Discovery of Novel Virus Sequences in a Flea Transcriptome

Fleas (Siphonaptera) are important blood-feeding parasites of mammals (including humans). Although they are known to be associated with bacteria (e.g. the plague bacterium *Yersinia pestis*), little is known about their association with viruses. This project concerns a virus discovery study on the recently released transcriptome of the European hedgehog flea (*Archaeopsylla erinacea*). Bioinformatics aided comparisons of the flea transcriptome against a curated virus genome database revealed the presence of various unknown virus sequences, along with some known viral pathogens (e.g. *Vaccinia* virus). BLAST searches and functional analyses in GenBank identified several transcripts as an RNA-dependent RNA polymerase (RdRP) related to those of *Tenuivirus*, an unclassified genus of group V (-)ssRNA viral pathogens of plants. Phylogenetic analyses of the sequences, including *Bunyaviridae* as an outgroup, confirm clustering with Rice Stripe Virus (RSV) and Rice Grassy Stunt Virus (RGSV) (both in *Tenuivirus*). Targeted sequencing of an *Archaeopsylla erinacea* genome, and 35 related flea species suggests that the presence of this viral RdRP in the transcriptome could be due to an active viral infection of this flea, rather than an integration of viral sequence into the host flea genome as has been observed with endogenous viral elements. Further studies are needed to resolve these competing hypotheses of viral-flea interaction. Results suggest the presence of undiscovered viral diversity associated with fleas in particular, and blood-feeding insects in general. Additionally, transcriptomes are useful resources for viral discovery.

Key Words: fleas, transcriptome, Tenuivirus, RNA-Dependent RNA Polymerase