



Journal of Experimental Biology and Agricultural Sciences

<http://www.jebas.org>

ISSN No. 2320 – 8694

COMPUTATIONAL ANALYSIS OF STRESS RESPONSE GENE AND OTHER RELATED GENES IN *Acyrtosiphon pisum* (PEA APHID)

Manaj Kumar Barman¹, Archana Kumari², Jatin Sarmah^{3*}

¹Manaj Kumar Barman, DBT (Govt. of India) sponsored Bioinformatics Infrastructure Facility, Bodoland University, Kokrajhar, Assam.

²Dr. Archana Kumari, RA, DBT (Govt. of India) sponsored Bioinformatics Infrastructure Facility, Bodoland University, Kokrajhar, Assam.

³Dr. Jatin Sarmah, Associate Professor, Department of Biotechnology and Coordinator, DBT (Govt. of India) sponsored Bioinformatics Infrastructure Facility, Bodoland University, Kokrajhar, Assam, India, PIN-783370

Received – June 27, 2018; Revision – August 22, 2018; Accepted – September 13, 2018

Available Online – October 5, 2018

DOI: [http://dx.doi.org/10.18006/2018.6\(5\).792.799](http://dx.doi.org/10.18006/2018.6(5).792.799)

KEYWORDS

Acyrtosiphon pisum

Stress response

EST

Transcriptome

Immune defense

Computational analysis

ABSTRACT

Acyrtosiphon pisum is a major crop pest found worldwide. It causes significant damage to legumes. The *A. pisum* is a model organism for biological studies. The aim of the work is to develop effective control methods for this pest species. A computational approach to identify genes for stress response, immune response, recognition and signaling was performed. The four unassembled ESTs libraries namely antennae, head, digestive tract and whole body were taken for the study. The functional annotation was done for four EST libraries by Blast2GO software. Local BLAST was performed with the assembled transcriptome sequence (GenBank: SRX316762) for gene expression analysis. The complete ORF of screened were predicted in ORF finder and validated by SmartBlast tool. Further, the proteins were characterized using several tools and servers: signal peptide by SignalP 4.0, physicochemical properties by the ProtParam tool, sub-cellular localization by ProtComp V9. The analysis revealed 26 genes expressed during microbial infection. Genes were further categorized into four ontologies: stress response, immune response, recognition and signaling. The signal peptide was present in 9 proteins. The aliphatic index, excitation index, instability index, theoretical points, molecular weight, amino acids residues, positively and negatively charged residues and grand average of hydrophobicity were the physicochemical properties studied in the proteins. From subcellular localization cytoplasmic, nuclear, mitochondrial, endoplasmic reticulum, plasma membrane and extracellular proteins were obtained.

* Corresponding author

E-mail: jatinsarmahindia@gmail.com (Dr. Jatin Sarmah)

Peer review under responsibility of Journal of Experimental Biology and Agricultural Sciences.

Production and Hosting by Horizon Publisher India [HPI]
(<http://www.horizonpublisherindia.in/>).
All rights reserved.

All the article published by Journal of Experimental Biology and Agricultural Sciences is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License Based on a work at www.jebas.org.



1 Introduction

Acyrtosiphon pisum (commonly called as pea aphid) is a sap-sucking insect of aphididae family. It is a model organism for biological study with sequenced genome. It has agronomical importance as it feeds on several species of legumes worldwide. The aphid is 4mm in length in adult stage. They ingest the phloem sap from the lower side of buds, leaves and pods of legumes and thus cause severe economic loss. *A. pisum* also act as a vector for many viral diseases of plants. The use of predators, parasitoids and chemical insecticides are applied to protect the plants against the pea aphids. The species reproduce by viviparous parthenogenesis. Each adult female gives birth to 4-12 female nymphs per day. They develop to mature females in 8-10 days and lifespan is 30 days.

Many aphid species including *A. pisum* causes serious physical and economic damage to agricultural crops of legumes family found in North-Eastern states of India. On many crops, insecticides provide simple solution for aphid control. The large-scale application of such insecticides is becoming increasingly unacceptable, however, and their large-scale application of pesticides needs to optimize in an environmentally acceptable way to maintain both farm income and adequate food supply. It is also important that the aphid species (more than 20) have developed resistance population against most of insecticides (Dixon, 1998; Javed et al., 2003). The use of plant verities resistance aphid species is an alternative to chemical control. Again, aphids have developed biotypes that able to overcome the few sources of aphid resistance in plants (Anstead et al., 2005). The biotype properties induced by its parasitoids *Aphidus ervi* (Trotta et al., 2018). Therefore, a number of studies have aimed to develop effective control methods for this pest species using different gene ontology such as stress response, immune etc.

Presently, great effort has been put to understand the mechanisms due to which the species is able to exploit different plant species by having ancestries adapted to many host plants which respond in different ways to chemical and biological control agents (Nagoshi et al., 2007; Machado et al., 2008; Groot et al., 2008; Barros et al., 2010; Carvalho et al., 2013; Jakka et al., 2014). This was possible after EST collection, which was selected in 2005 by the National Human Genome Research Institute for sequencing (Sabater-Munoz et al., 2006).

In pest control approaches, a suitable approach to identify genes with differential expression in different tissues or developmental stages is construction of cDNA libraries (Duan et al., 2013; Lenz et al., 2013; Robert et al., 2013; Kim et al., 2014; Nanoth-Vellichirammal et al., 2014; Benoit et al., 2014; Sinha et al., 2014).

Due to paucity of sufficient information from *in vivo* works, the *in silico* information acquired from transcriptome mining becomes handy for identification and characterization of proteins of interest. The present *in silico* work aims to identify the responsible genes for stress response, immune response, recognition and signaling response of *A. pisum* through comparison of expressed sequence tags (ESTs) from tissue specific one assembled reference library and four unassembled library to characterize their gene functions and sub-cellular locations. It is expected that the derived information shall be helpful in understanding defense mechanism of *A. pisum* as well as in formulating biological control mechanisms of the pest.

2 Materials and Methods

2.1 Retrieval of expressed sequences tag (EST) and transcriptome

The identification of REPAT genes includes stress response, signaling, recognition and immune response was based on the analysis of (SRX031648) assembled sequences downloaded from galaxy (<https://usegalaxy.org/>).

Additionally, four EST unassembled libraries were downloaded from NCBI (<https://www.ncbi.nlm.nih.gov/sites/batchentrez>). These four EST libraries were: ApAL3SD & ID0AEE (antennae), ApDT& ApHL3LD (digestive tract), ApHL3SD & ID0ACC (head), and ApMS;14419;14436(whole body, multistage). The libraries were assembled with SOAP2novo with the default settings available in the DDBJ website (Nagasaki et al., 2013).

2.2 Functional annotation of assembled Sequence and local blast with the transcriptome

The annotation of the assembled EST was predicted by using Blast2go software (Conesa et al. 2005). Blast2GO is an all in one bioinformatics solution for functional annotation of (novel) sequences and the analysis of annotation data.

A database was created from assembled transcriptome SRX031648 by using Blast2GO software. The annotated ESTs were local Blast with the assembled transcriptome database to detect the presence of stress and other related responses genes. The identification of stress response and other related response genes in *A. pisum* was based on homology analysis using Blastx searches against the SRX031648 assembled EST library, with an e-value <10⁻⁵ and other default parameters.

The statistical test suggested by Romualdi et al. (2001) was followed to analyze the frequency of the stress and other ESTs in each of the unassembled libraries as an indicator of their indicator level. The blast search was done by taking 1e⁻³⁰. The sequences with two or more matches were eliminated from the results.

Finally the best scores were selected. Thus the differences in gene expression in each library were evaluated and hits with similarity score of more than 95% were further considered.

2.3 ORF analysis of the detected (REPAT) stress response genes and other related genes

The ORF analysis of all the detected sequences was done by using ORF finder tool with significant Blast results (e-value $<10^{-5}$) to detect the presence of complete open reading frames (ORF). The sequences with the best e-value and high similarity with stress response and other related gene sequences were selected for the analysis.

2.4 Analysis of signal peptide for the stress response and other related proteins (REPAT)

The presence/absence of signal peptide was predicted by using the SignalP 4.0 server (<http://www.cbs.dtu.dk/services/SignalP/>).

2.5 Subcellular localization of detected 26 proteins

The potential subcellular localization of the 26 proteins was predicted using ProtComp Version 9.0 (<http://www.softberry.com>) in *A. pisum*.

2.6 Physico-chemical properties of the selected proteins

The physico-chemical properties of the 26 proteins predicted in *A. pisum* were calculated by ProtParam tool (<https://web.expasy.org/protparam/>). The parameters calculated were theoretical isoelectric point, extinction coefficients composition of amino acids, positively and negatively charged residues, aliphatic index, instability index and molecular weight.

3 Results

3.1 Molecular information and Blastx results of *A. pisum* REPAT genes

The blastx analysis showed the presence of 26 REPAT proteins in the cDNA library of *Acyrtosiphon pisum* (Table 1). Genes were distributed in different tissues; 6 REPAT members were found in antennae, 18 REPAT members in head, 4 REPAT members in digestive tract which is considered as primary site for pathogen entry and 24 REPAT members were found in non treated whole-body multistage.

The similarity with homologous REPAT genes found in *M. persicae* ranged from 68% to 99%. The similarity was higher than 95% in seven genes. In REPAT 745 and REPAT 1080 lowest similarity values were found. The number of amino acid residues present in *A. pisum* REPAT proteins were found to be same as those found in homologous in 17 cases. These are REPAT 192,

REPAT 509, REPAT 742, REPAT 898, REPAT 997, REPAT 1007, REPAT 1071, REPAT 1236, REPAT 1452, REPAT 1828, REPAT 2575, REPAT 3244, REPAT 3247, REPAT 4015 and REPAT 4134. There were differences of one residue in REPAT 2315 and two residues in REPAT 1045. Differences with more than four residues in the homologue of *M. persicae* were found in REPAT 145, REPAT 1080, REPAT 1561, REPAT 2688, REPAT 3811 and REPAT 3815.

3.2 Functional annotation and EST presence of *Acyrtosiphon pisum* REPAT proteins

Functional annotation of the 26 REPAT proteins was predicted by Blast2GO software (Conesa et al. 2005) to detect the presence of stress and other related responses genes in *A. pisum*. The analysis of Blast2GO results showed six proteins were detected in antennae, four proteins were detected in digestive tract and more than 15 proteins were present in two tissue libraries (Table 2). The frequency of the EST in each library showed high expression of some genes in different libraries. According to the number of ESTs found in each library, it was inferred that some genes were highly expressed: 20 hits in REPAT 1080 in whole-body multistage, 12 hits in REPAT 898 in head, 6 hits in REPAT 1007 in the digestive tract and 3 in antennae.

3.3 ORF analysis of the selected proteins of *A. pisum*

From the ORF finder analysis, we get the results of only 26 of these exhibited complete ORFs. Incomplete ORFs were found REPAT 65, REPAT 279, REPAT 1852, REPAT 2647, REPAT 3896, REPAT 4000 and REPAT 4152 of the *A. pisum* library. The complete ORF consist both initiation codon and termination codon but not in incomplete ORF.

3.4 Analysis of signal peptide for the stress response and other related proteins (REPAT)

The presence/absence of signal peptide was predicted by using the SignalP 4.0 server (<http://www.cbs.dtu.dk/services/SignalP/>). From this analysis we get 9 signal peptides present of our screened REPAT proteins shown in table 1.

3.5 Sub-cellular localization and physico-chemical properties of the predicted REPAT proteins

The subcellular localization of the REPAT proteins in *A. pisum* was predicted by using online tool ProtComp. The ProtComp analysis showed that all 26 proteins has the different functional category (Extracellular membrane bound, extracellular, nuclear, plasma membrane and Cytoplasmic) and can be associated with four ontology categories, namely, stress response (REPAT 145, REPAT 997, REPAT 1007, REPAT 1071, REPAT 1452, REPAT 1561, REPAT 2315, REPAT 2688, REPAT 3244, REPAT 3247,

Table 1 Molecular information and Blastx results of *A. pisum* REPAT genes

Gene	cDNA	Amino acids number <i>A. pisum</i>	Amino acids number homologous	Peptide signal	E-value	Similarity	Species
REPAT145	441	146	250	No	1.00E ⁻¹¹⁰	99%	<i>Myzus persicae</i> XP_022181545.1
REPAT192	471	156	156	Yes	2.00E ⁻¹¹⁹	99%	<i>Myzus persicae</i> XP_022168136.1
REPAT509	480	159	159	No	2.00E ⁻¹¹⁸	96%	<i>Myzus persicae</i> XP_022160420.1
REPAT650	633	220	220	No	3.00E ⁻¹⁶⁹	99%	<i>Diuraphis noxia</i> XP_01537959.1
REPAT742	474	157	157	Yes	2.00E ⁻⁸²	68%	<i>Myzus persicae</i> XP_022182500.1
REPAT874	468	155	194	Yes	1.00E ⁻⁸⁴	86%	<i>Diuraphis nixia</i> XP_015365446.1
REPAT898	738	245	245	Yes	0	97%	<i>Myzus persicae</i> XP_022167308.1
REPAT983	471	156	156	No	2.00E ⁻¹⁰⁹	94%	<i>Myzus persicae</i> XP_022169776.1
REPAT997	543	180	180	No	2.00E ⁻¹³²	97%	<i>Myzus persicae</i> XP_022168531.1
REPAT1007	819	272	272	Yes	0	93%	<i>Myzus persicae</i> XP_022177449.1
REPAT1045	465	154	156	No	4.00E ⁻¹⁰⁵	90%	<i>Myzus persicae</i> XP_022168261.1
REPAT1071	549	182	182	No	6.00E ⁻¹²⁰	88%	<i>Myzus persicae</i> XP_022180361.1
REPAT1080	546	181	161	No	3.00E ⁻⁸²	68%	<i>Myzus persicae</i> XP_022168131.1
REPAT1236	714	237	237	Yes	1.00E ⁻¹⁷⁸	97%	<i>Myzus persicae</i> XP_022177860.1
REPAT1452	381	126	126	No	8.00E ⁻⁸⁸	91%	<i>Diuraphis noxia</i> XP_015380029.1
REPAT1561	540	179	217	Yes	2.00E ⁻¹³⁰	99%	<i>Myzus persicae</i> XP_022172142.1
REPAT1828	726	241	241	No	3.00E ⁻¹⁶⁶	95%	<i>Myzus persicae</i> XP_022179937.1
REPAT2315	594	197	198	No	7.00E ⁻¹⁴¹	94%	<i>Myzus persicae</i> XP_022178118.1
REPAT2575	225	74	74	No	2.00E ⁻⁵⁴	95%	<i>Myzus persicae</i> XP_022160949.1
REPAT2688	651	216	231	No	4.00E ⁻¹⁶⁴	97%	<i>Myzus persicae</i> XP_022167065.1
REPAT3244	585	194	194	No	5.00E ⁻¹⁴⁵	99%	<i>Diuraphis noxia</i> XP_015370481.1
REPAT3247	648	215	215	Yes	4.00E ⁻¹⁶³	99%	<i>Diuraphis noxia</i> XP_015371593.1
REPAT3811	732	243	232	Yes	3.00E ⁻¹⁶⁷	94%	<i>Diuraphis pisum</i> XP_015364984.1
REPAT3815	720	239	294	No	4.00E ⁻¹⁶⁶	91%	<i>Myzus persicae</i> XP_022163681.1
REPAT4015	327	108	108	No	1.00E ⁻⁷⁷	96%	<i>Diuraphis noxia</i> XP_015371631.1
REPAT4134	306	101	101	No	1.00E ⁻⁷⁰	97%	<i>Myzus persicae</i> XP_022167054.1

REPAT 3815, REPAT 4015 and REPAT 4134), immune response (REPAT 192, REPAT 509, REPAT 898, REPAT 983, REPAT 1045, REPAT 1080, REPAT 1236 and REPAT 3811), signaling (REPAT 650, REPAT 874, REPAT 1828 and REPAT 2575) and recognition (REPAT 742). The stress response category was found in libraries from antennae, digestive tract, head and multiple stages of whole-body. The same functional and ontological categories were found for homologous REPAT proteins from *M. persicae*. Excepts REPAT 509 and REPAT 2315 were classified as secretory proteins by ProtComp.

The physico-chemical properties of REPAT proteins of *A. pisum* obtained by ProtParam tool are shown in Table 3. This illustrates the physico-chemical properties of stress response, immune response, signaling and recognition genes responsible for *A. pisum*. The result explains that the molecular weight is less than 30 Daltons in the proteins. The theoretical PIs indicate the presence of 7 basic proteins (509, 874, 1045, 1080, 1828, 2315 and 4015) and 19 acidic proteins (REPAT 145, 192, 650, 742, 898, 983, 997, 1007, 1071, 1236, 1452, 1561, 2575, 2688, 3244, 3247, 3815 and 4134). The proteins namely, REPAT 192, 509,

Table 2 Functional annotation and EST presence of *A. pisum* REPAT proteins

Protein	Functional category	Gene Ontology category (GO)	Enzyme	Gene Ontology Category (GO)	Antennae	Head	Digestive tract	Whole-body, multistage
REPAT145	Extracellular	Stress_response	No	Stress		X(1)		X(1)
REPAT192	Membrane bound Extracellular	Immune_response	No	Immune		X(7)		X(6)
REPAT509	Nuclear	Immune_response	No	Immune		X(1)		X(11)
REPAT650	Plasma membrane	Signaling_response	No	Signaling		X(1)		X(4)
REPAT742	Extracellular	Recognition_response	No	Recognition				X(4)
REPAT874	Extracellular	Signaling_response	Yes	Signaling		X(2)		X(5)
REPAT898	Extracellular	Immune_response	Yes	Immune	X(2)	X(12)		X(6)
REPAT983	Cytoplasmic	Immune_response	Yes	Immune		X(9)		X(11)
REPAT997	Extracellular	Stress_response	No	Stress				X(3)
REPAT1007	Plasma membrane	Stress_response	Yes	Stress			X(6)	
REPAT1045	Endoplasmic reticulum	Immune_response	No	Immune	X(1)	X(1)		X(10)
REPAT1071	Cytoplasmic	Stress_response	Yes	Stress				X(7)
REPAT1080	Extracellular	Immune_response	No	Immune				X(20)
REPAT1236	Extracellular	Immune_response	No	Immune	X(1)			X(5)
REPAT1452	Extracellular	Stress_response	Yes	Stress		X(6)	X(1)	X(8)
REPAT1561	Extracellular	Stress_response	Yes	Stress		X(3)		X(1)
REPAT1828	Endoplasmic reticulum	Signaling_response	Yes	Signaling				X(2)
REPAT2315	Nuclear	Stress_response	No	Stress	X(1)			X(2)
REPAT2575	Membrane bound Extracellular	Signaling_response	No	Signaling		X(1)		X(1)
REPAT2688	Cytoplasmic	Stress_response	Yes	Stress		X(2)	X(1)	X(2)
REPAT3244	Cytoplasmic	Stress_response	Yes	Stress	X(1)	X(2)		X(2)
REPAT3247	Membrane bound Mitochondrial	Stress_response	No	Stress		X(2)		X(2)
REPAT3811	Extracellular	Immune_response	Yes	Immune	X(3)	X(11)		X(4)
REPAT3815	Cytoplasmic	Stress_response	Yes	Stress		X(3)		
REPAT4015	Extracellular	Stress_response	Yes	Stress		X(1)		X(4)
REPAT4134	Extracellular	Stress_response	Yes	Stress		X(8)	X(2)	X(4)

874, 898, 997, 1007, 1045, 1080, 1452, 2575, 3244, 3247 and 3815 are found to be instable from the results of instability index. Almost all values obtained in computational physico-chemical analysis are similar to those found in homologous REPAT from *M. persicae*. The instability index below 40 were considered as stable and more than 40 were unstable.

4. Discussion and Conclusion

The genes described in the study based on features like similarity of nucleotides, numbers of amino acid residues and mol. wt. of the

amino acids as illustrated by Navarro-Cerrillo et al. (2013); Machado et al. (2016). Further, the proteins have same functional and ontological categories found in their homologous.

The study of the REPAT proteins is related to the tissue where the main expression and specific function occur. Earlier studies have shown the importance of REPAT proteins in immune response during infection by pathogens (Herrero et al., 2007; Hernandez-Rodriguez et al., 2009; Navarro-Cerrillo et al., 2012; Navarro-Cerrillo et al., 2013). However, we have detected 26 REPAT proteins from EST libraries of head, antennae, digestive tract and

Table 3 Physicochemical properties of REPAT proteins

Gene	Amino acids residues	Molecular weight (kDa)	Theoretical pI	Negatively charged residues (Asp+Glu)	Positively charged residues (Arg+Lys)	Extinction coefficients	Instability Index	Aliphatic Index	Grand average of hydrophobicity
REPAT145	146	16976.5	6.83	21	21	22588	39.91	86.1	-0.582
REPAT192	156	17835.26	4.75	20	14	25160	44.19	74.94	-0.202
REPAT509	159	18366.14	8.63	20	23	32680	51.49	70.57	-0.635
REPAT650	220	25139.25	5.31	34	27	38055	34.88	70.45	-0.53
REPAT742	157	17986.83	5.78	16	15	23420	24.99	80.13	-0.062
REPAT874	155	17341.77	9.34	9	13	13075	73.4	87.29	-0.39
REPAT898	245	27446.6	5.07	34	25	24045	45.86	52.08	-0.768
REPAT983	156	16619.09	5.12	18	15	8730	23.96	95.58	-0.012
REPAT997	180	20660.36	5.45	26	22	14940	59.8	68.78	-0.784
REPAT1007	272	30653.98	5.03	39	27	53190	40.54	92.46	-0.176
REPAT1045	154	16741.54	9.45	7	13	17545	40.23	103.31	0.321
REPAT1071	182	20223.32	4.99	28	18	7575	36.93	121.04	-0.227
REPAT1080	181	19848.99	7.45	17	18	18295	45.01	78.18	-0.017
REPAT1236	273	26735.08	6.1	25	23	53900	31.34	67.89	-0.459
REPAT1452	126	13818.91	6.71	11	11	6085	60.69	92.78	-0.148
REPAT1561	179	18357.6	6.48	13	11	1615	23.67	89.27	0.075
REPAT1828	241	27369.1	8.82	29	34	24200	26.35	112.32	-0.014
REPAT2315	197	22232.63	8.88	24	29	27180	32.67	90	-0.338
REPAT2575	74	8956.28	6.7	13	13	13200	48.99	47.43	-0.942
REPAT2688	216	24494.62	5.73	28	22	31985	33.49	80.42	-0.483
REPAT3244	194	21862.6	5.72	31	26	9970	43.76	74.85	-0.685
REPAT3247	215	24096.67	5.92	25	24	25565	50.49	67.12	-0.453
REPAT3811	244	26837.66	4.61	36	20	28515	37.8	53.11	-0.638
REPAT3815	239	27431.91	6.13	31	29	38390	44.09	88.58	-0.185
REPAT4015	108	12765.7	9.64	10	17	28420	35.87	74.91	-0.608
REPAT4134	101	10968.63	6.25	15	15	2980	36.5	98.42	-0.219

multistage of whole-body, showing different expression patterns based on EST frequency. The presence of these proteins in tissues other than midgut was mentioned previously (Navarro-Cerrillo et al., 2013). The presence of some specific REPAT proteins in head, antennae, digestive tract and whole-body multistage indicates a more widespread action of these proteins in *A. pisum*, which are likely involved in various processes, such as stress response, signaling and immune function, as indicated by

ProtParam analysis. The signaling system plays major roles in stress response (Patel & Chu, 2011; Li et al., 2018).

High expression level of genes related to immunity, detoxification, stress and maintenance of cell structure during analysis of the libraries (Schafer & Buettner, 2001; Bravo et al., 2011). The possibility of the existence of REPAT proteins with different ontological functions, based on differences in their

subcellular localization, expression pattern and interactions were forwarded by Navarro-Cerrillo et al. (2012).

The EST frequencies in the analyzed libraries may indicate high expression levels of REPAT 898, REPAT 938 and REPAT1080. These are involved in immune response in whole-body, multistage. Moreover, REPAT 1007 may be associated with stress response, as it showed high expression in digestive tract of larvae submitted to stress conditions and in the case of REPAT 1452 related to stress response that showed a high level of expression in whole-body, multistage.

The expression patterns indicate the existence of differences in expression rates of REPAT proteins in different tissues under different conditions. However, the interpretation of these results is difficult at the moment, as a library from whole-body, multistage of *A. pisum* larvae infected with bacteria or viruses is unavailable. Further, our analyses found the presence of 22 REPAT genes in the EST libraries of *A. pisum*. These results might be due to the fact that most REPAT proteins are over expressed in whole-body, multistage but only as a response to infection (Herrero et al., 2007; Hernandez-Rodriguez et al., 2009; Navarro-Cerrillo et al., 2013; Machado et al., 2016).

The present work is the first studies of REPAT family genes in *A. pisum* using bioinformatics approach based on transcriptome analysis using four EST libraries. Stress response and other related genes have an essential role in control of their development on crops. The analysis of the present work will help to increase the understanding of the genetic bases of stress response genes. This analysis is a key for further experiment about these pest species is pest management. Ultimately, it will increase the opportunity to develop effective control methods for this pest species that may have great importance in the field of agriculture and food production.

Acknowledgement

We gratefully acknowledge the Department of Biotechnology, Government of India sponsored Bioinformatics Infrastructure Facility at Bodoland University for offering the infrastructure for the work. All authors have contributed equally to this work.

Conflicts of interest: Nil

Reference

Anstead JA, Williamson MS, Denholm I (2005) Evidence for multiple origins of identical insecticide resistance mutations in the aphid *Myzus persicae*. *Insect Biochemistry and Molecular Biology* 35:249-256.

Barros EM, Torres JB, Ruberson JR, Oliveira MD (2010) Development of *Spodoptera frugiperda* on different hosts and damage to reproductive structures in cotton. *Entomologia Experimentalis. Applicata* 137: 237-245.

Benoit JB, Attardo GM, Michalkova V, Krause TB, Bohhva J, Zhang Q, Baumann AA, Mireji PO, Takae P, Denlinger DL, Ribeiro JM, Aksoy S (2014) A novel highly divergent protein family identified from a viviparous insect by RNA-seq analysis: A potential target for tsetse fly-specific abortifacients. *PLoS Genetics* 10: e1003874.

Bravo A, Likitvivanavong S, Gill SS, Soberon M (2011) *Bacillus thuringiensis*: A story of a successful bioinsecticide. *Insect Biochemistry and Molecular Biology*. 41: 423-431.

Carvalho RA, Omotoc, Field LM, Williamson MS, Bass C (2013) Investigating the Molecular Mechanisms of Organophosphate and Pyrethroid Resistance in the Fall Armyworm *Spodoptera frugiperda*. *PLoS ONE* 8: e62268.

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21: 3674-3676.

Dixon AFG (1998) *Aphid Ecology: An Optimization Approach*, second ed. Chapman and Hall, London.

Duan Y, Liu P, Li J, Li J, Chen P (2013) Immune gene discovery by expressed sequence tag (EST) analysis of hemocytes in the ridgetail white prawn *Exopalaemon carinicauda*. *Fish Shellfish Immunology* 34: 173-182.

Groot AT, Marr M, Scholf G, Lorenz S, Svatos A, Heckel DG (2008) Host strain specific sex pheromone variation in *Spodoptera frugiperda*. *Frontiers in Zoology* 5: 20.

Hernandez-Rodriguez CS, Ferre J, Herrero S (2009) Genomic structure and promoter analysis of pathogen induced REPAT genes from *Spodoptera exigua*. *Insect Molecular Biology* 18: 77-85.

Herrero S, Ansems M, Vanoers MM, Vlak JM, Bakker PL, and Demaagd RA (2007) REPAT, a new family of proteins induced by bacterial toxins and baculovirus infection in *Spodoptera exigua*. *Insect Biochemistry and Molecular Biology* 37: 1109-1118.

Jakka SR, Knight VR, Jurat-Fuentes JL (2014) Fitness costs associated with field-evolved resistance to Bt 28 V. Machado *et al.* maize in *Spodoptera frugiperda* (Lepidoptera: Noctuidae). *Journal of Economic Entomology* 107: 342-351.

Javed N, Viner R, Williamson MS, Field LM, Devonshire AL, Moores GD (2003) Characterization of acetylcholinesterases, and

- their genes, from the hemipteran species *Myzus persicae* (Sulzer), *Aphis gossypii* (Glover), *Bemisia tabaci* (Gennadius) and *Trialetrodes vaporariorum* (Westwood). *Insect Molecular Biology* 12:613-620.
- Kim J, Matsuba Y, Ning J, Schilimiller AL, Hamard, Jones AD, Pichersky E, Last RL (2014) Analysis of natural and induced variation in tomato glandular trichome flavonoids identifies a gene not present in the reference genome. *Plant Cell Online* 26: 3272-3285.
- Lenz TI, Eizaguirre C, Rotter B, Kable B, Milinski (2013) Exploring local immunological adaptation of two stickle back ecotypes by experimental infection and transcriptome-wide digital gene expression analysis. *Molecular Ecology* 22: 774-786.
- Li X, Qu MJ, Zhang Y, Li JW, Liu TX (2018) Expression of Neuropeptide F Gene and Its Regulation of Feeding Behavior in the Pea Aphid, *Acyrtosiphon pisum*. *Frontiers in Physiology* 9: 87.
- Machado JP, Johnson WE, Gilbert MTP, Zhang G, Jarvis ED, O'Brien SJ, Antunes A (2016) Bone-associated gene evolution and the origin of flight in birds. *BMC Genomics* 2016:17:371. DOI: <https://doi.org/10.1186/s12864-016-2681-7>
- Machado V, Wunder M, Baldissera VD, Oliveira JV, Fiuza LM, Nagoshi RN (2008) Molecular characterization of host strains of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) in Southern Brazil. *Annals of the Entomological Society of America* 101: 619-626.
- Nagasaki H, Mochizuki T, Kodama Y, Saruhashi S, Morizaki S, Sugawara H, Ohyanagi H, Kurata N, Okubo K, Takagi T, Kaminuma E, Nakamura Y (2013) DDBJ read annotation pipeline: a cloud computing based pipeline for high-throughput analysis of next generation sequencing data. *DNA Research* 20: 383-390.
- Nagoshi RN, Silvie P, Meagher RL, Lopez J, Machado V (2007) Identification and comparison of fall armyworm (Lepidoptera: Noctuidae) host strains in Brazil, Texas, and Florida. *Annals of the Entomological Society of America* 100: 394-402.
- Nanoth-Vellichirammal N, Zera AJ, Schilderr RJ Wehrkamp C, Riethoven JM, Brisson JA (2014) De novo transcriptome assembly from fat body and flight muscles transcripts to identify morph-specific gene expression profiles in *Gryllus firmus*. *PLoS ONE* 9: e82129.
- Navarro-Cerrillo G, Ferre J, Demaagd RA, Herrero S (2012) Functional interactions between members of the REPAT family of insect pathogen-induced proteins. *Insect Molecular Biology* 21: 335-342
- Navarro-Cerrillo G, Hernandez-Martinez P, Vogel H, Ferre J, Herrero S (2013) A new gene superfamily of pathogen-response (repat) genes in Lepidoptera: Classification and expression analysis. *Comparative Biochemistry and Physiology. Part-B: Biochemistry and Molecular Biology* 164: 10-17.
- Patel V, Chu CT (2011) Nuclear transport, oxidative stress, and neurodegeneration. *International Journal of Clinical and Experimental Pathology* 4 : 215-29.
- Robert JA, Pit C, Bonnett R, Yuen M, Keeling I, Bohlmann J Huber P (2013) Disentangling detoxification: gene expression analysis of feeding mountain pine beetle illuminates molecular-level host chemical defense detoxification mechanisms. *PLoS One* 8: e77777.
- Romualdi C, Bortoluzzi S, Danieli GA (2001) Detecting differentially expressed genes in multiple tag sampling experiments: comparative evaluation of statistical tests. *Human Molecular Genetics* 10:2133-2141.
- Sabater-Munoz B, Legeai F, Rispe C, Bonhomme J, Dearden P, Dossat C, Duclert A, Gauthier J-P, Ducray DG, Hunter W, Dang P, Kambhampati S, Martinez-Torres D, Cortes T, Moya A, Nakabachi A, Philippe C, Prunier-Leterme N, Rahbe Y, Simon J-C, Stern DL, Winker P, Tagu D (2006) Large-scale gene discovery in the pea aphid *Acyrtosiphon pisum* (Hemiptera). *Genome Biology* 7: R21.
- Schafer FQ, Buettner GR (2001) Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple". *Free Radical Biology and Medicine* 30 : 1191-212.
- Sinha JK, Ghosh S, Swain U, Girdharan NV, Raghunath M (2014) Increased macromolecular damage due to oxidative stress in the neocortex and hippocampus of WNIN/Ob, a novel rat model of premature aging. *Neuroscience* 269: 256-64.
- Trotta V, Forlano P, Falabella P, Battaglia D, Fanti P (2018) The aphid *Acyrtosiphon pisum* exhibits a greater survival after a heat shock when parasitized by the wasp *Aphidius ervi*. *Journal of Thermal Biology* 72: 53-58.