A THESIS

FOR THE DEGREE OF MASTER OF SCIENCE

Transcriptome analysis of small brown planthopper, *Laodelphax striatellus* carrying Rice Stripe Virus

벼 줄무늬잎마름병을 매개하는 애멸구의 전사체 분석

By

Joo-Hyun Lee

Major in Entomology

Department of Agricultural Biotechnology

Seoul National University

February, 2013

Transcriptome analysis of small brown planthopper, *Laodelphax striatellus* carrying Rice Stripe Virus

Major in Entomology

Department of Agricultural Biotechnology

Seoul National University

February, 2013

Joo-Hyun Lee

ABSTRACT

Rice stripe virus (RSV), the type member of the genus Tenuivirus which causes rice stripe disease, transmits by the feeding behavior of small brown planthopper (SBPH), *Laodelphax striatellus*. Because RSV replicates in SBPH persistently and SBPH has sap-sucking feeding behavior with its piercing-sucking mouthpart in the phloem, SBPH is one of the most notorious vectors of RSV. So far, most considerations have been focused on the protection of the rice from RSV and/or the SBPH rather than the interactions among the rice, RSV and SBPH. Hence, to investigate the interactions between the virus and vector insect, total RNA was extracted from RSV-viruliferous SBPH (RVLS) and non-viruliferous SBPH (NVLS) adults to construct expressed sequence tag databases for comparative transcriptome analysis. Over 30 million bases were sequenced by 454

pyrosequencing to construct 1,538 and 953 of isotigs of SBPH from the mRNA of RVLS and NVLS, respectively. The gene ontology (GO) analysis demonstrated that both transcriptomes of RVLS and NVLS have similar GO structure, however, the gene expression pattern analysis revealed that 108 genes were up-regulated and 28 genes were down-regulated significantly in the RVLS. These RSV-dependently regulated genes of SBPH possibly have important roles in the physiology and behavior of SBPH, transmission of RSV, and RSV and SBPH interaction.

Also, two kinds of picorna-like viruses were newly discovered from the EST libraries of each NVLS and RVLS, Which were named LsPV-1 and LsPV-2, respectively. The sequence analysis result revealed that LsPV-1 was a Korean isolate of Himetobi-P virus (HiPV) which belongs to *dicistorovidae* while the LsPV-2 genome structure resembled that of *iflaviridae*. A phylogenetic analysis of LsPV-2 RNA dependent RNA polymerase sequence showed that honeybee deformed wing virus was the most closely related virus. Interestingly, RSV and HiPV seem not to superinfect SBPH which already carrying the other virus, therefore, the incompatibility between RSV and HiPV might give us a clue to understand the viral microflora and the mechanism of persistency in SBPH. Also, The HiPV and LsPV-2 were incompatible each other in SBPH, suggesting that theses tow picorna-like viruses may have important functions in transmission of the RSV.

Key words: Rice stripe virus, *Laodelphax striatellus*, Transcriptome, picorna-like viruse, Himetobi P virus, iflavirus

Student Number: 2011-21307

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Rice stripe virus (RSV), the prototype virus of genus Tenuivirus, is a non-enveloped RNA virus which forms circular filamentous shape of ribonucleoprotein particles with four species of segmented single-stranded RNA genome of RNA1, 2, 3 and 4 according to the size of each RNA molecule (Hibino *et al.*, 1985; Ishikawa *et al.*, 1989). The largest segment, RNA1 encodes a 336.8 kDa protein which is thought to be a RNA dependent RNA polymerase (Toriyama *et al.*, 1994). RNA2 encodes 94 kDa glycoprotein and NS2 (Takahashi *et al.*, 1993), RNA3 encodes NS2 and CP (Zhu *et al.*, 1991), and RNA4 encodes SP and NS4 (Zhu *et al*., 1992). RNA1 is negative-sense single strand while RNA2, 3, and 4 are ambisense in their coding strategy (Hamamatsu *et al.*, 1993). RSV is one of the most serious viral rice pathogens which occur in South Korea, China, and Japan. The typical symptoms of RSV infected rice plant are chlorotic stripes or necrotic streaks on leaves, and premature wilting. Some of the leaves emerge abnormally unfolded, and then elongate to become twisted and droop. The RSV often causes significant reduction in yield. Especially the rice plants inoculated with RSV in 11-leaf stage or earlier produce no grain (Ling, 1972; Hibino, 1996).

The mechanisms of the insect vector mediated plant virus transmission are diverse based on the differences of the duration of which the vector insects can harbor infectious viral particles. The transmission mechanisms can be classified into three categories; nonpersistent (harboring viruses for a few seconds/minutes), semipersistent (a few hours/days, but lost upon molting), and persistent (often throughout the lifespan of the

vectors) (Ng and Falk, 2006). Most of the plant viruses are dependent on the vector insects for plant-to-plant spread, and one of the most important vector insects are hemipterans, which can transmit viruses during feeding with their piercing-sucking mouthpart in the phloem (Hogenhout *et al.*, 2008). The small brown planthopper (SBPH), *Laodelphax striatellus* Fallén (Hemiptera: Delphacidae), is one of the vectors of RSV and classified as a vector of persistent transmission because RSV replicates persistently in the SBPH ovary to transmit virus to the offspring by eggs transovarially (Toriyama, 1986), as well as circulatively to proliferate and escape from the midgut, salivary gland, and ovary barriers for host-to-host infection (Nault and Ammar, 1989). Many of the tenuiviruses rely on planthopper vectors for transmission, and the vector specificity is high (Falk and Tsai, 1998), and SBPH is one the most important insect vectors of RSV, however, our knowledge of the SBPH biology is limited and most of the interactions between RSV and SBPH have received less attention so far.

To investigate the RSV-SBPH interaction, we employed next generation sequencing (NGS) technology to construct expressed sequence tag (EST) libraries from SBPH because SBPH genome sequence is not available at this moment. The pyrosequencing technology (Ronaghi *et al.*, 1996) on the Roche 454 GS-FLX platform was used to create EST databases of RSV-viruliferous SBPH (RVLS) and non-viruliferous SBPH (NVLS), and the gene expression pattern of RVLS and NVLS were compared based on the read counts of each isotigs. We also conducted quantitative RT-PCR experiments with selected genes from the EST database to verify the sequencing results.

Also, on the other side, NGS technology was used for metagenomics. In this respect, a metagenomic approach of NGS technologies has revolutionized the methods for

discovery of new microorganisms and viruses. NGS is proven to be a very fast, inexpensive and accurate sequencing technology for identification of viral sequences derived from whole insects (Liu *et al*., 2011). In the process of analyzing the constructed EST libraries, two picorna-like viruses were discovered by chance from RVLS and NVLS, respectively. The sequence analysis results revealed that one of the picorna-like viruses from the NVLS named LsPV-1 was a Korean isolate of Himetobi P virus (HiPV), and the other one from the RVLS named LsPV-2 was identified with an iflavirus. Previously, a number of positive-strand RNA viruses had been described from various species of insects and named picorna-like viruses based on similarities to the biophysical properties of mammalian picornaviruses. These insect picorna-like viruses are subdivided into two genera by differences in their genomic organization, which is monocistronic or dicistronic (Christian *et al*., 2005a, 2005b). The HiPV is a dicistronic virus containing two non-overlapping open reading frames (ORFs) separated by an intergenic region which functions as an internal ribosomal entry site (IRES). The LsPV-2 seems to be a novel iflavirus because it has a unique large single ORF which encodes both structural and non-structural proteins. A phylogenetic analysis based on RNA-dependent RNA polymerase (RdRp) sequence of HiPV, LsPV-2 and other insect picorna-like viruses showed that LsPV-2 was most closely related to deformed wing virus (DWV), however, the identity between DWV and LsPV-2 was low. Here, we report the complete genomic nucleotide sequence of HiPV and partial genomic sequence of LsPV-2.

LITERATURE REVIEW

1. *Laodelphax striatellus*

Laodelphax striatellus, also known as the small brown planthopper (SBPH) is a polyphagous insect that has about 42 hosts species (yamada and yamamoto., 1955). The primary hosts of SBPH are oats, barely, wheat and rice. Especially, rice is the most favored hosts. SBPH is found in south Korea, China and Japan that is the most important vector of rice stripe viris (RSV). The population size of SBPH was relatively small in the middle parts of Korea, but it was much greater in the southern areas.

1) Biology

The SBPH is dimorphic, with fully winged 'macropterous' and truncate-winged 'brachypterous' forms. The macropterous forms are migrants and are responsible for colonizing new fields. Female lay about 400 eggs, most of them are laid in the middle and lower parts of rice, and can tolerate water exposure. The SBPH can overwinter as 3-4th nymphs in Korea. So, virus viruliferous SBPH constantly transmit RSV to next generations.

2) Life Cycle

A female SBPH lays about 400 eggs that are approximately 3-4 per lay within leaf sheath or middle and lower parts of rice.

Eggs usually hatch between 7-10 days and period of nymph is 18-20 days and period of adult is 20-25 days. The newly hatched nymphs are cottony white, and turn yellow brown within an hour. They feed on plant sap, and undergo five instars to become adults. The SBPH overwinter in fourth instar nymph in the soil, root of wheat, stubble, fallen leaves, weeds and other places. After overwintering, the fourth instar nymph wakes in at March, and feed mainly on wheat.

SBPH emerges five generations per year. The first population emerges from late April to early May. The peak of adult emergence of the first generation occurs from midMay to early June when SBPH moves out from wheat field to be reported the outbreak of SBPH in Korea. The forth instar nymph of fifth generation begins to overwinter in November.

3) Identification characteristics

The Adult female and male of SBPH were measured 3.6-3.8 mm and 3.4-3.6 mm in length, respectively. The body of adult SBPH is black to dark-brown, and coloration varies with season. The wings are hyaline and the two wings have forms of macropter and brachypter (Mori and Nakasuji 1990). The mesonotum is black to dark-brown, as the pterostigma on the fore wings, and the areas between the carinae of the frons are deep black. The body color of the SBPH varies with general body coloration. The shiny black mesonotum of the male, the pterostigma of the forewing and the male genitalia are diagnostic for the species. The form of the aedeagus and the parameres are unlike any other delphacid found on rice. Short, sharp anal tube appendages, and the shape of the parameres and the diaphragm are diagnostic.

4) Potential economic impact and description of damage

L. striatellus, a vector of Rice stripe virus and Rice black streaked dwarf virus, is one of the most important insect vectors of rice disease. And, it is able to overwinter in temperate regions including Japan, South Korea, and China (Otuka *et al*., 2012).

Recently, mass overseas migration of *L. striatellus* has been reported. Mass overseas migration and a subsequent outbreak of rice stripe disease were reported to have occurred in western Japan in 2008 (Otuka *et al*., 2010). After this migration, another mass immigration occurred in western Korea in 2009 (Kim 2009; Kim *et al*., 2009). Thus, Rice stripe disease is very important in Asia.

2. General characteristics of Rice Stripe Virus

Rice stripe virus (RSV) causes a severe disease of rice in Asian countries and transovarily transmitted in the circulated persistent manner by *L. striatellus* (Liang *et al*., 2005). In Asian, rice stripe disease mainly occurs from mid-May to early June. RSV is a plant pathogenic virus of the genus Tenuivirus. Other members of this genera are Maize stripe virus, Rice hoja blanca virus, and Iranian wheat stripe virus. These Tenuivirus causes diseases in their host plants with typical symptoms of chlorotic stripes on the affected leaves. And the infected rice plants produce no grain when they are inoculated earlier than the 11-leaf stage.

The genome of RSV comprises four species of segmented single-stranded RNA genome of RNA1, 2, 3 and 4 according to the size of each RNA molecule (Hibino *et al.*, 1985; Ishikawa *et al.*, 1989). The largest segment RNA1 encodes a 336.8 kDa protein

which is thought to be a RNA dependent RNA polymerase (Toriyama *et al.*, 1994). RNA2 encodes 94 kDa glycoprotein and NS2 (Takahashi *et al.*, 1993), RNA3 encodes NS2 and CP (Zhu *et al.*, 1991), and RNA4 encodes SP and NS4 (Zhu *et al.*, 1992). RNA1 is negative-sense single strand while RNA2, 3, and 4 are ambisense in their coding strategy (Hamamatsu *et al.*, 1993).

The RSV multiplies in the *L. striatellus* as shown by evidence of transovarial passage (yamada and yamamoto., 1955), and by serial transmission of non-viruliferous *L.striatellus* by injection of RSV viruliferous-*L.striatellus.* Also, evidences have revealed that amorphous or filamentous inclusions of RSV genes exist in most of the insect tissues (Zhang *et al*., 2010) by detection of RSV gene products by immunoflurescence light microscopy. Moreover, it has been confirmed that RSV particles exist in follicular cells of the ovarioles can be transmitted from female adults to their progeny via eggs (Liang *et al*., 2005).

In addition to rice, RSV has a wide host range. About 36 species of plants were compiled by Iida (Iida 1969). The representative hosts of RSV are *Avena sativa, Avena fatua*, *Alopecurus japonicas*, *Triticum aestivum*, and *Zea mays*.

In Korea, several granmineous plants including barley, foxtail, Italian ryegrass, milk vetch, and wheat have been found infected by RSV (Lee, 1969).

3. Next Generation Sequencing

After completed the human genome project in 2003, sequencing technology has been rapidly changed. The automated Sanger method is considered as a ‘first-generation’ technology, and newer methods are referred to as Next generation sequencing (NGS).

From 2007, term of Next generation sequencing (NGS) has been used. NGS technologies revolutionized genomics and their effects are becoming increasingly widespread. The use of NGS over the past five years has revolutionized the discovery of microorganism (Liu *et al*., 2011) and the technology allows saving the money and time. NGS not only changed the sequencing technology conveniently but also made a new platform of genome research.

The most common NGS platforms are three methods which are 454 GS FLX Titanium pyrosequencing from Roche, Illumina sequencing from Solexa, and ABI's SOLiD sequencing from Life science. 454 GS FLX Titanium pyrosequencing has merits that are long and accurate read length. It enables a range of applications including genome annotation, novel transcript identification, splice-variant detection, expression analysis, full-length gene assembly and SNP discovery. Solexa's technology is currently the most widely used platform in the field. It is able to find variant discovery by whole-genome resequencing or whole-exome capture (Metzker, 2010). ABI's SOLiD uses two-base-encoded probes, which has the primary advantage of improved accuracy (Valouev *et al*., 2008), but this technology needs long run times.

NGS is a non-Sanger-based and high-throughput methodology which allows for generation of millions of sequences at once. Therefore, NGS enables the metagenomics. Metagenomics is the study of metagenomes, genetic material recovered directly from diverse samples. 454 GS FLX Titanium pyrosequencing typically produces ~400 bp reads, Illumina and SOLiD produce 25-27 bp reads (Schuster, 2007). Short read lengths of Illumina and SOLiD are compensated by the much larger number of sequence reads. The metagenomics using NGS has the potential to advance knowledge in a wide variety of

fields. It can also be applied to solve practical challenges in medicine, engineering, agriculture and ecology.

4. Picorna-like virus

The invertebrate picorna-like viruses are small, non-enveloped viruses that form as isometric particle. The positive-sense, single-stranded RNA genome is either mono- of dicistronic (van Regenmorter *et al*., 2000). By the difference in their genomic organization, these insect picorna-like viruses are subdivided into two genera which are monocistronic of dicistronic.

1) The monocistronic viruses

The monocistronic viruses have been grouped in the genus *Iflavirus*. *Iflavirus* belong to family *Iflaviridae* which has a seven species. This family is a member of the order '*Picornavirales*'. The seven species are Deformed wing virus (DWV), *Ectropis obliqua* virus (EoV), Infectious flacherie virus (IFV), *Perina nuda* virus (PnV), Sacbrood virus (SBV), Slow bee paralysis virus (SBPV), *Varroa destructor* virus-1 (VDV-1). The type species of this genus is infectious flacherie virus (IFV), isolated from *bombyx mori* (Isawa *et al*., 1998) and characterized by a unique large open reading frame (ORF) that encodes both structural and non structural proteins. The name (Ifla) is derived from the type species name IFV.

Some of the insects including aphids, leafhopper, flies, ants and silkworms are commonly infected by the *iflavirus*. SBV infects larvae of the honeybee, *Apis mellifera*, resulting in failure of pupation, and death (Grabensteiner *et al*., 2001), and it is also

estimated the cause of colony collapse disorder (CCD). Infected larvae change in color from pearly white to pale yellow, and shortly after death they dry out, forming a dark brown gondola-shaped scale. VDV-1 and EoV cause a lethal granulose infection in larvae of the tea looper, *Ectropis oblique* which spread widely in China to cause severe damages of the tea output (Lin et al. 2010).

2) The dicistronic viruses

The dicistronic viruses have been grouped into the genus *Aparavirus* and *Cripavirus*. *Aparavirus* and *Cripavirus* belong to family *Dicistroviridae* which has a six species and nine species, respectively. This family is a member of the order '*Picornavirales*'. The six species of *Aparavirus* are Acute bee paralysis virus (ABPV), Israeli acute paralysis virus (IAPV), Kashmir bee virus (KBV), Mud crab virus, *Solenopsis invicta* virus-1 (SINV-1), Taura syndrome virus (TSV). The nine species of *Cripavirus* are Aphid lethal paralysis virus (ALPV), Black queen cell virus (BQCV), Cricket paralysis virus (CrPV), *Drosophila* C virus (DCV), Himetobi P virus (HiPV), *Homalodisca coagulata* virus- 1 (HoCV-1), *Plautia stali* intestine virus (PSIV), *Rhopalosiphum padi* virus (RhPV), Triatoma virus (TrV). In contrast to *Iflavirus*, the genome of *Aparavirus* and *Cripavirus* contain two non-overlapping ORFs, which are separated by an intergenic region which functions as an internal ribosome entry site (IRES).

The BQCV of *Cripavirus* isolated from adult worker bees of drone pupae (Leat *et al*., 2000) and HiPV isolated from *Laodelphax striatellus* or *Nilaparvata lugens* and CrPV were initially discovered in Australia field crickets.

The type species of *Cripavirus* is CrPV (Wilson *et al*., 2000). The paralytic disease spread rapidly through a breeding colony as well as through a laboratory population causing about 95% of mortality. Infected crickets die without any visible symptoms and at all stages of development, but death rates are highest dring the pre-wing stage just before adult stage.

MATERIALS AND METHODS

1) RSV-viruliferous and non-viruliferous *L. striatellus.*

The RVLS and NVLS were collected from the RSV infected rice field and healthy rice seedlings, respectively. The RVLS and NVLS colonies were reared in the laboratory on 2-3 cm tall rice seedlings in glass vessels under a light:dark cycle of 16:8 h at 25°C and transferred to fresh seedlings every 10-14 days to assure sufficient nutrition. After a few generations, we conducted RT-PCR with extracted RNA in order to confirm the RVLS or NVLS.

2) Total RNA extraction and RT-PCR

Total RNA for nucleotide sequencing was extracted from RVLS and NVLS with TRI Reagent (Invitrogen, USA) according to the manufacturer's instruction. RT-PCR was performed to confirm RSV in the RVLS and NVLS. Two pairs of primers for specific amplification of RSV sequences were designed as follows; RSV-CP-F: 5'-GCT CAT AAA GCT GAT GTT G-3' and RSV-CP -R: 5'-GTC TTC ACT TTC CCA TTG-3' correspond to the nucleotide sequence of coat protein gene in RNA3 with 969 base pair (bp) of expected PCR product. RSV-NCP-F: 5'-GAA AGA TCC AGA GAG AGT CA-3' and RSV-NCP-R: 5'-CAA AAT AAG TTG GCA TCT TC-3' correspond to the nucleotide sequence of non-capsid protein gene in RNA4 with 537 bp of expected PCR product. Another two pairs of primers for specific amplification of SBPH sequences were designed as follows; CCOS-F: 5'-ATT GCT CAT AAA GCT GAT GTT G-3' and CCOS-R primer: 5'-CCG TAC CAG GAC GAA TTA-3' correspond to the nucleotide sequence of

Cytochrome c oxidase subunit II with 684 bp of expected PCR product. ATP-F: 5'-CCC TCA AGT CAC CTC TCT-3' and ATP-R: 5'- ATG GAA TTT CTG AGG AA-3' correspond to the nucleotide sequence of ATP synthase F0 subunit 6 with 655 bp of expected PCR product. cDNAs were synthesized with the reverse primer of each genes, and RT-PCR reactions were performed with AccuPower one-step RT-PCR premix (Bioneer, Korea).

3) Transciptome sequencing and EST library construction

To conduct the transcriptome sequencing of RVLS and NVLS, total RNA samples were extracted with TRI Reagent from approximately 500 SBPHs of RVLS and NVLS colonies, respectively. 200 ng of mRNA samples were eluted from the total RNA samples by using oligo-dT hybridization followed by heat fragmentation, and reverse-transcription to synthesize cDNA with random hexamer. Second strand was synthesized with second strand enzyme followed by 2 hour incubation in buffer for recovery of double stranded cDNA, 5'-phosphorylation and the addition of 3'-dA overhang. Finally, adaptor was ligated and the samples were amplified to be prepared for sequencing. Approximately 5 μg of each amplified cDNA samples from RVLS and NVLS were subjected to two of 1/8 plate sequencing runs with a Roche 454-GS-FLX Titanium sequencer. The sequence data were processed with GS-FLX software v2.5.3 to remove low quality sequences, and the resulting sequences were then screened against the NCBI UniVec database and *E. coli* genome sequences to remove contaminant sequences, then the sequences shorter than 100 bp were discarded. The processed sequences were assembled into isotigs using Newbler v

2.5.3 software to construct two EST libraries of RVLS and NVLS. Also, after combining the RVLS and NVLS reads, all the sequence reads were assembled as one library.

4) Gene ontology analysis and Comparison of gene expression

The isotigs of RVLS and NVLS were subjected to GenBank nr database with BLASTX and gene ontology (GO) analysis to annotate the isotigs into three main categories; biological process, cellular component, and molecular function (Ashburner *et al.*, 2000). And the 2,555 isotigs of the combined dataset were analyzed by comparing with *Drosophila* proteins using BLASTX with a cut-off value of 1e-05 to identify their gene functions. The resulting data set was used to reconstruct functional class profiles and pathways using the KEGG (Kyoto Encyclopedia of Genes and Genomes).

5) Analysis of relative transcription levels by qPCR

To validate the pyrosequencing results of the SBPH transcriptomes, the transcription levels measured based on the read counts were compared with those of qPCR. For qPCR, total of seven translation-related isotigs (translational initiation factor 1, 2, and 3, and ribosomal protein 1, 2, 3, and 4), heat shock protein 1 and 2, and vitellogenin genes were selected for qPCR along with elongation factor 2 (EF2) gene as a reference for normalization. The cDNAs were synthesized from the total RNA of RVLS and NVLS using QuantiTect Reverse Transcription Kit (QIAGEN, Germany) according to the manufacturer's instructions, followed by qPCR conducted using EvaGreen qPCR Mastermix (Applied Biological Materials Inc, Canada) and CFX96™ Real-Time System (BIO-RAD, USA). The cycling profile used for qPCR was as follows: a preheating step

for enzyme activation at 95°C for 10 min, followed by 40 cycles of 95°C for 15 sec, 58°C for 15 sec, and 72°C for 30 sec. The relative transcription levels were calculated using the $2^{-\Delta Ct}$ method (Pfaffl, 2001). The primers used for the qPCR are listed in Table 1.

6) cDNA sequencing of two picorna-like viruses

Based on the isotig sequences of the picorla-like viruses which were obtained by analysis of each EST library, RT-PCT primer sets were designed to amplify the virus genome. The viral genomes were amplified into 2.5 kb of subgenomic fragments to clone into pGem-T easy vector for Sanger sequencing. The PCR was carried out using Phusion High-Fidelity DNA polymerase (Thermo scientific) in a 50 μl volume. The PCR was performed under the following temperature

cycles: one cycle at 98℃ for 30 sec, 33 cycles at 98℃ for 10 sec, 60℃ for 30 sec, 72℃ for 90 sec, and one cycle at 72℃ for 10 min. In order to walk towards the 5' and 3' end of the viral genomes, 5' and 3' RACE was used repeatedly by using a SMARTer PCR cDNA Synthesis Kit (Clontech Laboratories, Inc. CA, USA). The products of RACE were cloned into the TA cloning plasmid vector, pGEM-T Easy (Promega, Madison, WI, USA) and sequenced in both direction.

7) Sequence and phylogenetic analysis

The complete nucleotide and amino acid sequences of the two picorna-like viruses were compared with the GenBank/EMBL databases using the BLAST (Altschul *et al*., 1990) and FASTA (Pearson and Lipman, 1988) programs. Multiple alignments of amino

acid sequences were obtained using the CLUSTAL W (Thompson *et al*., 1994) program, and the GeneDoc program was used to edit the alignments. Molecular phylogeny of the two picorna-like viruses was performed with the RdRp deduced amino acid sequences from a total of 22 virus sequences which belong to *Picornaviridae*, *Dicistrociridae*, and *Iflaviridae* along with two out group virus sequences, parsnip yellow fleck virus (PYFV, GenBank ID: D14066) in *Sequiviridae* and cowpea mosaic virus (CPMV, GeneBank ID: X00206) in *Comoviridae*. Phylogenetic analyses were conducted using the neighbor-joining method (Saitou and Nei, 1987) incorporated in MEGA software ver. 5. The GenBank sequence accession numbers of the viruses used in this study are as follows : encephalomyocarditis virus (NC_001479), foot-and-mouth disease virus (X00871), human poliovirus (PV, NC_002058), hepatitis A virus (M14707), parsnip yellow fleck virus (PYFV, D14066), cowpea mosaic virus (CPMV, X00206), ABPV (AF486073), TSV (AF277675), TrV (AF178440), HiPV (AB017037), PSIV (AB006531), ALPV (AF536531), CrPV (AF218039), DCV (AF014388), BQCV (AF183905), RhPV (AF022937), DWV (AJ489744), KV (NC_005876), VDV-1 (NC_006494), SBV (AF092924), VcSRV (AY534885), IFV (AB000906), PnV (AF323747), and EoPV (NC_005092).

Table1 . Primers used for quantitative real-time PCR expression profiling of putative transcripts generated from 454-pyrosequencing.

Target gene	Sequence of the primer pair
Vitellogenin	Fw: 5' GGACGAATACTGAAGCACACTTGG 3' Re: 5' CAGCAGCAGCAGCAGCAAC 3'
Heat shock protein 1	Fw: 5' GAAGAGAGCGTTTATTGAGTGTTGTG 3' Re: 5' CTCCTAAATGGGCGATGACTATGATG 3'
Heat shock protein 2	Fw: 5' GTCCTCCTCGTCAGCCTTGTC 3' Re: 5 GAAGAAGCACTCGCAGTTTATCGG '3'
Translation initiation factor 1	Fw: 5' GCAAGGAAATCACGCCATACTGTC 3' Re: 5' CAATGACTGGAGTGTGTTCTGTGG 3'
Translation initiation factor 2	Fw: 5' TGTGTACTTGGTAGCCGATGACTG 3' Re: 5' CGATGCCGAGAAGCGAAACG 3'
Translation initiation factor 3	Fw: 5' TGCGTGAGATGTCGTTTGATGTG 3' Re: 5' CGAGAGGTTTGGAAAGTGTGGTG 3'
Ribosomal protein 1	Fw: 5' TGTCACAAAGCGTGCCAAGAAAC 3' Re: 5' TGGTAAGGTCGGTCCTGTAATGC 3'
Ribosomal protein 2	Fw: 5' CGGGCGTGATGAATGGATTTGG 3' Re: 5' GCTCTCCTGCGTCGTCTCG 3'
Ribosomal protein 3	Fw: 5' GTCACAAACGGAAGTATCGCTCAC 3' Re: 5' ATCGGCTCGTCACAGTTCTCG 3'
Ribosomal protein 4	Fw: 5' GTAGAAGCGAAACAGCCCAACTC 3' Re: 5' GTGACCCTTCCTACCGAATCCG 3'

RESULTS

1. EST Analysis of Small Brown Planthopper, *Laodelphax striatellus* Carrying Rice Stripe Virus

1) RT-PCR for confirmation of RVLS or NVLS

RT-PCR was performed with two sets of RSV and SBPH specific primers, respectively. CCOS-F/CCOS-R and ATP-F/ATP-R yielded expected products of 684 bp and 655 bp, respectively not only from NVLS samples but also from RVLS samples (Fig. 1A). RSV-CP-F/RSV-CP-R, RSV-NCP-F/RSV-NCP-R yielded an expected product of 969 bp, 537 bp, respectively from RVLS samples only (Fig. 1B). As a result of this RT-PCR, the samples were confirmed whether RVLS or NVLS.

2) Transciptome sequencing results and EST library analysis

As previously described, the cDNA samples of RVLS and NVLS were subjected to two of the 1/8 production runs on the 454 GS-FLX sequencing instrument, respectively. After filtering adaptors and low-quality sequences, the cDNA samples of RVLS and NVLS generated 90,147 and 90,009 reads, respectively. These sequence data of RVLS and NVLS were deposited in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) with accession numbers of SRR538840 and SRR538841, respectively. Both data sets of assembled isotigs and singletons from RVLS and NVLS were pooled to construct a combined sequence dataset which was supposed to

have more and longer SBPH isotigs. The combined dataset was used to build a new reference ETS library to compare with RVLS and NVLS, and to analyze differential gene expression pattern. The features of the three EST libraries are summarized in Table 2. As show in Table 2, a total of 1,538 and 699 isotigs were obtained from the RVLS and NVLS datasets, with average length of 720 bp and 699 bp, respectively. The statistics of the combined dataset were 180,156 reads, 2,555 isotigs, and 745 bp of average length.

3) Gene ontology analysis

To obtain an overview of the transcriptome of SBPH, the isotigs of RVLS and NVLS were subjected to gene ontology (GO) analysis. The GO distributions of the isotigs and singletons of both RVLS and NVLS were similar to each other (Table 3), and the detailed GO annotations of each subcategory also showed similar trends of distributions. This result implies that the RSV did not alter the gene expression attributes of SBPH significantly. The 2,555 isotigs of the combined dataset were analyzed by comparing with *Drosophila* proteins using BLASTX with a cut-off value of 1e-05 to identify their gene functions. Total of 1,467 isotigs of the BLASTX hits were subjected to Kyoto Encyclopedia of Genes and Genomes (KEGG) orthology analysis (Kanehisa *et al.*, 2004) to categorize their functional classes (Fig. 2A)

4) Comparison of gene expression

To compare the gene expression pattern of RVLS and NVLS, the total read counts of SBPH isotigs from RVLS and NVLS were normalized arithmetically, and then total of the 2,543 SBPH isotigs from the combined EST library were plotted based on the

normalized read counts in RVLS and NVLS sequencing results. Significant differences of the transcription levels were determined by comparing the read counts of each isotig (binominal probability of < 0.1) to identify the isotigs which were up- or down-regulated by RSV. The distribution of isotigs showed that majority of the host genes were not significantly affected while 453 (17.8%) and 428 (16.8%) of isotigs were up- and down-regulated in RVLS, respectively (Fig. 3). Especially, the numbers of isotigs which showed more than 10-fold of up- or down-regulation in RVLS were 108 and 28, respectively. The isotigs which showed significant differences of the transcription levels were also subjected to KEGG orthology anaylsis to show their functional classes. The isotigs related to folding, sorting and degradation, and cellular processes were up-regulated while the genes related to translation and xenobiotics metabolism were down-regulated. In particular, the genes related to translation were significantly down-regulated (Fig. 2B).

5) Comparison of transcriptom profiles between RVLS and NVLS using qPCR

The qPCR results confirmed that the ten isotigs were indeed transcribed in both RVLS and NVLS, and the calculation of the relative transcription levels of the isotigs using $2^{-\Delta Ct}$ method (Pfaffl, 2001) demonstrated that the seven genes (vitellogenin, heat shock protein 1 and 2, and ribosomal protein 1, 2, 3 and 4) and the other three genes (translation initiation factor 1, 2 and 3) were up- and down-regulated in RVLS, respectively. To calculate fold differences of the isotigs, the read counts of the isotigs in RVLS and NVLS which are the corresponding isotigs

of the qPCR experiment were measured in reads per kilobase of exon per million mapped sequence reads (RPKM) (Mortazavi *et al.*, 2008). The sequence reads of the five isotigs (heat shock protein 1 and 2, translation initiation factor 1, 2 and 3) were confirmed in both RVLS and NVLS while the other five isotigs (vitellogenin, ribosomal protein 1, 2, 3 and 4) were confirmed in RVLS exclusively. The expression levels measured by qPCR were in a good accordance with the ratios of read counts of the EST libraries (Fig. 4).

6) Rice stripe virus transcripts in RVLS

The RSV transcripts of three of the seven viral ORFs, NS3, CP and RdRP were confirmed in RVLS, and their sequence read counts were 293, 224 and 5, respectively. The total RSV read counts were 522, and it was 0.63% of the total of 83,041 valid sequence reads of RVLS isotigs and singletons. The low read count of RdRp suggests that the replication of RSV is not highly active in SBPH while NS3 was the most abundant RSV transcript in RVLS which codes for viral suppressor of siRNA (Hemmes *et al.*, 2009). CP was the second most abundant RSV transcript which codes for capsid protein (Hayano *et al.*, 1990). Both NS3 and CP are the genes located on genomic and antigenomic strand of RSV RNA3, respectively.

Therefore the similar levels of NS3 and CP transcripts possibly suggest that the genes in RNA3 are under the same gene expression regulation mechanism, however, qPCR experiments to quantify the RSV transcripts in RVLS need to be performed to prove the hypothesis.

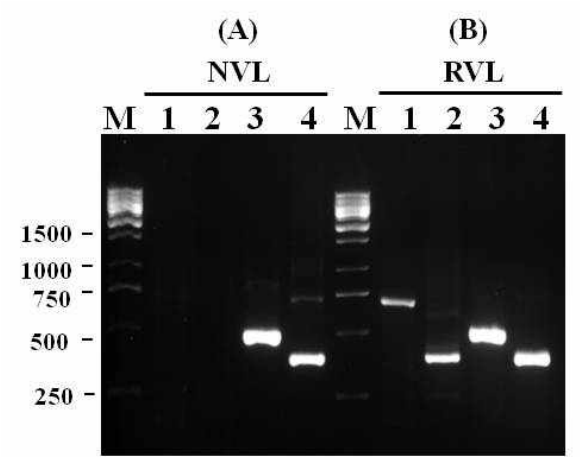


Fig. 1. Agarose gel analysis of RT-PCR assays. RSV specific primers (lane 1 and 2) and SBPH specific primers (lane 3 and 4). Only the RNA samples from the RVLS amplified viral sequences (A), while both RVLS and NVLS samples amplified SBPH specific sequences (B). M, 1 kb DNA ladder.

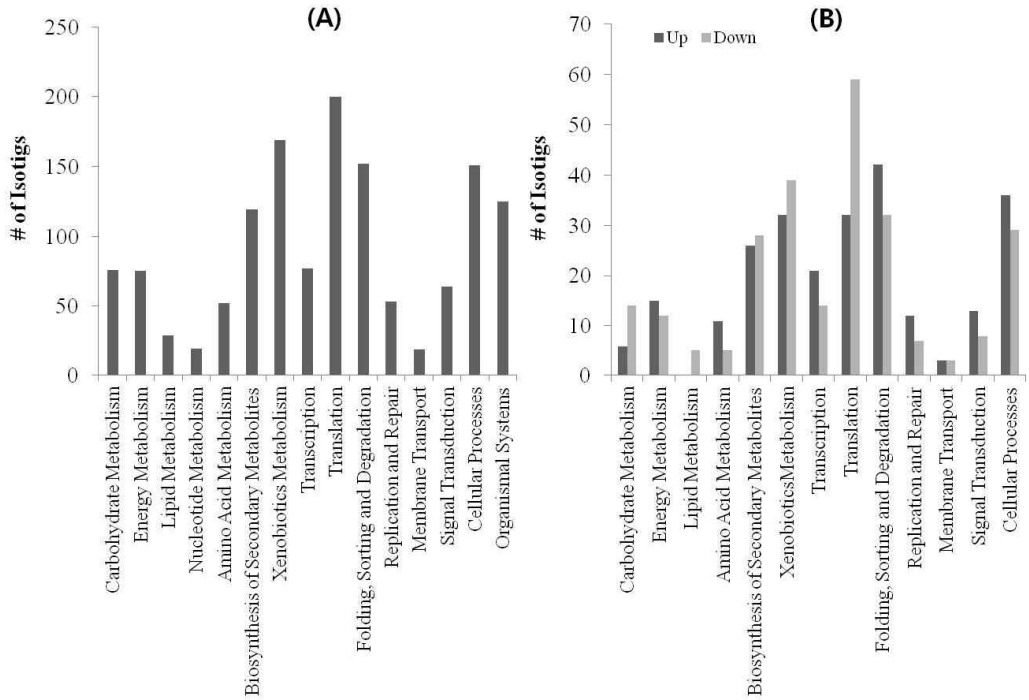


Fig. 2. KEGG analysis of total (A), and UP and DOWN (B) isotigs. (A) Functional groups of the combined ETS database defined by KEGG orthology were classified as carbohydrate metabolism, energy metabolism, lipid metabolism, nucleotide metabolism, amino acid metabolism, biosynthesis of other secondary metabolites, xenobiotics biodegradation and metabolism, transcription, translation, folding sorting and degradation, peplication and repair, membrane transport, signal transduction, cellular processes, and organismal systems. (B) The isotigs which showed significantly different transcription level based on the read number (a binominal probability of < 0.1) were classified as UP or DOWN. UP and DOWN indicates the genes up- and down-regulated in RVLS, respectively.

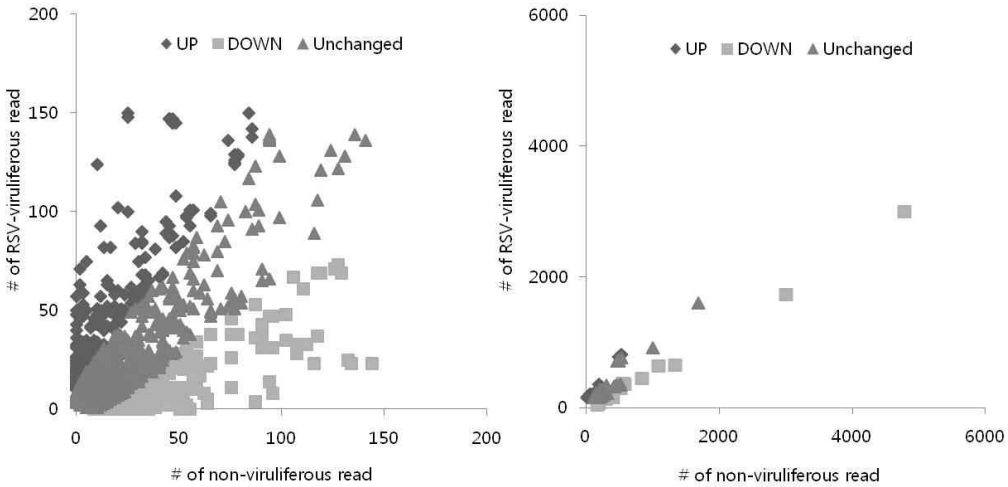


Fig. 3. Graph of the read counts of RVLS and NVLS for each isotig. The sequence read counts of RVLS and NVLS of each isotig were graphed on an x-y plot. For convenience, isotigs were ploted on two separated graphs. For the contigs which have less than 200 read counts are shown in the left panel, and the isotigs with more than 200 read counts were graphed in the write panel. Total of 2543 SBPH isotigs were subjected to this plot, and 1662 of the isotigs were not significantly affected ($p>0.1$) while 453 and 428 of isotigs were up- and down-regulated, respectively. Up- and down-regulated isotigs are indicated as triangles and squares, respectively.

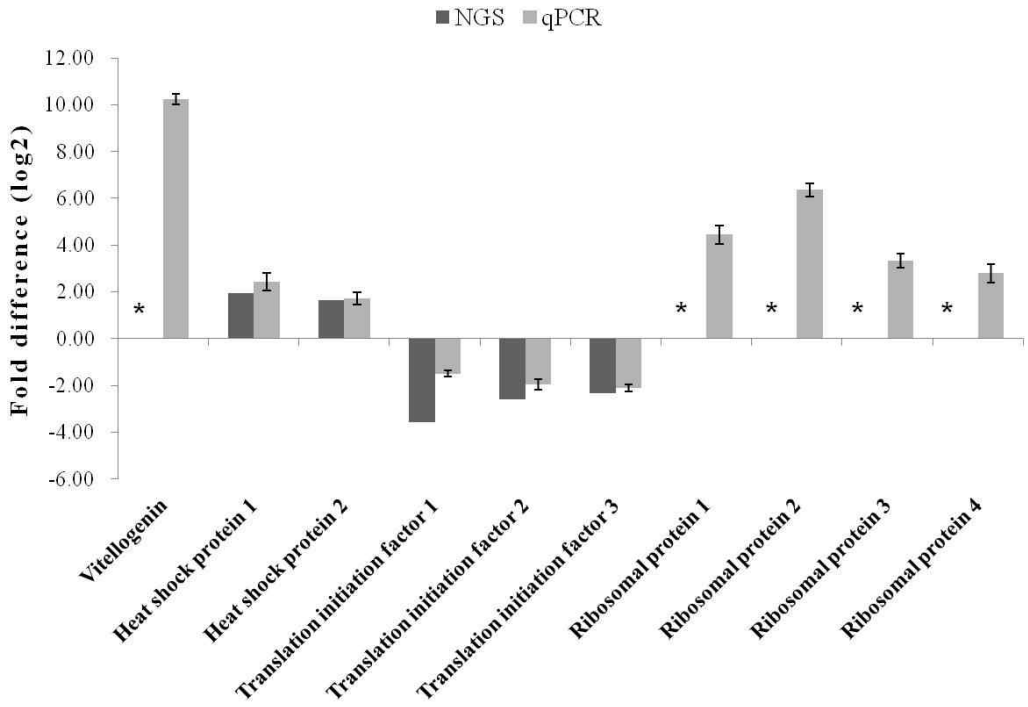


Fig. 4. Validation of the 454 pyrosequencing results by qPCR. The expression profiles of the ten isotigs were analyzed by qPCR to validate the 454 pyrosequencing results. The relative transcription levels are expressed in fold differences of RVLS to NVLS calculated by qPCR using $2^{-\Delta Ct}$ method.

Table 2. Analysis of *L.striatellus* EST libraries.

	RSV-viluliferous *L. strialellus*	non-viluliferous *L. strialellus*	Combined two samples
Total bases (bp)	30,071,326	30,566,488	60,637,814
Number of reads	90,147	90,009	180,156
Average reads length (bp)	336	342	342
Number of contigs	1,538	952	2,555
-Average contig length (bp)	720	699	745
-Ranage of contig length (bp)	15~8,596	69~9,265	82~9,179
Number of singletons	25,855	14,733	33,808

Table 3. The GO distributions of the isotigs and singletons

	RSV-viruliferous *L. strialellus*		non-viruliferous *L. striatellus*	
	Isotig	Singleton	Isotig	Singleton
Total	1,538	25,855	952	14,733
NR BLASTX hit ($e<10^{-3}$)	927 (60.3%)	9,846 (61.9%)	595 (62.5%)	5,690 (38.6%)
GO hit total	829 (53.9%)	8,232 (31.8%)	521 (54.7%)	4,684 (31.8%)
- Biological precess	295 (19.2%)	2,800 (10.8%)	182 (19.1%)	1,629 (11.1%)
- Cellular component	203 (13.2%)	2,137 (8.3%)	137 (14.4%)	1,182 (8%)
- Molecular function	331 (21.5%)	3,295 (12.7%)	202 (21.2%)	1,873 (12.7%)
GO no-hit	709 (46.1%)	17,623 (68.2%)	431 (45.3%)	10,049 (68.2%)

2. Identification and Genome Sequence of Two Picorna-like Viruses in *Laodelphax striatellus*

1) Sequence analysis of the two picorna-like viruses' genome

Two picorna-like viruses were isolated from NVLS and RVLS, and named LsPV-1 and LsPV-2, respectively. The LsPV-1 cDNA sequence found in NVLS EST library showed covering 86% of Himetobi P virus (HiPV) genome. The LsPV-1 genome was 9,272 nt in length excluding the poly (A) tail, and contained two open reading frames (ORFs) which were separated by a 176 nt intergenic region that functions as an internal ribosome entry site (IRES). The 5' ORF encodes the non-structural proteins and the 3' ORF encodes the capsid proteins. The result of LsPV-1 sequencing showed 99% sequence identity to that of HiPV reported in Japan (Toriyama *et al*., 1992).

The LsPV-2 cDNA sequence found in RVLS EST library showed 28% sequence identity to that of slow bee paralysis virus (SBPV). The LsPV-2 genome was 8,769 nt in length excluding the 5' and 3' end regions and the poly(A) tail, and contained a single, large open reading frame (nt 1–8,535) encoding a polyprotein of 2,845 aa. Total of six clones of the 5' end of the viral genome were sequenced, and the 5' terminal nucleotide sequence was determined by comparison of the sequences from the six clones. The partial nucleotide sequence of the LsPV-2 identified no homologue in the BLAST search result, however, the RdRp of LsPV-2 showed 52% identity to that of Deformed wing virus (DWV). In terms of sequence homology and genome organization, LsPV-2 resembled insect picorna-like viruses belonging to *Iflaviridae*.

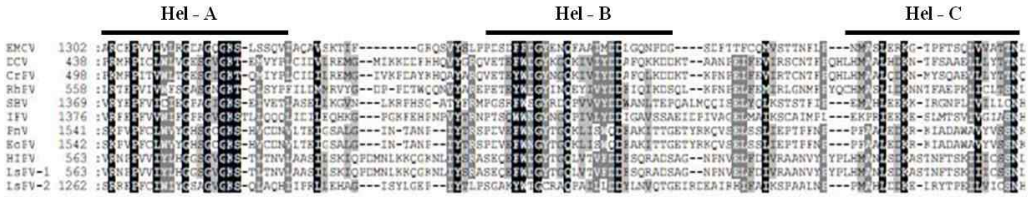
2) Predicted amino acid sequence and phylogenetic analysis

Based on the RNA genome sequence, non-structural proteins of the two picorna-like viruses had conserved functional motifs that are characteristics of viral helicase, protease, and RdRps in the picorna-like superfamily (Koonin and Dolja, 1993). Non-structural polyprotein of LsPV-1 and LsPV-2 located at the 5' and 3' part of the RNA genome, respectively. The three conserved domains in helicase of picorna-like viruses were found in putative helicase coding sequences of LsPV-1 and LsPV-2 (Fig. 5A). Also, the putative protease of LsPV-1 and LsPV-2 shared a conserved region in the proteases of other picorna-like viruses (Fig. 5B). Eight conserved domains have been identified in RdRp amino acid sequences (Koonin and Dolja, 1993) and they also exist in LsPV-1 and LsPV-2 (Fig. 5C).

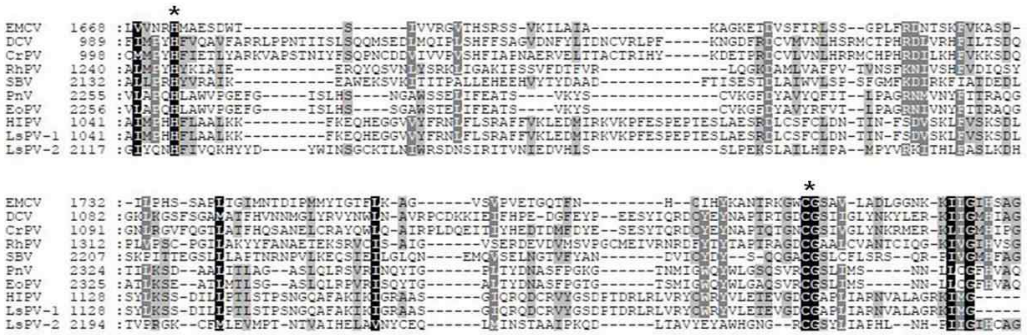
The phylogenetic tree was constructed for the RdRp amino acid sequences of the LsPV-1, LsPV-2 and other 23 viruses of the picorna-like superfamily. The phylogenetic tree segregated the three virus taxa, *Dicistroviridae*, *Picornaviridae* and *Iflaviridae* (fig. 6). The LsPV-1 showed the closest relationship to HiPV, hence we concluded that LsPV-1 is a Korean isolate of Himetobi-P virus (HiPV). Also, the phylogenetic tree showed that LsPV-2 is closely related to VDV, DWV and KV. These results suggest that LsPV-2 is a new member of *Iflaviridae.*

Fig. 5. Multiple sequence alignment of the LsPV-1 and LsPV-2 with picorna-like viruses for the helicase (A), protease (B) and RNA-dependent RNA polymerase (RdRp) (C) sequences. Conserved regions corresponding to those recognized by Koonin and Dolja (1993) are indicated by bars above the protein alignment. Residues forming the catalytic triad in the protease (Koonin and Dolja, 1993) are marked with asterisks. Black, dark-gray, and light-gray shading indicates 100%, 80%, and 60% sequence identity, respectively. The numbers at the beginning of the sequences represent the amino acid position from the start of the open reading frame (ORF).

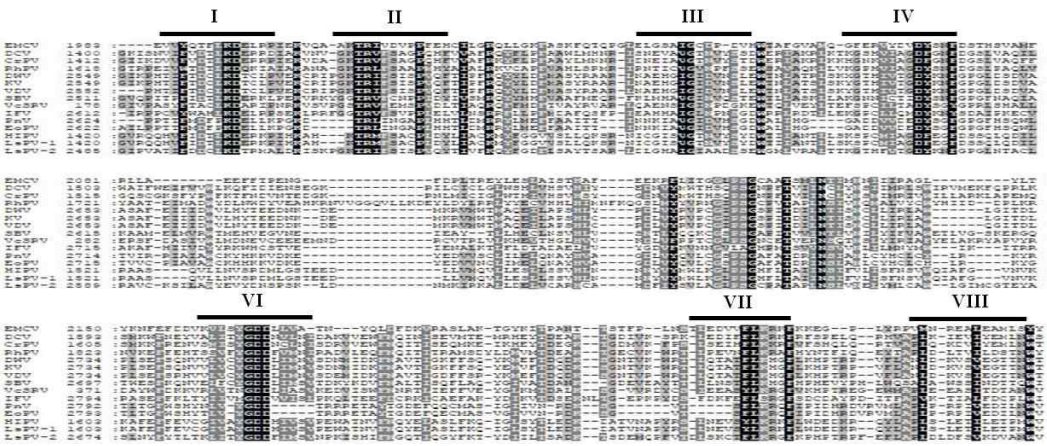
A. Helicase



B. Protease



C. RNA-dependent RNA polymerase



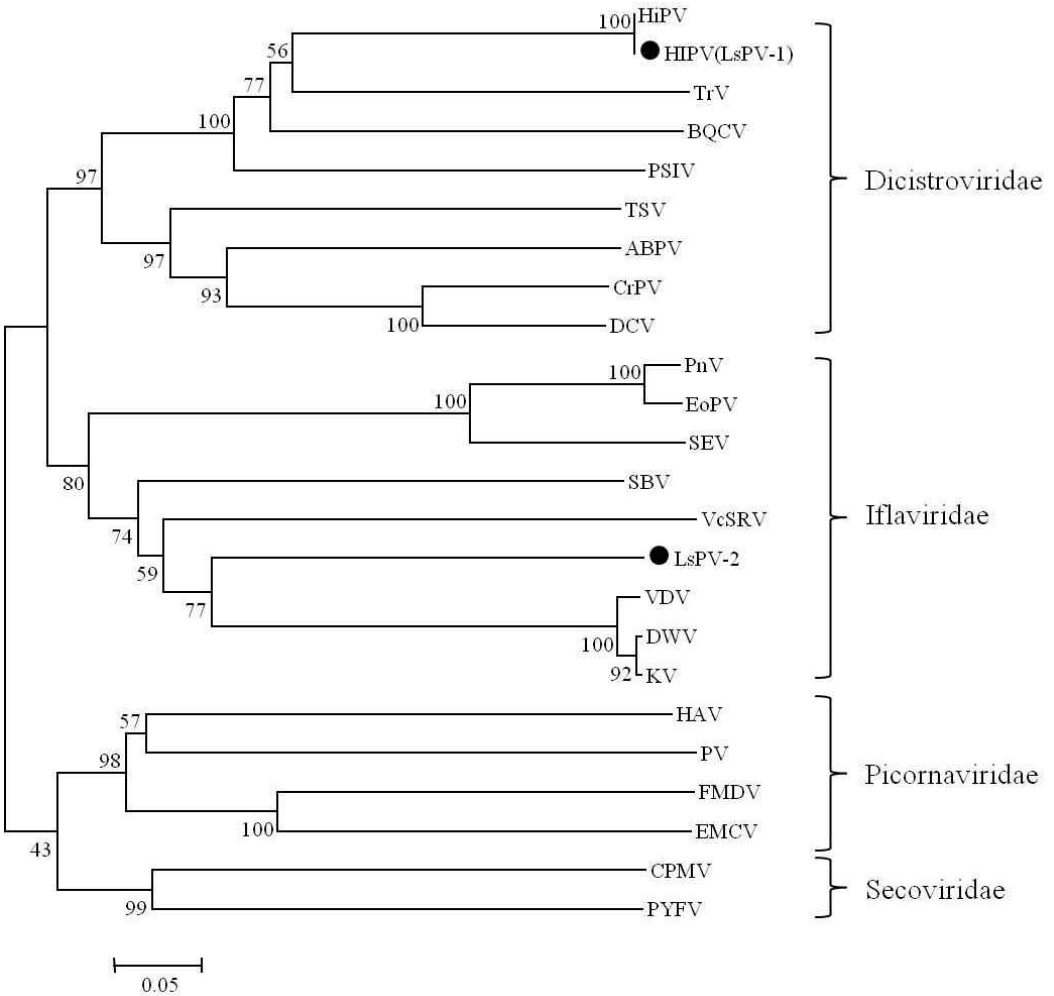


Fig. 6. The Phylogenetic tree obtained from the RNA-dependent RNA polymerase (RdRp) sequence of 24 viruses belonging to *Picornaviridae*, *Dicistroviridae*, and *Iflaviridae* including that of LsPV-1 and LsPV-2. The parsnip yellow virus (PYFV) in *Sequiviridae* and cowpea mosaic virus (CPMV) in *Comoviridae* were used as out groups. Numbers at each node specify bootstrap percentage of 1000 replications. The scale bar indicates the number of substitutions per site. The evolutionary distances were computed using the p-distance method and are in the units of the number of amino acid differences per site. There were a total of 171 positions in the final dataset. Evolutionary analyses were conducted in MEGA5.

DISCUSSION

The 454 pyrosequencing results demonstrated that the next generation sequencing (NGS) technique was able to reflect the characteristics of the transcriptome of RVLS. The genes in the RVLS which showed different transcription patterns indicated that the RSV may alter the gene expression of the SBPH which possibly affect the physiology, fitness and behavior of RVLS. The up-regulated heat shock protein genes might be one of the examples that the RSV may take the advantage of altering gene transcription to improve fitness of the vector because heat shock proteins play a significant role in generating tolerance to stresses (Krebs, 1999; Sorensen *et al.*, 2003). It is also interesting to note that the transcription level of vitellogenin in RVLS was highly up-regulated, which means RSV might modulated the host vitellogenin pathway genes to promote the oogenesis, or altered sex ratio of RVLS to facilitate its persistent transovarial transmission via the offsprings of RVLS. The highly up-regulated ribosomal protein genes were also the possible consequence of the up-regulation of vitellogenin because the ovarian cells need more amount of translational machinery, however, the down-regulated translation initiation factors are paradoxical because the recruitment of transcription initiation factors is a crucial step of protein synthesis. Many viruses can inhibit host translation initiation factors to take over host gene expression by which are usually viral protein mediated protein level interactions, rather than transcription regulation (Walsh and Mohr, 2011). It suggests that RSV possibly can alter the transcription of the genes for cellular translational machinery direct or indirect manner. Another example of viral modulation of cell physiology is the major viral transcript NS3, a RNAi suppressor which binds 21 nt

siRNA, and it might be a viral strategy to evade defense system of SBPH and rice plant (Hemmes *et al.*, 2007; Alvarado and Scholthof, 2009). The RVLS and NVLS showed unique aspects of viral microflora with incompatibilities between RSV and LsPV-2 of RVLS and LsPV-1 of NVLS. Also, these incompatibilities of picorna-like viruses possibly exist in the field population of SBPH which may become a novel approach of RSV control in future. In this study, we laid the groundwork for the further study of RSV and SBPH interaction by characterization and genome analysis of the two picorna-like viruses. Early studies on vector transmission of plant viruses demonstrated the specificity of the virus-vector interaction, which suggested that the instances of the biological properties associated with the transmission process may differ significantly (Ng and Falk., 2006). However, the genomics approach to study the virus-vector interaction using NGS technique may contribute to investigate the gene regulation mechanism in the persistent vector insects and the upstream regulation factors, which will be a new approach to study the interaction among the virus, plant, and vector.

LITERATURE CITED

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Alvarado, V. and Scholthof, H.B. 2009. Plant responses against invasive nucleic acids: Rna silencing and its suppression by plant viral pathogens. *Semin Cell Dev Biol*, 20: 1032-1040.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G. and Consortium, G.O. 2000. Gene ontology: Tool for the unification of biology. *Nat Genet*, 25: 25-29.

Christian, P., Carstens, E., Domier, L., Johnson, K., Nakashima, N., Scotti, P., van der Wilk, F., 2005a. Dicistroviridae. In: Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. (Eds.), Virus Taxonomy: Eighth Report of the International Committee on the Taxonomy of Viruses. Academic Press, San Diego, pp. 783–788.

Christian, P., Carstens, E., Domier, L., Johnson, K., Nakashima, N., Scotti, P., van der Wilk, F., 2005b. Iflavirus. In: Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. (Eds.), Virus Taxonomy: Eighth Report of the International Committee on the Taxonomy of Viruses. Academic Press, San Diego, pp. 779–782.

Falk, B.W. and Tsai, J.H. 1998. Biology and molecular biology of viruses in the genus tenuivirus. *Annual Review of Phytopathology*, 36: 139-163.

Hamamatsu, C., Toriyama, S., Toyoda, T. and Ishihama, A. 1993. Ambisense coding strategy of the rice stripe virus genome: In vitro translation studies. *J Gen Virol*, 74 (Pt 6): 1125-1131.

Hayano, Y., Kakutani, T., Hayashi, T. and Minobe, Y. 1990. Coding strategy of rice stripe virus: Major nonstructural protein is encoded in viral rna segment 4 and coat protein in rna complementary to segment 3. *Virology*, 177: 372-374.

Hemmes, H., Kaaij, L., Lohuis, D., Prins, M., Goldbach, R. and Schnettler, E. 2009. Binding of small interfering rna molecules is crucial for rna interference suppressor activity of rice hoja blanca virus ns3 in plants. *J Gen Virol*, 90: 1762-1766.

Hemmes, H., Lakatos, L., Goldbach, R., Burgyan, J. and Prins, M. 2007. The ns3 protein of rice hoja blanca tenuivirus suppresses rna silencing in plant and insect hosts by efficiently binding both sirnas and mirnas. *RNA*, 13: 1079-1089.

Hibino, H. 1996. Biology and epidemiology of rice viruses. *Annu Rev Phytopathol*, 34: 249-274.

Hibino, H., Usugi, T., Omura, T., Tsuchizaki, T., Shohara, K. and Iwasaki, M. 1985. Rice grassy stunt virus - a planthopper-borne circular filament. *Phytopathology*, 75: 894-899.

Hogenhout, S.A., Ammar el, D., Whitfield, A.E. and Redinbaugh, M.G. 2008. Insect vector interactions with persistently transmitted viruses. *Annu Rev Phytopathol*, 46: 327-359.

Iida, T.T., 1969. Dwarf, yellow dwarf, stripe, and black-streaked dwarf diseases of rice., p. 3-11., in proceedings of a symposium on the virus diseases of the rice plan., 25-28 April, 1967, Los banos, Philippines, Johns Hopkines Press, Baltimore.

Ishikawa, K., Omura, T. and Hibino, H. 1989. Morphological-characteristics of rice stripe virus. *Journal of General Virology*, 70: 3465-3468.

Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. and Hattori, M. 2004. The kegg resource for deciphering the genome. *Nucleic Acids Res*, 32: D277-D280.

Kim CS, Lee GS, Choi HS ., 2009. Virus-insect-plant interaction at RSV outbreak regions in Korea. In: Proceedings of APEC workshop on the epidemics of migratory insect pests and associated virus diseases in rice and their impact on food security in APEC member economies, Seoul, pp 61–74

Kim JS ., 2009. Forecasting and occurrence of SBPH and RSV in Korea. In: Proceedings of APEC workshop on the epidemics of migratory insect pests and associated virus diseases in rice and their impact on food security in APEC member economies, Seoul, pp 89–106

Krebs, R.A. 1999. A comparison of hsp70 expression and thermotolerance in adults and larvae of three drosophila species. *Cell Stress Chaperon*, 4: 243-249.

Ling, K.C., 1972. Rice virus diseases. Los Banos, Philippines: International rice research institute.

Lee, S.C., 1969. Rice stripe disease in Korea, p.67-73., in proceedings of a symposium on the virus diseases of the rice plan., 25-28 April, 1967, Los banos, Philippines, Johns Hopkines Press, Baltimore.

Mortazavi, A., Williams, B.A., Mccue, K., Schaeffer, L. and Wold, B. 2008. Mapping and quantifying mammalian transcriptomes by rna-seq. *Nat Methods*, 5: 621-628.

Nault, L.R. and Ammar, E. 1989. Leafhopper and planthopper transmission of plant-viruses. *Annu Rev Entomol*, 34: 503-529.

Ng, J.C. and Falk, B.W. 2006. Virus-vector interactions mediating nonpersistent and semipersistent transmission of plant viruses. *Annu Rev Phytopathol*, 44: 183-212.

Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real-time rt-pcr. *Nucleic Acids Res*, 29.

Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M. and Nyren, P. 1996. Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem*, 242: 84-89.

Sorensen, J.G., Kristensen, T.N. and Loeschcke, V. 2003. The evolutionary and ecological role of heat shock proteins. *Ecol Lett*, 6: 1025-1037.

Takahashi, M., Toriyama, S., Hamamatsu, C. and Ishihama, A. 1993. Nucleotide sequence and possible ambisense coding strategy of rice stripe virus rna segment 2. *J Gen Virol*, 74 (Pt 4): 769-773.

Tong, Z.G., Gao, Z.H., Wang, F., Zhou, J. and Zhang, Z. 2009. Selection of reliable reference genes for gene expression studies in peach using real-time pcr. *Bmc Mol Biol*, 10.

Toriyama, S. 1986. Rice stripe virus: Prototype of a new group of viruses that replicate in plants and insects. *Microbiol Sci*, 3: 347-351.

Toriyama, S., Takahashi, M., Sano, Y., Shimizu, T. and Ishihama, A. 1994. Nucleotide sequence of rna 1, the largest genomic segment of rice stripe virus, the prototype of the tenuiviruses. *J Gen Virol*, 75 (Pt 12): 3569-3579.

Walsh, D. and Mohr, I. 2011. Viral subversion of the host protein synthesis machinery. *Nat Rev Microbiol*, 9: 860-875.

Zhu, Y., Hayakawa, T. and Toriyama, S. 1992. Complete nucleotide sequence of rna 4 of rice stripe virus isolate t, and comparison with another isolate and with maize stripe virus. *J Gen Virol*, 73 (Pt 5): 1309-1312.

Zhu, Y., Hayakawa, T., Toriyama, S. and Takahashi, M. 1991. Complete nucleotide sequence of rna 3 of rice stripe virus: An ambisense coding

strategy. *J Gen Virol*, 72 (Pt 4): 763-767.

벼 줄무늬잎마름병을 매개하는 애멸구의 전사체 분석

서울대학교

농생명공학부 곤충학전공

이주현

초록

벼 줄무늬잎마름병바이러스 (Rice stripe virus, RSV)는 Tenuivirus에 속하며 애멸구에 의해서 벼에 전파되어 벼 줄무늬잎마름병을 일으킨다. 벼 체관부를 뚫고 흡즙하는 애멸구의 먹이 행동과, RSV가 애멸구 체내에서 복제와 증식이 가능하다는 점 때문에 애멸구는 RSV의 지속적인 감염을 가능하게 하는 가장 심각한 RSV 매개충이다. 현재까지는 RSV와 애멸구, 벼 사이의 상호관계에 집중하기 보다는 RSV나 애멸구로부터 벼를 보호하는 것에 집중하였다. 이러한 이유로 이 논문에서는 RSV와 애멸구 사이의 상호관계를 조사하는 차원에서 RSV보독 애멸구와 비보독 애멸구 사이에서의 차이를 밝히기 위해 각 각의 애멸구에서 RNA를 추출하여 전사체 분석을 실시하였다. 454 pyrosequencing을 통해서 약 3천만 개의 염기를 읽었으며 RSV 보독

애멸구에서는 1,538개의 isotig를, RSV 비보독 애멸구에서는 953개의 isotig를 만들었다. 유전자를 기능에 따라 분류했을 때, RSV 보독 애멸구와 비보독 애멸구 사이에서는 비슷한 성향을 나타내었다. 그러나 유전자의 발현 경향을 보았을 때, RSV 보독 애멸구에서 비보독 애멸구에 비해 108개의 유전자가 10배 이상 많이 발현 되었으며, 28개의 유전자가 10배 이하 낮게 발현 되는 차이를 보였다. 이처럼 RSV에 의해서 발현이 증가되거나 억제되는 유전자들은 애멸구가 RSV를 전파하는데 있어서 애멸구의 생리학적인 부분이나 행동적인 부분에 중요한 역할을 할 것으로 예상된다.

또한, 전사체 분석 과정에서 RSV 비보독 애멸구와 RSV 보독 애멸구에서 각각 새로운 두 종류의 picorna-like virus 를 발견하였다. RSV 비보독 애멸구에서 발견한 바이러스를 LsPV-1, RSV 보독 애멸구에서 발견한 바이러스를 LsPV-2 라고 명명하였다. 시퀀스를 분석 한 결과 LsPV-1 은 *dicistroviridae* 에 속하는 일본에서 보고된 Himetobi-P virus (HIPV)와 매우 유사한 시퀀스를 가지고 있었기 때문에 한국에서 분리한 HiPV 로 추정되며, LsPV-2 는 *iflaviridae* 에 속하는 새로운 iflavirus 로 추정된다. LsPV-2 를 RNA 의존성 RNA 중합효소 (RNA dependent RNA polymerase, RdRp) 시퀀스를 바탕으로 계통 분석한 결과 deformed wing virus 와 유사한 바이러스라는 결과를 보였다. 흥미로운 것은 RSV 와 HiPV 가 서로 중복 감염되지 않는다는 다는 점이고, 이러한 상호 양립하지 않는 현상은 애멸구 내에서 지속적으로 복제하고 증식하는 RSV 의 전염 메커니즘과 그 바이러스 상을 이해하는 데에

단서가 될 것이다. 또한 RSV 비보독 애멸구와 보독 애멸구에서 각각 발견된 HiPV 와 LsPV-2 의 상호 양립하지 않는 현상 역시, RSV 의 전염 메커니즘에 영향을 줄 것 이라고 예측된다.

검색어: **Rice stripe virus, *Laodelphax striatellus*, Transcriptome, picorna-like viruse, Himetobi P virus, iflavirus**

Student Number: 2011-21307