6

Summary and Conclusion

Sessile organisms like plants are often found to become victim of a huge number of abiotic as well as biotic stress conditions that includes algal, fungal and bacterial infections, insect attacks. Out of all the insects that target tea plant as their host, *Helopeltis theivora* or tea mosquito bug (TMB) establishes itself to be the major and most destructive insect pest of tea. TMB feeding results in more or less complete loss of yield of the crop. The infested shoots become unsuitable for plucking and further processing. As such, use of certain pesticides and chemical formulations have resulted in partial control of TMB but complete eradication of this pest has been impossible to attain till date as because this pest feeds on a wide range of alternate hosts. During unfavorable conditions, TMB maintains its life cycle on its alternate hosts until the tea plant is available for feeding. TMB has become the most dangerous enemy of the tea growers all over the world. In this study we have endeavored to study certain biochemical and enzymatic changes that are a result of TMB feeding. Our objective was also to perform a detailed RNA-seq analysis including transcriptome, lncRNA, cirRNA and ceRNA network analysis.

A decrease in total phenolic content and total flavonoid content in all eleven clones of the plant as a result of TMB attack on these plants was reported. This decrease in total phenolic content, is, however, not significant. It is noteworthy to mention that the Darjeeling clones (P312 and AV2) showed higher phenolic and flavonoid content than rest of the Assam clones. Essay of enzymatic antioxidants revealed that the infested plants exhibited significantly higher POX, APX, PAL, PPO and CAT activities.

This study uncovered the transcriptomic and ncRNA-mediated response of the tea plant to TMB feeding. We obtained an average of 39902058 number of clean reads in the control samples and 45566217 reads in the TMB-infested samples through the high-throughput sequencing. For the RNA-seq analysis, an average of 14.7 GB of clean data was obtained and in total, we obtained 154463 transcripts in the final transcriptome of our study. Approximately 80% alignment rate was achieved with the reference genome of *C. sinensis*. In total, 9502 transcripts were identified as lncRNAs (average length 432 nucleotides), and among them, majority (70.5%) was reported to derive from intergenic/unknown regions of the genome. A very less percentage of the identified lncRNAs (0.86%) was found to be evolutionarily conserved among lncRNAs of other plant species reported and deposited in public databases. Most BLAST hits corresponded to the lncRNAs from *Triticum aestivum.* LncRNA expression analysis revealed the selective expression of 1656 lncRNAs in healthy tissues and 1142 lncRNAs in TMB infested tissues. Differential expression of several protein coding genes and lncRNAs was documented in tea plant in response to TMB infestation. Out of 80 DELs, 46 were down-regulated and 34 were up-regulated in expression. Differential expression pattern of five selected lncRNAs was validated by qRT-PCR experiment which correlated with the expression pattern of lncRNAs analysed through RNA-seq analysis. The *cis*-acting method of gene regulation by lncRNAs was analysed and 27 genes were found to be *cis*-targets of the DELs and in total, 5804 genes were found to be targets of lncRNAs. The lncRNA-target genes were enriched in 378 GO terms including 310 under biological process, 30 under cellular component and 38 under molecular function. Target genes were suppressed in GO terms related to cell cycle, nuclear division, organelle fission while they were upregulated in GO terms associated

with “cellular response to phosphate starvation”, “transferase, oxidoreductase and dioxygenase activity”, “response to chemical/oxygen-containing compound”. Through GSEA of KEGG pathway analysis we found that lncRNA-target genes were significantly enriched in 20 KEGG pathways. Pathways like “biosynthesis of N-glycans”, “endocytosis”, “amino sugar and nucleotide sugar metabolism” were suppressed while pathways of “photosynthesis”, “biosynthesis of secondary metabolites”, “terpenoids biosynthesis”, metabolic pathways of certain amino acids like “beta-alanine”, “tryptophan”, “cysteine”, “methionine”, “tyrosine” got activated.

The healthy/control plants showed the selective expression of 2375 genes and the infested plants exhibited the selective expression of 1599 genes out of total 36644 expressed genes. Differential expression analysis was performed and 3665 genes were found out to be differentially expressed in infested condition. Transcript accumulation of 1767 genes and decrease in expression of 1898 genes was noted. Differential expression of six genes were validated using qRT-PCR. GSEA analysis indicated that the DEGs were associated with 771 GO terms (481 under biological process, 103 under cellular compartment and 187 under molecular function). Groups of DEGs were up-regulated in GO terms like “secondary metabolic process”, “response to external biotic stimulus”, “response to other organism”, “oxidoreductase activity”, “dioxygenase activity”. While groups of DEGs belonging to GO terms like “mitotic cell cycle process”, “nuclear division”, “microtubule motor activity and binding”, “organelle fission” got down-regulated during TMB feeding. In addition, significant set of genes were activated in 17 KEGG pathways including those for “terpenoids metabolism”, “cell signalling pathways (MAPK)”, “tryptophan metabolism”, “cyanoamino acid metabolism”, “zeatin biosynthesis”, “betalain biosynthesis”. Pathways involving DEGs

getting suppressed are “carotenoids and amino acids biosynthesis”, “base excision repair (BER)”, “amino and nucleotide sugar metabolism” etc. Length distribution and expression comparison between lncRNAs and mRNAs revealed consistent results with previous literature. LncRNAs were comparatively shorter in length and lower in terms of their FPKM values than mRNAs. GSEA analysis of DEGs and DELs suggested the involvement of these RNA molecules in regulating pathways associated with primary and secondary metabolism in tea plant.

Through this high-throughput data generated, a total of 709 circRNAs were identified. A greater percentage of circRNAs (62%) were derived from the intergenic regions. 285 circRNAs were only reported to express in the healthy/control plants, while 209 were found to be specific to treated samples. 60% of identified circRNAs of this study were found to be homologous with circRNAs deposited in public databases. Differential expression pattern of cirRNAs showed that out of 80 DECs, 19 got down-regulated and 15 got up-regulated in expression. Sequence complementarity analysis and miRNA-circRNA hybrid forming ability assessment analysis revealed that 17 DECs could be potential targets of 55 *C. sinensis* miRNAs. Eleven genes were annotated as parental genes of the DECs. Eight of these circRNAs were predicted to be derived from introns of the parental genes and others were predicted to be derived from exons. Functional annotation of DEC-parental genes and DEC-correlated genes showed the annotation of these genes under 407 GO terms, of which 322 were under “Biological Processes”, 37 under “Cellular Component” and 48 GO terms were found to be under “Molecular function”. The DECs are predicted to regulate genes belonging to various important biological functional pathways like “diterpenoid biosynthesis”, “nucleocytoplasmic transport”, “tryptophan metabolism”, “ascorbate and aldarate

metabolism", "biosynthesis of secondary metabolites" and "biosynthesis of cofactors", "N-glycan biosynthesis" and "porphyrin metabolism". In addition, the DECs could also function as eTMs for 9 genes such as aspartyl protease, lectin receptor, phospholipase etc.

NcRNAs like lncRNAs and circRNAs have the capacity to act as miRNA sponges and can sequester miRNA active level. They compete with each other to bind with miRNAs possessing complementary binding sites thus making them unavailable for binding with target mRNAs. In this study, TMB-responsive competing endogenous network (ceRNA network) was constructed. 28 DELs and 17 DECs were found to be probable targets of *C. sinensis* miRNAs. Notably, conserved miRNAs that are well-known regulators of plant-insect interaction like miR156, miR171 and miR395 were predicted to target 6 lncRNAs of this study. Additionally, 7 lncRNAs were found to act as eTMs. The final ceRNA network consisted of 5 lncRNAs, 17 circRNAs, 33 mRNAs and 37 miRNAs.

This study on TMB-elicited lncRNAs might serve as useful genomic resource for further study relating to TMB infestation and will also strengthen the scientific knowledge regarding *H. theivora* infestation on tea plant. The involvement of conserved miRNAs viz. miR156, miR171, miR395 in plant's response to phloem feeders and chewing insects is well-studied and reported in literature. The present study indicated specific lncRNAs that could be potential targets for these significant miRNAs. TCONS_00083891, TCONS_00053665 and TCONS_00032903 were predicted to act as targets for miR156, which is also known to target LRR-RLKs. Additionally, TCONS_00099260 was seen to get targeted by 12 miRNAs including miR171 that also targets the GRAS family TFs. Therefore, developing mutants overexpressing or

knocking out the particular lncRNAs might be an initial start to develop TMB resilience in tea plant. Despite having extremely little evolutionary conservation between species, lncRNA-targets may nonetheless be conserved. Therefore, tailoring or increasing crop defense responses against specific herbivores may be accomplished by ectopically expressing specific lncRNAs in other plant species. LncRNAs exhibit low expression and therefore future research should concentrate on the discovery of cell type specific lncRNAs. Future research can also be focused on in-depth transcriptome analysis of a single cell type or tissue type to identify unique lncRNAs. Another fascinating area that requires more research is how lncRNAs interact with other molecular components in the cell. Although there are a number of *in-silico* studies related to plant lncRNA research, yet reports on functional characterization of plant lncRNAs are scanty. However, new developments, such as large-scale mutant library screening using sophisticated imaging, should speed up the process of functional validation of lncRNAs in the near future. The results of this study also provided clues for involvement of circRNAs during insect feeding and tea plant defense response. This study also pointed out some specific circRNAs to act as potential eTMs of genes like aspartyl protease. Therefore, it would be interesting to see if transgenics exhibiting an elevated or demoted expression of these circRNAs would or wouldn't have a consequence on TMB resistance in the tea plant. The circular RNAs are persistent as they are not destroyed as effectively as linear RNAs due to the absence of 5′ and 3′ ends. Consequently, stable circRNAs may be employed as molecular markers. Certain studies revealed that circRNAs might be sometimes translated to proteins. Therefore, it may be worthwhile to conduct further research into the questions of how circRNAs are translated and what roles circRNA-coding proteins play in plants during insect infestation.