

## Sequence Analysis of Segment 8 of Five Chinese Isolates of Rice Gall Dwarf Virus and Expression of a Main Outer Capsid Protein in *Escherichia coli* \*

Ming-rong DENG, Xiao-lei RUAN, Fu-xiu LIU, Qin ZHAO and Hua-ping LI \*\*

(*Laboratory of Plant Virology, South China Agricultural University, Guangzhou, Guangdong, 510642, China*)

**Abstract:** The rice gall dwarf disease, caused by the *Rice gall dwarf virus* (RGDV) is a serious disease occurring in rice in many regions of Guangdong province. As a basis to control the disease we have studied the genomic diversity of a variety of isolates from different locations. Genome segment 8 (S8), encoding a main outer capsid protein (Pns8) of RGDV five isolates (BL, CH, DQ, GZ, XY) from Guangdong province was cloned and sequenced. The results revealed that all the S8 segments of the five isolates consisted of 1 578 nucleotides and had a single open reading frame (ORF) extending for 1 301 nucleotides from nucleotide 21 which encoded a polypeptide of 426 amino acids with an estimated molecular weight of 47.4 kDa. The S8 full-length sequence and the ORF sequence shared 97.3%-98.8% and 97.3%-99.1% nucleotide sequence identities within the five Chinese isolates, and shared 94.8%-95.6% and 95.0%-96.0% identities with those of the Thailand isolate respectively. The deduced amino acid sequence of Pns8 in GZ isolate was identical to that in the Thailand isolate, while the amino acid sequence variability of Pns8 within five Chinese isolates ranged from 0.5% to 2.1%. These results indicate that the S8 segment of RGDV is highly conserved in different isolates from different locations. The S8 cDNA from the XY isolate was cloned into the plasmid vector pET-28b(+) and a fused expression protein with an apparent molecular mass of 51kDa was specifically detected in an analysis of *Escherichia coli* Rosetta (DE3) II cells. To our knowledge, this is the first report on analysis of the RGDV segment 8 sequence and genetic comparison of different RGDV isolates and their protein expression.

**Key words:** *Rice gall dwarf virus*; Segment 8; Sequence analysis; Prokaryotic expression

Rice gall dwarf disease is one of the most serious diseases in rice in Southern China, Thailand, Korea and other Southeastern Asian countries (2, 11, 13). It was first discovered in Thailand in 1979 (11). It has erupted several times in Southern China since 1980s,

and is still represents one of the main rice diseases in these regions. The symptoms include stunting of the plants, gall formation along leaf blades and sheaths, and dark green discoloration (2, 11).

*Rice gall dwarf virus* (RGDV), the pathogen of this

Received: 2007-01-09, Accepted: 2007-05-09

\* Foundation item: National natural science foundation of China (30370929) and Guangdong province natural science foundation (C036845)

\*\* Corresponding author. Tel: +86-20-85281107, Email: huaping@scau.edu.cn

disease, is grouped in the genus *Phytoreovirus* (11, 12) in the family of *Roeviridae* (12). The virus multiplies in leafhopper vectors such as *Nephrotettix nigropictus* and *N. cinticeps*. It is transmitted in a persistent manner and restricted within the phloem cells. RGDV, as well as *Rice dwarf virus* (RDV) and *Wound tumor virus* (WTV), is an icosahedral double-shelled particle approximately 65-70 nm in diameter that contains 12 genome segments of dsRNA (4, 10, 12), which have been named S1 to S12 according to their electrophoretic mobility in polyacrylamide gels.

RNAs of all 12 genome segments of RDV and WTV have been sequenced (8,9,15). But in the case of RGDV, only the nucleotide sequences of genome segments S2 (8), S3 (15), S5 (3), S8 (10), S9 (4), S10 (10), and S11 (9) of Thailand isolate have been reported.

In China, this disease broke out first in Xinyin County, Guangdong Province. Fortunately, it was successfully controlled by controlling its principal vectors at the early stage of rice growth. But in recent years, the presence of RGDV has also been discovered in other Counties, such as GaoZhou and DeQin, in Guangdong Province.

To determine the epidemiology of the rice gall dwarf disease, an investigation was carried out in China. This disease was found to be present in other two counties of Guangdong Province, BoLuo and CongHua, which revealed that RGDV had been continuing to spread. How RGDV has been introduced to these areas remains unknown since the leafhopper could not migrate for long distances and raised the additional question as to why is RGDV sporadic in Thailand while epidemic in South China? Previous reports showed that some differences were found between isolates from China and from Thailand in

their principal vectors and artificial inoculation hosts (2). To understand the relationships between Chinese and Thailand isolates, the structure and function of RGDV isolate, and viral transmission in South China, the RGDV segment S8, encoding the outer major capsid protein (Pns8) from five different locations in China, was sequenced and Pns8 of one isolate (XY) was analyzed in this study. To our knowledge, this is the first report on sequence analysis and genetic comparison of segment 8 from difference RGDV isolates, and the expression of the capsid protein.

## MATERIALS AND METHODS

### Sample collection and preparation

Leaf samples of infected rice showing typical RGDV symptoms were collected from five Counties - XinYi (XY), GaoZhou (GZ), DeQin (DQ), BoLuo (BL), CongHua (CH) - in Guangdong Province, China. XY and GZ are neighbors, while the others are hundreds of kilometers away from each other. For control purposes, samples of uninfected plants were also collected. All samples were preserved at -80°C until total RNA extraction.

Total RNAs of diseased or healthy leaves were extracted using a Total RNA Extraction (Plant Leaf) Kit (Watson, China), according to the manufacturer's instructions. Total RNAs were re-suspended in diethylpyrocarbonate (DEPC)-treated double distilled water and used as templates for RT-PCR.

The entire segment 8 (S8) of these five isolates were amplified respectively by RT-PCR using a one-step RT-PCR Kit (TaKaRa, China). The primers used were S8-P1 (5'-GGTATTTTGTACCAACACG-3') and S8-P2 (5'-ATCATTGGTGACCACAC-3') designed according to the S8 sequence of the Thailand

isolate (TL) (*GenBank* Accession number: D13410). The detailed reaction conditions were as follows: reverse transcription at 50°C for 30 min, pre denaturation at 94°C for 5 min, followed by a standard PCR procedure with 30 cycles of denaturation at 94°C for 1 min, annealing at 53°C for 1 min, and extension at 72°C for 2 min. A final extension was performed at 72°C for 10 min. PCR products were analyzed on a 0.8% agarose gel.

### Cloning and sequencing

Gel slices containing PCR products were excised from 0.8% agarose gels and purified using the AxyPrep DNA Gel Extraction Kit (Axygen, China). Purified PCR products were ligated into the pMD18-T Vector (TaKaRa, China) and recombined plasmids were transformed into *E. coli* competent cells. Cells carrying recombinant plasmids were selected on Luria–Bertani agar plates containing ampicillin, X-gal, and IPTG. Plasmid DNA was purified using the alkaline lysis protocol and sequenced using an ABI PRISM 3730 DNA analyzer (Applied Biosystems, USA). For each RGDV isolate, 3 or 4 independent clones originating from different RT-PCR reactions were selected.

### Data Analysis

Sequence data were compiled and analyzed using BioEdit (Isis Pharmaceuticals, Inc., Canada). The theoretical isoelectric point (pI) of the deduced peptide was predicted by the online tool “Compute pI/Mw Tool” at [http://us.expasy.org/tools/pi\\_tool.html](http://us.expasy.org/tools/pi_tool.html). Phylogenetic trees were constructed with MEGA (version 3.0) by using maximum parsimony analysis with 1000 bootstrap replicates. Sequence of Thailand isolate used for comparison and phylogenetic analysis in this study was obtained from the *GenBank* database

(accession number: D13410).

### Prokaryotic expression of RGDV segment 8 cDNA (Pns8)

RGDV segment 8 cDNA encoding a main capsid protein (Pns8) was PCR-amplified using a pair of primers 5'-GAATTCATGTCGCCAGCTTGG AT-3' and 5'-GTCGACTTAGTTACTGTGTAATAC C-3' flanked by *Eco*RI and *Sal*I restriction sites. The PCR product was subcloned into pET-28b (+) (Invitrogen, USA) at the *Eco*RI/*Sal*I sites to yield an expression plasmid pET.CP. *E. coli* Rosetta (DE3) II cells harboring pET.CP were incubated at 37°C in a 100-mL flask containing 10 mL LB medium supplemented with 30 µg kanamycin/ml and 34 µg chloramphenicol/mL. When the OD<sub>600</sub> reached 0.6, isopropylthio-α-D-galactopyranoside (IPTG) was added into the medium to a final concentration of 1 mmol/L, and *E. coli* Rosetta (DE3) II cells were subsequently cultured at 42 °C to induce the protein expression. After 1, 3, 5, and 7 hours of culture respectively, the cells were harvested and total cell proteins were extracted and analyzed by SDS-PAGE.

## RESULTS

### Gene cloning and sequence analysis

The genomic segment 8 of five isolates from China were amplified and sequenced successfully. The results showed that they were all 1578 bp in length and each contained one long open reading frame which extended for 1301 nt from 21 nt, and encoded the RGDV major outer capsid protein. The nucleotide sequence data reported in this paper were deposited in *Genbank* with accession numbers of AY999077 (BL), AY999078 (CH), AY999079 (DQ), AY999080 (GZ) and AY999081 (XY), respectively.

The genomic segment 8 and the coding region of these five Chinese isolates shared 94.8% ~95.6% and 95.0%~96.0% nucleotide sequence identities with that of the Thailand isolate respectively, 97.3%~98.8% and 97.3%~99.1% within these five isolates respectively (Table1). The variations of nucleotide sequence between Chinese isolates and the Thailand isolate were greater than that within the Chinese isolates. Variation seemed to be related to the geographic distance at the nucleotide level. But at the amino acid level, the situation changed. The deduced amino acid (Pns8) of GZ isolate was identical to that of the Thailand isolate, while the variability's of Pns8 within these five Chinese isolates ranged from 0.5% to 2.1% (Table2). The residue variation position appeared to be random. It is of note that XY and GZ, the two neighbor isolates, still had a variation of 2 residues. The sequence divergence resulted in apparent changes in the theoretical isoelectric point (pI) of the deduced peptides. The pIs of the deduced peptides of both BL

and DQ isolates were 7.62, while the others were 6.74. The effect of the changes remained unknown since no remarkable phenotype variability was observed within the five Chinese isolates.

Two phylogenetic trees were built and shown in rectangular cladogram on the basis of the nucleotide sequences of the S8 and its coding region (Fig. 1). The two phylogenetic trees were similar and all the five Chinese isolates were grouped together, which indicated that they were more closely related to each other than to Thailand isolate. It is interesting that the two isolates, BL and CH, first reported in this paper, were clustered with XY isolate in one clade. This suggested that RGDV in BoLuo County and CongHua County might be newly introduced from Xinyi County, the original region, where the disease occurred first in China. The phylogenetic tree could not be constructed on the basis of amino acid sequences by maximum parsimony method because there were insufficient parsimony-informative sites.

Table 1. Comparison of the full-length and coding region sequences of genomic segment 8 of RGDV five Chinese isolates (BL, CH, DQ, GZ, XY) and Thailand isolate (TL)

Isolates	TL	BL	CH	DQ	GZ	XY
TL	-	95.2 <sup>a</sup>	94.8	95.3	95.6	95.2
BL	95.6 <sup>b</sup>	-	98.4	98.0	98.2	98.8
CH	95.0	98.4	-	97.4	97.5	98.2
DQ	95.4	98.1	97.3	-	97.3	97.8
GZ	96.0	98.8	98.0	97.9	-	98.0
XY	95.6	99.1	98.2	98.0	98.7	-

<sup>a</sup> Similarity percentage of the full-length of S8 is indicated in upper right triangle. <sup>b</sup> Similarity percentage of the coding region of S8 is indicated in lower left triangle.

Table 2. Divergence of the deduced amino acids of the outer main capsid protein gene of RGDV five Chinese isolates (BL, CH, DQ, GZ, XY) and Thailand isolate (TL)

	TL	BL	CH	DQ	GZ	XY
TL	-	0.5 <sup>a</sup>	1.2	0.9	0.0	0.5
BL	2 <sup>b</sup>	-	1.7	1.4	0.5	0.9
CH	5	7	-	2.1	1.2	1.7
DQ	4	6	9	-	0.9	1.4
GZ	0	2	5	4	-	0.5
XY	2	4	7	6	2	-

<sup>a</sup> Divergence percentage in upper right triangle. <sup>b</sup> Sum of variation residues in lower left triangle.

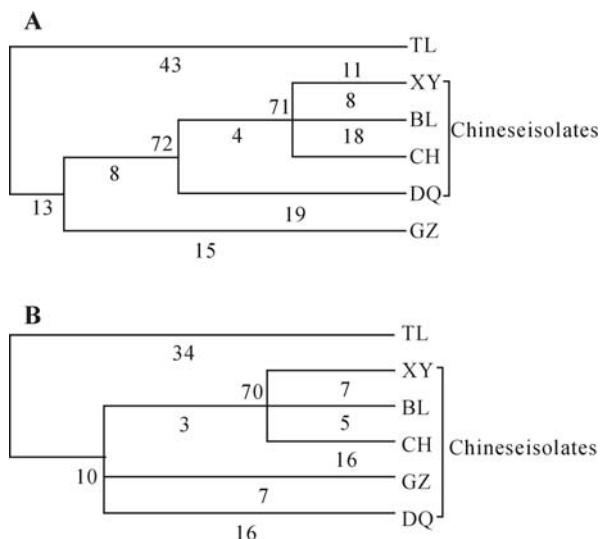


Fig.1. Phylogenetic trees based on the nucleotide sequences of the genomic segment 8 (A) and its coding region (B) of RGDV five Chinese isolates (BL, CH, DQ, GZ, XY) and Thailand isolate (TL).

#### Expression of Pns8 of RGDV XY isolate

Segment 8 cDNA encoding Pns8 of XY isolate was amplified and digested with *Eco*RI/*Sall*, ligated into linear vector pET-28b (+), and the recombinant pET.CP vector was obtained. Sequencing data showed that protein expression codons were completely correct. After *E. coli* Rosetta (DE3) II cells containing pET.CP were grown at 37 °C and the OD<sub>600</sub> reached 0.6, followed by further induction culture with IPTG for 3, 5, 7 hours at 42 °C respectively, the total cell fractions from growing 3-7 hours showed that the recombinant fusion protein was produced. This was further confirmed by SDS-PAGE analysis (Fig. 2). Molecular mass (51kDa) of the protein band was in good agreement with the size deduced from the whole amino acid sequence. On the contrary, no similar protein bands were detected in 3 control samples (Fig. 2). This preliminarily indicated that Pns8 was successfully expressed in *E. coli* Rosetta (DE3) II.

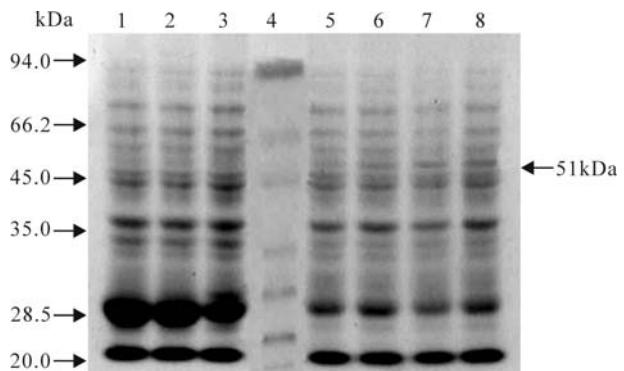


Fig. 2. Analysis of prokaryotic expression of recombinant RGDV pET.CP plasmid with SDS-PAGE. 1, *E.coli* Rosetta(DE3) II without induction; 2, *E.coli* Rosetta(DE3) II induced with IPTG; 3, *E.coli* Rosetta(DE3) II containing pET.CP without induction; 4, Protein Marker; 5-8, *E.coli* Rosetta (DE3) II containing pET.CP was induced with IPTG for 1, 3, 5 and 7 hours, respectively.

#### DISCUSSION

RGDV has only been reported in China and in a few other Southeastern Asian Countries. Some previous reports have shown that there were some apparent differences in biological characteristics between the Chinese and Thailand isolates (2, 11), and revealed that there were different virus strains or strain differentiation in RGDV. Furthermore, RGDV was present in different regions in China, and could infect different rice cultivars. It is interesting therefore to determine whether these isolates are identical to each other. RGDV fragment S8, encoding the RGDV main outer coat protein, which might play some important roles in viral particle assembly, movement in host, transmission and the interaction between virus and host, was chosen for cloning and analyzing in this study.

Nucleotide sequence analysis showed that RGDV S8 shared 97.3%-98.8% identities within five Chinese isolates, and 94.8%-95.6% identities between the Chinese and Thailand isolates. These indicated that

the genetic variation of RGDV S8 was more likely to be related to the geographic locations. However, the deduced amino acid sequences of the encoded protein (Pns8) showed a contradictory result. The Pns8 sequence of the GZ isolate from China was completely identical to that of the Thailand isolate, while the variability within five Chinese isolates ranged from 0.5% to 2.1%, and even a 2-residue change was found between two neighbor isolates (XY and GZ). This seemed to indicate that Pns8 amino acid sequence of RGDV was highly conserved, although there was higher variability over the entire nucleotide sequence of the S8 segment, which might be a consequence of adapting to the respective ecological environment in different geographical locations.

RGDV consists of 12 dsRNA segments. Only the nucleotide sequences of genome segments S2 (8), S3 (15), S5 (3), S8 (10), S9 (4), S10 (10), and S11 (9) from the Thailand isolate, and S9 from the Chinese XY isolate (1) have been reported. S8 sequence analysis indicated that RGDV S8 might not be sufficient to entirely determine the differences in biological characteristics between Chinese and Thailand isolates. Therefore, it is necessary to make an analysis and comparison of other RGDV segments between the isolates from two countries, which will be helpful to understand the relationship between different isolates of RGDV and the evolution of RGDV.

RGDV is restricted to plant phloem cells and has very low concentrations in rice plants, so it is difficult to get high purified yields from rice plants for preparation of antiserum (12). The protein expression in *E. coli* provides an alternative way to produce acid sequences between RGDV and RDV.

enough protein in high quality for production of antiserum. Successful expression of the main outer capsid protein (Pns8) of RGDV in *E. coli* Rosetta (DE3) II may provide a foundation for virus detection and analysis of virus functions.

Although different induction conditions were tested, Pns8 could not be expressed in all selected bacterial strains, such as *E. coli* BL21 (DE3), pLysS, BL21 (DE3) and JM109 (DE3). The protein could be detected in *E. coli* Rosetta (DE3) II strain, though the expression amount was still low. This may be due to the existences of a number of rare codons (22.08%) of amino acids in Pns8 and several continuous and stable stem-loop structures in the translation initiation region of Pns8 mRNA (data not shown), which were not of benefit for the ribosome to contact and conjugate for enhancing translation start (5,16). Suzuki et al. (14) propose that the *E. coli* expression system is not suitable for the protein expressions from the viruses of the family *Reoviridae* including RDV, WTV and RGDV. However, the main outer capsid protein from RDV was cloned into vector pTrcHisA and was efficiently expressed in *E. coli* (6). Lu et al. (7) cloned several segments of the second outer capsid protein gene of RDV into vector pTrcHisA, and found that the C-terminal segment could be expressed while the N-terminal segment could not. However, N-terminal segment could be successfully expressed if it was cloned into vector pET11d instead of pTrcHisA (7). These results indicated that the protein expression was significantly influenced by different expression strategies, selection of vectors and bacterial strains. Such situations will occur in the protein expression of RGDV, because there are high homologies of amino

Although Pns8 of RGDV in this study was

expressed in the analysis of SDS-PAGE, it was not further confirmed with Western blot because of antiserum deficiency. Pns8 purification and preparation of the antiserum are currently under investigation.

### Acknowledgement

This research was financially supported by National natural science foundation of China in a grant entitled “Strain genetic diversity and molecular characterization of *Rice gall dwarf virus*” with grant No. 30370929 and Guangdong Natural science foundation in a grant entitled “Genomic analysis and strain diversity of *Rice gall dwarf virus*” with grant No. C036845.

### References

1. Fan G C, Wu J J, Huang M H , et al. 2005, Genetic structure of Rice gall dwarf virus segment 9. *Virologica Sinica*, 20 (5): 539-542. (in Chinese)
2. Fan H Z, Zhang S G , He X Z. 1983, Rice gall dwarf disease-a new rice viral disease in Zhanjaing of Guangdong province. *Acta Phytopathologica Sinica*, 13 (4) : 1-6. (in Chinese)
3. Ichimi K, Kikuchi A, Moriyasu Y, et al. 2002, Sequence analysis and GTP-binding ability of the minor core protein P5 of *Rice Gall Dwarf Virus*. *Japan Agr Res Quart*, 36: 83-87.
4. Koganezawa H, Hibino H, Motoyoshi F, et al. 1990, Nucleotide sequence of segment S9 of the genome of *Rice Gall Dwarf Virus*. *J Gen Virol*, 71: 1861-1863.
5. Kornitzer D, Teff D, Altuvia S, et al. 1989, Genetic analysis of bacteriophage cIII gene: mRNA structural requirements for translation initiation. *J Bacteriol*, 171 (5) : 2563-2572.
6. Li Wei, Li Yi, Zhang Xu, et al. 1995, Cloning, sequence and expression of Rice dwarf virus segment 8 in *Escherichia coli*. *Chin J Virol*, 11 (1): 56-62. (in Chinese)
7. Lu Rui-fang, Li Yi, Yang Cong-lin, et al. 1999, Cloning, sequence and expression of the outer capsid protein encoded by rice dwarf virus S2 in *Escherichia col*. *Acta Microbiol Sinica*, 39 (4) : 305-314. (in Chinese)
8. Maruyama W, Ichimi K, Fukui Y, et al. 1997, The minor outer capsid protein P2 of *Rice Gall Dwarf Virus* has a primary structure conserved with, yet is chemically dissimilar to *Rice Dwarf Virus* P2, a protein associated with virus infectivity. *Arch Virol*, 142: 2011-2019.
9. Moriyasu Y, Ishikawa K, Kikuchi A, et al. 2000, Sequence analysis of Pns11, a nonstructural protein of *Rice Gall Dwarf Virus*, and its expression and detection in infected rice plants and vector insects. *Virus Genes*, 20: 237-241.
10. Noda H, Ishikawa K, Hibino H, et al. 1991, Nucleotide sequence of genome segment S8, encoding a capsid protein, and S10, encoding a 36K protein, of *Rice Gall Dwarf Virus*. *J Gen Virol*, 72: 2837-2842.
11. Omura T, Inoue H, Morinaka T, et al. 1980, Rice gall dwarf, a new virus disease. *Plant Dis*, 64: 795-797.
12. Omura T, Inoue H, Saito Y. 1982, Purification and some properties of *Rice Gall Dwarf Virus*, a new Phytoreovirus. *Phytopathology*, 72: 1246-1249.
13. Ong C A, Ong T B. 1982, *Rice Gall Dwarf Virus* occurrence in Peninsular Malaysia. *Internatl Rice Res Newsletter*, 7 (2): 7.
14. Suzuki N, Sugawara M, Kusano T, et al. 1994, Immunodetection of rice dwarf phytoreoviral proteins in both insect and plant hosts. *Virology*, 202 (1): 41-48.
15. Takahashi Y, Tomiyama M, Hibino H, et al. 1994, Conserved primary structures in core capsids proteins and reassembly of core particles and outer capsids between rice gall dwarf and rice dwarf phytoreo-viruses. *J Virol*, 75: 269-275.
16. Wikstrom P M, Lind L K, Berg D E, et al. 1992, Importance of mRNA folding and start codon accessibility in the expression of gene in a ribosomal protein operon of *Escherichia coli*. *J Mol Biol*, 224: 949-966.