. S4 Validation of *Crispr-OsAP* homozygous lines.

Fig. S5 The agronomic trait of *Crispr-OsAP* homozygous lines.

Fig. S6 Validation of *Crispr-OsAP* quintuple mutant lines.

Fig. S7 Validation of rice aspartic proteinase 47 (OsAP47) over-expressing transgenic lines.

Fig. S8 Comparative genomic analysis of rice aspartic proteinase 47 (OsAP47).

Fig. S9 Geographic distribution of rice aspartic proteinase 47 (OsAP47) two haplotypes in international rice accessions.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.