ISJ 12: 214-224, 2015 ISSN 1824-307X

RESEARCH REPORT

Paralogous gene conversion, allelic divergence of *attacin* genes and its expression profile in response to *Bm*NPV infection in silkworm *Bombyx mori*

G Lekha, T Gupta, K Trivedy, K Ponnuvel

Genomics Division, Seri Biotech Research Laboratory, Carmelaram Post, Kodathi, Bangalore 560 035, India

Accepted July 27, 2015

Abstract

The genomic organization, structure and polymorphism of *attacin* gene within the mulberry silkworm *Bombyx mori* strains have been analyzed. Genomic contig (AADK01007556) of *B. mori attacin* gene contains locus with two transcribed basic *attacin* genes, which were designated as *attacin* I and *attacin* II. Survey of the naturally occurring genetic variation in different strains of silkworm *B. mori* at the promoter and coding regions of two *attacin* genes revealed high levels of silent nucleotide variations (1- 4 % per nucleotide heterozygosity) without polymorphism at the amino acid level (non-Synonymous substitution). We also investigated variations in gene expression of *attacin* I and *attacin* II in silkworm *B. mori* infected with nucleopolyhedrovirus (*Bm*NPV). Two B*. mori* strains, Sarupat, CSR-2 which were resistant and susceptible to *BmNPV* infection respectively were used in this study. Expression profiles of *B. mori* genes were analyzed using microarray technique and results revealed that the immune response genes including *attacin* were selectively up regulated in virus invaded midguts of both races. Microarray data and real-time qPCR results revealed that *attacin* I gene was significantly up-regulated in the midgut of Sarupat following *Bm*NPV infection, indicating its specific role in the anti-viral response. Our results imply that these up-regulated *attacin* genes were not only involved in anti-bacterial mechanism, but are also involved in *B. mori* immune response against *Bm*NPV infection.

Key Words: *Bombyx mori; attacin*; microarray; genomic organization; differential expression

Introduction

Insects fight bacterial infection, in part, through the extra cellular circulation of a variety of short, general antibacterial peptides (Iwanaga and Lee, 2005). Although over 400 different innate immune peptides have been identified in eukaryotes (Hoffmann *et al*., 1999), most insects produce relatively small number, fewer than 10 peptide classes and these have to effectively combat a wide range of potential pathogens. Among the different antibacterial proteins produced in insects, *attacin*, a high molecular weight protein has a major role in insect innate immunity. The amino acid sequence deduced from cloned *attacin* cDNA of *B. mori* revealed that the cDNA encodes an *attacin* precursor protein (Sugiyama *et al.,* 1995). The putative mature protein of *Bombyx mori attacin* revealed varying levels of identity in amino acid sequences with those of *Hyalophora cecropia* acidic (70.4 %) and basic (68.3 %) *attacin*s and *Sarcophaga peregrina* sarcotoxin IIA (18.8 %). Injection of *Escherichia coli* cells into *B. mori* larvae resulted in rapid induction of the expression of *B. mori attacin* gene that continued at least for 48 h mainly in fat bodies and hemocytes (Sugiyama *et al.,* 1995). Taniai *et al.* (1996) isolated a genomic clone encoding *attacin* from genomic library of *B. mori*, and determined the nucleotide sequence of the 5'-upstream region. Mature *attacin* peptides are typically 190 - 214 amino acids in length (*Sarcophaga* peptides are longer) and adopt a "random coil" structure in solution (Gunne *et al*., 1990). This loose, flexible structure is devoid of disulfide bonds and does not take a rigid conformational shape which may allow relatively free amino acid substitution, explaining the low level of amino acid identity between *attacin* homologs in distant taxa.

Dushay *et al.* (2000) reported cloning of two closely linked *attacin* genes from *D. melanogaster*. A comparison of their protein coding sequences

Corresponding author:
Kangayam M Ponnuvel
Genomics Division
Seri Biotech Research Laboratory
Carmelaram Post, Kodathi
Bangalore 560 035, India
E-mail: kmpvel@yahoo.com

revealed that the amino acid sequences were more highly conserved than the nucleotide sequences, suggesting expression of both the genes (Wheelan *et al*., 2001). In this paper we present data on the quantity of polymorphism in the *B. mori attacin* genes and their expression profile in resistant and susceptible race. Further, the genomic structure of the *attacin* gene was analyzed and compared with *attacin* sequences of selected Indian silkworm strains. The structure of exon and intron as well as the phylogenetic relation of *B. mori attacin* gene to that in other insects were also compared and analyzed.

It is also reported that *attacin* gene has two paralogous genes *i.e., attacin* I as well as *attacin* II, both the genes are found to be expressed after bacterial infection (Tanaka *et al*., 2008). The organization of both *attacin* genes and its position are explained in this report which are found to be located on the 6th chromosome.

There are few antibacterial proteins such as *gloverin, lebocin, serpin* and these genes have been found to be involved in the immune response against the viral infection, especially against *Bm*NPV infection (Cheng *et al.,* 2014). There is not much information about the role of *attacin* gene against *Bm*NPV infection. A microarray analysis was carried out to identify the genes associated with *Bm*NPV resistance. There are many antibacterial proteins found to be upregulated after *Bm*NPV infection. Among those antibacterial genes, the expression of *attacin* gene was significantly upregulated after *Bm*NPV infection in the microarray analysis, indicating its prominent role in antiviral immunity. In the present study the differential expression of both *attacin* I and *attacin* II genes has also been analyzed after *Bm*NPV infection to know the role of these genes in the antiviral immune response in silkworm *B*. *mori*.

Materials and Methods

Selection of silkworm races

The silkworm B*ombyx mori* races *viz.,* Sarupat and CSR-2 were selected for the study, as these are known to be most resistant and most susceptible to *Bm*NPV. These two silkworm races were used for the microarray as well as for quantification and gene expression analysis using qPCR.

Virus and inoculations

B. mori multiple nucleopolyhedrovirus stock was maintained at this laboratory and used as viral inoculum. The viral inoculum was prepared by counting the number of viral polyhedra in a Neubauer chamber. The oral inoculation of *Bm*NPV occlusion bodies was carried out in healthy newly moulted '0 day' fifth instar larvae (first day after 4th moult) of Sarupat and CSR-2 races with viral dosage of 40,000 polyhedral inclusion bodies (PIB) per larva. Three replications containing twenty-five silkworms were maintained for each silkworm race. Similarly, the uninoculated control batches were reared separately under disease free environment. Silkworms feeding on *Bm*NPV-free mulberry leaves were placed in labelled boxes until feeding was complete and then transferred to a controlled room where they remained until the end of the experiment.

Collection of tissue

*Bm*NPV-infected fifth instar larvae (n = 6) were dissected and the midgut tissues was removed at different (6, 12, 18, 24, 30) h after post infection (hpi). They were quickly washed in diethylpyrocarbonate (DEPC)-treated solution and immediately frozen at -80 $^{\circ}$C for further analysis.

RNA isolation and cDNA synthesis

The RNA was extracted from different tissues like hemocytes, midgut, fat body and cuticle with TRIzol reagent (Invitrogen, USA), and then denatured in formaldehyde, formamide and electrophoresed in 2.0 % agarose gels. The first strand cDNA was synthesized using DNase treated RNA sample (2 µg) along with 1µl oligo (dT) (0.01mM) (Eurofin India Pvt Ltd, Bangalore) was added followed by incubation at 70 ºC for 3 min. Finally, 1X reverse transcriptase buffer (4µl), 10 mM dNTP (2 µl), 5 mM DTT (2µl) and M-MLV Superscript III reverse transcriptase (Invitrogen, USA) (0.5 µl) was added to obtain a final volume of 20 µl. The reaction mixture was incubated at 42 ºC for 60 min and terminated by heating at 75 ºC for 10 min according to the manufacturer's protocol.

Identification of attacin gene and genomic contig

The cDNA of *attacin* gene was already identified and deposited. The *attacin* cDNA sequence was blast (BLAST) searched with *B. mori* genomic DNA database (Xia *et al.,* 2004), for identification of corresponding contig homologous sequence for *attacin* gene. The genomic DNA sequence showing homologous sequence to *B.mori attacin* gene was identified and subsequently translated to determine putative amino acid sequence. The amino acid sequence was further analyzed through conserved domain search for the presence of the two functional domains in *attacin* (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb/cgi).

EST expression in different tissues

The specific expression of *attacin* I and *attacin* II in different tissues were identified by performing blastn followed by retrieval of the EST from the library. The different tissues selected are hemocytes, midgut, fat body and cuticle.

Microarray experiment and data analysis

A genome wide oligonucleotide microarray containing 24,924 probes were used to investigate the gene expression profiles of *Bm*NPV infected as well as control midguts of Sarupat and CSR-2 silkworm *B. mori* at 12 h after post infection. The complete sets of raw and normalized data from this study have been deposited in the NCBI Gene Expression Omnibus (GEO) repository.

Amplification of attacin gene in different silkworm races

The genomic DNA isolated from silk moths using standard protocols (Nagaraja and Nagaraju,

1995) was used as template in the PCR reaction. The up and down gene specific primers for *attacin* gene in the *B. mori* genomic contig were designed using the software program Primer3 (http://frodo.wi.mit.edu/cgibin/ primer3). The forward primer used was 5'-GGCTGGAAAGCTGGAACTAA-3' and the reverse was 5'-AGTCCATAGCCTGGGAACCT-3'. The reaction was done in an Eppendorf thermal cycler, PTC200, using 20 μl reaction mixture containing 50-100 ng of genomic DNA as template, 2.0 μl of 10X PCR buffer, 0.2 mM dNTPs, 1.5 mM $MgCl_2$, 1 μl each of forward and reverse primers and 0.3 U of *Taq* DNA polymerase (MBI fermentas). The PCR schedule was 94 °C for 3 min followed by 30 cycles of 94 °C for 30 s, 50 °C for 30 s, 72 °C for 2 min and a final extension of 7 min at 72 °C. The PCR amplified products were purified through Gel-spin column (Bangalore Genei) and M13 primer was used for the sequencing reaction. The amplified PCR product 699 bp length (Fig. 3) was cloned in TA cloning vector with M13 sequences flanking the 5' and 3' region and sequenced with gene specific primer as sequencing primer.

Tissue specific expression profile

Tissues were collected from haemocytes, fat body, midgut and cuticle of fifth instar third day larvae for tissue specific expression analysis. The *attacin* I and *attacin* II expression was analyzed using forward primer 5'- GCAGGCAAGGTCAATTTGTT-3' and reverse 5'-CGGTTGATGACGTCAGAGTG-3' for *attacin* I. Forward primer-5'TCGAGGTCGTATTGCAGACA-3' and 5'GGCTCCCACGAAGATCTGTA-3' of reverse primer for *attacin* II. The reactions were conducted on a Stratagene MxPro-Mx3005P Real-Time PCR system (Agilent technologies) using the SYBR Premix Ex Taq Kit (TaKaRa), according to the manufacturer's protocol. Each amplification reaction was performed using a 20 μl reaction mixture, under the following conditions: denaturation at 95 °C for 30 s, followed by 40 cycles of 95 °C for 10 s and at 55 °C for 30 s. The experiment was performed in triplicate and results were standardized to the expression level of the constitutive β actin gene. A non-template control (NTC) sample was also run to detect contamination, if any.

Expression in Sarupat and CSR-2 midgut

The *Bm*NPV infected and control midgut samples were collected at different intervals of post infection from 0 to 30 h. The RNA was isolated from the midgut tissues and cDNA was synthesized. The cDNA was used as template to quantify *attacin* I and *attacin* II gene expression by qPCR.

Results

The microarray analysis was carried out to investigate the gene expression profiles in silkworm *B. mori* against *Bm*NPV infection. The results indicated that some of the antibacterial proteins including *attacin* gene were upregulated after *Bm*NPV infection and thus indicating their potential role in the antiviral immune response (Sagisaka *et al.,* 2010). Therefore, an attempt has been made to study the differential expression of *attacin* gene in *Bm*NPV resistance. In addition to that the organization of paralogous *attacin* genes, their tissue specific expression and variation in the promoter and coding regions were analyzed. The *attacin* cDNA sequence (accession no. S78369) was blast (BLAST) searched with *B. mori* genomic DNA database and a single contig (accession no AADK01007556) possessing the *attacin* gene sequence was identified. Two paralogous sequences similar to *attacin* gene sequence were present in the locus. Further, there were three exon with two intron regions of 91 bp and 79 bp length respectively. These two gene sequences were arranged in a direction opposite to the single contig with a gap of 3.4 kb length and the paralogous genes were designated as *attacin* I and *attacin* II (Fig.1). The conserved domain analysis showed the presence of two sub domains the in G domain of *attacin*, similar to that of other insect *attacin*s. The data indicates that both *attacin* I and II genes are conserved across the taxa. *Attacin* I revealed full length cDNA of 831 bp length similar to the original cDNA sequence (accession no. S78369), while *attacin* II sequence matched partially with a length of 683 bp.

Analysis of 5' regulatory sequences of both the *attacin* genes indicated that *attacin* I and II possess

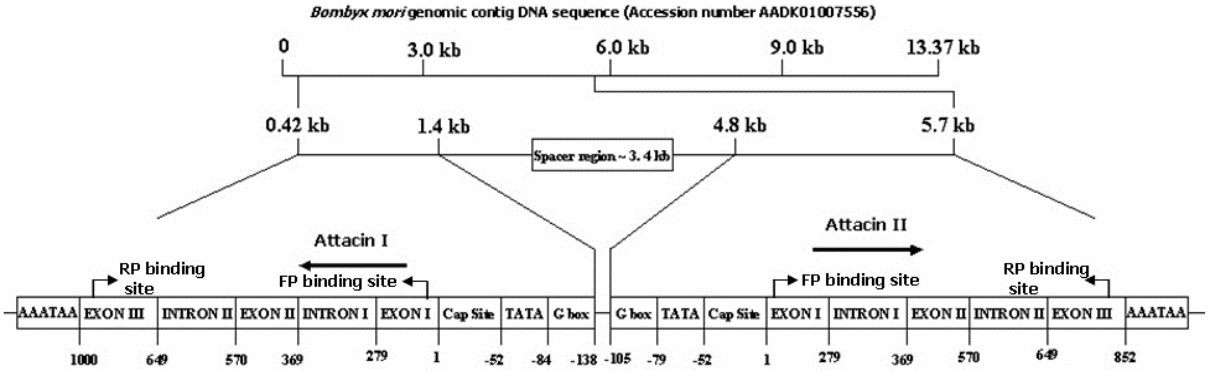

Fig. 1 The *B. mori attacin* genomic organization. The position of the genes and their transcriptional directions (→) are shown underneath. The overall structure of *attacin* I and II and the distance from the start codon (+1) to the functional parts shown in base pairs.

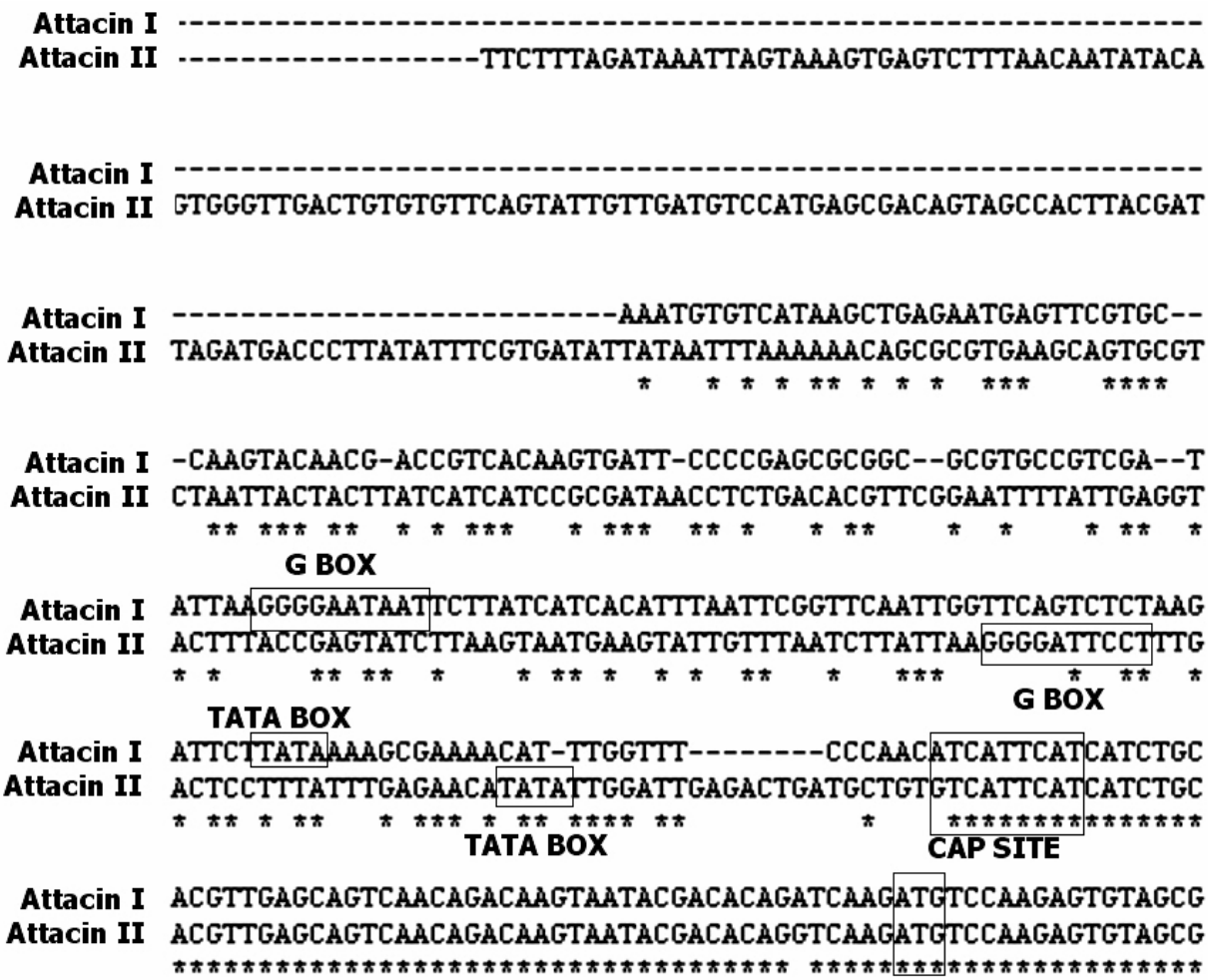

Fig. 2 Comparison of promoter regions of *attacin* I and II of silkworm *B. mori.* The G-box, TATA box and Cap site are boxed.

G-box, TATA box and cap site followed by exon/intron region. The G-box and TATA box were found to be located at two different positions on the 5' region in both the *attacin* genes. The G-box was located in 105 bp region of *attacin* I gene, whereas, it was in 138 bp region in *attacin* II gene. Similarly, the TATA box location was at 79 bp region in *attacin* I and at 84 bp region in *attacin* II of the upstream region from the starting codon. Interestingly, no difference was found in the cap site location (position at 52 bp) of both *attacin* I and *attacin* II upstream region (Fig. 2).

The *attacin* gene sequences of Pure Mysore, Daizo, Nistari, NB4D2, CSR19 and cDNA sequence (Acc. No. S78369) were compared using multiple alignment program (Clustal W). The results indicate single nucleotide variations at 22 places (Fig. 3). The sequences were clustered through phylogenetic analysis and the dendogram obtained indicated that all the multivoltine races formed a separate cluster, while, the bivoltine races formed another (Fig. 4).

Further, nucleotide variation was analyzed for possible changes in the amino acid level. These genes exhibit high levels of silent nucleotide variations (synonymous substitution), but within the silkworm races they are not excessively polymorphic at the amino acid level, as most of the nucleotide variation did not yield changes in the amino acid sequence. Among the 14-nucleotide variations observed within the amplified region, most of the changes caused no change in the amino acid sequence. The amino acid serine was coded by various degenerative codons due to which there were no changes at amino acid level. Similar findings were observed in the other amino acids such as valine, isoleucine, leucine and alanine.

EST expression in different tissues

The complete sequences of *attacin* I and II were blast searched in the silk base database and the EST sequences with maximum homology expressed in different tissues were retrieved. A total of 37 transcripts were retrieved from the EST library, of which 34 transcripts belonging to *attacin* I and 3 transcripts were that of *attacin* II. Out of 34 transcripts, *attacin* I expressed maximum (44 %) in fat body tissues (Fig. 5) followed by corpora allata (17 %). In case of *attacin* II, among the three transcripts, two were expressed in the fat body while one was from corpora allata.

```
PUREMYSORE      ACAAGAACGTACTCAGCGCCATCGGGTCCGCCGACTTCAACGACCGCCACAAGCTCAGCG
DAIZO           ACAAGAACGTACTCAGCGCCATCGGGTCCGCCGACTTCAACGACCGCCACAAGCTCAGCG
NISTARI         ACAAGAACGTACTCAGCGCCATCGGGTCCGCCGACTTCAACGACCGCCACAAGCTCAGCG
NB4D2           ACAAGAACGTGCTCAGCGCCATCGGGTCCGCCGACTTCAACGACCGCCATAAGCTCAGCG
CSR19           ACAAGAACGTGCTCAGCGCCATCGGGTCCGCCGACTTCAACGACCGCCATAAGCTCAGCG
cDNA            ACAAGAACGTGCTCAGCGCCATCGGGTCCGCCGACTTCAACGACCGCCACAAGCTCAGCG
                ********** ************************************* **********

PUREMYSORE      CCGCCTCCGCAGGGCTCGCTCTGGACAATGTGTGAGTACCTACCTACTGGCAGCTAGTAT
DAIZO           CCGCCTCCGCAGGGCTCGCTCTGGACAATGTGTGAGTACCTACCTACTGGCAGCTAGTAT
NISTARI         CCGCCTCCGCAGGGCTCGCTCTGGACAACGTGTGAGTACCTACCTACTGGCGGCTAGTAT
NB4D2           CCGCCTCTGCAGGGCTCGCTCTGGACAACGTGTGAGTACCTGCCTAGCGGCGGCTAATAT
CSR19           CCGCCTCTGCAGGGCTCGCTCTGGACAACGTGTGAGTACCTGCCTAGCGGCGGCTAATAT
cDNA            CCGCCTCCGCAGGGCTCGCTCTAGACAACGT----------------------------
                ******* ************* ***** **

PUREMYSORE      ACAATATACTGCTCATTGTGCTCCGGGCTGGTGGCTTCATGTTCGTTTGTCCATGTTGCA
DAIZO           ACAATATACTGCTCATTGTGCTCCGGGCTGGTGGCTTCATGTTCGTTTGTCCATGTTGCA
NISTARI         ACAATATACTGCTCATTGTGCTCCGGGCTGGTGGCTTCATGTTCGTTTGTCCATGTTGCA
NB4D2           ACATTATAATGCTCATTGTGCTCCGGGCTGGTGGCTTCATGCCCGTTTGTCCATGTTGCA
CSR19           ACATTATAATGCTCATTGTGCTCCGGGCTGGTGGCTTCATGCCCGTTTGTCCATGTTGCA
cDNA            ------------------------------------------------------------

PUREMYSORE      GAAACGGGCACGGGCTGAGCCTGACCGGGACCCGCATTCCCGGCTTCGGGGAGCAGCTCG
DAIZO           GAAACGGGCACGGGCTGAGCCTGACCGGGACCCGCATTCCCGGCTTCGGGGAGCAGCTCG
NISTARI         AAAACGGGCACGGGCTGAGCCTGACCGGGACCCGCATTCCCGGCTTCGGGGAGCAGCTCG
NB4D2           AAAACGGGCACGGGCTGAGCCTGACCGGGACCCGCATTCCCGGCTTCGGGGAGCAGCTCG
CSR19           AAAACGGGCACGGGCTGAGCCTGACCGGGACCCGCATTCCCGGCTTCGGGGAGCAGCTCG
cDNA            -AAACGGGCACGGGCTGAGTCTGACCGGGACCCGCATTCCCGGCTTCGGGGAGCAGCTCG
                 ***************** ****************************************

PUREMYSORE      GCGTCGCAGGCAAGGTCAACTTGTTCCACAATAACAACCACGACCTGAGCGCCAAGGCGT
DAIZO           GCGTCGCAGGCAAGGTCAACTTGTTCCACAATAACAACCACGACCTGAGCGCCAAGGCGT
NISTARI         GCGTCGCAGGCAAGGTCAACTTGTTCCACAATAACAACCACGACCTGAGCGCCAAGGCGT
NB4D2           GCGTCGCAGGCAAGGTCAACTTGTTCCACAATAACAACCACGACCTGAGCGCCAAGGCGT
CSR19           GCGTCGCAGGCAAGGTCAACTTGTTCCACAATAACAACCACGACCTGAGCGCCAAGGCGT
cDNA            GCGTCGCAGGCAAGGTCAACTTGTTCCACAATAACAACCACGACCTGAGCGCCAAGGCGT
                ************************************************************
```

Fig. 3 Multiple alignment of *attacin* gene sequences from different silkworm races of *B. mori*. The intron regions are boxed.

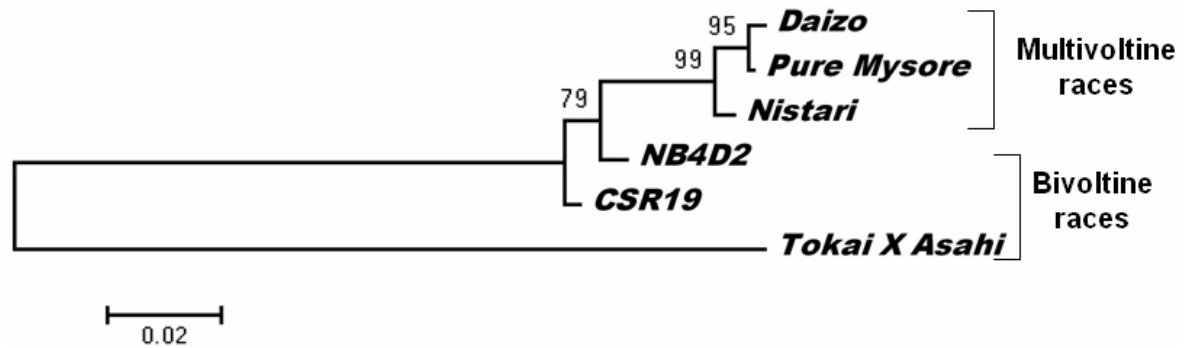

Fig. 4 The evolutional tree was obtained by the neighbor-joining method based on the multiple alignments of *attacin* DNA sequences of different silkworm strains. The numbers on each branch indicate the percentage of the most parsimonious trees, which were found in 1000 bootstrap replications performed with *MEGA5* programme.

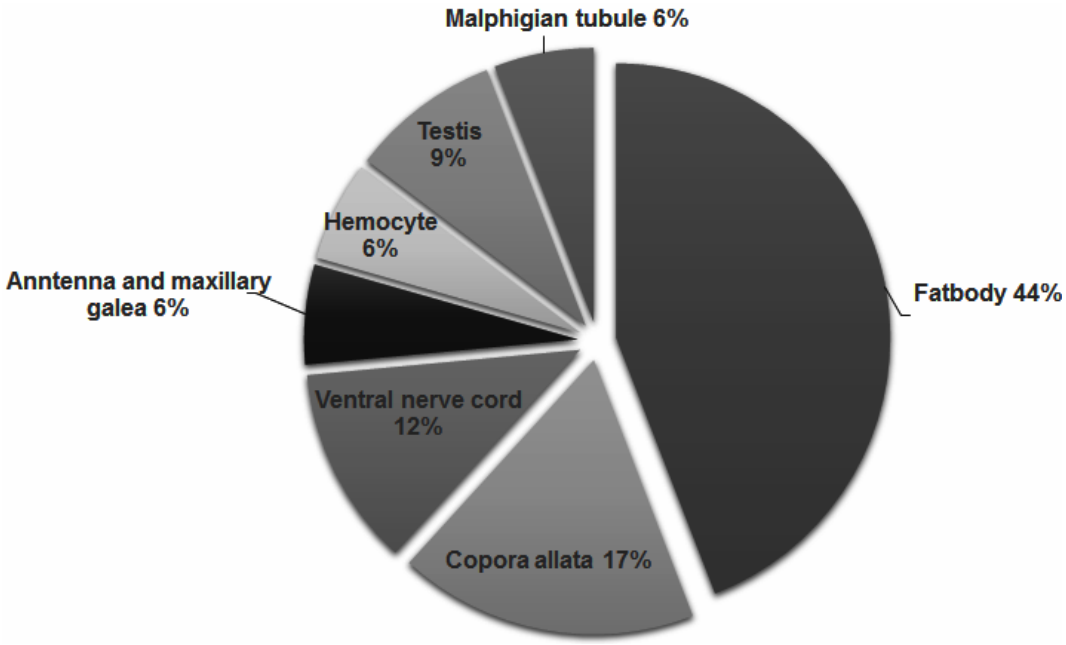

Fig. 5 The pie chart indicating the number of *attacin* I transcripts in different tissues.

Microarray experiment and data analysis

The genes associated with *Bm*NPV infection were identified in Sarupat (resistant) and CSR-2(susceptible) *B. mori* silkworm races since these races reveal divergent responses with respect to *Bm*NPV infection. Oligonucleotide microarray containing 24,924 probes were used to investigate the gene expression profiles in the midgut tissue of *Bm*NPV infected and uninfected silkworms after 12 hours post infection (hpi). Results revealed that, 735 and 589 genes were up-regulated and downregulated, respectively, at 12 hpi, in Sarupat, whereas, 2183 genes were up-regulated and 2115 down-regulated in CSR-2 (data not shown). It was observed that immune related proteins showed higher expression in *Bm*NPV infected tissues, of which *attacin* I and *attacin* II had a significantly up and down regulation in resistant and susceptible silkworm races, respectively (Fig. 6). Based on this data, it was concluded that *attacin* I was upregulated in the *Bm*NPV infected Sarupat midgut tissue, however, higher expression of *attacin* II was found in *Bm*NPV infected CSR-2 being a susceptible race. To validate the expression of these genes, the primers were designed for *attacin* I and *attacin* II for further qPCR analysis.

Tissue specific expression profile

The qPCR analysis revealed that the *attacin* I expression was higher in the fat body followed by midgut, cuticle and hemocytes (Fig. 7). The decrease in the expression in the hemocytes possibly occurred because the viral infection damaged physical functions, resulting in the reduction of the gene expression (Cheng *et al*., 2014). None of the earlier report indicate expression of *attacin* in the mid gut tissues. In the present study the significant amount of transcripts were found to be expressed in the midgut tissues. It has already been reported that the *attacin* gene is expressed in the fat body and similar findings has been observed in the present study.

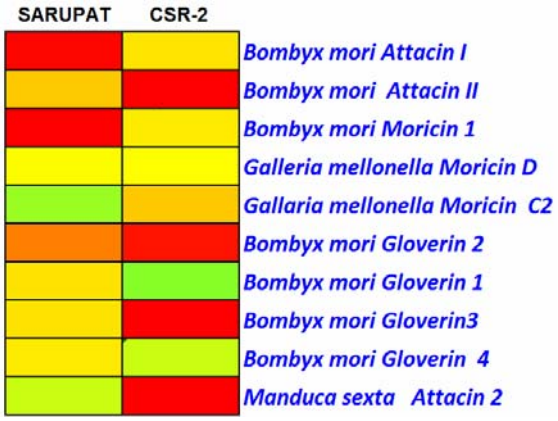

Fig. 6 Heat map of hierarchical clustering of differentially expressed genes in *Bm*NPV infected and uninfected midguts at 12 h of post infection in Sarupat and CSR-2 (clustering type: hierarchical clustering, Distance metric: Pearson correlation). The colors in the heat map display the relative values of all tiles; green indicates the lowest expression, yellow indicates the intermediate expression, and red indicates the highest expression. The numerical values give the actual values on a log 2 scale, which were associated with each color.

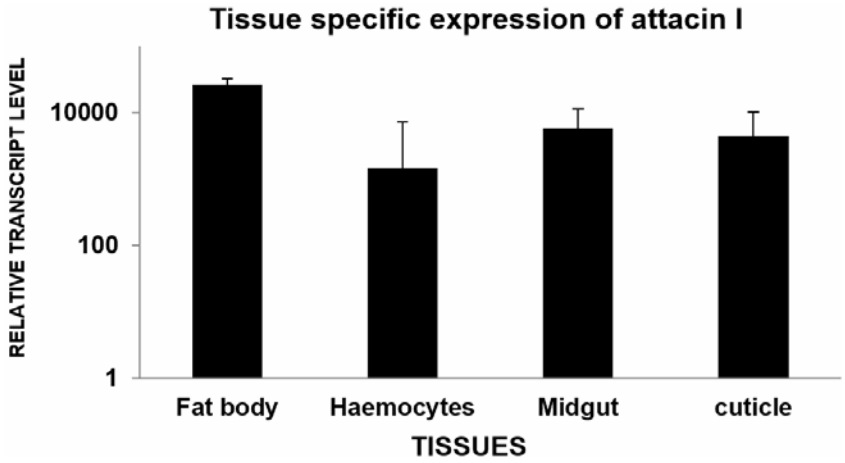

Fig. 7 Relative gene expression patterns of *attacin* I upregulated during *Bm*NPV infection in Sarupat. RNA was isolated at 12 hours post infection. The relative expression levels of each gene was normalized using the Ct values that were obtained for the housekeeping gene β actin amplifications run on the same plate. In each assay, the expression level is shown relative to the lowest expression level, which is arbitrarily set at one. All samples were tested in triplicate. The mean value ±SD was used for analysis of relative transcript levels for each time point using the ΔΔCt method. A Non-template control (NTC) sample was also run to detect contamination if any. ■- *Bm*NPV infected □-uninfected

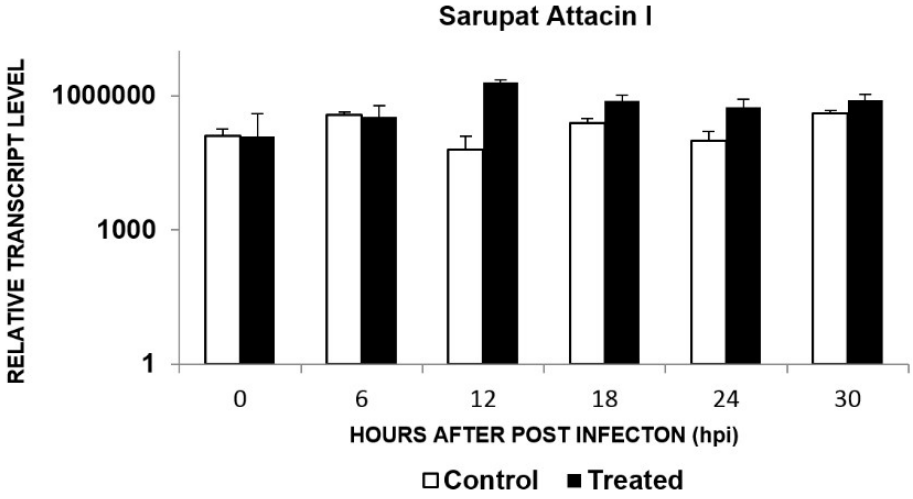

Fig. 8 Relative gene expression patterns of gene expression in *Bm*NPV infected and uninfected samples of Sarupat with *attacin* I. RNA was isolated at 6 hourly intervals from 0 to 30h. The relative expression levels of each gene at different time points were normalized using the Ct values that were obtained for the housekeeping gene β actin amplifications run on the same plate. In each assay, the expression level is shown relative to the lowest expression level, which is arbitrarily set at one. All samples were tested in triplicate. The mean value ±SD was used for analysis of relative transcript levels for each time point using the ΔΔCt method. A Non-template control (NTC) sample was also run to detect contamination if any. ■- *Bm*NPV infected □-uninfected

Quantification of attacin I and II gene expression in BmNPV infected larvae

In addition to microarray data, the expression of *attacin* I and *attacin* II in the midgut tissues was confirmed through qPCR analysis in the control as well as *Bm*NPV infected silkworm races *i.e.,* Sarupat and CSR-2 at different time intervals. In the *Bm*NPV resistant race (Sarupat), the expression of *attacin* I has gradually increased from 6 hpi to 18 hpi and then the expression was maintained steadily up to 30 hpi (Fig. 8). The expression of *attacin* II was lesser in Sarupat, when compared with *attacin* I expression, which also showed a gradual decrease in the expression up to 18 hpi (Fig. 9). In CSR-2, which is a *Bm*NPV susceptible race, the expression of *attacin* I in *Bm*NPV infected samples were found to be lesser when compared to that of the control samples (Fig. 10). On the contrary to Sarupat, the expression of *attacin* II increased from 0 h to 30 h in *Bm*NPV infected larvae of CSR-2. The highest level of expression of *attacin* II gene was observed in CSR-2 up to 18 hpi, thereafter the expression gradually decreased and steadily maintained up to 30 h of post infection (Fig. 11).

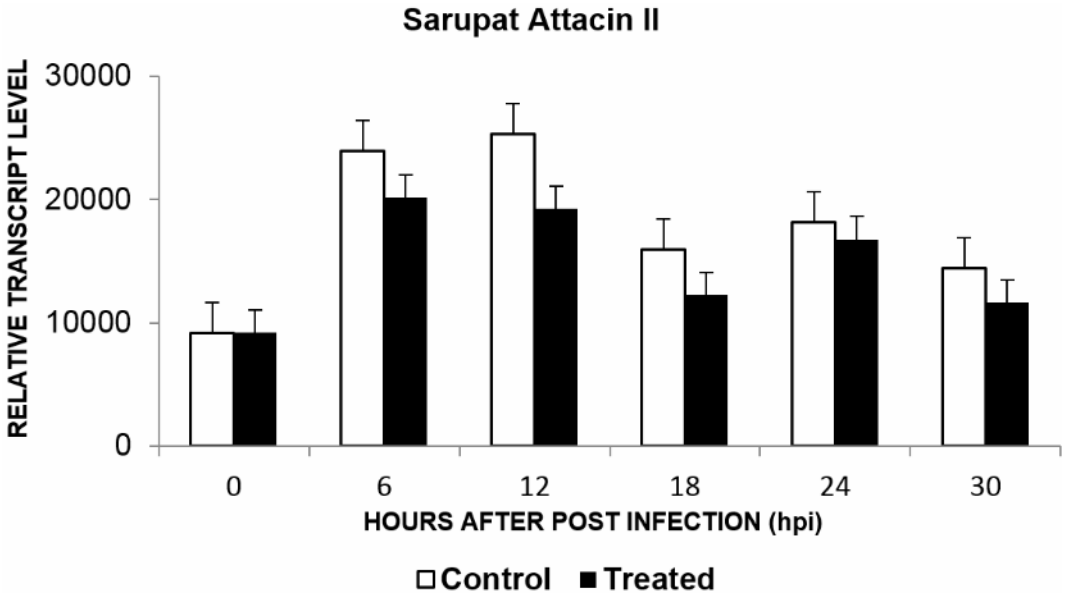

Fig. 9 Relative gene expression patterns of gene expression in *Bm*NPV infected and uninfected samples of Sarupat with *attacin* II. RNA was isolated at 6 hourly intervals from 0 to 30h. The relative expression levels of each gene at different time points were normalized using the Ct values that were obtained for the housekeeping gene β actin amplifications run on the same plate. In each assay, the expression level is shown relative to the lowest expression level, which is arbitrarily set at one. All samples were tested in triplicate. The mean value ±SD was used for analysis of relative transcript levels for each time point using the ΔΔCt method. A Non-template control (NTC) sample was also run to detect contamination if any. ■- *Bm*NPV infected □-uninfected.

Discussion

It is already known that *B. mori* genome consists of two *attacin* genes viz. *attacin* I and *attacin* II which play an anti-bacterial role. However, the genome-wide microarray analysis followed by real time PCR revealed that *attacin* gene not only functions as an anti-bacterial gene but also has an anti-viral role in silkworm *B. mori*. In order to identify the specific role against *Bm*NPV infection and to analyze differential expression of the *attacin* genes, an attempt has been made to analyze the expression profiles of these genes in *Bm*NPV infected Sarupat (*Bm*NPV resistant) and CSR-2 (*Bm*NPV susceptible) silkworm races. The up regulation of *attacin* I gene in the *Bm*NPV infected midgut samples of Sarupat indicated that *attacin* I was specifically being expressed in response to *Bm*NPV infection in the resistant race whereas the *attacin* II gene expression was comparatively lesser in Sarupat. In case of CSR-2, the expression of *attacin* I in *Bm*NPV infected samples were found to be lesser than that in the control samples. However, the expression of *attacin* II was comparatively higher in the *Bm*NPV infected samples of silkworm race CSR-2. This observation indicates that among the two *attacin* genes, *attacin* I has association with *Bm*NPV resistance.

Families of *attacin*-like peptides (usually two to four functional genes per haploid genome) have been identified in the lepidopteran species *B. mori* (Sugiyama *et al*., 1995), *H. cecropia* (Hultmark *et al.,* 1983), *Hyphantria cunea* (Shin *et al.*, 1998), *Trichoplusiani* (Kang *et al.,* 1996), and *Heliothis virescens* (Ourth *et al.,* 1994), as well as in the dipteran species *S. peregrina* (Ando *et al.,* 1987) and *D. melanogaster* (Asling *et al.,* 1995; Dushay *et al.,* 2000; Hedengren *et al.,* 2000). The antibacterial peptides are often conserved across evolutionarily distance taxa but, little is known about the level and structure of the polymorphism within different species (Choe *et al.,* 2002). Sugiyama *et al.* (1995) cloned the *attacin* gene and its 5' upstream regulatory region was characterized. In the present study it was observed that the two *attacin* genes of *B.mori* were arranged in opposite directions in a single contig with a gap of 4.2 kb length and these two genes were designated as *attacin* I and *attacin* II. These genes are transcribed in opposite directions and interrupted at homologous position by two introns. The major difference between these two *attacin* genes is the size of exon III. In *attacin* I it is 351 nt while in *attacin* II it is only 203 nt. Similar findings have also been observed in giant silk moth *H. cecropia* (Sun and Faye, 1995) indicating that the duplication of *attacin* gene as well as gene synteny is conserved within the insect taxa.

Kadalayil *et al.* (1999) showed that the promoters of several inducible insect immune genes possess GATA sites 0 - 12 bp away from NF-kappaB binding site (NF-kB site) in functionally important promoter regions. Clusters of GATA and NF-kB sites are also observed in the promoters of two important mammalian immune genes, namely IL-6 and IL-3. In *B. mori* also the nucleotide sequence of both the *attacin* gene 5'-upstream region

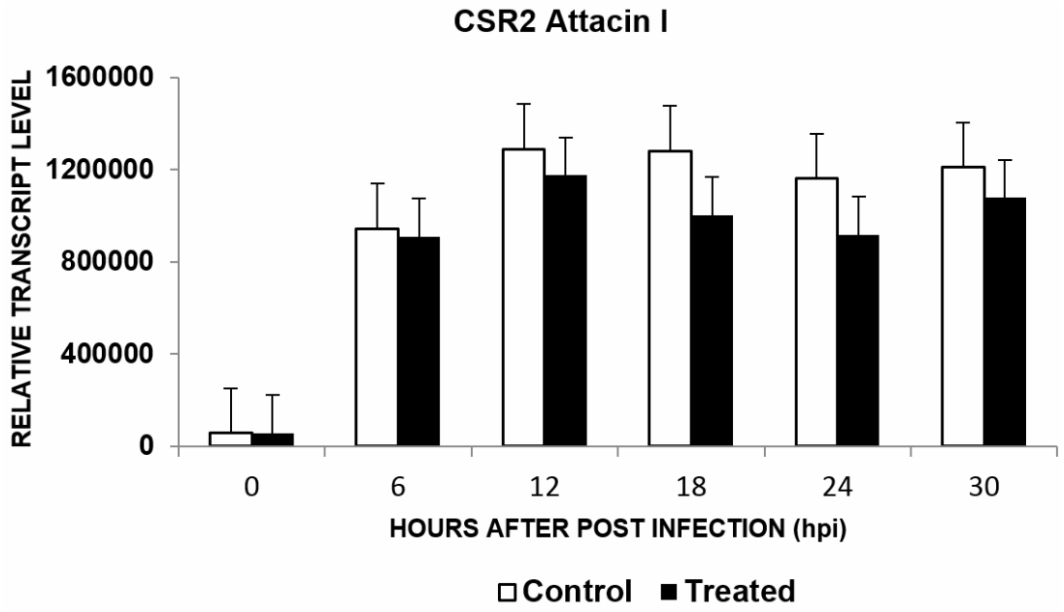

Fig. 10 Relative gene expression patterns of gene expression in *Bm*NPV infected and uninfected samples of CSR-2 with *attacin* I. RNA was isolated at 6 hourly intervals from 0 to 30 h. The relative expression levels of each gene at different time points were normalized using the Ct values that were obtained for the housekeeping gene β actin amplifications run on the same plate. In each assay, the expression level is shown relative to the lowest expression level, which is arbitrarily set at one. All samples were tested in triplicate. The mean value ±SD was used for analysis of relative transcript levels for each time point using the ΔΔCt method. A Non-template control (NTC) sample was also run to detect contamination if any. ■- *Bm*NPV infected □-uninfected.

contains a lipopolysaccharide (LPS) response element (NF-kB site), CAAT box and TATA box. The cap site of both *attacin* I and *attacin* II genes is located at a position of 52 bp in the upstream region. Tanaka *et al.* (2008) reported that the cap site is very important for the active transcription of *attacin* genes thereby indicating that both *B. mori attacin* I and II genes are actively transcribed after microbial infection. However, there is a variation in the position of different promoter elements of *attacin* genes which may affect transcription efficiency.

Lazzaro and Clark (2001) analyzed natural genetic variation in alleles of *D. melanogaster attacin*s A and B and observed that, the overall level of nucleotide diversity is high in each of these, but there is no excess of amino acid polymorphism. They also observed that, *attacin*s A and B have experienced multiple paralogous gene conversion events. Our results also revealed *attacin* gene polymorphisms at nucleotide level, but not at amino acid level, in the different silkworm strains studied.

Gloverin and *lebocin* seems to be lepidopteran-specific antibacterial peptides (Axen *et al.,* 1997) and have expression levels that were strongly induced in *B. mori* larval fat body by *Escherichia coli* immune challenge. *Bm*NPV infection also caused a strong induction of *B. mori gloverin* and *lebocin* gene expressions in larval midguts and this induction occurred in both *B. mori* strains for *gloverin*-3 and *lebocin* genes. The antiviral mechanism that occurs in the resistant *B. mori* strain is not due to resistance against the *Bm*NPV invasion but rather due to the inhibition of *Bm*NPV proliferation in the larval midgut (Bao *et al.,* 2009). The defense processes against *Bm*NPV infection that occur in the resistant larvae might be regulated via interactions involving multiple genes (Liu *et al*., 2000).

Huang *et al*. (2013) observed that antimicrobial peptides have an antiviral role in response to Alphavirus replication in arthropods. They focused their study on the antiviral response of *D. melanogaster* innate immune system induced by RNA replication of Sindbis virus (SINV). Further, they carried out microarray analysis in search for SINV replication sensitive genes. Out of the 95 SINV replication genes identified, two of the genes were found to be antimicrobial peptides viz. *attC* and *dptB*. Knocking out these genes either led to an increase in viral RNA synthesis or defects in development in the presence of SINV replication complex. These findings clearly demonstrate the antiviral role of *attacin* in *Drosophila*. Choi *et al.* (2012) also demonstrated that, genes like *attacin* were significantly up regulated during viral infection. In the present study, similar attempts have been made to identify immune response genes in *Bm*NPV infected silkworms using microarray techniques. *Attacin* I and II genes of *B.mori* were found to be differentially expressed after *Bm*NPV infection thereby proving these genes to be *Bm*NPV responsive.

Further, the work carried out by Huang *et al.* (2013) also demonstrates the phenomenon of allelic

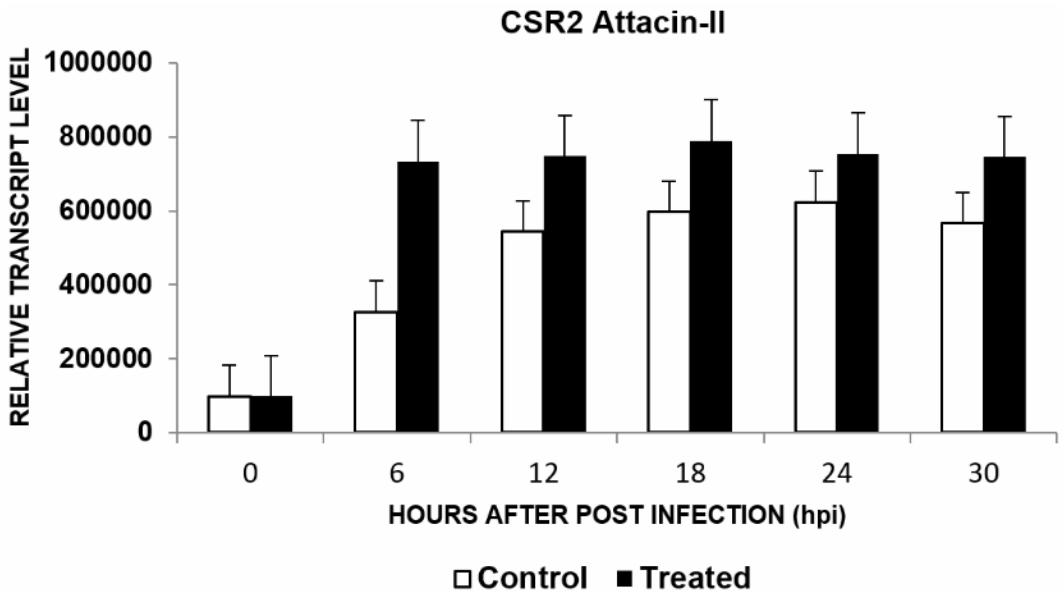

Fig. 11 Relative gene expression patterns of gene expression in *Bm*NPV infected and uninfected samples of CSR-2 with *attacin* II. RNA was isolated at 6 hourly intervals from 0 to 30 h. The relative expression levels of each gene at different time points were normalized using the Ct values that were obtained for the housekeeping gene β actin amplifications run on the same plate. In each assay, the expression level is shown relative to the lowest expression level, which is arbitrarily set at one. All samples were tested in triplicate. The mean value ± SD was used for analysis of relative transcript levels for each time point using the ΔΔCt method. A Non-template control (NTC) sample was also run to detect contamination if any. ■- *Bm*NPV infected □-uninfected.

divergence and the functional diversity of *attacin* genes. Drosophila consists of four *attacin* genes. The *attaA, attB* and *attC* genes are located on chromosome 2 while *attD* is located on chromosome 3. The AttA and AttB proteins share 98% identity with each other while AttC is 73% identical to AttA and AttB. In spite of such close similarity the expressions of *attA*, *attB* and *attD* were not found to be SINV replication sensitive. Among the four paralogous *attacin* genes only *att*C was found to be SINV replication responsive which suggests that only a single gene (*attC*) acquired the antiviral role while the other genes remained devoid of such function. In our study it was found that *attacin* I gene specifically was upregulated during *Bm*NPV infection, in the resistant race. On the contrary *attacin* II was found to be devoid of such expression.

A list of immune protein genes has been identified from the microarray analysis that is *Bm*NPV responsive and regulated by the innate immune pathways of *B. mori*. Among these genes *attacin* I was found to demonstrate an anti-viral role which otherwise has always been reported for anti-bacterial activity. Among the two reported *attacin* paralogous genes, *attacin* I acquired antiviral role which is unique when compared to its ancestral gene. However, more work is needed to determine the mode of action of *attacin* I as well as its molecular mechanism of effector molecules including upstream signaling cascades which will provide insight into the role of innate immunity response to viral infection in silkworm *B. mori*.

Acknowledgement

This study was supported by grants from the Central Silk Board, Ministry of Textiles, Government of India, Bangalore, India.

References

Ando K, Okada M, Natori S. Purification of *Sarcotoxin* II, antibacterial proteins of *Sarcophaga peregrina* (flesh fly) larvae. Biochem. 26: 226-230, 1987.

Asling B, Dushay MS, Hultmark D. Identification of early genes in the *Drosophila* immune response by PCR-based differential display: the *Attacin* A gene and the evolution of *attacin*-like proteins. Insect. Biochem. Mol. Biol. 25: 511-518, 1995.

Axen A, Carlsson A, Engstrom A, Bennich H. *Gloverin*, an antibacterial protein from the immune haemolymph of *Hyalophora* pupae. Eur. J. Biochem. 247: 614-619, 1997.

Bao YY, Tang XD, Lv ZY, Wang XY, Tian HC, Xu YP. Gene expression profiling of resistant and susceptible *Bombyx mori* strains reveals nucleopolyhedrovirus-associated variations in host gene transcript levels. Genomics, doi:10.1016/j.ygeno.2009.04.003, 2009.

Cheng Y, Wang XY, Du C, Gao J, Xu JP. Expression analysis of several antiviral related genes to *Bm*NPV in different resistant strains of Silkworm, *Bombyx mori.* J. Insect Sci. 14: 76, 2014.

Choe KM, Werner T, Stoven S, Hultmark D, Anderson KV. Requirement for a peptidoglycan recognition protein (PGRP) in Relish activation

and antibacterial immune responses in *Drosophila*. Science 296: 359-362, 2002.

Choi YY, Roh JY, Wang Y, Zhen Z, Tao XY, Lee JH, *et al*. Analysis of Genes Expression of *Spodoptera exigua* Larvae upon AcMNPV Infection. PLoS ONE 7: e42462, 2012.

Dushay MS, Roethele JB, Chaverri JM, Dulek DE, Syed SK, Kitami T, *et al.* Two *attacin* antibacterial genes of *Drosophila melanogaster*. Gene 246: 49-57, 2000.

Gunne H, Hellers M, Steiner H. Structure of pre pro*attacin* and its processing in insect cells infected with a recombinant baculovirus. Eur. J. Biochem. 187: 699-703, 1990.

Hedengren M, Borge K, Hultmark D. Expression and evolution of the *Drosophila attacin*/diptericin gene family. Biochem. Biophys. Res. Commun. 279: 574-581, 2000.

Hoffmann JA, Kafatos FC, Janeway CA, Ezekowitz RAB. Phylogenetic perspectives in innate immunity. Science 284: 1313-1318, 1999.

Huang Z, Kingsolver MB, Avadhanula V, Hardy RW. An Antiviral Role for Antimicrobial Peptides during the Arthropod Response to Alphavirus Replication. J. Virol. 87: 4272-4280, 2013.

Hultmark D, Engstrom A, Anderson K, Steiner H, Bennich H, Boman HG. Insect immunity. *Attacin*s, a family of antibacterial proteins from *Hyalophora cecropia.* EMBO J. 2: 571-576, 1983.

Iwanaga S, Lee BL. Recent Advances in the Innate Immunity of Invertebrate Animals. J. Biochem. Mol. Biol. 38: 128-150, 2005.

Kadalayil L, Petersen UM, Engstrom Y. Adjacent GATA and [kappa] B-like motifs regulate the expression of a *Drosophila* immune gene. Nucleic Acids Res. 25: 1233-1239, 1999.

Kang D, Liu G, Gunne H, Steiner H. PCR differential display of immune gene expression in *Tricoplusiani.* Insect Biochem. Mol. Biol. 26: 177-184, 1996.

Lazzaro BP, Clark AG. Evidence for recurrent paralogous gene conversion and exceptional allelic divergence in the *attacin* genes of *Drosophila melanogaster*. Genetics 159: 659-671, 2001.

Liu GX, Kang G, Steiner H. *Trichoplusia* in *lebocin*, an inducible immune gene with a downstream insertion element. Biochem. Biophys. Res. Commun. 269: 803-807, 2000.

Nagaraja GM, Nagaraju J. Genome fingerprinting of the silkworm *Bombyx mori* using random arbitrary primers. Electrophoresis 16: 1633-1638, 1995.

Ourth DD, Lockey TD, Renis HE. Induction of *cercopin* like and *attacin* like antibacterial but not antiviral activity in *Heliothis virescenes* larvae. Biochem. Biophys. Res. Commun. 200: 34-44, 1994.

Sagisaka A, Fujita K, Nakamura Y, Ishibashi J, Noda H, Imanishi S. Genome-wide analysis of host gene expression in the silkworm cells infected with *Bombyx mori* nucleopolyhedrovirus. Virus Res. 147: 166-175, 2010.

Shin SW, Park SS, Park DS, Kim MG, Kim SC. Isolation and characterization of immune-related genes from the fall webworm, *Hyphantria cunea*, using PCR based differential display and subtractive cloning. Insect. Biochem. Mol. Biol. 28: 827-837, 1998.

Sugiyama M, Kuniyoshi H, Kotani E, Taniai K, Kadono-Okuda K, Kato Y, *et al*. Characterization of a *Bombyx mori* cDNA encoding a novel member of the *attacin* family of insect antibacterial proteins. Insect Biochem. Mol. Biol. 25: 385-392, 1995.

Sun SC, Faye I. Transcription of immune genes in the giant silk moth, *Hyalophora cecropia*, is augmented by H_2O_2 and diminished by thiol reagents. Eur. J. Biochem. 231: 93-98, 1995.

Tanaka H, Ishibashi J, Fujita K, Tomimoto K. A genome wide analysis of genes and gene families involved in innate immunity of *Bombyx mori*. Insect Biochem. Mol. Biol. 38: 1087-1110, 2008.

Taniai K, Ishii T, Sugiyama M, Miyanoshita A, Yamakawa M. Nucleotide sequence of 5'-upstream region and expression of a silkworm gene encoding a new member of the *attacin* family. Biochem. Biophys. Res. Commun. 220: 594-599, 1996.

Wheelan SJ, Church MD, Ostell JM. Spidey: A Tool for mRNA-to-Genomic Alignments. *Genome Research online:* http://www.genome.org/cgi/content/full/GR-1953Rv1#otherarticles, 2001.

Xia Q, Zhou Z, Lu C, Cheng D, Dai F, Li B, *et al*. A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). Science 306: 1937-1940, 2004.