SERENDIPITOUS DISCOVERY OF AN RNA VIRUS FROM THE CRICKET, *ACHETA DOMESTICUS*

STEVEN M. VALLES[1] AND YANPING CHEN[2]
[1]Corresponding author. USDA-ARS, Center for Medical, Agricultural and Veterinary Entomology
1600 SW 23rd Drive, Gainesville, FL 32608

[2]USDA-ARS, Bee Research Laboratory, 10300 Baltimore Blvd., Beltsville, MD 20705

Bioinformatic analysis of approximately 2000 expressed sequence tags (ESTs) from a monogyne *Solenopsis invicta* expression library identified 8 clones exhibiting significant homology to single-stranded RNA viruses (Valles et al. 2004). Three of these clones (3F6, 14D5, and 24C10) yielded a contiguous sequence fragment from which 3' and 5' rapid amplification of cDNA ends (RACE) reactions were conducted and ultimately led to the discovery of a new RNA virus (*Solenopsis invicta* virus-1 [SINV-1]) infecting the red imported fire ant, *S. invicta* (Valles et al. 2004; Valles & Strong 2005). Subsequent examination of fire ant ESTs revealed that clone 11F1 exhibited significant homology with the capsid polyprotein of the *Drosophila* C RNA virus and was distinct from the SINV-1 genome sequence. Thus, we were hopeful that EST 11F1 might similarly lead to a new fire ant virus discovery. However, when we conducted extensive field surveys ($n > 50$) of *S. invicta* nests by RT-PCR using oligonucleotide primers designed to the sequence of clone 11F1, all samples proved negative. Up to that point, sampling was conducted as had been done for the SINV-1; a scintillation vial was plunged into an ant nest and adults that fell into the vial were collected and used for RNA extraction and RT-PCR analysis. Failure to detect the sequence prompted a change in the collection protocol. We considered that perhaps if the 11F1 sequence was from a virus in the fire ant, that it could be infecting a stage other than adults. To retrieve a sample of brood for this analysis required collection of entire colonies. Our first collection was taken on a Friday, and because the floating method (Jouvenaz et al. 1977) to remove the ants and brood from the collection bucket would require several days to complete, we postponed the extraction process until the following Monday. This delay necessitated feeding the colonies, standard practice for fire ant nest collections. Each colony was provided with water, 10% sucrose solution, and approximately 30 frozen-killed crickets, *Acheta domesticus*, purchased from a cricket bait farm (Jerry's Bait Farm, Slocomb, AL).

RT-PCR analysis was conducted later the following week with RNA extracted from adults and brood (separately) of these colonies; every single sample tested positive. Although not immediately, we realized that perhaps the crickets were the actual source of the suspected virus, and ants were consuming the infected crickets and consequently testing positive. One-step RT-PCR (Invitrogen, Carlsbad, CA) analysis (1 cycle at 45°C for 30 min, 1 cycle at 94°C for 2 min, 35 cycles of 94°C for 15 s, 55°C for 15 s, 68°C for 30 s) of RNA extracted directly from crickets ($n = 26$ from three different lots) with clone 11F1-specific oligonucleotide primers (primer 60: 5' CAGATGGGTGCGAATAACTTCAAATC, primer 61: 5' CACTTCGAAAAACAACTCAGTCTCCTG) produced an amplicon of anticipated size (311 bp) from 13 of the crickets (50% positive). Three of the amplicons were gel-purified and ligated into the pCR4-TOPO vector, transformed into TOP10 competent cells (Invitrogen) and sequenced by the Interdisciplinary Center for Biotechnology Research (University of Florida). The sequence was 99.7% identical (310/311) to the corresponding sequence of EST clone 11F1.

A 3'RACE reaction was conducted with RNA prepared from one of the crickets testing positive for the 11F1 sequence (GeneRacer kit, Invitrogen). cDNA was synthesized from 1 µg of total RNA with the GeneRacer Oligo dT primer (42°C for 50 min, 70°C for 15 min, held on ice for 2 min, RNase H added and incubated at 37°C for 20 min). The cDNA was subsequently used as template for PCR with a gene-specific primer, p393 (5'GCATCTACTGTACCCAATGTTCACCAGCGGTACAC) and GeneRacer 3' primer. The reaction produced a single amplicon of approximately 1300 bp that was gel purified, ligated, transformed, and sequenced as described above. Sequences were assembled with the NTI Vector software (Invitrogen) to produce a 1565 nucleotide contiguous fragment consistent with the corresponding regions of fire ant EST 11F1. The sequence was polyadenylated at the 3' end. The consensus assembly was deposited in the GenBank database under accession number DQ112164. Analysis of the sequence revealed a single large open reading frame (ORF) beginning at nucleotide 4 (start codon) and terminating at nucleotide 1439 (stop codon). The translated ORF yielded a predicted polypeptide comprised of 478 amino acids with a molecular mass of 52600 Da.

Protein-protein BLAST (Altschul et al. 1997) analysis showed that the predicted amino acid sequence exhibited significant similarity with the capsid polyproteins of positive-stranded RNA viruses in the GenBank database. Table 1 summarizes the level of sequence identity realized from the BLAST search. Specifically, the new sequence